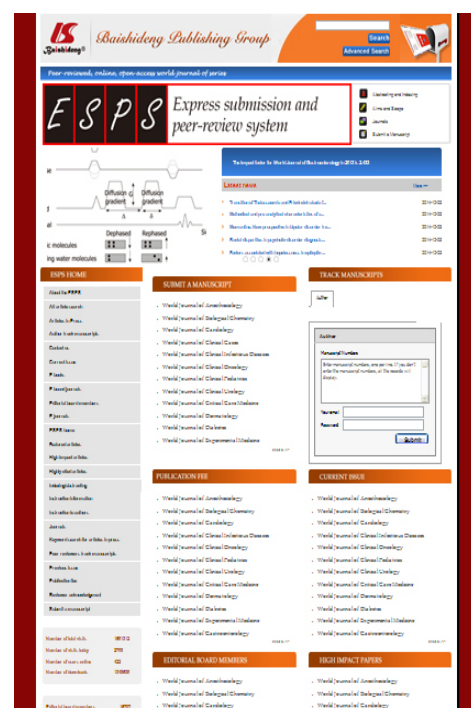
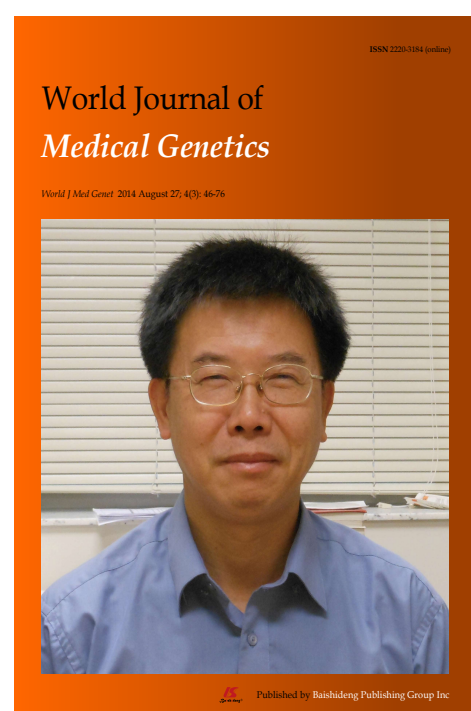
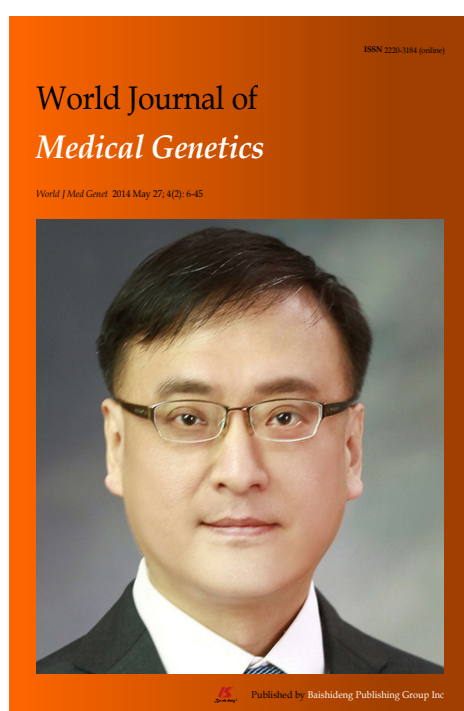


# World Journal of *Medical Genetics*

2014 Bound Volume 4 Issue 1-4: 1-109





## Editorial Board

2011-2015

The *World Journal of Medical Genetics* Editorial Board consists of 345 members, representing a team of worldwide experts in medical genetics. They are from 51 countries, including Argentina (1), Armenia (1), Australia (5), Austria (2), Belgium (2), Brazil (5), Bulgaria (2), Canada (6), Chile (1), China (41), Croatia (1), Czech Republic (1), Denmark (3), Ecuador (1), Egypt (2), France (10), Germany (4), Greece (2), Hungary (1), India (12), Iran (7), Ireland (1), Israel (2), Italy (22), Japan (24), Latvia (1), Malaysia (2), Mexico (2), Netherlands (6), New Zealand (2), Norway (2), Pakistan (1), Peru (1), Poland (1), Portugal (1), Russia (2), Saudi Arabia (3), Serbia (1), Singapore (3), Slovenia (2), South Korea (7), Spain (17), Sweden (2), Switzerland (2), Thailand (3), Turkey (8), Ukraine (1), United Arab Emirates (1), United Kingdom (11), United States (103), and Uruguay (1).

### EDITOR-IN-CHIEF

Hans van Bokhoven, *Nijmegen*

### GUEST EDITORIAL BOARD MEMBERS

Chia-Hsiang Chen, *Zhunan*  
Ji-Yih Chen, *Taoyuan*  
Bor-Luen Chiang, *Taipei*  
Yu-Chen Hu, *Hsinchu*  
Hai-Gwo Hwu, *Taipei*  
Suh-Hang H Juo, *Kaohsiung*  
Hsien-Hsiung Lee, *Taichung*  
Yueh-Lun Lee, *Taipei*  
Chin-San Liu, *Changhua*  
Pei-Jung Lu, *Tainan*  
Chien-Kuo Tai, *Chiayi*  
Po-Nien Tsao, *Taipei*  
Wendy W H Weng, *Taipei*  
Jeng-Hsien Yen, *Kaohsiung*

### MEMBERS OF THE EDITORIAL BOARD



**Argentina**

Jose Carlos Pirola, *Buenos Aires*



**Armenia**

Hasmik Mkrtchyan, *Yerevan*



**Australia**

Simon Easteal, *Canberra*  
Manuel B Graeber, *Camperdown*

Jeremy Jowett, *Melbourne*  
Karen Mather, *Randwick*  
Grant Morahan, *Perth*



**Austria**

Harald Aschauer, *Vienna*  
Anton Wutz, *Vienna*



**Belgium**

Teresinha Leal, *Brussels*  
Dirk van West, *Antwerp*



**Brazil**

Jeane E Lvisentainer, *São Paulo*  
Sharbel Weidner Maluf, *Porto Alegre*  
Milton O Moraes, *Rio de Janeiro*  
Fernando R Vargas, *Rio de Janeiro*  
Isabela Jube Wastowski, *Goiás*



**Bulgaria**

Spaska Stanilova, *Stara Zagora*  
Albena Todorova, *Sofia*



**Canada**

Regen Drouin, *Sherbrooke*  
Gonzalo Hortelano, *Ontario*  
William Jia, *Vancouver*

Sean Li, *Ottawa*  
Shiva M Singh, *Ontario*  
Li-Ting Song, *Toronto*



**Chile**

Jara S Lilian, *Santiago*



**China**

George G Chen, *Hong Kong*  
Volodymyr Dvornyk, *Hong Kong*  
Ning-Han Feng, *Nanjing*  
Shi-Ang Huang, *Wuhan*  
Wei Huang, *Shanghai*  
Chun-Yan Ji, *Jinan*  
Hong-Chuan Jin, *Hangzhou*  
Ke-Shen Li, *Zhanjiang*  
Zhuo-Zhuang Lu, *Beijing*  
Xiao-Ping Luo, *Wuhan*  
Wen-Bin Qian, *Hangzhou*  
Rong Rui, *Nanjing*  
Lun-Quan Sun, *Changsha*  
Qian Tao, *Hong Kong*  
George Sai-Wah Tsao, *Hong Kong*  
Ming-Rong Wang, *Beijing*  
Shi-Jie Wang, *Shijiazhuang*  
Wei Wang, *Beijing*  
Ji Wu, *Shanghai*  
Hua-Xi Xu, *Zhenjiang*  
Rui-An Xu, *Xiamen*  
Wei-Hua Yan, *Linhai*  
Rong-Cun Yang, *Tianjin*  
Bang-Ce Ye, *Shanghai*  
Rui-Xing Yin, *Nanning*  
Yi Zeng, *Beijing*

Xiao-Feng Zhu, *Guangzhou*



**Croatia**

Gordan Lauc, *Zagreb*



**Czech Republic**

Marie Lipoldova, *Prague*



**Denmark**

Vibeke Andersen, *Aabenraa*

Thomas V O Hansen, *Copenhagen*

Goutam Sahana, *Tjele*



**Ecuador**

Cesar Paz-Y-Miño, *Quito*



**Egypt**

Menha M A Swellam, *Giza*

Samia A Temtamy, *Cairo*



**France**

Christophe Chevillard, *Marseille*

Johanna Chluba, *Dijon*

Zdenko Herceg, *Lyon*

Enzo Lalli, *Valbonne*

James Lespinasse, *Chambéry*

Bernard S Lopez, *Fontenay aux Roses*

Eric Pasmant, *Clichy*

Kay-D Wagner, *Nice*

Qing Wang, *Lyon*

Georges Vassaux, *Nice Cedex*



**Germany**

Stefan Böhringer, *Essen*

Anibh Martin Das, *Hannover*

Gudrun A Rappold, *Heidelberg*

Hermona Soreq, *Jerusalem*



**Greece**

Roxani Angelopoulou, *Athens*

Voula Velissariou, *Athens*



**Hungary**

Bela Meleg, *Pecs*



**India**

Arvind Kumar Arya, *Meerut*

Prakash Sadashiv Gambhir, *Pune*

Katta MK Girisha, *Manipal*

Narendra Joshi, *Navi Mumbai*

Chinmay KumarPanda, *Kolkata*

Rajni Rani, *New Delhi*

Susanta Roychoudhury, *Kolkata*

Sudeep R Shah, *Mumbai*

Ranbir Chander Sobti, *Chandigarh*

Mandava Swarna, *Mumbai*

Meera Vaswani, *New Delhi*

Suresh P Vyas, *Sagar*



**Iran**

Yahya Daneshbod, *Shiraz*

Fariborz Ghaffarpasand, *Shiraz*

Mohammad H Modarressi, *Tehran*

Shirin Hasani Ranjbar, *Tehran*

Nader Tajik, *Tehran*

DM Kordi Tamandani, *Zahedan*

Alireza Zomorodipour, *Tehra*



**Ireland**

Ross McManus, *Dublin*



**Israel**

Aliza Shlomit Amiel, *Kfar-Saba*

Rachel Ben-Shlomo, *Tivon*



**Italy**

Francesco Acquati, *Varese*

Nenad Bukvic, *Foggia*

Gabriele Candiani, *Milan*

Antonio Cao, *Cagliari*

Marco Castori, *Rome*

Massimo Collino, *Torino*

Teresa Esposito, *Naples*

Antonia Follenzi, *Novara*

Tommasina Guglielmelli, *Turin*

Lidia Larizza, *Milan*

Fortunato Lonardo, *Benevento*

Marco Lucarelli, *Rome*

Laura Mandelli, *Bologna*

Antonio Musio, *Pisa*

Pier Giuseppe Pelicci, *Milan*

Daniela Perotti, *Milano*

Antonio M Persico, *Rome*

Enrico Pola, *Rome*

Giovanni Romeo, *Bologna*

Emanuela Signori, *Rome*

Alessio Squassina, *Cagliari*

Anna Stornaiuolo, *Milano*



**Japan**

Yutaka Hata, *Tokyo*

Mizukami Hiroaki, *Shimotsuke*

Johji Inazawa, *Tokyo*

Yasutomi Kamei, *Tokyo*

Tetsufumi Kanazawa, *Osaka*

Hiroyuki Kanzaki, *Miyagi*

Koichi Kawakami, *Shizuoka*

Akinori Kimura, *Tokyo*

Hiroshi Kunugi, *Tokyo*

Alexander Lezhava, *Yokohama*

Hisao Masai, *Tokyo*

Ryuichi Morishita, *Osaka*

Hitoshi Osaka, *Yokohama*

Masa-Aki Shibata, *Osaka*

Kazumi Suzukawa, *Tsukuba*

Masatoshi Tagawa, *Chiba*

Shin'ichi Takeda, *Tokyo*

Keizo Takenaga, *Shimane*

Hiroshi Tanooka, *Tokyo*

Takafumi Uchida, *Sendai*

Sho-ichi Yamagishi, *Kurume*

Toshiyuki Yamamoto, *Tokyo*

Norio Yasui-Furukori, *Hirosaki*

Kiyotsugu Yoshida, *Tokyo*



**Latvia**

Laima Tihomirova, *Riga*



**Malaysia**

Ravindran Ankathil, *Kelantan*

Hua Siew Gan, *Kelantan*



**Mexico**

Jose FMunoz-Valle, *Zapopan*

Gilberto Vargas-Alarcon, *Tlalpan*



**Netherlands**

Annemieke Aartsma-Rus, *Leiden*

Raoul CM Hennekam, *Amsterdam*

Anke H Maitland-van der Zee, *Utrecht*

Paul J Van Diest, *Utrecht*

Mannis van Oven, *Leiden*



**New Zealand**

Ryuji Fukuzawa, *Dunedin*

Christine M Morris, *Christchurch*



**Norway**

Kristian Tambs, *Oslo*

Martin Tesli, *Oslo*



**Pakistan**

Muhammad Naeem, *Islamabad*



**Peru**

Gustavo F Gonzales, *Lima*



**Poland**

Piotr J Wysocki, *Poznan*

**Portugal**

Manuel R Teixeira, *Porto*

**Russia**

Anton V Kiselev, *Saint-Petersburg*  
Yurov Yuri B, *Moscow*

**Saudi Arabia**

Khaled K Abu-Amro, *Riyadh*  
Khawla S Al-Kuraya, *Riyadh*  
Wael Mohamed Swelam, *Medina*

**Serbia**

Lidija Radenovic, *Belgrade*

**Singapore**

N Gopalakrishna Iyer, *Singapore*  
Peter E Lobie, *Singapore*  
Eng-King Tan, *Singapore*

**Slovenia**

Borut Peterlin, *Ljubljana*  
Uros Potocnik, *Maribor*

**South Korea**

Byung-Hoon Jeong, *Anyang*  
Jung Jin Kim, *Seoul*  
Yonggoo Kim, *Seoul*  
Taeg Kyu Kwon, *Taegu*  
Seong-Wook Lee, *Yongin*  
Taesung Park, *Seoul*  
Han-Wook Yoo, *Seoul*

**Spain**

Salvador F Aliño, *Valencia*  
David Araújo-Vilar, *Santiago De Compostela*  
Victor Asensi, *Oviedo*  
Ignacio Blanco-Guillermo, *Barcelona*  
Judith Sanz i Buxo, *Catalonia*  
Juan I Perez Calvo, *Zaragoza*  
Javier S Castresana, *Pamplona*  
Cristina Fillat, *Barcelona*  
Karen Heath, *Madrid*  
Adrián Llerena, *Badajoz*  
Jose F Marti Masso, *San Sebastian*  
Jose M Millan, *Valencia*  
Juan Pie, *Zaragoza*  
Jesus Prieto, *Pamplona*  
Juan A Rey, *Madrid*  
Miguel Chillon Rodriguez, *Bellaterra*  
Maria E Saez, *Seville*

**Sweden**

Karin Klinga-Levan, *Skovde*  
Mef Christina Nilbert, *Lund*

**Switzerland**

Angela Ciuffi, *Lausanne*  
Daniel Schorderet, *Lausanne*

**Thailand**

Vorasuk Shotelersuk, *Bangkok*  
Tewin Tencomnao, *Bangkok*  
Viroj Wiwanitkit, *Bangkok*

**Turkey**

Mehmet Necdet Akkus, *Mersin*  
Julide Altinisik, *Balikesir*  
Emrah Caylak, *Cankiri*  
Merih Cetinkaya, *Bursa*  
Ali Karaman, *Erzurum*  
Mustafa Sahin, *Ankara*  
Ahter Dilsad Sanlioglu, *Antalya*  
Salih Sanlioglu, *Antalya*

**Ukraine**

Ludmila Livshits, *Kyiv*

**United Arab Emirates**

Bassam R Ali, *Al-Ain*

**United Kingdom**

Kristin Becker, *Cardiff*  
Constantinos Demonacos, *Manchester*  
David Jeremy Galton, *London*  
Elaine Green, *Cardiff*  
Wen G Jiang, *Cardiff*  
Patrick J Morrison, *Belfast*  
Elisabeth Nacheva, *London*  
Mary Callaghan Nicol, *Cardiff*  
James S Owen, *London*  
David J Timson, *Belfast*  
Eugene M Tulchinsky, *Leicester*

**United States**

Praveen R Arany, *Cambridge*  
Nedal Arar, *San Antonio*  
Richard G Boles, *Los Angeles*  
Merlin G Butler, *Kansas*  
Hai-Feng Chen, *Hayward*  
J Don Chen, *Piscataway*  
James Chen, *Jefferson*  
Xiang-Ning Chen, *Richmond*  
Paola Costa-Mallen, *Kenmore*  
Qi Dai, *Nashville*

Shuo Dong, *Houston*  
Yao-Shan Fan, *Miami*  
Bing-Liang Fang, *Houston*  
Peter J Francis, *Portland*  
Xiao-Yi Gao, *Los Angeles*  
Yu-Bin Ge, *Detroit*  
Antonio Giordano, *Philadelphia*  
Thomas J Giordano, *Ann Arbor*  
Stephen J Glatt, *Syracuse*  
WT Godbey, *New Orleans*  
Dennis R Grayson, *Chicago*  
Dong-Sheng Gu, *Indianapolis*  
Zong-Sheng Guo, *Pittsburgh*  
Hakon Hakonarson, *Philadelphia*  
Wayne W Hancock, *Philadelphia*  
David W Hein, *Louisville*  
Hui-Xiao Hong, *Jefferson*  
Ji-Fan Hu, *Palo Alto*  
Shi-Le Huang, *Shreveport*  
Ying Huang, *Syracuse*  
Johnny Huard, *Pittsburgh*  
Barbara H Iglewski, *Rochester*  
Yurij Ionov, *Buffalo*  
Jing-Fang Ju, *Stony Brook*  
Berit Kerner, *Los Angeles*  
Alisa E Koch, *Ann Arbor*  
Paul C Kuo, *Maywood*  
Robert R Langley, *Houston*  
Eduardo C Lau, *Milwaukee*  
K-H William Lau, *Loma Linda*  
Mong-Hong Lee, *Houston*  
Dawei Li, *New Haven*  
Feng-Zhi Li, *Buffalo*  
Shi-Bo Li, *Oklahoma*  
Ming-Fong Lin, *Omaha*  
Steven R Lindheim, *Cincinnati*  
Feng Liu, *Chapel Hill*  
Xiao-Qi Liu, *West Lafayette*  
Yao-Zhong Liu, *New Orleans*  
Yong-Jun Liu, *New Orleans*  
Bo Lu, *Nashville*  
Qun Lu, *Greenville*  
Xing-Guang Luo, *West Haven*  
James L Manley, *New York*  
Viraj Master, *Atlanta*  
Richard C McEachin, *Ann Arbor*  
Jian-Feng Meng, *Kansas*  
Duane A Mitchell, *Durham*  
Yin-Yuan Mo, *Springfield*  
Viktor Morozov, *Bethesda*  
Srinivas Mummidi, *San Antonio*  
Mayumi Naramura, *Omaha*  
Swapan K Nath, *Oklahoma*  
Muthu Periasamy, *Columbus*  
Gwendolyn P Quinn, *Tampa*  
Rajalingam Raja, *Los Angeles*  
Rajagopal Ramesh, *Houston*  
Jasti S Rao, *Peoria*  
Charles J Rosser, *Orlando*  
Mark A Rothstein, *Louisville*  
Ananda L Roy, *Boston*  
Dharambir K Sanghera, *Oklahoma*  
Thomas Scholl, *Westborough*  
Rong Shao, *Springfield*  
Yuenian Eric Shi, *Roslyn Heights*  
Shree Ram Singh, *Frederick*  
Si-Hong Song, *Seminole*  
Constantine A Stratakis, *Bethesda*  
Manjunath N Swamy, *El Paso*  
Ming Tan, *Mobile*  
Bakhos A Tannous, *Charlestown*



Flora Tassone, *Sacramento*  
Mustafa Tekin, *Miami*  
Brad Therrell, *Austin*  
Barry Trink, *Baltimore*  
Tibor Valyi-Nagy, *Chicago*  
Andre van Wijnen, *Worcester*  
Chia-Yeng Wang, *Chicago*  
Jean Yin Jen Wang, *La Jolla*  
Pin Wang, *Los Angeles*

Li Wei, *Cleveland*  
Qing-Yi Wei, *Houston*  
Qi-Ze Wei, *Manhattan*  
David A Weinstein, *Gainesville*  
Mary Wilson, *Iowa*  
Xiaoling Xuei, *Indianapolis*  
Edward Kuo-Liang Yang, *Chandler*  
Guang Yang, *Augusta*  
Zeng-Quan Yang, *Detroit*

Lin Yao, *Augusta*  
Si-Ming Zhang, *Albuquerque*  
Xiaoliu Shaun Zhang, *Houston*  
Tong Zhu, *Durham*



**Uruguay**

Jose Luis Badano, *Montevideo*

# World Journal of *Medical Genetics*

*World J Med Genet* 2014 February 27; 4(1): 1-5





**EDITORIAL**

- 1 Genetic counseling in post-genomic era: Don't pretend to know the meaning of a gene mutation if you don't know  
*Song L*

## Contents

*World Journal of Medical Genetics*  
Volume 4 Number 1 February 27, 2014

**APPENDIX** I-V Instructions to authors

**ABOUT COVER** Editorial Board Member of *World Journal of Medical Genetics*, Liting Song, MD, MSc, Scientist, Hope Biomedical Research, 809-50 Ruddington Drive, Toronto, ON M2K 2J8, Canada

**AIM AND SCOPE** *World Journal of Medical Genetics* (*World J Med Genet*, *WJMG*, online ISSN 2220-3184, DOI: 10.5496) is a peer-reviewed open access academic journal that aims to guide clinical practice and improve diagnostic and therapeutic skills of clinicians.

*WJMG* covers topics concerning genes and the pathology of human disease, molecular analysis of simple and complex genetic traits, cancer genetics, epigenetics, gene therapy, developmental genetics, regulation of gene expression, strategies and technologies for extracting function from genomic data, pharmacological genomics, genome evolution. The current columns of *WJMG* include editorial, frontier, diagnostic advances, therapeutics advances, field of vision, mini-reviews, review, topic highlight, medical ethics, original articles, case report, clinical case conference (Clinicopathological conference), and autobiography.

We encourage authors to submit their manuscripts to *WJMG*. We will give priority to manuscripts that are supported by major national and international foundations and those that are of great basic and clinical significance.

**INDEXING/ABSTRACTING** *World Journal of Medical Genetics* is now indexed in Digital Object Identifier.

**FLYLEAF** I-IV Editorial Board

## EDITORS FOR THIS ISSUE

Responsible Assistant Editor: *Xin-Xin Che*  
Responsible Electronic Editor: *Su-Qing Liu*  
Proofing Editor-in-Chief: *Lian-Sheng Ma*

Responsible Science Editor: *Yuan Qi*

**NAME OF JOURNAL**  
*World Journal of Medical Genetics*

**ISSN**  
ISSN 2220-3184 (online)

**LAUNCH DATE**  
December 27, 2011

**FREQUENCY**  
Quarterly

**EDITOR-IN-CHIEF**  
**Hans van Bokhoven, Professor, PhD**, Department of Human Genetics and Cognitive Neurosciences, Radboud university Nijmegen Medical centre, PO Box 9101, 6500 HB Nijmegen, The Netherlands

**EDITORIAL OFFICE**  
Jin-Lei Wang, Director  
Xiu-Xia Song, Vice Director

*World Journal of Medical Genetics*  
Room 903, Building D, Ocean International Center,  
No. 62 Dongsihuan Zhonglu, Chaoyang District,  
Beijing 100025, China  
Telephone: +86-10-85381891  
Fax: +86-10-85381893  
E-mail: [bpgoffice@wjgnet.com](mailto:bpgoffice@wjgnet.com)  
<http://www.wjgnet.com>

**PUBLISHER**  
Baishideng Publishing Group Co., Limited  
Flat C, 23/F, Lucky Plaza,  
315-321 Lockhart Road, Wan Chai,  
Hong Kong, China  
Fax: +852-6555-7188  
Telephone: +852-3177-9906  
E-mail: [bpgoffice@wjgnet.com](mailto:bpgoffice@wjgnet.com)  
<http://www.wjgnet.com>

**PUBLICATION DATE**  
February 27, 2014

## COPYRIGHT

© 2014 Baishideng. Articles published by this Open Access journal are distributed under the terms of the Creative Commons Attribution Non-commercial License, which permits use, distribution, and reproduction in any medium, provided the original work is properly cited, the use is non commercial and is otherwise in compliance with the license.

## SPECIAL STATEMENT

All articles published in this journal represent the viewpoints of the authors except where indicated otherwise.

## INSTRUCTIONS TO AUTHORS

Full instructions are available online at [http://www.wjgnet.com/2220-3184/g\\_info\\_20100722180909.htm](http://www.wjgnet.com/2220-3184/g_info_20100722180909.htm)

## ONLINE SUBMISSION

<http://www.wjgnet.com/esp/>



## Genetic counseling in post-genomic era: Don't pretend to know the meaning of a gene mutation if you don't know

Liting Song

Liting Song, Hope Biomedical Research, Toronto, ON M2K 2J8, Canada

Author contributions: Song L solely contributed to this paper.

Correspondence to: Liting Song, MD, MSc, Scientist, Hope Biomedical Research, 809-50 Ruddington Drive, Toronto, ON M2K 2J8, Canada. [ltsong@yahoo.com](mailto:ltsong@yahoo.com)

Telephone: +1-416-7331573 Fax: +1-416-7331573

Received: June 13, 2013 Revised: August 23, 2013

Accepted: December 12, 2013

Published online: February 27, 2014

### Abstract

In this post-genomic era, more and more susceptibility loci of many possible genetic diseases are published. As our knowledge about these susceptibility loci is limited and partial, we should be very careful and responsible when patients seek genetic counseling about these possible genetic diseases. We should apply Confucius's principle about knowledge and information to genetic counseling, and tell the truth to our patients about what we know and what we do not know. Like many other cancers, breast cancer is a very complicated, multifactorial disease; genetic factors, lifestyles and eating habits, environmental factors, and viral infections might be involved in breast cancer; hence, it is difficult to figure out the real etiology of breast cancer. It is not crystal clear that a person who carries mutations of the breast cancer 1, early onset and/or breast cancer 2, early onset genes would eventually get breast cancer in her/his lifetime. No person should undergo a preventive double mastectomy, unless we know the etiology of breast cancer someday.

© 2014 Baishideng Publishing Group Co., Limited. All rights reserved.

**Key words:** Genetic counseling; Genetic disease; Susceptibility loci; Breast cancer 1, early onset and breast cancer 2, early onset genes; Preventive double mastectomy

**Core tip:** Many susceptibility loci of possible genetic diseases are published. As our knowledge about these susceptibility loci is limited and partial, we should be very careful and responsible when patients seek genetic counseling about these possible genetic diseases. Currently, I have not seen any solid evidence in support of the linkage between breast cancer and breast cancer 1, early onset (*BRCA1*) or/and breast cancer 2, early onset (*BRCA2*); and it is not crystal clear that a person who carries mutations of the *BRCA1* and/or *BRCA2* genes would eventually get breast cancer in her/his lifetime. No person should undergo a preventive double mastectomy, unless we know the etiology of breast cancer someday.

Song L. Genetic counseling in post-genomic era: Don't pretend to know the meaning of a gene mutation if you don't know. *World J Med Genet* 2014; 4(1): 1-5 Available from: URL: <http://www.wjgnet.com/2220-3184/full/v4/i1/1.htm> DOI: <http://dx.doi.org/10.5496/wjmg.v4.i1.1>

### INTRODUCTION

You either know it or you don't, don't pretend to know it if you don't, this is the proper attitude toward knowledge-Confucius (551–479 BC).

Nenad *et al*<sup>[1]</sup> raised concerns about genetic counseling in post-genomic era. Nowadays, it seems we have known so many susceptibility loci of many possible genetic diseases already, and more and more such susceptibility loci will be published in the future. More than 2000 years ago Confucius told us that we should handle information and knowledge very carefully and responsibly, the same attitude and spirit should be applied to genetic counseling in post-genomic era.

Firstly, we should indicate and list all those clearly established genetic diseases, such as Down syndrome<sup>[2]</sup>, sickle cell anemia<sup>[3,4]</sup>, Fanconi anemia<sup>[5]</sup>, and hemophilia<sup>[6]</sup>.

We know these genetic diseases relatively well, patients can have their better choices after receiving genetic counseling.

Secondly, we should also analyze and study the literature, in order to indicate and list all possible genetic diseases, such as some certain breast cancers<sup>[7-12]</sup>, autism<sup>[13,14]</sup>, some types of obesity<sup>[15]</sup>, and diabetes<sup>[16-18]</sup>. As our knowledge about these susceptibility loci is limited and partial, we should be very careful and responsible when patients seek genetic counseling about these possible genetic diseases. We should tell them frankly what we know and what we do not know, and try our best to help them to make their hard decisions.

Thirdly, we should be able to tell patients that what diseases are definitely not genetic diseases, such as bacterial, parasitic, and viral infections, even though many congenital infections like toxoplasmosis, rubella, hepatitis B, syphilis, herpes, cytomegalovirus, and human immunodeficiency virus can be transmitted from mother to child during pregnancy, delivery or breastfeeding<sup>[19]</sup>.

Recently, actress Jolie<sup>[20]</sup> underwent a preventive double mastectomy. I personally disagree with this action.

## IS BREAST CANCER A GENETIC DISEASE DUE TO MUTATIONS OF BREAST CANCER 1, EARLY ONSET OR/AND BREAST CANCER 2, EARLY ONSET GENE?

Breast cancer 1, early onset (*BRCA1*) gene is located on chromosome 17q<sup>[7-9]</sup>, and breast cancer 2, early onset (*BRCA2*) gene is located on chromosome 13q<sup>[10-12]</sup>. Besides *BRCA1* and *BRCA2*, there are some other susceptibility loci of breast cancers<sup>[21-25]</sup>.

A phenotype (trait) can be determined by one allele (dominant) or by two alleles (recessive). As we know, there are different types of genetic diseases. Some are autosomal-recessive genetic diseases like cystic fibrosis, sickle-cell anemia, and fanconi anemia (except Fanconi anemia subtype B); some are autosomal-dominant genetic diseases like Huntington's disease; Fanconi anemia subtype B, Duchenne muscular dystrophy, and Wiskott-Aldrich syndrome are X-linked recessive genetic diseases<sup>[26]</sup>, and Rett syndrome<sup>[27,28]</sup>, X-linked vitamin D-resistant rickets<sup>[29]</sup> are X-linked dominant genetic diseases.

If some breast cancers were autosomal-complete dominant genetic diseases, then all of the carriers of *BRCA1* or/and *BRCA2* gene mutations should have breast cancer too. But from the published data, we could not conclude that was true, because only some of those carriers developed breast cancer in their lifetimes<sup>[8,30-32]</sup>. We need to know whether breast cancer is an autosomal-incomplete dominant genetic disease. On the other hand, if some breast cancers were autosomal-recessive genetic diseases, then we need to confirm this is true: both of her/his parents of a breast cancer patient who carries the homozygous genes of *BRCA1* or/and *BRCA2* mutations

should be carriers or patients of *BRCA1* or/and *BRCA2* gene mutations. Homozygosity mapping should be performed to clarify this unclear problem<sup>[33-40]</sup>.

Currently, I have not seen any solid evidence in support of the linkage between breast cancer and *BRCA1* or/and *BRCA2*; and it is not crystal clear that a person who carries mutations of the *BRCA1* and/or *BRCA2* genes would eventually get breast cancer in her/his lifetime<sup>[31,32]</sup>.

## CONCLUSION

Like many other cancers, breast cancer is a very complicated, multifactorial disease; genetic factors, lifestyles and eating habits, environmental factors such as radiation, toxic chemicals<sup>[41,42]</sup>, and viral infections<sup>[43]</sup> might be involved in breast cancer; hence, it is difficult to figure out the real etiology of breast cancer. For a complex disease, a true linkage is so hard to find<sup>[44]</sup>, therefore, we should perform honest and strict epidemiological and genetic studies to evaluate the real risk of *BRCA1* and *BRCA2* gene mutations. While Angelina Jolie's natural breasts are gone already, we do not want to see more persons follow her to undergo such a preventive double mastectomy in the future, unless we are very clear and sure about the genetic link between some types of breast cancer and the mutations of *BRCA1* and *BRCA2* genes.

## REFERENCES

1. Nenad B, Maurizio M. Genetic counseling in post-genomic era-to be or not to be. *World J Med Genet* 2013; 3: 9-13 [DOI: 10.5496/wjmg.v3.i3.9]
2. Rahmani Z, Blouin JL, Créau-Goldberg N, Watkins PC, Mattei JF, Poissonnier M, Prieur M, Chettouh Z, Nicole A, Aurias A, Sinet P-M, Delabar J-M. Down syndrome critical region around D21S55 on proximal 21q22.3. *Am J Med Genet Suppl* 1990; 7: 98-103 [PMID: 2149984 DOI: 10.1002/ajmg.1320370720]
3. Ingram VM. A specific chemical difference between the globins of normal human and sickle-cell anaemia haemoglobin. *Nature* 1956; 178: 792-794 [PMID: 13369537]
4. Ingram VM. Gene mutations in human haemoglobin: the chemical difference between normal and sickle cell haemoglobin. *Nature* 1957; 180: 326-328 [PMID: 13464827]
5. Bogliolo M, Schuster B, Stoepker C, Derkunt B, Su Y, Raams A, Trujillo JP, Minguillón J, Ramírez MJ, Pujol R, Casado JA, Baños R, Rio P, Knies K, Zúñiga S, Benítez J, Bueren JA, Jaspers NG, Schärer OD, de Winter JP, Schindler D, Surrallés J. Mutations in ERCC4, encoding the DNA-repair endonuclease XPF, cause Fanconi anemia. *Am J Hum Genet* 2013; 92: 800-806 [PMID: 23623386 DOI: 10.1016/j.ajhg.2013.04.002]
6. Nilsson IM. [Haemophilia--then and now]. *Sydsven Medicinhist Sallsk Arsskr* 1994; 31: 33-52 [PMID: 11640407]
7. Hall JM, Lee MK, Newman B, Morrow JE, Anderson LA, Huey B, King MC. Linkage of early-onset familial breast cancer to chromosome 17q21. *Science* 1990; 250: 1684-1689 [PMID: 2270482]
8. Hall JM, Friedman L, Guenther C, Lee MK, Weber JL, Black DM, King MC. Closing in on a breast cancer gene on chromosome 17q. *Am J Hum Genet* 1992; 50: 1235-1242 [PMID: 1598904]
9. Miki Y, Swensen J, Shattuck-Eidens D, Futreal PA, Harshman K, Tavtigian S, Liu Q, Cochran C, Bennett LM, Ding W, Bell R, Rosenthal J, Hussey C, Tran T, McClure M, Frye C, Hattier T, Phelps R, Haugen-Strano A, Katcher H, Yakumo K, Gholami Z, Shaffer D, Stone S, Bayer S, Wray C, Bogden

- R, Dayananth P, Ward J, Tonin P, Narod S, Bristow P, Norris FH, Helvering L, Morrison P, Rosteck P, Lai M, Barrett C, Lewis C, Neuhausen S, Cannon-Albright L, Goldgar D, Wiseman R, Kamb A, Skolnick MH. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science* 1994; **266**: 66-71 [PMID: 7545954]
- 10 **Wooster R**, Neuhausen S, Mangion J, Quirk Y, Ford D, Collins N, Nguyen K, Seal S, Tran T, Averill D, Fields P, Marshall G, Narod S, Lenoir GM, Lynch H, Feunteun J, Devilee P, Cornelisse CJ, Menko FH, Daly PA, Ormiston W, McManus R, Pye C, Lewis CM, Cannon-Albright LA, Peto J, Ponder BAJ, Skolnick MH, Easton DF, Goldgar DE, Stratton MR. Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12-13. *Science* 1994; **265**: 2088-2090 [PMID: 8091231]
  - 11 **Wooster R**, Bignell G, Lancaster J, Swift S, Seal S, Mangion J, Collins N, Gregory S, Gumbs C, Micklem G. Identification of the breast cancer susceptibility gene BRCA2. *Nature* 1995; **378**: 789-792 [PMID: 8524414]
  - 12 **Tavtigian SV**, Simard J, Rommens J, Couch F, Shattuck-Eidens D, Neuhausen S, Merajver S, Thorlacius S, Offit K, Stoppa-Lyonnet D, Belanger C, Bell R, Berry S, Bogden R, Chen Q, Davis T, Dumont M, Frye C, Hattier T, Jammulapati S, Janecki T, Jiang P, Kehrer R, Leblanc JF, Mitchell JT, McArthur-Morrison J, Nguyen K, Peng Y, Samson C, Schroeder M, Snyder SC, Steele L, Stringfellow M, Stroup C, Swedlund B, Swense J, Teng D, Thomas A, Tran T, Tranchant M, Weaver-Feldhaus J, Wong AK, Shizuya H, Eyfjord JE, Cannon-Albright L, Tranchant M, Labrie F, Skolnick MH, Weber B, Kamb A, Goldgar DE. The complete BRCA2 gene and mutations in chromosome 13q-linked kindreds. *Nat Genet* 1996; **12**: 333-337 [PMID: 8589730]
  - 13 **Alarcón M**, Cantor RM, Liu J, Gilliam TC, Geschwind DH; Autism Genetic Research Exchange Consortium. Evidence for a language quantitative trait locus on chromosome 7q in multiplex autism families. *Am J Hum Genet* 2002; **70**: 60-71 [PMID: 11741194]
  - 14 **Cantor RM**, Kono N, Duvall JA, Alvarez-Retuerto A, Stone JL, Alarcón M, Nelson SF, Geschwind DH. Replication of autism linkage: fine-mapping peak at 17q21. *Am J Hum Genet* 2005; **76**: 1050-1056 [PMID: 15877280]
  - 15 **Stone S**, Abkevich V, Russell DL, Riley R, Timms K, Tran T, Trem D, Frank D, Jammulapati S, Neff CD, Iliev D, Gress R, He G, Frech GC, Adams TD, Skolnick MH, Lanchbury JS, Gutin A, Hunt SC, Shattuck D. TBC1D1 is a candidate for a severe obesity gene and evidence for a gene/gene interaction in obesity predisposition. *Hum Mol Genet* 2006; **15**: 2709-2720 [PMID: 16893906]
  - 16 **Pociot F**, McDermott MF. Genetics of type 1 diabetes mellitus. *Genes Immun* 2002; **3**: 235-249 [PMID: 12140742]
  - 17 **Concannon P**, Chen WM, Julier C, Morahan G, Akolkar B, Erlich HA, Hilner JE, Nerup J, Nierras C, Pociot F, Todd JA, Rich SS; Type 1 Diabetes Genetics Consortium. Genome-wide scan for linkage to type 1 diabetes in 2,496 multiplex families from the Type 1 Diabetes Genetics Consortium. *Diabetes* 2009; **58**: 1018-1022 [PMID: 19136655 DOI: 10.2337/db08-1551]
  - 18 **Elbein SC**, Gamazon ER, Das SK, Rasouli N, Kern PA, Cox NJ. Genetic risk factors for type 2 diabetes: a trans-regulatory genetic architecture? *Am J Hum Genet* 2012; **91**: 466-477 [PMID: 22958899 DOI: 10.1016/j.ajhg.2012.08.002]
  - 19 **Del Pizzo J**. Focus on diagnosis: congenital infections (TORCH). *Pediatr Rev* 2011; **32**: 537-542 [PMID: 22135424 DOI: 10.1542/pir.32-12-537]
  - 20 **Jolie A**. My medical choice. Available from: URL: <http://www.nytimes.com/2013/05/14/opinion-medical-choice.html?ref=opinion>
  - 21 **Couch FJ**, Wang X, McGuffog L, Lee A, Olswold C, Kuchenbaecker KB, Soucy P, Fredericksen Z, Barrowdale D, Dennis J, Gaudet MM, Dicks E, Kosel M, Healey S, Sinilnikova OM, Lee A, Bacot F, Vincent D, Hogervorst FB, Peock S, Stoppa-Lyonnet D, Jakubowska A, Radice P, Schmutzler RK, Domchek SM, Piedmonte M, Singer CF, Friedman E, Thomassen M, Hansen TV, Neuhausen SL, Szabo CI, Blanco I, Greene MH, Karlan BY, Garber J, Phelan CM, Weitzel JN, Montagna M, Olah E, Andrulis IL, Godwin AK, Yannoukakos D, Goldgar DE, Caldes T, Nevanlinna H, Osorio A, Terry MB, Daly MB, van Rensburg EJ, Hamann U, Ramus SJ, Toland AE, Caligo MA, Olopade OI, Tung N, Claes K, Beattie MS, Southey MC, Imyanitov EN, Tischkowitz M, Janavicius R, John EM, Kwong A, Diez O, Balmaña J, Barkardottir RB, Arun BK, Rennert G, Teo SH, Ganz PA, Campbell I, van der Hout AH, van Deurzen CH, Seynaeve C, Gómez García EB, van Leeuwen FE, Meijers-Heijboer HE, Gille JJ, Ausems MG, Blok MJ, Ligtenberg MJ, Rookus MA, Devilee P, Verhoef S, van Os TA, Wijnen JT, Frost D, Ellis S, Fineberg E, Platte R, Evans DG, Izatt L, Eeles RA, Adlard J, Eccles DM, Cook J, Brewer C, Douglas F, Hodgson S, Morrison PJ, Side LE, Donaldson A, Houghton C, Rogers MT, Dorkins H, Eason J, Gregory H, McCann E, Murray A, Calender A, Hardouin A, Berthet P, Delnatte C, Nogues C, Lasset C, Houdayer C, Leroux D, Rouleau E, Prieur F, Damiola F, Sobol H, Coudier I, Venat-Bouvet L, Castera L, Gauthier-Villars M, Léoné M, Pujol P, Mazoyer S, Bignon YJ, Złowocka-Perłowska E, Gronwald J, Lubinski J, Durda K, Jaworska K, Huzarski T, Spurdle AB, Viel A, Peissel B, Bonanni B, Melloni G, Ottini L, Papi L, Varesco L, Tibiletti MG, Peterlongo P, Volorio S, Manoukian S, Pensotti V, Arnold N, Engel C, Deissler H, Gadzicki D, Gehrig A, Kast K, Rhiem K, Meindl A, Niederacher D, Ditsch N, Plendl H, Preisler-Adams S, Engert S, Sutter C, Varon-Mateeva R, Wappenschmidt B, Weber BH, Arver B, Stenmark-Askmal M, Loman N, Rosenquist R, Einbeigi Z, Nathanson KL, Rebbeck TR, Blank SV, Cohn DE, Rodriguez GC, Small L, Friedlander M, Bae-Jump VL, Fink-Retter A, Rappaport C, Gschwandler-Kaulich D, Pfeiler G, Tea MK, Lindor NM, Kaufman B, Shimon Paluch S, Laitman Y, Skytte AB, Gerdes AM, Pedersen IS, Moeller ST, Kruse TA, Jensen UB, Vijai J, Sarrel K, Robson M, Kauff N, Mulligan AM, Glendon G, Ozcelik H, Ejlersen B, Nielsen FC, Jønson L, Andersen MK, Ding YC, Steele L, Foretova L, Teulé A, Lázaro C, Brunet J, Pujana MA, Mai PL, Loud JT, Walsh C, Lester J, Orsulic S, Narod SA, Herzog J, Sand SR, Tognazzo S, Agata S, Vaszko T, Weaver J, Stavropoulou AV, Buys SS, Romero A, de la Hoya M, Aittomäki K, Muranen TA, Duran M, Chung WK, Lasa A, Dorfling CM, Miron A, Benitez J, Senter L, Huo D, Chan SB, Sokolenko AP, Chiquette J, Tihomirova L, Friebel TM, Agnarsson BA, Lu KH, Lejbkiewicz F, James PA, Hall P, Dunning AM, Tessier D, Cunningham J, Slager SL, Wang C, Hart S, Stevens K, Simard J, Pastinen T, Pankratz VS, Offit K, Easton DF, Chenevix-Trench G, Antoniou AC; CIMBA. Genome-wide association study in BRCA1 mutation carriers identifies novel loci associated with breast and ovarian cancer risk. *PLoS Genet* 2013; **9**: e1003212 [PMID: 23544013 DOI: 10.1371/journal.pgen.1003212]
  - 22 **Antoniou AC**, Wang X, Fredericksen ZS, McGuffog L, Tarrell R, Sinilnikova OM, Healey S, Morrison J, Kartsonaki C, Lesnick T, Ghoussaini M, Barrowdale D, Peock S, Cook M, Oliver C, Frost D, Eccles D, Evans DG, Eeles R, Izatt L, Chu C, Douglas F, Paterson J, Stoppa-Lyonnet D, Houdayer C, Mazoyer S, Giraud S, Lasset C, Remenieras A, Caron O, Hardouin A, Berthet P, Hogervorst FB, Rookus MA, Jager A, van den Ouweland A, Hoogerbrugge N, van der Luijt RB, Meijers-Heijboer H, Gómez García EB, Devilee P, Vreeswijk MP, Lubinski J, Jakubowska A, Gronwald J, Huzarski T, Byrski T, Górski B, Cybulski C, Spurdle AB, Holland H, Goldgar DE, John EM, Hopper JL, Southey M, Buys SS, Daly MB, Terry MB, Schmutzler RK, Wappenschmidt B, Engel C, Meindl A, Preisler-Adams S, Arnold N, Niederacher D, Sutter C, Domchek SM, Nathanson KL, Rebbeck T, Blum JL, Piedmonte M, Rodriguez GC, Wakeley K, Boggess JF, Basil J, Blank SV, Friedman E, Kaufman B, Laitman Y, Milgrom R, Andrulis IL, Glendon G, Ozcelik H, Kirchhoff T, Vijai J, Gaudet MM, Altshuler D, Guiducci C, Loman N, Harbst K, Rantala J, Eh-



- rencrona H, Gerdes AM, Thomassen M, Sunde L, Peterlongo P, Manoukian S, Bonanni B, Viel A, Radice P, Caldes T, de la Hoya M, Singer CF, Fink-Retter A, Greene MH, Mai PL, Loud JT, Guidugli L, Lindor NM, Hansen TV, Nielsen FC, Blanco I, Lazaro C, Garber J, Ramus SJ, Gayther SA, Phelan C, Narod S, Szabo CI, Benitez J, Osorio A, Nevanlinna H, Heikinen T, Caligo MA, Beattie MS, Hamann U, Godwin AK, Montagna M, Casella C, Neuhausen SL, Karlan BY, Tung N, Toland AE, Weitzel J, Olopade O, Simard J, Soucy P, Rubinstein WS, Arason A, Rennert G, Martin NG, Montgomery GW, Chang-Claude J, Flesch-Janys D, Brauch H, Severi G, Baglietto L, Cox A, Cross SS, Miron P, Gerty SM, Tapper W, Yannoukakos D, Fountzilas G, Fasching PA, Beckmann MW, Dos Santos Silva I, Peto J, Lambrechts D, Paridaens R, Rüdiger T, Försti A, Winqvist R, Pylkäs K, Diasio RB, Lee AM, Eckel-Passow J, Vachon C, Blows F, Driver K, Dunning A, Pharoah PP, Offit K, Pankratz VS, Hakonarson H, Chenevix-Trench G, Easton DF, Couch FJ. A locus on 19p13 modifies risk of breast cancer in BRCA1 mutation carriers and is associated with hormone receptor-negative breast cancer in the general population. *Nat Genet* 2010; **42**: 885-892 [PMID: 20852631 DOI: 10.1038/ng.669]
- 23 **Park DJ**, Lesueur F, Nguyen-Dumont T, Pertesi M, Odefrey F, Hammet F, Neuhausen SL, John EM, Andrulis IL, Terry MB, Daly M, Buys S, Le Calvez-Kelm F, Lonie A, Pope BJ, Tsimiklis H, Voegelé C, Hilbers FM, Hoogerbrugge N, Barroso A, Osorio A, Giles GG, Devilee P, Benitez J, Hopper JL, Tavtigian SV, Goldgar DE, Southey MC. Rare mutations in XRCC2 increase the risk of breast cancer. *Am J Hum Genet* 2012; **90**: 734-739 [PMID: 22464251 DOI: 10.1016/j.ajhg.2012.02.027]
- 24 **Michailidou K**, Hall P, Gonzalez-Neira A, Ghoussaini M, Dennis J, Milne RL, Schmidt MK, Chang-Claude J, Bojesen SE, Bolla MK, Wang Q, Dicks E, Lee A, Turnbull C, Rahman N; Breast and Ovarian Cancer Susceptibility Collaboration, Fletcher O, Peto J, Gibson L, Dos Santos Silva I, Nevanlinna H, Muranen TA, Aittomäki K, Blomqvist C, Czene K, Irwanto A, Liu J, Waisfisz Q, Meijers-Heijboer H, Adank M; Hereditary Breast and Ovarian Cancer Research Group Netherlands (HEBON), van der Luijt RB, Hein R, Dahmen N, Beckman L, Meindl A, Schmutzler RK, Müller-Mysok B, Lichtner P, Hopper JL, Southey MC, Makalic E, Schmidt DF, Uitterlinden AG, Hofman A, Hunter DJ, Chanock SJ, Vincent D, Bacot F, Tessier DC, Canisius S, Wessels LF, Haiman CA, Shah M, Luben R, Brown J, Luccarini C, Schoof N, Humphreys K, Li J, Nordestgaard BG, Nielsen SF, Flyger H, Couch FJ, Wang X, Vachon C, Stevens KN, Lambrechts D, Moisse M, Paridaens R, Christiaens MR, Rudolph A, Nickels S, Flesch-Janys D, Johnson N, Aitken Z, Aaltonen K, Heikinen T, Broeks A, Veer LJ, van der Schoot CE, Guénel P, Truong T, Laurent-Puig P, Menegaux F, Marme F, Schneeweiss A, Sohn C, Burwinkel B, Zamora MP, Perez JL, Pita G, Alonso MR, Cox A, Brock IW, Cross SS, Reed MW, Sawyer EJ, Tomlinson I, Kerin MJ, Miller N, Henderson BE, Schumacher F, Le Marchand L, Andrulis IL, Knight JA, Glendon G, Mulligan AM; kConFab Investigators; Australian Ovarian Cancer Study Group, Lindblom A, Margolin S, Hooning MJ, Hollestelle A, van den Ouweland AM, Jager A, Bui QM, Stone J, Dite GS, Apicella C, Tsimiklis H, Giles GG, Severi G, Baglietto L, Fasching PA, Haerle L, Ekici AB, Beckmann MW, Brenner H, Müller H, Arndt V, Stegmaier C, Sewardlow A, Ashworth A, Orr N, Jones M, Figueroa J, Lissowska J, Brinton L, Goldberg MS, Labrèche F, Dumont M, Winqvist R, Pylkäs K, Jukkola-Vuorinen A, Grip M, Brauch H, Hamann U, Brüning T; GENICA (Gene Environment Interaction and Breast Cancer in Germany) Network, Radice P, Peterlongo P, Manoukian S, Bonanni B, Devilee P, Tollenaar RA, Seynaeve C, van Asperen CJ, Jakubowska A, Lubinski J, Jaworska K, Durda K, Mannermaa A, Kataja V, Kosma VM, Hartikainen JM, Bogdanova NV, Antonenkova NN, Dörk T, Kristensen VN, Anton-Culver H, Slager S, Toland AE, Edge S, Fostira F, Kang D, Yoo KY, Noh DY, Matsuo K, Ito H, Iwata H, Sueta A, Wu AH, Tseng CC, Van Den Berg D, Stram DO, Shu XO, Lu W, Gao YT, Cai H, Teo SH, Yip CH, Phuah SY, Cornes BK, Hartman M, Miao H, Lim WY, Sng JH, Muir K, Lophatananon A, Stewart-Brown S, Siriwanarangsang P, Shen CY, Hsiung CN, Wu PE, Ding SL, Sangrajrang S, Gaborieau V, Brennan P, McKay J, Blot WJ, Signorello LB, Cai Q, Zheng W, Deming-Halverson S, Shrubsole M, Long J, Simard J, Garcia-Closas M, Pharoah PD, Chenevix-Trench G, Dunning AM, Benitez J, Easton DF. Large-scale genotyping identifies 41 new loci associated with breast cancer risk. *Nat Genet* 2013; **45**: 353-361 [PMID: 23535729 DOI: 10.1038/ng.2563]
- 25 **Garcia-Closas M**, Couch FJ, Lindstrom S, Michailidou K, Schmidt MK, Brook MN, Orr N, Rhie SK, Riboli E, Feigelson HS, Le Marchand L, Buring JE, Eccles D, Miron P, Fasching PA, Brauch H, Chang-Claude J, Carpenter J, Godwin AK, Nevanlinna H, Giles GG, Cox A, Hopper JL, Bolla MK, Wang Q, Dennis J, Dicks E, Howat WJ, Schoof N, Bojesen SE, Lambrechts D, Broeks A, Andrulis IL, Guénel P, Burwinkel B, Sawyer EJ, Hollestelle A, Fletcher O, Winqvist R, Brenner H, Mannermaa A, Hamann U, Meindl A, Lindblom A, Zheng W, Devilee P, Goldberg MS, Lubinski J, Kristensen V, Sewardlow A, Anton-Culver H, Dörk T, Muir K, Matsuo K, Wu AH, Radice P, Teo SH, Shu XO, Blot W, Kang D, Hartman M, Sangrajrang S, Shen CY, Southey MC, Park DJ, Hammet F, Stone J, Veer LJ, Rutgers EJ, Lophatananon A, Stewart-Brown S, Siriwanarangsang P, Peto J, Schrauder MG, Ekici AB, Beckmann MW, Dos Santos Silva I, Johnson N, Warren H, Tomlinson I, Kerin MJ, Miller N, Marme F, Schneeweiss A, Sohn C, Truong T, Laurent-Puig P, Kerbrat P, Nordestgaard BG, Nielsen SF, Flyger H, Milne RL, Perez JL, Menéndez P, Müller H, Arndt V, Stegmaier C, Lichtner P, Lochmann M, Justenhoven C, Ko YD; Gene ENvironmental Interaction and breast CANcer (GENICA) Network, Muranen TA, Aittomäki K, Blomqvist C, Greco D, Heikinen T, Ito H, Iwata H, Yatabe Y, Antonenkova NN, Margolin S, Kataja V, Kosma VM, Hartikainen JM, Balleine R, kConFab Investigators, Tseng CC, Berg DV, Stram DO, Neven P, Dieudonné AS, Leunen K, Rudolph A, Nickels S, Flesch-Janys D, Peterlongo P, Peissel B, Bernard L, Olson JE, Wang X, Stevens K, Severi G, Baglietto L, McLean C, Coetzee GA, Feng Y, Henderson BE, Schumacher F, Bogdanova NV, Labrèche F, Dumont M, Yip CH, Taib NA, Cheng CY, Shrubsole M, Long J, Pylkäs K, Jukkola-Vuorinen A, Kauppila S, Knight JA, Glendon G, Mulligan AM, Tollenaar RA, Seynaeve CM, Kriege M, Hooning MJ, van den Ouweland AM, van Deurzen CH, Lu W, Gao YT, Cai H, Balasubramanian SP, Cross SS, Reed MW, Signorello L, Cai Q, Shah M, Miao H, Chan CW, Chia KS, Jakubowska A, Jaworska K, Durda K, Hsiung CN, Wu PE, Yu JC, Ashworth A, Jones M, Tessier DC, Gonzalez-Neira A, Pita G, Alonso MR, Vincent D, Bacot F, Ambrosone CB, Bandera EV, John EM, Chen GK, Hu JJ, Rodriguez-Gil JL, Bernstein L, Press MF, Ziegler RG, Millikan RM, Deming-Halverson SL, Nyante S, Ingles SA, Waisfisz Q, Tsimiklis H, Makalic E, Schmidt D, Bui M, Gibson L, Müller-Mysok B, Schmutzler RK, Hein R, Dahmen N, Beckmann L, Aaltonen K, Czene K, Irwanto A, Liu J, Turnbull C; Familial Breast Cancer Study (FBCS), Rahman N, Meijers-Heijboer H, Uitterlinden AG, Rivadeneira F; Australian Breast Cancer Tissue Bank (ABCTB) Investigators, Olsowd C, Slager S, Pilarski R, Ademuyiwa F, Konstantopoulou I, Martin NG, Montgomery GW, Slamon DJ, Rauh C, Lux MP, Jud SM, Bruning T, Weaver J, Sharma P, Pathak H, Tapper W, Gerty S, Durcan L, Trichopoulos D, Tumino R, Peeters PH, Kaaks R, Campa D, Canzian F, Weiderpass E, Johansson M, Khaw KT, Travis R, Clavel-Chapelon F, Kolonel LN, Chen C, Beck A, Hankinson SE, Berg CD, Hoover RN, Lissowska J, Figueroa JD, Chasman DI, Gaudet MM, Diver WR, Willett WC, Hunter DJ, Simard J, Benitez J, Dunning AM, Sherman ME, Chenevix-Trench G, Chanock SJ, Hall P, Pharoah PD, Vachon C, Easton DF, Haiman CA, Kraft P. Genome-wide association studies identify four ER

- negative-specific breast cancer risk loci. *Nat Genet* 2013; **45**: 392-398 [PMID: 23535733 DOI: 10.1038/ng.2561]
- 26 **Song L.** Chapter 6: Gene therapy of some genetic diseases by transferring normal human genomic DNA into somatic cells and stem cells from patients. In *Non-viral Gene Therapy*. In Tech-Open Access Publisher. Rijeka, Croatia, 2011: 105-146
  - 27 **Trappe R,** Laccone F, Cobilanschi J, Meins M, Huppke P, Hanefeld F, Engel W. MECP2 mutations in sporadic cases of Rett syndrome are almost exclusively of paternal origin. *Am J Hum Genet* 2001; **68**: 1093-1101 [PMID: 11309679]
  - 28 **Amir RE,** Van den Veyver IB, Wan M, Tran CQ, Francke U, Zoghbi HY. Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat Genet* 1999; **23**: 185-188 [PMID: 10508514]
  - 29 **Francis F,** Hennig S, Korn B, Reinhardt R, de Jong P, Poustka A, Lehrach H, Rowe PSN, Goulding JN, Summerfield T, Mountford R, Read AP, Popowska E, Pronicka E, Davies KE, O'Riordan JLH, Econs MJ, Nesbitt T, Drezner MK, Oudet C, Pannetier S, Hanauer A, Strom TM, Meindl A, Lorenz B, Cagnoli B, Mohnike KL, Murken J, Meitinger T. A gene (PEX) with homologies to endopeptidases is mutated in patients with X-linked hypophosphatemic rickets. The HYP Consortium. *Nat Genet* 1995; **11**: 130-136 [PMID: 7550339]
  - 30 **Newman B,** Austin MA, Lee M, King MC. Inheritance of human breast cancer: evidence for autosomal dominant transmission in high-risk families. *Proc Natl Acad Sci USA* 1988; **85**: 3044-3048 [PMID: 3362861]
  - 31 **Anglian Breast Cancer Study Group.** Prevalence and penetrance of BRCA1 and BRCA2 mutations in a population-based series of breast cancer cases. Anglian Breast Cancer Study Group. *Br J Cancer* 2000; **83**: 1301-1308 [PMID: 11044354]
  - 32 **Malone KE,** Daling JR, Thompson JD, O'Brien CA, Francisco LV, Ostrander EA. BRCA1 mutations and breast cancer in the general population: analyses in women before age 35 years and in women before age 45 years with first-degree family history. *JAMA* 1998; **279**: 922-929 [PMID: 9544766]
  - 33 **Lander ES,** Botstein D. Strategies for studying heterogeneous genetic traits in humans by using a linkage map of restriction fragment length polymorphisms. *Proc Natl Acad Sci USA* 1986; **83**: 7353-7357 [PMID: 2876423]
  - 34 **Lander ES,** Botstein D. Homozygosity mapping: a way to map human recessive traits with the DNA of inbred children. *Science* 1987; **236**: 1567-1570 [PMID: 2884728]
  - 35 **Feinstein M,** Markus B, Noyman I, Shalev H, Flusser H, Shelef I, Liani-Leibson K, Shorer Z, Cohen I, Khateeb S, Sivan S, Birk OS. Pelizaeus-Merzbacher-like disease caused by AIMP1/p43 homozygous mutation. *Am J Hum Genet* 2010; **87**: 820-828 [PMID: 21092922 DOI: 10.1016/j.ajhg.2010.10.016]
  - 36 **Bandah-Rozenfeld D,** Mizrahi-Meissonnier L, Farhy C, Obolensky A, Chowers I, Pe'er J, Merin S, Ben-Yosef T, Ashery-Padan R, Banin E, Sharon D. Homozygosity mapping reveals null mutations in FAM161A as a cause of autosomal-recessive retinitis pigmentosa. *Am J Hum Genet* 2010; **87**: 382-391 [PMID: 20705279 DOI: 10.1016/j.ajhg.2010.07.022]
  - 37 **Tukel T,** Šošić D, Al-Gazali LI, Erazo M, Casasnovas J, Franco HL, Richardson JA, Olson EN, Cadilla CL, Desnick RJ. Homozygous nonsense mutations in TWIST2 cause Setleis syndrome. *Am J Hum Genet* 2010; **87**: 289-296 [PMID: 20691403 DOI: 10.1016/j.ajhg.2010.07.009]
  - 38 **Walsh T,** Shahin H, Elkan-Miller T, Lee MK, Thornton AM, Roeb W, Abu Rayyan A, Loulus S, Avraham KB, King MC, Kanaan M. Whole exome sequencing and homozygosity mapping identify mutation in the cell polarity protein GPM2 as the cause of nonsyndromic hearing loss DFN82. *Am J Hum Genet* 2010; **87**: 90-94 [PMID: 20602914 DOI: 10.1016/j.ajhg.2010.05.010]
  - 39 **Christiansen HE,** Schwarze U, Pyott SM, AlSwaied A, Al Balwi M, Alrasheed S, Pepin MG, Weis MA, Eyre DR, Byers PH. Homozygosity for a missense mutation in SERPINH1, which encodes the collagen chaperone protein HSP47, results in severe recessive osteogenesis imperfecta. *Am J Hum Genet* 2010; **86**: 389-398 [PMID: 20188343 DOI: 10.1016/j.ajhg.2010.01.034]
  - 40 **Morales J,** Al-Sharif L, Khalil DS, Shinwari JM, Bavi P, Al-Mahrouqi RA, Al-Rajhi A, Alkuraya FS, Meyer BF, Al Tassan N. Homozygous mutations in ADAMTS10 and ADAMTS17 cause lenticular myopia, ectopia lentis, glaucoma, spherophakia, and short stature. *Am J Hum Genet* 2009; **85**: 558-568 [PMID: 19836009 DOI: 10.1016/j.ajhg.2009.09.011]
  - 41 **Brody JG,** Rudel RA, Michels KB, Moysich KB, Bernstein L, Attfield KR, Gray S. Environmental pollutants, diet, physical activity, body size, and breast cancer: where do we stand in research to identify opportunities for prevention? *Cancer* 2007; **109**: 2627-2634 [PMID: 17503444]
  - 42 **Martin AM,** Weber BL. Genetic and hormonal risk factors in breast cancer. *J Natl Cancer Inst* 2000; **92**: 1126-1135 [PMID: 10904085]
  - 43 **Glenn WK,** Heng B, Delprado W, Iacopetta B, Whitaker NJ, Lawson JS. Epstein-Barr virus, human papillomavirus and mouse mammary tumour virus as multiple viruses in breast cancer. *PLoS One* 2012; **7**: e48788 [PMID: 23183846 DOI: 10.1371/journal.pone.0048788]
  - 44 **Altmüller J,** Palmer LJ, Fischer G, Scherb H, Wjst M. Genomewide scans of complex human diseases: true linkage is hard to find. *Am J Hum Genet* 2001; **69**: 936-950 [PMID: 11565063 DOI: 10.1086/324069]

**P- Reviewers:** Asensi V, Fillat C, Sherer TB

**S- Editor:** Gou SX **L- Editor:** A **E- Editor:** Liu SQ







## INSTRUCTIONS TO AUTHORS

### GENERAL INFORMATION

*World Journal of Medical Genetics* (*World J Med Genet*, *WJMG*, online ISSN 2220-3184, DOI: 10.5496) is a peer-reviewed open access (OA) academic journal that aims to guide clinical practice and improve diagnostic and therapeutic skills of clinicians.

#### Aim and scope

*WJMG* covers topics concerning genes and the pathology of human disease, molecular analysis of simple and complex genetic traits, cancer genetics, epigenetics, gene therapy, developmental genetics, regulation of gene expression, strategies and technologies for extracting function from genomic data, pharmacological genomics, genome evolution. The current columns of *WJMG* include editorial, frontier, diagnostic advances, therapeutics advances, field of vision, mini-reviews, review, topic highlight, medical ethics, original articles, case report, clinical case conference (Clinicopathological conference), and autobiography.

We encourage authors to submit their manuscripts to *WJMG*. We will give priority to manuscripts that are supported by major national and international foundations and those that are of great basic and clinical significance.

*WJMG* is edited and published by Baishideng Publishing Group (BPG). BPG has a strong professional editorial team composed of science editors, language editors and electronic editors. BPG currently publishes 43 OA clinical medical journals, including 42 in English, has a total of 15471 editorial board members or peer reviewers, and is a world first-class publisher.

#### Columns

The columns in the issues of *WJMG* will include: (1) Editorial: The editorial board members are invited to make comments on an important topic in their field in terms of its current research status and future directions to lead the development of this discipline; (2) Frontier: The editorial board members are invited to select a highly cited cutting-edge original paper of his/her own to summarize major findings, the problems that have been resolved and remain to be resolved, and future research directions to help readers understand his/her important academic point of view and future research directions in the field; (3) Diagnostic Advances: The editorial board members are invited to write high-quality diagnostic advances in their field to improve the diagnostic skills of readers. The topic covers general clinical diagnosis, differential diagnosis, pathological diagnosis, laboratory diagnosis, imaging diagnosis, endoscopic diagnosis, biotechnological diagnosis, functional diagnosis, and physical diagnosis; (4) Therapeutics Advances: The editorial board members are invited to write high-quality therapeutic advances in their field to help improve the therapeutic skills of readers. The topic covers medication therapy, psychotherapy, physical therapy, replacement therapy, interventional therapy, minimally invasive therapy, endoscopic therapy, transplantation therapy, and surgical therapy; (5) Field of Vision: The editorial board members are invited to write commentaries on classic articles, hot topic articles, or latest articles to keep readers at the forefront of research and increase their levels of clinical research. Classic articles refer to papers that are included in Web of Knowledge and have received a large number of citations (ranking in the top 1%) after being published for more than years, reflecting the quality and impact of papers. Hot topic articles refer to papers that are included in Web of Knowledge and have received a large number of citations after being published for no more than

2 years, reflecting cutting-edge trends in scientific research. Latest articles refer to the latest published high-quality papers that are included in PubMed, reflecting the latest research trends. These commentary articles should focus on the status quo of research, the most important research topics, the problems that have now been resolved and remain to be resolved, and future research directions. Basic information about the article to be commented (including authors, article title, journal name, year, volume, and inclusive page numbers); (6) Minireviews: The editorial board members are invited to write short reviews on recent advances and trends in research of molecular biology, genomics, and related cutting-edge technologies to provide readers with the latest knowledge and help improve their diagnostic and therapeutic skills; (7) Review: To make a systematic review to focus on the status quo of research, the most important research topics, the problems that have now been resolved and remain to be resolved, and future research directions; (8) Topic Highlight: The editorial board members are invited to write a series of articles (7-10 articles) to comment and discuss a hot topic to help improve the diagnostic and therapeutic skills of readers; (9) Medical Ethics: The editorial board members are invited to write articles about medical ethics to increase readers' knowledge of medical ethics. The topic covers international ethics guidelines, animal studies, clinical trials, organ transplantation, etc.; (10) Clinical Case Conference or Clinicopathological Conference: The editorial board members are invited to contribute high-quality clinical case conference; (11) Original Articles: To report innovative and original findings in medical genetics; (12) Brief Articles: To briefly report the novel and innovative findings in medical genetics; (13) Meta-Analysis: Covers the systematic review, mixed treatment comparison, meta-regression, and overview of reviews, in order to summarize a given quantitative effect, e.g., the clinical effectiveness and safety of clinical treatments by combining data from two or more randomized controlled trials, thereby providing more precise and externally valid estimates than those which would stem from each individual dataset if analyzed separately from the others; (15) Letters to the Editor: To discuss and make reply to the contributions published in *WJMG*, or to introduce and comment on a controversial issue of general interest; (16) Book Reviews: To introduce and comment on quality monographs of medical genetics; and (17) Autobiography: The editorial board members are invited to write their autobiography to provide readers with stories of success or failure in their scientific research career. The topic covers their basic personal information and information about when they started doing research work, where and how they did research work, what they have achieved, and their lessons from success or failure.

#### Name of journal

*World Journal of Medical Genetics*

#### ISSN

ISSN 2220-3184 (online)

#### Launch date

December 27, 2011

#### Frequency

Quarterly

#### Editor-in-Chief

Hans van Bokhoven, Professor, PhD, Department of Human

## Instructions to authors

Genetics and Cognitive Neurosciences, Radboud university Nijmegen  
Medical centre, PO Box 9101, 6500 HB Nijmegen, The Netherlands

### Editorial office

Jin-Lei Wang, Director  
Xiu-Xia Song, Vice Director  
*World Journal of Medical Genetics*  
Room 903, Building D, Ocean International Center,  
No. 62 Dongsihuan Zhonglu, Chaoyang District,  
Beijing 100025, China  
Telephone: +86-10-85381891  
Fax: +86-10-85381893  
E-mail: [bpgooffice@wjgnet.com](mailto:bpgooffice@wjgnet.com)  
<http://www.wjgnet.com>

### Publisher

Baishideng Publishing Group Co., Limited  
Flat C, 23/F, Lucky Plaza, 315-321 Lockhart Road,  
Wan Chai, Hong Kong, China  
Telephone: +852-58042046  
Fax: +852-31158812  
E-mail: [bpgooffice@wjgnet.com](mailto:bpgooffice@wjgnet.com)  
<http://www.wjgnet.com>

### Production center

Beijing Baishideng BioMed Scientific Co., Limited  
Room 903, Building D, Ocean International Center,  
No. 62 Dongsihuan Zhonglu, Chaoyang District,  
Beijing 100025, China  
Telephone: +86-10-85381892  
Fax: +86-10-85381893

### Representative office

USA Office  
8226 Regency Drive,  
Pleasanton, CA 94588-3144, United States

### Instructions to authors

Full instructions are available online at [http://www.wjgnet.com/2220-3184/g\\_info\\_20100722180909.htm](http://www.wjgnet.com/2220-3184/g_info_20100722180909.htm).

### Indexed and Abstracted in

Digital Object Identifier.

## SPECIAL STATEMENT

All articles published in this journal represent the viewpoints of the authors except where indicated otherwise.

### Biostatistical editing

Statistical review is performed after peer review. We invite an expert in Biomedical Statistics to evaluate the statistical method used in the paper, including *t*-test (group or paired comparisons), chi-squared test, Redit, probit, logit, regression (linear, curvilinear, or stepwise), correlation, analysis of variance, analysis of covariance, *etc.* The reviewing points include: (1) Statistical methods should be described when they are used to verify the results; (2) Whether the statistical techniques are suitable or correct; (3) Only homogeneous data can be averaged. Standard deviations are preferred to standard errors. Give the number of observations and subjects (*n*). Losses in observations, such as drop-outs from the study should be reported; (4) Values such as ED50, LD50, IC50 should have their 95% confidence limits calculated and compared by weighted probit analysis (Bliss and Finney); and (5) The word "significantly" should be replaced by its synonyms (if it indicates extent) or the *P* value (if it indicates statistical significance).

### Conflict-of-interest statement

In the interests of transparency and to help reviewers assess any potential bias, *WJMG* requires authors of all papers to declare any competing commercial, personal, political, intellectual, or religious interests

in relation to the submitted work. Referees are also asked to indicate any potential conflict they might have reviewing a particular paper. Before submitting, authors are suggested to read "Uniform Requirements for Manuscripts Submitted to Biomedical Journals: Ethical Considerations in the Conduct and Reporting of Research: Conflicts of Interest" from International Committee of Medical Journal Editors (ICMJE), which is available at: [http://www.icmje.org/ethical\\_4conflicts.html](http://www.icmje.org/ethical_4conflicts.html).

Sample wording: [Name of individual] has received fees for serving as a speaker, a consultant and an advisory board member for [names of organizations], and has received research funding from [names of organization]. [Name of individual] is an employee of [name of organization]. [Name of individual] owns stocks and shares in [name of organization]. [Name of individual] owns patent [patent identification and brief description].

### Statement of informed consent

Manuscripts should contain a statement to the effect that all human studies have been reviewed by the appropriate ethics committee or it should be stated clearly in the text that all persons gave their informed consent prior to their inclusion in the study. Details that might disclose the identity of the subjects under study should be omitted. Authors should also draw attention to the Code of Ethics of the World Medical Association (Declaration of Helsinki, 1964, as revised in 2004).

### Statement of human and animal rights

When reporting the results from experiments, authors should follow the highest standards and the trial should conform to Good Clinical Practice (for example, US Food and Drug Administration Good Clinical Practice in FDA-Regulated Clinical Trials; UK Medicines Research Council Guidelines for Good Clinical Practice in Clinical Trials) and/or the World Medical Association Declaration of Helsinki. Generally, we suggest authors follow the lead investigator's national standard. If doubt exists whether the research was conducted in accordance with the above standards, the authors must explain the rationale for their approach and demonstrate that the institutional review body explicitly approved the doubtful aspects of the study.

Before submitting, authors should make their study approved by the relevant research ethics committee or institutional review board. If human participants were involved, manuscripts must be accompanied by a statement that the experiments were undertaken with the understanding and appropriate informed consent of each. Any personal item or information will not be published without explicit consents from the involved patients. If experimental animals were used, the materials and methods (experimental procedures) section must clearly indicate that appropriate measures were taken to minimize pain or discomfort, and details of animal care should be provided.

## SUBMISSION OF MANUSCRIPTS

Manuscripts should be typed in 1.5 line spacing and 12 pt. Book Antiqua with ample margins. Number all pages consecutively, and start each of the following sections on a new page: Title Page, Abstract, Introduction, Materials and Methods, Results, Discussion, Acknowledgements, References, Tables, Figures, and Figure Legends. Neither the editors nor the publisher are responsible for the opinions expressed by contributors. Manuscripts formally accepted for publication become the permanent property of Baishideng Publishing Group Co., Limited, and may not be reproduced by any means, in whole or in part, without the written permission of both the authors and the publisher. We reserve the right to copy-edit and put onto our website accepted manuscripts. Authors should follow the relevant guidelines for the care and use of laboratory animals of their institution or national animal welfare committee. For the sake of transparency in regard to the performance and reporting of clinical trials, we endorse the policy of the ICMJE to refuse to publish papers on clinical trial results if the trial was not recorded in a publicly-accessible registry at its outset. The only register now available, to our knowledge, is <http://www.clinicaltrials.gov> sponsored by the United States National Library of Medicine and we encourage all potential contributors to register with it. However, in the case that other registers become available you will be duly notified.

A letter of recommendation from each author's organization should be provided with the contributed article to ensure the privacy and secrecy of research is protected.

Authors should retain one copy of the text, tables, photographs and illustrations because rejected manuscripts will not be returned to the author(s) and the editors will not be responsible for loss or damage to photographs and illustrations sustained during mailing.

### Online submissions

Manuscripts should be submitted through the Online Submission System at: <http://www.wjgnet.com/esps/>. Authors are highly recommended to consult the ONLINE INSTRUCTIONS TO AUTHORS ([http://www.wjgnet.com/2220-3184/g\\_info\\_20100722180909.htm](http://www.wjgnet.com/2220-3184/g_info_20100722180909.htm)) before attempting to submit online. For assistance, authors encountering problems with the Online Submission System may send an email describing the problem to [bpgoffice@wjgnet.com](mailto:bpgoffice@wjgnet.com), or by telephone: +86-10-85381892. If you submit your manuscript online, do not make a postal contribution. Repeated online submission for the same manuscript is strictly prohibited.

## MANUSCRIPT PREPARATION

All contributions should be written in English. All articles must be submitted using word-processing software. All submissions must be typed in 1.5 line spacing and 12 pt. Book Antiqua with ample margins. Style should conform to our house format. Required information for each of the manuscript sections is as follows:

### Title page

**Title:** Title should be less than 12 words.

**Running title:** A short running title of less than 6 words should be provided.

**Authorship:** Authorship credit should be in accordance with the standard proposed by ICMJE, based on (1) substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data; (2) drafting the article or revising it critically for important intellectual content; and (3) final approval of the version to be published. Authors should meet conditions 1, 2, and 3.

**Institution:** Author names should be given first, then the complete name of institution, city, province and postcode. For example, Xu-Chen Zhang, Li-Xin Mei, Department of Pathology, Chengde Medical College, Chengde 067000, Hebei Province, China. One author may be represented from two institutions, for example, George Sgourakis, Department of General, Visceral, and Transplantation Surgery, Essen 45122, Germany; George Sgourakis, 2nd Surgical Department, Korgialenio-Benakio Red Cross Hospital, Athens 15451, Greece

**Author contributions:** The format of this section should be: Author contributions: Wang CL and Liang L contributed equally to this work; Wang CL, Liang L, Fu JF, Zou CC, Hong F and Wu XM designed the research; Wang CL, Zou CC, Hong F and Wu XM performed the research; Xue JZ and Lu JR contributed new reagents/analytic tools; Wang CL, Liang L and Fu JF analyzed the data; and Wang CL, Liang L and Fu JF wrote the paper.

**Supportive foundations:** The complete name and number of supportive foundations should be provided, *e.g.*, Supported by National Natural Science Foundation of China, No. 30224801

**Correspondence to:** Only one corresponding address should be provided. Author names should be given first, then author title, affiliation, the complete name of institution, city, postcode, province, country, and email. All the letters in the email should be in lower case. A space interval should be inserted between country name and email address. For example, Montgomery Bissell, MD, Professor of Medicine, Chief, Liver Center, Gastroenterology Division, Uni-

versity of California, Box 0538, San Francisco, CA 94143, United States. [montgomery.bissell@ucsf.edu](mailto:montgomery.bissell@ucsf.edu)

**Telephone and fax:** Telephone and fax should consist of +, country number, district number and telephone or fax number, *e.g.*, Telephone: +86-10-85381892 Fax: +86-10-85381893

**Peer reviewers:** All articles received are subject to peer review. Normally, three experts are invited for each article. Decision on acceptance is made only when at least two experts recommend publication of an article. All peer-reviewers are acknowledged on Express Submission and Peer-review System website.

### Abstract

There are unstructured abstracts (no less than 200 words) and structured abstracts. The specific requirements for structured abstracts are as follows:

An informative, structured abstract should accompany each manuscript. Abstracts of original contributions should be structured into the following sections: AIM (no more than 20 words; Only the purpose of the study should be included. Please write the Aim in the form of "To investigate/study/..."), METHODS (no less than 140 words for Original Articles; and no less than 80 words for Brief Articles), RESULTS (no less than 150 words for Original Articles and no less than 120 words for Brief Articles; You should present *P* values where appropriate and must provide relevant data to illustrate how they were obtained, *e.g.*,  $6.92 \pm 3.86$  vs  $3.61 \pm 1.67$ ,  $P < 0.001$ ), and CONCLUSION (no more than 26 words).

### Key words

Please list 5-10 key words, selected mainly from *Index Medicus*, which reflect the content of the study.

### Core tip

Please write a summary of less than 100 words to outline the most innovative and important arguments and core contents in your paper to attract readers.

### Text

For articles of these sections, original articles and brief articles, the main text should be structured into the following sections: INTRODUCTION, MATERIALS AND METHODS, RESULTS and DISCUSSION, and should include appropriate Figures and Tables. Data should be presented in the main text or in Figures and Tables, but not in both.

### Illustrations

Figures should be numbered as 1, 2, 3, *etc.*, and mentioned clearly in the main text. Provide a brief title for each figure on a separate page. Detailed legends should not be provided under the figures. This part should be added into the text where the figures are applicable. Keeping all elements compiled is necessary in line-art image. Scale bars should be used rather than magnification factors, with the length of the bar defined in the legend rather than on the bar itself. File names should identify the figure and panel. Avoid layering type directly over shaded or textured areas. Please use uniform legends for the same subjects. For example: Figure 1 Pathological changes in atrophic gastritis after treatment. A: ...; B: ...; C: ...; D: ...; E: ...; F: ...; G: ... *etc.* It is our principle to publish high resolution-figures for the E-versions.

### Tables

Three-line tables should be numbered 1, 2, 3, *etc.*, and mentioned clearly in the main text. Provide a brief title for each table. Detailed legends should not be included under tables, but rather added into the text where applicable. The information should complement, but not duplicate the text. Use one horizontal line under the title, a second under column heads, and a third below the Table, above any footnotes. Vertical and italic lines should be omitted.

### Notes in tables and illustrations

Data that are not statistically significant should not be noted. \**P* <



## Instructions to authors

0.05, <sup>b</sup> $P < 0.01$  should be noted ( $P > 0.05$  should not be noted). If there are other series of  $P$  values, <sup>c</sup> $P < 0.05$  and <sup>d</sup> $P < 0.01$  are used. A third series of  $P$  values can be expressed as <sup>e</sup> $P < 0.05$  and <sup>f</sup> $P < 0.01$ . Other notes in tables or under illustrations should be expressed as <sup>1</sup>F, <sup>2</sup>F, <sup>3</sup>F; or sometimes as other symbols with a superscript (Arabic numerals) in the upper left corner. In a multi-curve illustration, each curve should be labeled with ●, ○, ■, □, ▲, △, etc., in a certain sequence.

### Acknowledgments

Brief acknowledgments of persons who have made genuine contributions to the manuscript and who endorse the data and conclusions should be included. Authors are responsible for obtaining written permission to use any copyrighted text and/or illustrations.

## REFERENCES

### Coding system

The author should number the references in Arabic numerals according to the citation order in the text. Put reference numbers in square brackets in superscript at the end of citation content or after the cited author's name. For citation content which is part of the narration, the coding number and square brackets should be typeset normally. For example, "Crohn's disease (CD) is associated with increased intestinal permeability<sup>[1,2]</sup>". If references are cited directly in the text, they should be put together within the text, for example, "From references<sup>[19,22-24]</sup>, we know that..."

When the authors write the references, please ensure that the order in text is the same as in the references section, and also ensure the spelling accuracy of the first author's name. Do not list the same citation twice.

### PMID and DOI

Please provide PubMed citation numbers to the reference list, e.g., PMID and DOI, which can be found at <http://www.ncbi.nlm.nih.gov/sites/entrez?db=pubmed> and <http://www.crossref.org/SimpleTextQuery/>, respectively. The numbers will be used in E-version of this journal.

### Style for journal references

Authors: the name of the first author should be typed in bold-faced letters. The family name of all authors should be typed with the initial letter capitalized, followed by their abbreviated first and middle initials. (For example, Lian-Sheng Ma is abbreviated as Ma LS, Bo-Rong Pan as Pan BR). The title of the cited article and italicized journal title (journal title should be in its abbreviated form as shown in PubMed), publication date, volume number (in black), start page, and end page [PMID: 11819634 DOI: 10.3748/wjg.13.5396].

### Style for book references

Authors: the name of the first author should be typed in bold-faced letters. The surname of all authors should be typed with the initial letter capitalized, followed by their abbreviated middle and first initials. (For example, Lian-Sheng Ma is abbreviated as Ma LS, Bo-Rong Pan as Pan BR) Book title. Publication number. Publication place: Publication press, Year: start page and end page.

### Format

#### Journals

English journal article (list all authors and include the PMID where applicable)

- 1 **Jung EM**, Clevert DA, Schreyer AG, Schmitt S, Rennert J, Kubale R, Feuerbach S, Jung F. Evaluation of quantitative contrast harmonic imaging to assess malignancy of liver tumors: A prospective controlled two-center study. *World J Gastroenterol* 2007; **13**: 6356-6364 [PMID: 18081224 DOI: 10.3748/wjg.13.6356]

Chinese journal article (list all authors and include the PMID where applicable)

- 2 **Lin GZ**, Wang XZ, Wang P, Lin J, Yang FD. Immunologic effect of Jianpi Yishen decoction in treatment of Pixu-diarhoea. *Shijie Huaren Xiaohua Zazhi* 1999; **7**: 285-287

In press

- 3 **Tian D**, Araki H, Stahl E, Bergelson J, Kreitman M. Signature

of balancing selection in Arabidopsis. *Proc Natl Acad Sci USA* 2006; In press

Organization as author

- 4 **Diabetes Prevention Program Research Group**. Hypertension, insulin, and proinsulin in participants with impaired glucose tolerance. *Hypertension* 2002; **40**: 679-686 [PMID: 12411462 PMID:2516377 DOI:10.1161/01.HYP.0000035706.28494.09]

Both personal authors and an organization as author

- 5 **Vallancien G**, Emberton M, Harving N, van Moorselaar RJ; Alf-One Study Group. Sexual dysfunction in 1, 274 European men suffering from lower urinary tract symptoms. *J Urol* 2003; **169**: 2257-2261 [PMID: 12771764 DOI:10.1097/01.ju.0000067940.76090.73]

No author given

- 6 21st century heart solution may have a sting in the tail. *BMJ* 2002; **325**: 184 [PMID: 12142303 DOI:10.1136/bmj.325.7357.184]

Volume with supplement

- 7 **Geraud G**, Spierings EL, Keywood C. Tolerability and safety of frovatriptan with short- and long-term use for treatment of migraine and in comparison with sumatriptan. *Headache* 2002; **42** Suppl 2: S93-99 [PMID: 12028325 DOI:10.1046/j.1526-4610.42.s2.7.x]

Issue with no volume

- 8 **Banit DM**, Kaufer H, Hartford JM. Intraoperative frozen section analysis in revision total joint arthroplasty. *Clin Orthop Relat Res* 2002; (**401**): 230-238 [PMID: 12151900 DOI:10.1097/00003086-200208000-00026]

No volume or issue

- 9 Outreach: Bringing HIV-positive individuals into care. *HRS-A Careaction* 2002; 1-6 [PMID: 12154804]

### Books

Personal author(s)

- 10 **Sherlock S**, Dooley J. Diseases of the liver and biliary system. 9th ed. Oxford: Blackwell Sci Pub, 1993: 258-296

Chapter in a book (list all authors)

- 11 **Lam SK**. Academic investigator's perspectives of medical treatment for peptic ulcer. In: Swabb EA, Azabo S. Ulcer disease: investigation and basis for therapy. New York: Marcel Dekker, 1991: 431-450

Author(s) and editor(s)

- 12 **Breedlove GK**, Schorffheide AM. Adolescent pregnancy. 2nd ed. Wiczorek RR, editor. White Plains (NY): March of Dimes Education Services, 2001: 20-34

Conference proceedings

- 13 **Harnden P**, Joffe JK, Jones WG, editors. Germ cell tumours V. Proceedings of the 5th Germ cell tumours Conference; 2001 Sep 13-15; Leeds, UK. New York: Springer, 2002: 30-56

Conference paper

- 14 **Christensen S**, Oppacher F. An analysis of Koza's computational effort statistic for genetic programming. In: Foster JA, Lutton E, Miller J, Ryan C, Tettamanzi AG, editors. Genetic programming. EuroGP 2002: Proceedings of the 5th European Conference on Genetic Programming; 2002 Apr 3-5; Kinsdale, Ireland. Berlin: Springer, 2002: 182-191

Electronic journal (list all authors)

- 15 Morse SS. Factors in the emergence of infectious diseases. *Emerg Infect Dis* serial online, 1995-01-03, cited 1996-06-05; 1(1): 24 screens. Available from: URL: <http://www.cdc.gov/ncidod/eid/index.htm>

Patent (list all authors)

- 16 **Pagedas AC**, inventor; Ancel Surgical R&D Inc., assignee. Flexible endoscopic grasping and cutting device and positioning tool assembly. United States patent US 20020103498. 2002 Aug 1

### Statistical data

Write as mean  $\pm$  SD or mean  $\pm$  SE.

### Statistical expression

Express *t* test as *t* (in italics), *F* test as *F* (in italics), chi square test as  $\chi^2$  (in Greek), related coefficient as *r* (in italics), degree of freedom as *v* (in Greek), sample number as *n* (in italics), and probability as *P* (in italics).

### Units

Use SI units. For example: body mass, *m* (B) = 78 kg; blood pressure, *p* (B) = 16.2/12.3 kPa; incubation time, *t* (incubation) = 96 h; blood glucose concentration, *c* (glucose)  $6.4 \pm 2.1$  mmol/L; blood CEA mass concentration, *p* (CEA) = 8.6 24.5  $\mu$ g/L; CO<sub>2</sub> volume fraction, 50 mL/L CO<sub>2</sub>, not 5% CO<sub>2</sub>; likewise for 40 g/L formaldehyde, not 10% formalin; and mass fraction, 8 ng/g, *etc.* Arabic numerals such as 23, 243, 641 should be read 23 243 641.

The format for how to accurately write common units and quantums can be found at: [http://www.wjgnet.com/2220-3184/g\\_info\\_20100725073806.htm](http://www.wjgnet.com/2220-3184/g_info_20100725073806.htm).

### Abbreviations

Standard abbreviations should be defined in the abstract and on first mention in the text. In general, terms should not be abbreviated unless they are used repeatedly and the abbreviation is helpful to the reader. Permissible abbreviations are listed in Units, Symbols and Abbreviations: A Guide for Biological and Medical Editors and Authors (Ed. Baron DN, 1988) published by The Royal Society of Medicine, London. Certain commonly used abbreviations, such as DNA, RNA, HIV, LD50, PCR, HBV, ECG, WBC, RBC, CT, ESR, CSF, IgG, ELISA, PBS, ATP, EDTA, mAb, can be used directly without further explanation.

### Italics

Quantities: *t* time or temperature, *c* concentration, *A* area, *l* length, *m* mass, *V* volume.

Genotypes: *gyrA*, *arg 1*, *c myc*, *c fos*, *etc.*

Restriction enzymes: *EcoRI*, *HindI*, *BamHI*, *Kho I*, *Kpn I*, *etc.*

Biology: *H. pylori*, *E. coli*, *etc.*

### Examples for paper writing

All types of articles' writing style and requirement will be found in the link: <http://www.wjgnet.com/esps/NavigationInfo.aspx?id=15>

## RESUBMISSION OF THE REVISED MANUSCRIPTS

Authors must revise their manuscript carefully according to the revision policies of Baishideng Publishing Group Co., Limited. The revised version, along with the signed copyright transfer agreement,

responses to the reviewers, and English language Grade A certificate (for non-native speakers of English), should be submitted to the online system *via* the link contained in the e-mail sent by the editor. If you have any questions about the revision, please send e-mail to [esps@wjgnet.com](mailto:esps@wjgnet.com).

### Language evaluation

The language of a manuscript will be graded before it is sent for revision. (1) Grade A: priority publishing; (2) Grade B: minor language polishing; (3) Grade C: a great deal of language polishing needed; and (4) Grade D: rejected. Revised articles should reach Grade A.

### Copyright assignment form

Please download a Copyright assignment form from [http://www.wjgnet.com/2220-3184/g\\_info\\_20100725073726.htm](http://www.wjgnet.com/2220-3184/g_info_20100725073726.htm).

### Responses to reviewers

Please revise your article according to the comments/suggestions provided by the reviewers. The format for responses to the reviewers' comments can be found at: [http://www.wjgnet.com/2220-3184/g\\_info\\_20100725073445.htm](http://www.wjgnet.com/2220-3184/g_info_20100725073445.htm).

### Proof of financial support

For papers supported by a foundation, authors should provide a copy of the approval document and serial number of the foundation.

### Statement about anonymous publication of the peer reviewers' comments

In order to increase the quality of peer review, push authors to carefully revise their manuscripts based on the peer reviewers' comments, and promote academic interactions among peer reviewers, authors and readers, we decide to anonymously publish the reviewers' comments and author's responses at the same time the manuscript is published online.

## PUBLICATION FEE

WJMG is an international, peer-reviewed, OA online journal. Articles published by this journal are distributed under the terms of the Creative Commons Attribution Non-commercial License, which permits use, distribution, and reproduction in any medium and format, provided the original work is properly cited. The use is non-commercial and is otherwise in compliance with the license. Authors of accepted articles must pay a publication fee. Publication fee: 600 USD per article. All invited articles are published free of charge.



# World Journal of *Medical Genetics*

*World J Med Genet* 2014 May 27; 4(2): 6-45





Contents

Quarterly Volume 4 Number 2 May 27, 2014

- |                        |    |   |
|------------------------|----|---|
| <b>TOPIC HIGHLIGHT</b> | 6  | Structure-function relationship in viral RNA genomes: The case of hepatitis C virus<br><i>Romero-López C, Berzal-Herranz A</i>  |
| <b>REVIEW</b>          | 19 | Molecular epidemiology of hepatitis B virus in Asia<br><i>Utsumi T, Yano Y, Hotta H</i>   |
|                        | 27 | Acquisition and dissemination mechanisms of CTXΦ in <i>Vibrio cholerae</i> : New paradigm for <i>dif</i> residents<br><i>Das B, Nair GB, Bhadra RK</i>                  |
| <b>MINIREVIEWS</b>     | 34 | Osteopathia striata with cranial sclerosis, Wilms' tumor and the <i>WTX</i> gene<br><i>Cattaneo E, Ciceri S, Liberati N, Radice P, Tarani L, Selicorni A, Perotti D</i> |
| <b>RESEARCH REPORT</b> | 39 | Earlier onset and multiple primaries in familial as opposed to sporadic esophageal cancer<br><i>Wen XD, Wen DG, Yang Y, Shan BE, Wang SJ</i>                            |

## Contents

*World Journal of Medical Genetics*  
Volume 4 Number 2 May 27, 2014

**APPENDIX** I-V Instructions to authors

**ABOUT COVER** Editorial Board Member of *World Journal of Medical Genetics*, Daniela Perotti, PhD, Scientist, Unit of Molecular bases of genetic risk and genetic testing, Department of Preventive and Predictive Medicine, Fondazione IRCCS Istituto Nazionale dei Tumori, AMADEOLAB, Via Amadeo 42, 20133 Milano, Italy

**AIM AND SCOPE** *World Journal of Medical Genetics* (*World J Med Genet*, *WJMG*, online ISSN 2220-3184, DOI: 10.5496) is a peer-reviewed open access academic journal that aims to guide clinical practice and improve diagnostic and therapeutic skills of clinicians.

*WJMG* covers topics concerning genes and the pathology of human disease, molecular analysis of simple and complex genetic traits, cancer genetics, epigenetics, gene therapy, developmental genetics, regulation of gene expression, strategies and technologies for extracting function from genomic data, pharmacological genomics, genome evolution. The current columns of *WJMG* include editorial, frontier, diagnostic advances, therapeutics advances, field of vision, mini-reviews, review, topic highlight, medical ethics, original articles, case report, clinical case conference (Clinicopathological conference), and autobiography.

We encourage authors to submit their manuscripts to *WJMG*. We will give priority to manuscripts that are supported by major national and international foundations and those that are of great basic and clinical significance.

**INDEXING/ABSTRACTING** *World Journal of Medical Genetics* is now indexed in Digital Object Identifier.

**FLYLEAF** I-IV Editorial Board

## EDITORS FOR THIS ISSUE

Responsible Assistant Editor: *Xiang Li*  
Responsible Electronic Editor: *Su-Qing Liu*  
Proofing Editor-in-Chief: *Lian-Sheng Ma*

Responsible Science Editor: *Fang-Fang Ji*  
Proofing Editorial Office Director: *Xiu-Xia Song*

**NAME OF JOURNAL**  
*World Journal of Medical Genetics*

**ISSN**  
ISSN 2220-3184 (online)

**LAUNCH DATE**  
December 27, 2011

**FREQUENCY**  
Quarterly

**EDITOR-IN-CHIEF**  
**Hans van Bokhoven, Professor, PhD**, Department of Human Genetics and Cognitive Neurosciences, Radboud university Nijmegen Medical centre, PO Box 9101, 6500 HB Nijmegen, The Netherlands

**EDITORIAL OFFICE**  
Jin-Lei Wang, Director  
Xiu-Xia Song, Vice Director  
*World Journal of Medical Genetics*

Room 903, Building D, Ocean International Center, No. 62 Dongsihuan Zhonglu, Chaoyang District, Beijing 100025, China  
Telephone: +86-10-59080039  
Fax: +86-10-85381893  
E-mail: editorialoffice@wjgnet.com  
Help Desk: <http://www.wjgnet.com/esp/helpdesk.aspx>  
<http://www.wjgnet.com>

**PUBLISHER**  
Baishideng Publishing Group Inc  
8226 Regency Drive,  
Pleasanton, CA 94588, USA  
Telephone: +1-925-223-8242  
Fax: +1-925-223-8243  
E-mail: bpgoffice@wjgnet.com  
Help Desk: <http://www.wjgnet.com/esp/helpdesk.aspx>  
<http://www.wjgnet.com>

**PUBLICATION DATE**  
May 27, 2014

## COPYRIGHT

© 2014 Baishideng Publishing Group Inc. Articles published by this Open Access journal are distributed under the terms of the Creative Commons Attribution Non-commercial License, which permits use, distribution, and reproduction in any medium, provided the original work is properly cited, the use is non commercial and is otherwise in compliance with the license.

## SPECIAL STATEMENT

All articles published in journals owned by the Baishideng Publishing Group (BPG) represent the views and opinions of their authors, and not the views, opinions or policies of the BPG, except where otherwise explicitly indicated.

## INSTRUCTIONS TO AUTHORS

Full instructions are available online at [http://www.wjgnet.com/2220-3184/g\\_info\\_20100722180909.htm](http://www.wjgnet.com/2220-3184/g_info_20100722180909.htm)

## ONLINE SUBMISSION

<http://www.wjgnet.com/esp/>

Alfredo Berzal-Herranz, Granada Alfredo Berzal-Herranz, PhD, Series Editor

## Structure-function relationship in viral RNA genomes: The case of hepatitis C virus

Cristina Romero-López, Alfredo Berzal-Herranz

Cristina Romero-López, Alfredo Berzal-Herranz, Instituto de Parasitología y Biomedicina López-Neyra, IPBLN-CSIC, PTS Granada, Armilla, 18016 Granada, Spain

Author contributions: Romero-López C and Berzal-Herranz A wrote the paper.

Supported by Spanish Ministry of Economy and Competitiveness, No. BFU2012-31213; Junta de Andalucía, No. CVI-7430; and FEDER funds from the EU

Correspondence to: Cristina Romero-López, PhD, Instituto de Parasitología y Biomedicina López-Neyra, IPBLN-CSIC, PTS Granada, Avda. del Conocimiento s/n, Armilla, 18016 Granada, Spain. [cristina\\_romero@ipb.csic.es](mailto:cristina_romero@ipb.csic.es)

Telephone: +34-958-181648 Fax: +34-958-181632

Received: December 10, 2013 Revised: January 23, 2014

Accepted: April 3, 2014

Published online: May 27, 2014

### Abstract

The acquisition of a storage information system beyond the nucleotide sequence has been a crucial issue for the propagation and dispersion of RNA viruses. This system is composed by highly conserved, complex structural units in the genomic RNA, termed functional RNA domains. These elements interact with other regions of the viral genome and/or proteins to direct viral translation, replication and encapsidation. The genomic RNA of the hepatitis C virus (HCV) is a good model for investigating about conserved structural units. It contains functional domains, defined by highly conserved structural RNA motifs, mostly located in the 5'-untranslatable regions (5'UTRs) and 3'UTR, but also occupying long stretches of the coding sequence. Viral translation initiation is mediated by an internal ribosome entry site located at the 5' terminus of the viral genome and regulated by distal functional RNA domains placed at the 3' end. Subsequent RNA replication strongly depends on the 3'UTR folding and is also influenced by the 5' end

of the HCV RNA. Further increase in the genome copy number unleashes the formation of homodimers by direct interaction of two genomic RNA molecules, which are finally packed and released to the extracellular medium. All these processes, as well as transitions between them, are controlled by structural RNA elements that establish a complex, direct and long-distance RNA-RNA interaction network. This review summarizes current knowledge about functional RNA domains within the HCV RNA genome and provides an overview of the control exerted by direct, long-range RNA-RNA contacts for the execution of the viral cycle.

© 2014 Baishideng Publishing Group Inc. All rights reserved.

**Key words:** Functional RNA domain; Cis-acting replicating element; Hepatitis C virus; Internal ribosome entry site; RNA-RNA interaction; Untranslatable region

**Core tip:** This review summarizes the main aspects of structurally conserved genomic RNA elements in the hepatitis C virus (HCV) genome and their role in the viral cycle. The genome of RNA viruses is a dynamic genetic entity endorsed with an information storage system defined by highly conserved, complex structural units, termed functional RNA domains. The genome of HCV contains several well-studied functional RNA domains that control essential viral processes, as well as transitions between them, by recruiting protein factors and also by establishing a complex, direct and long-range RNA-RNA interaction network.

Romero-López C, Berzal-Herranz A. Structure-function relationship in viral RNA genomes: The case of hepatitis C virus. *World J Med Genet* 2014; 4(2): 6-18 Available from: URL: <http://www.wjmgnet.com/2220-3184/full/v4/i2/6.htm> DOI: <http://dx.doi.org/10.5496/wjmg.v4.i2.6>



## INTRODUCTION

The genomes of RNA viruses are not passive elements. The inherent high error rate of the viral polymerase during replication provides an important evolutive advantage by the generation of genotypically and phenotypically different virus pools on which natural selection operates<sup>[1]</sup>. By using this strategy, viruses have got RNA genomes with numerous signals overlapping protein coding sequences, thus achieving multiple levels of regulation throughout the infectious cycle. All this information is compactly packed in a minimal size for optimal propagation. Viral RNA genomes use a information storage system beyond the nucleotide sequence, defined by highly conserved regions that exhibit complex folding and play direct, functional roles in the viral cycle<sup>[2-4]</sup>. Two levels of structure or folding can be distinguished within an RNA molecule: (1) the secondary structure involves double and single stranded regions arrangements; and (2) the tertiary structure is determined by the relationships established between secondary structure elements. The combination of both conformational levels establishes the final shape of the RNA to generate the so-called functional RNA domains. These are dynamic elements since their structure can be selectively adopted from a wide variety of possible foldings to execute a specific function by recruiting protein factors, or modulating the conformation and function of distant regulatory elements<sup>[5]</sup>. These mechanisms achieve an active control of the gene expression. Therefore, RNA folding acts as a regulatory machine to diversify RNA genome functions with a minimal size.

Functional RNA domains are typically identified as one or more stem-loops with highly conserved sequence motifs located in the loops. These elements were initially described located in the 5'-untranslatable regions (5' UTRs) and the 3'UTRs of viral genomes, but now evidences are accumulating for their widespread distribution throughout the entire genomic RNA<sup>[5]</sup>. They can be organized, either as well-defined, phylogenetically conserved RNA structural motifs, or as sets of extensive folded regions throughout the whole viral genome [genome-scale ordered RNA structures (GORS)], following a clear structural pattern that may change even between closely related viruses<sup>[6,7]</sup>.

The recent advent of novel bioinformatic tools and experimental techniques to probe and study RNA structure has provided high-resolution pictures of numerous viral RNA molecules. Among them, structural elements of the hepatitis C virus (HCV) genomic RNA are one of the best characterized from many different viruses. HCV infection affects to more than 3% of the world population, with high incidence of cirrhosis, hepatic steatosis and hepatocellular carcinoma. To date, no efficient vaccines have been developed against HCV and current treatments based on pegylated-interferon  $\alpha$  and ribavirin are the standard of care (SOC) regimen with a limited efficacy of around 40% of the patients. Additionally, this therapy has important side effects. Recently, two direct-acting antiviral drugs targeting the viral protease

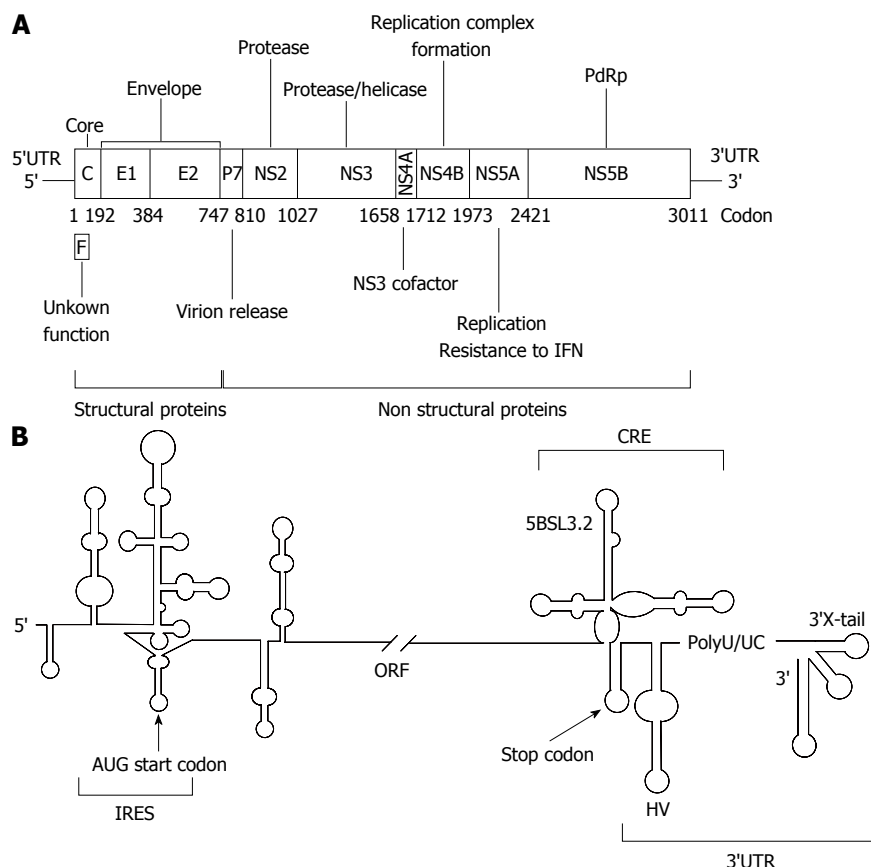
NS3, telaprevir and boceprevir, have been approved by the United States Food and Drug Administration<sup>[8]</sup>. These compounds can be administered in conjunction with pegylated-interferon  $\alpha$  and ribavirin for a short period of time to achieve an improved sustained virological response<sup>[8]</sup> with respect to the SOC. Unfortunately, prolonged treatments lead to the appearance of resistant variants. Other drugs targeting either the protease NS3 (simeprevir) or the viral polymerase NS5B (sofosbuvir) are currently being tested in Phase II / III clinical trials.

HCV belongs to the *Flaviviridae* family, which includes yellow fever virus, bovine diarrhea virus and dengue virus. The HCV genome shows such a variability that up to six different genotypes, with hundreds of subtypes and isolates, have been identified<sup>[9,10]</sup>. Viral genotype clearly affects the success of interferon therapy, although no clear correlation with virulence exists. Further, the HCV population infecting a patient is structured in terms of quasi-species. This term defines the closely related sequences of a heterogeneous viral population infecting a single individual<sup>[11]</sup>. Quasispecies structure has been associated with the failure of infected people to clear the virus and the subsequent development of a chronic infection<sup>[12]</sup>. Therefore, the identification of conserved therapeutic targets and the search for fully effective antiviral compounds is a major goal of HCV research. The functional importance of genomic structural elements for virus persistence and their high conservation rate suggests they might make good therapeutic targets. This review focuses in the main structural features of the HCV genomic RNA functional domains and their roles in the viral cycle.

## HCV RNA GENOME ACTIVELY CONTROLS THE INFECTIVE CYCLE

The HCV genome is about 9600 nucleotides-long, single stranded positive RNA molecule<sup>[13-15]</sup> that encodes for a single open reading frame (ORF) flanked by two highly conserved UTRs (5'UTR and 3'UTR) (Figure 1A). The viral genome controls important processes of the infective cycle. During early infection, the genome acts as mRNA to generate the viral structural (core protein C, p7 and the envelope proteins E1 and E2) and non-structural proteins (NS2, NS3, NS4A, NS4B, NS5A and NS5B). HCV translation is initiated by an internal ribosome entry site (IRES)-dependent mechanism<sup>[16,17]</sup> different to the cap-dependent method used for most cellular mRNAs. The IRES element is mostly located at the 5'UTR and spans a short stretch of the core coding sequence<sup>[18,19]</sup> (Figure 1B). Both the initiation translation step and the subsequent elongation phase are influenced by the presence of domains located at the 3' end of the HCV genome<sup>[20-25]</sup>. This process is dependent on the acquisition of a circular topology resembling the closed-loop structure adopted by cellular cap-mRNAs. Such architecture is achieved by both the recruitment of protein factors, able to simultaneously bind to the 5'UTRs and 3'UTRs of the genomic HCV RNA<sup>[20-23,26-28]</sup>, and also by the establish-





**Figure 1 Genetic organization of the hepatitis C virus genomic RNA.** A: A schematic view of the hepatitis C virus (HCV) genome, showing the 5' and 3' untranslated regions (UTRs) and the genes encoding for the different viral proteins. Numbers allude to codon positions; B: A detailed diagram of the secondary structure proposed for the 5' and 3' ends is pictured. The region required for internal ribosome entry site (IRES) activity is indicated. The 3' end of the viral genomic RNA is organized into two well-defined structural elements: the cis-acting replicating element region containing the essential domain 5BSL3.2, and the 3'X-tail, separated by a hypervariable sequence (HV) and a polyU/UC tract. Start and stop translation codons, placed at positions 342 and 9371, respectively, are indicated by arrows. Numbers refer to the aminoacid positions according to the HCV Con1 isolate (GenBank accession number AJ238799). RdRp: RNA-dependent RNA polymerase; IFN: interferon; ORF: Open reading frame.

ment of direct, long-range RNA-RNA interactions<sup>[29-31]</sup>. Once viral proteins levels have reached a certain threshold, the genomic RNA serves as a template to initiate replication at the 3'UTR in a structure dependent manner. This process is also influenced by the 5' end of the HCV RNA<sup>[32,33]</sup>. The accumulation of viral genomes enhances the formation of homodimers by the interaction of two viral RNA molecules in the presence of the core chaperone protein<sup>[34-37]</sup>. Packaged genomic RNA is finally enveloped and released to the extracellular environment.

The maintenance of a proper balance between these processes involves fine regulation mechanisms, which involve the interplay of functional RNA domains located throughout the entire ORF<sup>[38,39]</sup>. The 5' core coding sequence helps in the preservation of structures important for IRES activity and replication (Figure 1B)<sup>[19,40-43]</sup>. Within the 3' end of the NS5B coding sequence, the stem-loop 5BSL3.2 is embedded in a cruciform structure that has been identified as a *cis*-essential element for viral RNA synthesis [cis-acting replicating element (CRE)] (Figure 1B)<sup>[44,45]</sup> and as a regulatory partner of the IRES function<sup>[25]</sup>.

An interesting feature of all these functional RNA domains is that they do not operate only by recruiting

protein factors. Instead, they establish a complex and dynamic network of contacts, which fits viral necessities to promote the consecution of different steps of the viral cycle, as well as the switch between them. Furthermore, this interacting web provides important benefits, such as minimizing protein requisites.

Next sections will outline the current knowledge about different HCV functional RNA domains and their involvement in the complex interaction network that governs the initiation of essential viral events and the transitions between them.

## THE HCV IRES REGION

The initiation of the HCV protein synthesis is driven by the high affinity interaction IRES-40S<sup>[46-48]</sup>. This primary contact promotes conformational changes that directly clamp the viral RNA to the ribosomal subunit and thus position the appropriate start codon in the P site<sup>[49]</sup>. The further binding of eIF3 aids the incorporation of the ternary complex eIF2-GTP-tRNA<sup>Met</sup> to yield the 48S particle<sup>[48,50]</sup>. The formation of the active translation complex 80S is assessed by the GTP hydrolysis for the concurrent release of eIF2 and eIF3<sup>[51]</sup> and the final joining of the

60S subunit. It is noteworthy that this mechanism is primarily accomplished by functional RNA domains, thus minimizing protein factor requirements and simplifying the pathway for the assembly of the fully active ribosome.

The secondary structure of the HCV IRES region was originally proposed by Brown *et al.*<sup>[52]</sup> and latter refined to include several new motifs and interactions. Under physiological magnesium conditions, the HCV IRES folds into two major domains with well defined functions (II and III; Figure 2)<sup>[53]</sup>, plus a short stem-loop containing the start codon (domain IV)<sup>[54]</sup>. Rather than forming a tightly packed element, domains II and III are extended and aligned at both sides of a complex double pseudoknot structure (PK1 and PK2; Figure 2)<sup>[49,55]</sup>. The 3D architecture and several single RNA structural elements are highly conserved among other closely related viruses from the *Flaviviridae* family<sup>[46,56,57]</sup>.

Domain II is an autonomously folded module composed of two short helical segments, the basal subdomain IIa and the apical subdomain IIb, separated by a highly conserved internal E-loop<sup>[56,58]</sup> and capped by an apical loop (Figure 2). Domain II adopts an overall distorted L-shape conformation<sup>[59]</sup> because of the twist forced by the internal E-loop. This folding is conserved in HCV and related viruses<sup>[51]</sup>.

While domain II is not essential for 40S recruitment<sup>[47,60,61]</sup>, it has been shown that its deletion decreases viral protein synthesis yield up to five-fold by blocking the formation of the translationally active 80S complex<sup>[46,48,60,62,63]</sup>. Analysis by cryo-EM have demonstrated that the bend in domain II is a requisite for changing the conformation of the 40S ribosomal subunit<sup>[49,64]</sup>, in a reminiscent manner to that shown by eIF1 in the canonical cap-dependent translation initiation mechanism<sup>[65]</sup>. The apical loop placed in subdomain IIb would also contribute to this structural rearrangement in the ribosome<sup>[66]</sup>. Remarkably, all these conformational reorganization events do not only account on ribosomal proteins but also on the 18S rRNA. This could be the result of the coordinated action mediated by long-distant contacts established between domains II and IV<sup>[67,68]</sup>. Ribosome folding rearrangements further induce eIF2-GTP hydrolysis, triggering the release of protein factors and the recruitment of the 60S subunit to constitute the 80S complex<sup>[48,51,63,64]</sup>.

The large, highly branched domain III consists of six hairpins (designated subdomains IIIa to IIIf) organized around three- and four-way junctions (Figure 2)<sup>[52]</sup>, which can be identified as recruiting centers for the translational machinery. The apical IIIabc junction is the platform for the binding of eIF3<sup>[50,69]</sup>. The main goal of this interaction seems to be the relief of the competition between eIF3 and the IRES for a common site in the 40S ribosomal subunit, as well as avoiding the formation of canonical 43S translational complexes<sup>[70]</sup>. This assesses that HCV mRNA translation is specially favored over that of host mRNAs.

The middle section of domain III is defined by a three-way junction that contains the essential G-rich

subdomain III d (Figure 2). This element is the core 40S binding center<sup>[60,71,72]</sup>. Its structure is that of a dynamic stem-loop with an internal E loop motif and an apical loop with typical U-turn geometry<sup>[73,74]</sup>. This architecture exposes the bases placed in the apical loop and favors their interaction with viral and host ligands, both nucleic acids and proteins. Further, the subdomain III d seems to be a determinant partner in the acquisition of the functional folding of the surrounding domains<sup>[53]</sup>.

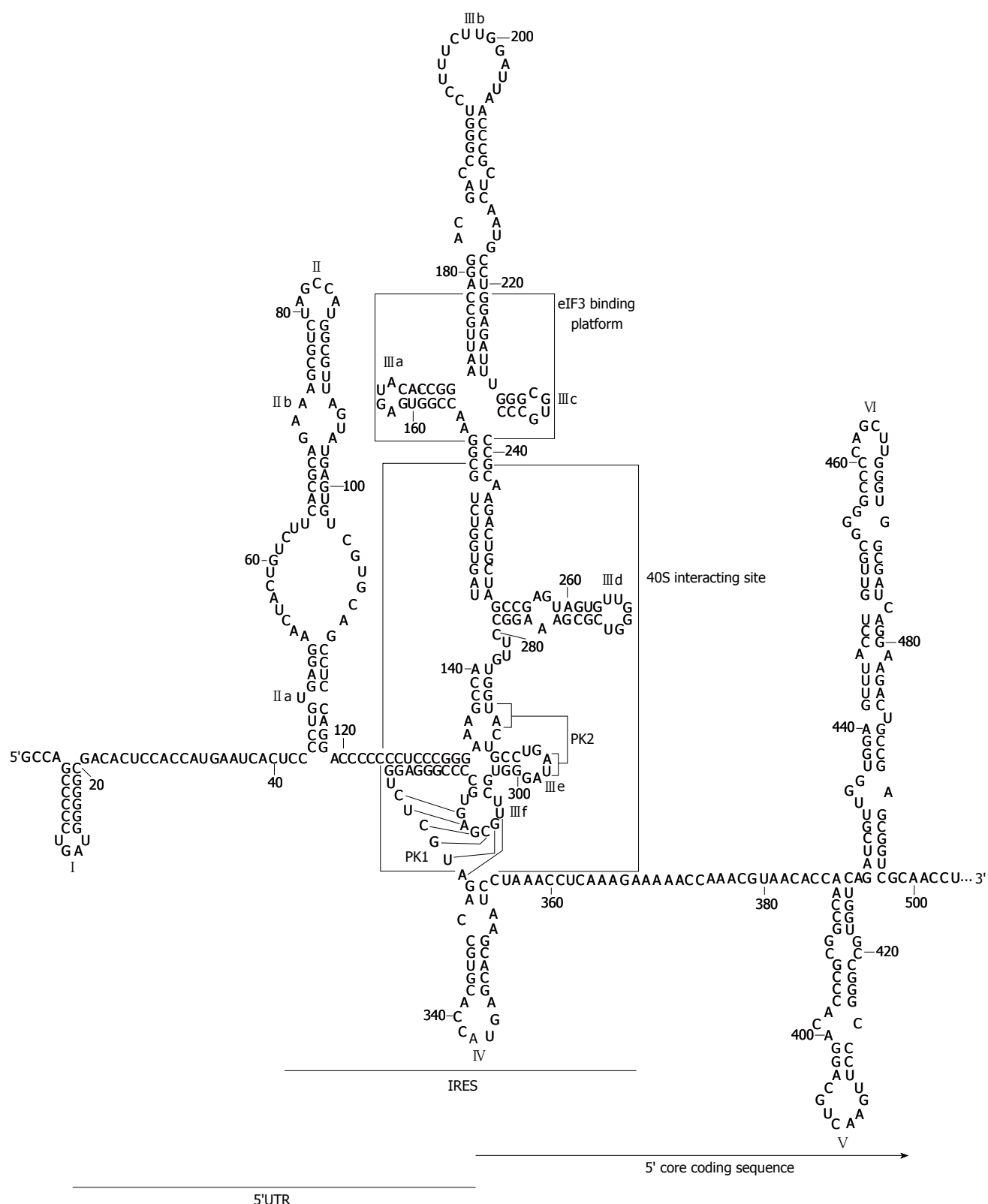
The basal fragment of domain III (subdomains IIIe and IIIf) includes the highly conserved, complex double-pseudoknot motif (PK1 and PK2; Figure 2)<sup>[55]</sup>, which defines a four-way junction to constrain the position of the AUG codon at the P-site of the 40S ribosomal subunit. Remarkably, the spatial distance between the pseudoknot and the AUG firmly resembles to that observed between the canonical Shine-Dalgarno motif and the initiation codon in prokaryotic mRNAs<sup>[75]</sup>. As noted above, the structural element PK1-PK2 also guides domains II and IIIabc in an extended conformation to get the easy access of protein factors.

Domain IV exposes the AUG start codon, at nucleotide 342, in an apical loop enclosing a helical motif (Figure 2). This structure is not conserved in other HCV-like IRESs<sup>[76]</sup>. In fact, the stem must be unwound to allow for the recognition of the AUG codon, which could entail some disadvantages. This is in good agreement with data demonstrating that the stability of stem-loop IV is inversely correlated to IRES translational efficiency<sup>[54]</sup>.

Therefore, the HCV IRES is defined by a set of RNA domains that replace the functions played by many host factors to provide a simplified way for the initiation of the viral proteins synthesis. Moreover, these functional domains are able to manipulate the translational machinery to assess the preferential reading of the HCV mRNA.

## ESSENTIAL 3'UTR

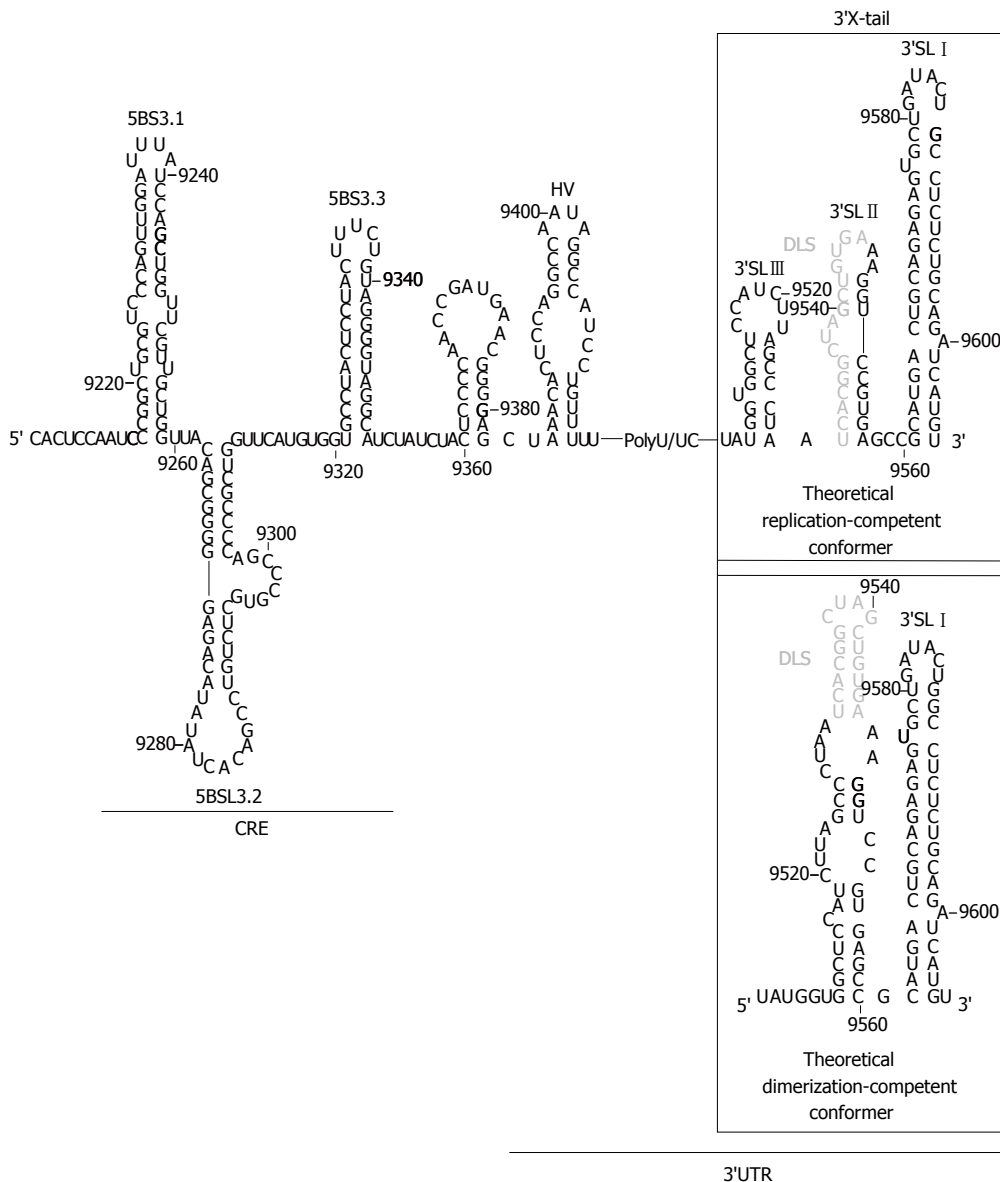
The HCV 3'UTR is of primary importance for the initiation of the minus RNA strand synthesis during the viral replication step<sup>[33,77,78]</sup> and also may act as enhancer of the IRES function<sup>[20-24]</sup>. It is about 240 nts long sequence placed at the 3' end of the viral genome<sup>[79]</sup>, with evolutionarily conserved secondary structure elements that define three functionally and conformationally independent modules (Figure 3). From 5' to 3': (1) A poorly conserved sequence of around 40 nts, termed hypervariable region at the 5' end of the 3'UTR. It folds as a single stem-loop, which is not completely required for viral replication<sup>[33,78]</sup>; (2) A polyU/UC tract, whose length and composition is a critical determinant of efficient HCV replication in cell culture<sup>[80]</sup>. It has been proved that a minimum of 26 U nts homopolymer is enough for efficient amplification of the viral RNA<sup>[33,78]</sup>. Further, it can act outside of its usual molecular context, thus suggesting that this is not only a linker region<sup>[80]</sup>. The polyU/UC stretch also interacts with host factors related to cellular protein synthesis, such as polypyrimidine tract-binding protein<sup>[81,82]</sup>, the La autoantigen<sup>[83]</sup>, heterogeneous nuclear ribonucleoprotein C (Gon-



**Figure 2** Secondary structure proposed for the hepatitis C virus internal ribosome entry site. The 5' untranslatable region (UTR) plus domains V and VI located at the core coding sequence are included. Minimum region for internal ribosome entry site (IRES) activity is depicted. Domains involved in the interaction with eIF3 factor and ribosomal subunit 40S are marked in boxes. Pseudoknot elements are indicated as PK1 and PK2. The translation start codon is shown in bold. Nucleotide numbering corresponds to hepatitis C virus con1 isolate.

tarek, 1999 #1925) and glyceraldehyde-3-phosphate dehydrogenase<sup>[84]</sup>, among others<sup>[85,86]</sup>. It seems likely that the recruitment of these factors could contribute to regulate viral translation mediated by the IRES region<sup>[20,21]</sup>; and (3)

The 3'X-tail is a highly conserved, 98-nts long sequence, located at the 3' termini of the HCV genome. It theoretically folds into two alternate and mutually exclusive conformations<sup>[35]</sup> (Figure 3). Both predicted structures pre-



**Figure 3 Theoretical folding of the 3' end of the hepatitis C virus genomic RNA.** The figure shows the entire 3' end containing the cis-acting replicating element (CRE) region plus the 3' untranslated region (UTR). The 3'X-tail folds into two different conformers with distinct functional roles. Dimer linkage sequence (DLS) is shown in grey. Translation stop codon in position 9371 is indicated in bold. Nucleotide numbering is as noted in Figure 2.

serve the essential 3'SL I placed at the very 3' end, which has important implications for the initiation and specificity of the viral RNA replication<sup>[87-89]</sup>. The 55 nts segment placed upstream of 3'SL I folds either as two stem-loops, named 3'SL II and 3'SL III, or as a single stem-loop exposing a 16 nts long palindromic sequence [dimer linkage sequence (DLS)] (Figure 3). Both conformers assume different functionalities during the HCV cycle and are therefore related to transitions between different steps of the viral infection.

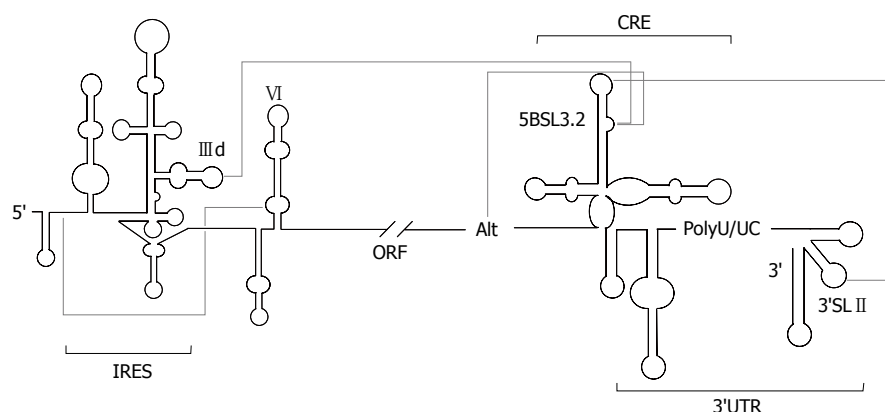
The molecular basis of the 3'UTR functioning are not well understood. Several reports have described the binding of both viral and host factors to the different structural elements of the 3'UTR<sup>[81,90-97]</sup>, but these findings do not provide completely satisfactory explanations for many of the experimental observations. The involvement of the 3'UTR in a long range RNA-RNA interaction

network with other genomic elements would likely fill the gaps in the complex functioning of this region<sup>[30,31,33,98,99]</sup>.

## FUNCTIONAL RNA DOMAINS WITHIN THE CODING SEQUENCE

### *Cis-acting elements within the core coding sequence*

Advances in novel bioinformatic tools have allowed for the extensive search of evolutionarily conserved RNA domains, resulting in the identification of domains distinct from those present in the UTRs. Comparative analyses of numerous HCV isolates sequences revealed an unusual high degree of conservation in the 5' end of the core protein coding sequence<sup>[100]</sup>. Interestingly, this conservation could not be explained only by the preservation of the amino acid sequence since synonymous substitutions were suppressed. This finding entails a functional



**Figure 4 Long-range RNA-RNA interactions in the hepatitis C virus genome.** Detailed diagram of the interacting network in the genomic hepatitis C virus RNA. The minimum region for internal ribosome entry site (IRES) activity is marked. The 3' end of the viral genome, containing both the cis-acting replicating element (CRE) and the 3' untranslatable region (UTR), is included. Functional RNA domains involved in the establishment of long-distant contacts are indicated. Figure adapted from<sup>[6]</sup>.

constrain that was related to the presence of an alternative ORF coding for the so-called protein F<sup>[101,102]</sup> (Figure 1A); and to the existence of structural RNA domains with functional roles in the HCV cycle<sup>[38]</sup> (Figure 2). While the production and biological role of protein F is still a controversial issue<sup>[103]</sup>, it has been demonstrated that the 5' core coding sequence folds as two stem-loops (domains V and VI; Figure 2) important for IRES activity and viral replication<sup>[19,40,42,43]</sup>.

The mechanism of action of domains V and VI is unclear. It has been proposed their participation in a long-range RNA-RNA interaction involving nucleotides 24-38 of the linker region between domains I and II in the 5'UTR, and 428-442 placed in domain VI (Figure 4)<sup>[40,41]</sup>. This contact would render a locked conformation of the IRES, which could be released by the interaction of the liver-specific microRNA miR-122 with nucleotides 22-28 of the HCV RNA<sup>[104]</sup>. This hypothesis provides a mechanism for the involvement of domain VI in viral translation, as well as supporting the essential role of miR-122 in HCV infection<sup>[105,106]</sup>. Alternatively, Roberts *et al*<sup>[107]</sup> found that viral translation regulation mediated by miR-122 is strictly dependent on Argonaute proteins and does not involve the structural transition previously proposed. It should be noted that investigations were performed with different experimental tools and model systems. Hence, it is not possible to discard any of the proposals; neither they are mutually exclusive in a cellular context.

#### **Cis-acting replicating element in the viral RNA polymerase coding sequence**

In addition to the core coding sequence, the 3' end of the HCV ORF also harbors evolutionarily conserved structural RNA elements. Up to six different stem-loop motifs have been identified by using a combination of sequence alignment and thermodynamic folding softwares, as well as classical comparative analysis<sup>[38,39,108,109]</sup>. One of these structural elements, the so-called domain 5BSL3.2 or SL9266, is embedded into a cruciform structure delimit-

ed by two adjacent stem-loops, 5BSL3.1 and 5BSL3.3 (CRE, Figure 3). While the essentiality of 5BSL3.2 for virus replication has been largely demonstrated<sup>[44,45,98,109]</sup>, the role of the two additional domains 5BSL3.1 and 3.3 is still unclear<sup>[80]</sup>.

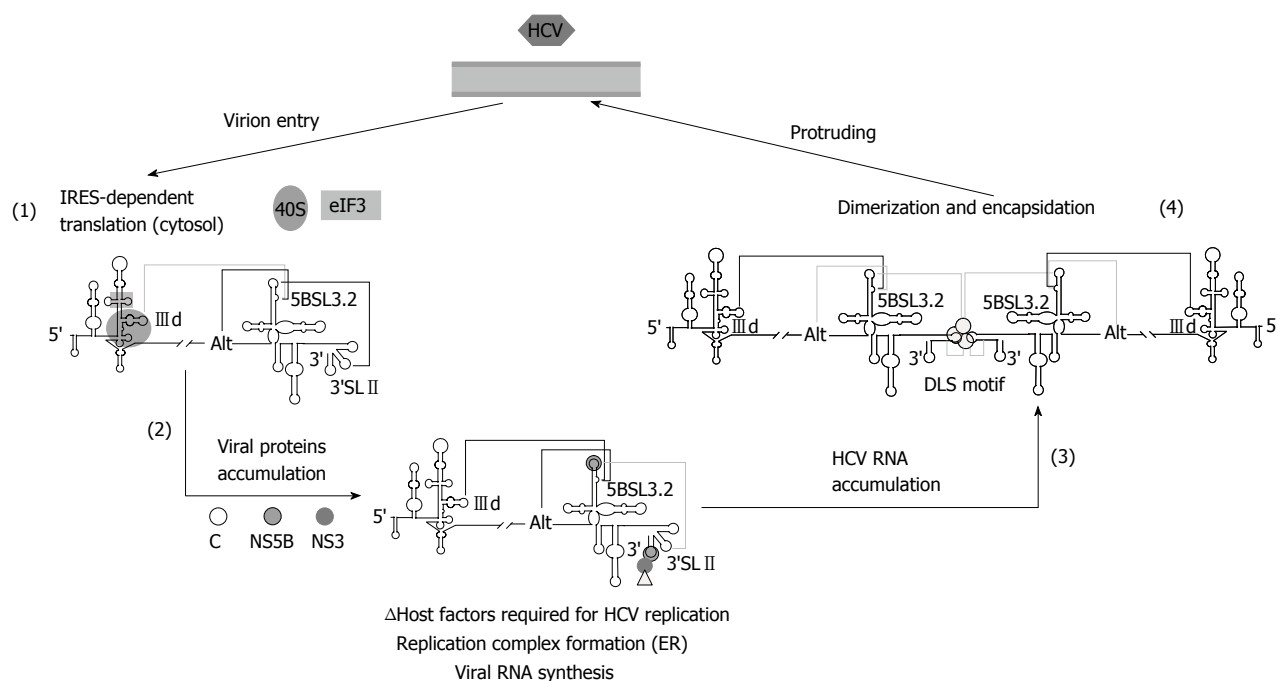
The 5BSL3.2 stem-loop consists of two G-C rich helices connected by an eight-base bulge, and capped by a 12-base apical loop (Figure 3)<sup>[45,98]</sup>. Disruptions in either the sequence or its folding lead to replication-incompetent HCV genomes<sup>[45,98]</sup>. Moreover, subtle changes in the apical loop prevent RNA replication, indicating that sequence specificity is required for interaction with protein factors, such as the NS5B protein (viral RNA dependent RNA polymerase)<sup>[110]</sup> and, more likely, distal RNA functional elements<sup>[29,98,99,109]</sup>. Relocation of 5BSL3.2 was only possible to the 3' variable region preceding the poly(U/UC) tract, involving a functional link with the 3'UTR<sup>[98]</sup>. Domain 5BSL3.2 has been also shown to act as an inhibitory element of the viral IRES function<sup>[25]</sup>, even in the presence of a translational enhancer such as the HCV 3'UTR. This action is strictly dependent on the sequence and the structural integrity of the bulge, pointing again to the existence of interactions with distant functional RNA domains of the viral genome.

### **LONG-RANGE RNA-RNA INTERACTION NETWORK IN THE HCV GENOME**

As it has been mentioned, the preservation of a proper equilibrium among different viral process and the adequate transitions between them must be assessed for reaching adaptive fitness and virus persistence. To accomplish this, the available functional genomic RNA domains establish an intricate and dynamic interacting web that is mediated, not only by the well-known protein-related 5'UTR-3'UTR bridges<sup>[21-23,27,28,111]</sup>, but more importantly by the formation of direct RNA-RNA contacts that minimize protein requisites.

The domain 5BSL3.2 is a good example of an all-RNA-based mechanism. This element participates in





**Figure 5** Proposed model for the participation of long-distant RNA-RNA interactions in the consecution of the hepatitis C virus infective cycle. Virions penetrate inside the cell and the genomic RNA is released in the cytoplasm to initiate viral translation (1). The 40S ribosomal subunit binds to the III d subdomain of the internal ribosome entry site (IRES) region and avoids its interaction with domain 5BSL3.2, thus favoring the contacts Alt-5BSL3.2-3'SL II. Once protein levels have surpassed a certain concentration (2), replication complexes are fixed on the surface of the endoplasmic reticulum (ER) to initiate the genomic RNA amplification. The recruitment of protein factors at the 3'X-tail would hide the 3'SL II domain. This would displace the interactions balance toward the III d-5BSL3.2 contact, thus impeding efficient translation. Importantly, a considerable pool of molecules should alternatively display a preferred folding defined by the interaction Alt-5BSL3.2, which is indispensable for replication. The accumulation of newly naked viral genomes (3) could induce a replication-defective state by the acquisition of a favored conformation involving the interaction III d-5BSL3.2, which would yield the dimer linkage sequence (DLS) exposed in an apical loop. The presence of the viral core chaperone protein would participate in the formation of dimeric genomes (4), which would be finally encapsidated, enveloped and released to the extracellular medium. Improbable contacts are indicated by grey lines at each step. Figure adapted from<sup>[31]</sup>.

viral translation and replication by its integration in a complex network of interactions with distant regions of the HCV genomic RNA (Figure 4). The apical loop of 5BSL3.2 is complementary to the apical loop of the 3'SL II within the 3'X-tail<sup>[98,99,112]</sup>. The resulting kissing loop contact contributes to the structural organization of the 3'X-tail and is essential for HCV replication<sup>[98]</sup>. The 8-nts bulge may establish two different interactions: (1) one with the apical loop of the subdomain III d of the IRES region<sup>[29,112]</sup>, which is related to the aforementioned translational inhibitory effect<sup>[25]</sup>; (2) the second with the Alt sequence, centred around position 9110, upstream of the CRE element<sup>[99,109,112]</sup>. This interaction is again critical for the synthesis of the viral genomic RNA<sup>[109]</sup>. Analyses by different biochemical techniques have proved that the complex interplay IRES-5BSL3.2-3'X-tail influences the global architecture of the affected regions and the surrounding functional RNA elements<sup>[30,31,99]</sup>. Thus, the 3' end of the HCV RNA genome, which contains both the CRE and the 3'UTR elements, fine-tunes the three dimensional structure of the IRES region<sup>[30]</sup>, which could be associated to the regulation of viral translation<sup>[25]</sup>. Conversely, the interaction III d-5BSL3.2 induce structural rearrangements in the 3'X-tail that finally lead to the conformational transition of the essential domains 3'SL II and 3'SL III, which switch to a single stem-loop folding that exposes the DLS motif in an apical loop<sup>[31]</sup> (Figure 3).

Importantly, it has been recently reported that all these interactions are equally probable<sup>[112]</sup>. Therefore, choosing between different contacts might depend on the presence of additional host and/or viral proteins.

Based on these findings, it has been recently proposed a working model<sup>[31]</sup>, which integrates current knowledge concerning to RNA-RNA interactions in the HCV genome, and their implications for the consecution of the viral cycle (Figure 5). In the first stage of the infection, the HCV IRES would be occupied by the translational machinery, thus avoiding any contact with the 5BSL3.2 domain. This would favor the establishment of the interactions 5BSL3.2-3'SL II, which occludes the DLS motif, and 5BSL3.2-Alt. After protein synthesis, the CRE and the 3'X-tail would recruit the viral polymerase (NS5B) and other replication complex factors (both RNA and proteins)<sup>[92,109,111,113-117]</sup>. In this context, both the 5BSL3.2-III d and 5BSL3.2-Alt interactions could be equally feasible. Swapping between them could contribute to the creation of a translational repressed state<sup>[25]</sup> and an enhanced replicative process<sup>[109]</sup>. The subsequent amplification of viral RNA molecules would displace the structural equilibrium between the 5BSL3.2-III d and 5BSL3.2-Alt interactions toward the long-range IRES-CRE contact. This would increase the proportion of RNA genomes exposing the DLS motif in the apical loop of the dimerizable conformation, leading to the formation of dimeric

genomic particles in the presence of the core chaperone protein<sup>[34]</sup>.

Therefore, domain 5BSL3.2 would occupy the central position in a complex and dynamic interacting web that would help to bring the ends of the HCV genome into close proximity to support the formation of a biologically favoured close-loop topology. Swapping between different RNA structural partners through the viral cycle would thus control the course of the infectious process.

## EXTENSIVE STRUCTURED REGIONS IN VIRAL RNA GENOMES

The search for novel conserved RNA structural units in viral genomes has been prompted in recent years by the appearance of bioinformatic tools that allow the study of the secondary structure of whole RNA genomes. Initial investigations based on the study of folding free energies in many positive stranded animal and plant viral RNA genomes identified extensive secondary structure regions that followed well-defined patterns<sup>[6,39]</sup>-the so-called GORS. They were initially related to different mechanisms for controlling viral replication, yet their prevalence appeared to be quite variable among different genera. For example, extensive base-pairing within the coding sequence was thermodynamically predicted for the hepacivirus genome, while in the closely related *Pestivirus* and *Flavivirus* genera this pattern was clearly absent. Since replication strategies are usually conserved among the members of a same family, it is unlikely that GORS work as fundamental base for the execution of the viral cycle. Remarkably, GORS are strongly associated to viral persistence<sup>[7]</sup>, thus raising the question whether they can be involved in the suppression of innate intracellular defence mechanisms. Thermodynamic predictions and phylogenetic studies based on base-pairing rules have recently been combined with oligonucleotide probe accessibility and atomic force microscopy studies to investigate the link between theoretical predictions and the 3D conformation of viral genomes with and without GORS in solution<sup>[7]</sup>. The results showed that the HCV genome is a tightly compact molecule, in contrast to RNA genomes that lack GORS, such as that of poliovirus which folds pleomorphically, commonly involving long single stranded stretches. These studies have contributed to understand how RNA conformation could be related with a virus defence system. Thus, it seems likely that extensive folding areas could interfere with the antiviral cellular pathways triggered by double-stranded RNA, such as the interferon production during the initial infection, in an analogous manner to the expression of structured RNA transcripts by large DNA viruses<sup>[118]</sup>. Though much remains to be investigated about this phenomenon, it has an undeniable relevance for virus-host interactions.

## CONCLUSION

During last years, the great advances in the fields of RNA

structure determination by high-throughput techniques and bioinformatic tools have enabled the first pictures for the structural organization of the eukaryotic transcriptome, the so-called structure. In molecular virology, these advances have gained special relevance for their implication in the identification of functional RNA domains. These structurally conserved RNA elements interact with protein factors and other RNA domains to direct and regulate essential viral functions as well as switching between different steps of the viral cycle. Interfering with the functioning of these structural domains offers a potential means of treating viral infections, such as that caused by the HCV. Further implementations of the current methodologies will undoubtedly improve the identification and validation of functional RNA domains in the near future, thus extending our knowledge of RNA-mediated regulation not only in viral systems, but also in many cellular processes.

## REFERENCES

- 1 **Eigen M.** Error catastrophe and antiviral strategy. *Proc Natl Acad Sci USA* 2002; **99**: 13374-13376 [PMID: 12370416 DOI: 10.1073/pnas.212514799]
- 2 **Wagner A, Stadler PF.** Viral RNA and evolved mutational robustness. *J Exp Zool* 1999; **285**: 119-127 [PMID: 10440723 DOI: 10.1002/(SICI)1097-010X(19990815)285: 2<119::AID-JEZ4>3.0.CO;2-D]
- 3 **Contreras AM, Hiasa Y, He W, Terella A, Schmidt EV, Chung RT.** Viral RNA mutations are region specific and increased by ribavirin in a full-length hepatitis C virus replication system. *J Virol* 2002; **76**: 8505-8517 [PMID: 12163570 DOI: 10.1128/JVI.76.17.8505-8517.2002]
- 4 **Churkin A, Cohen M, Shemer-Avni Y, Barash D.** Bioinformatic analysis of the neutrality of RNA secondary structure elements across genotypes reveals evidence for direct evolution of genetic robustness in HCV. *J Bioinform Comput Biol* 2010; **8**: 1013-1026 [PMID: 21121024 DOI: 10.1142/S0219720010005087]
- 5 **Romero-López C, Berzal-Herranz A.** Unmasking the information encoded as structural motifs of viral RNA genomes: a potential antiviral target. *Rev Med Virol* 2013; **23**: 340-354 [PMID: 23983005 DOI: 10.1002/rmv.1756]
- 6 **Simmonds P, Tuplin A, Evans DJ.** Detection of genome-scale ordered RNA structure (GORS) in genomes of positive-stranded RNA viruses: Implications for virus evolution and host persistence. *RNA* 2004; **10**: 1337-1351 [PMID: 15273323 DOI: 10.1261/rna.7640104]
- 7 **Davis M, Sagan SM, Pezacki JP, Evans DJ, Simmonds P.** Bioinformatic and physical characterizations of genome-scale ordered RNA structure in mammalian RNA viruses. *J Virol* 2008; **82**: 11824-11836 [PMID: 18799591 DOI: 10.1128/JVI.01078-08]
- 8 **Zhu Y, Chen S.** Antiviral treatment of hepatitis C virus infection and factors affecting efficacy. *World J Gastroenterol* 2013; **19**: 8963-8973 [PMID: 24379621 DOI: 10.3748/wjg.v19.i47.8963]
- 9 **Simmonds P, McOmish F, Yap PL, Chan SW, Lin CK, Dushenko G, Saeed AA, Holmes EC.** Sequence variability in the 5' non-coding region of hepatitis C virus: identification of a new virus type and restrictions on sequence diversity. *J Gen Virol* 1993; **74** (Pt 4): 661-668 [PMID: 8385694 DOI: 10.1099/0022-1317-74-4-661]
- 10 **Thomson BJ, Finch RG.** Hepatitis C virus infection. *Clin Microbiol Infect* 2005; **11**: 86-94 [PMID: 15679481 DOI: 10.1111/j.1469-0691.2004.01061.x]
- 11 **Martell M, Esteban JI, Quer J, Genescà J, Weiner A, Esteban**

- R, Guardia J, Gómez J. Hepatitis C virus (HCV) circulates as a population of different but closely related genomes: quasi-species nature of HCV genome distribution. *J Virol* 1992; **66**: 3225-3229 [PMID: 1313927]
- 12 **Farci P.** Hepatitis C virus. The importance of viral heterogeneity. *Clin Liver Dis* 2001; **5**: 895-916 [PMID: 11685800 DOI: 10.1016/S1089-3261(05)70200-2]
  - 13 **Choo QL**, Kuo G, Weiner AJ, Overby LR, Bradley DW, Houghton M. Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. *Science* 1989; **244**: 359-362 [PMID: 2523562 DOI: 10.1126/science.2523562]
  - 14 **Kato N**, Hijikata M, Ootsuyama Y, Nakagawa M, Ohkoshi S, Sugimura T, Shimotohno K. Molecular cloning of the human hepatitis C virus genome from Japanese patients with non-A, non-B hepatitis. *Proc Natl Acad Sci USA* 1990; **87**: 9524-9528 [PMID: 2175903 DOI: 10.1073/pnas.87.24.9524]
  - 15 **Takamizawa A**, Mori C, Fuke I, Manabe S, Murakami S, Fujita J, Onishi E, Andoh T, Yoshida I, Okayama H. Structure and organization of the hepatitis C virus genome isolated from human carriers. *J Virol* 1991; **65**: 1105-1113 [PMID: 1847440]
  - 16 **Tsukiyama-Kohara K**, Iizuka N, Kohara M, Nomoto A. Internal ribosome entry site within hepatitis C virus RNA. *J Virol* 1992; **66**: 1476-1483 [PMID: 1310759]
  - 17 **Wang C**, Sarnow P, Siddiqui A. Translation of human hepatitis C virus RNA in cultured cells is mediated by an internal ribosome-binding mechanism. *J Virol* 1993; **67**: 3338-3344 [PMID: 8388503]
  - 18 **Reynolds JE**, Kaminski A, Kettinen HJ, Grace K, Clarke BE, Carroll AR, Rowlands DJ, Jackson RJ. Unique features of internal initiation of hepatitis C virus RNA translation. *EMBO J* 1995; **14**: 6010-6020 [PMID: 8846793]
  - 19 **Wang TH**, Rijnbrand RC, Lemon SM. Core protein-coding sequence, but not core protein, modulates the efficiency of cap-independent translation directed by the internal ribosome entry site of hepatitis C virus. *J Virol* 2000; **74**: 11347-11358 [PMID: 11070035 DOI: 10.1128/JVI.74.23.11347-11358.2000]
  - 20 **Ito T**, Tahara SM, Lai MM. The 3'-untranslated region of hepatitis C virus RNA enhances translation from an internal ribosomal entry site. *J Virol* 1998; **72**: 8789-8796 [PMID: 9765423]
  - 21 **McCaffrey AP**, Ohashi K, Meuse L, Shen S, Lancaster AM, Lukavsky PJ, Sarnow P, Kay MA. Determinants of hepatitis C translational initiation in vitro, in cultured cells and mice. *Mol Ther* 2002; **5**: 676-684 [PMID: 12027551 DOI: 10.1006/mthe.2002.0600]
  - 22 **Bradrick SS**, Walters RW, Gromeier M. The hepatitis C virus 3'-untranslated region or a poly(A) tract promote efficient translation subsequent to the initiation phase. *Nucleic Acids Res* 2006; **34**: 1293-1303 [PMID: 16510853 DOI: 10.1093/nar/gkl019]
  - 23 **Song Y**, Friebe P, Tzima E, Jünemann C, Bartenschlager R, Niepmann M. The hepatitis C virus RNA 3'-untranslated region strongly enhances translation directed by the internal ribosome entry site. *J Virol* 2006; **80**: 11579-11588 [PMID: 16971433 DOI: 10.1128/JVI.00675-06]
  - 24 **Bung C**, Bochkava Z, Terenin I, Zinovkin R, Shatsky IN, Niepmann M. Influence of the hepatitis C virus 3'-untranslated region on IRES-dependent and cap-dependent translation initiation. *FEBS Lett* 2010; **584**: 837-842 [PMID: 20079737 DOI: 10.1016/j.febslet.2010.01.015]
  - 25 **Romero-López C**, Berzal-Herranz A. The functional RNA domain 5BSL3.2 within the NS5B coding sequence influences hepatitis C virus IRES-mediated translation. *Cell Mol Life Sci* 2012; **69**: 103-113 [PMID: 21598019 DOI: 10.1007/s00018-011-0729-z]
  - 26 **Ito T**, Lai MM. An internal polypyrimidine-tract-binding protein-binding site in the hepatitis C virus RNA attenuates translation, which is relieved by the 3'-untranslated sequence. *Virology* 1999; **254**: 288-296 [PMID: 9986795 DOI: 10.1006/viro.1998.9541]
  - 27 **Fang JW**, Moyer RW. The effects of the conserved extreme 3' end sequence of hepatitis C virus (HCV) RNA on the in vitro stabilization and translation of the HCV RNA genome. *J Hepatol* 2000; **33**: 632-639 [PMID: 11059869 DOI: 10.1016/S0168-8278(00)80016-1]
  - 28 **Kong LK**, Sarnow P. Cytoplasmic expression of mRNAs containing the internal ribosome entry site and 3' noncoding region of hepatitis C virus: effects of the 3' leader on mRNA translation and mRNA stability. *J Virol* 2002; **76**: 12457-12462 [PMID: 12438571 DOI: 10.1128/JVI.76.24.12457-12462.2002]
  - 29 **Romero-López C**, Berzal-Herranz A. A long-range RNA-RNA interaction between the 5' and 3' ends of the HCV genome. *RNA* 2009; **15**: 1740-1752 [PMID: 19605533 DOI: 10.1261/rna.1680809]
  - 30 **Romero-López C**, Barroso-Deljesus A, García-Sacristán A, Briones C, Berzal-Herranz A. The folding of the hepatitis C virus internal ribosome entry site depends on the 3'-end of the viral genome. *Nucleic Acids Res* 2012; **40**: 11697-11713 [PMID: 23066110 DOI: 10.1093/nar/gks927]
  - 31 **Romero-López C**, Barroso-Deljesus A, García-Sacristán A, Briones C, Berzal-Herranz A. End-to-end crosstalk within the hepatitis C virus genome mediates the conformational switch of the 3'X-tail region. *Nucleic Acids Res* 2014; **42**: 567-582 [PMID: 24049069 DOI: 10.1093/nar/gkt841]
  - 32 **Friebe P**, Lohmann V, Krieger N, Bartenschlager R. Sequences in the 5' nontranslated region of hepatitis C virus required for RNA replication. *J Virol* 2001; **75**: 12047-12057 [PMID: 11711595 DOI: 10.1128/JVI.75.24.12047-12057.2001]
  - 33 **Friebe P**, Bartenschlager R. Genetic analysis of sequences in the 3' nontranslated region of hepatitis C virus that are important for RNA replication. *J Virol* 2002; **76**: 5326-5338 [PMID: 11991961 DOI: 10.1128/JVI.76.11.5326-5338.2002]
  - 34 **Cristofari G**, Ivanyi-Nagy R, Gabus C, Boulant S, Laverne JP, Penin F, Darlix JL. The hepatitis C virus Core protein is a potent nucleic acid chaperone that directs dimerization of the viral (+) strand RNA in vitro. *Nucleic Acids Res* 2004; **32**: 2623-2631 [PMID: 15141033 DOI: 10.1093/nar/gkh579]
  - 35 **Ivanyi-Nagy R**, Kanevsky I, Gabus C, Laverne JP, Fichoux D, Penin F, Fossé P, Darlix JL. Analysis of hepatitis C virus RNA dimerization and core-RNA interactions. *Nucleic Acids Res* 2006; **34**: 2618-2633 [PMID: 16707664 DOI: 10.1093/nar/gkl240]
  - 36 **Shetty S**, Kim S, Shimakami T, Lemon SM, Mihailescu MR. Hepatitis C virus genomic RNA dimerization is mediated via a kissing complex intermediate. *RNA* 2010; **16**: 913-925 [PMID: 20360391 DOI: 10.1261/rna.1960410]
  - 37 **Sharma KK**, de Rocquigny H, Darlix JL, Laverne JP, Pénin F, Lessinger JM, Mély Y. Analysis of the RNA chaperoning activity of the hepatitis C virus core protein on the conserved 3'X region of the viral genome. *Nucleic Acids Res* 2012; **40**: 2540-2553 [PMID: 22127859 DOI: 10.1093/nar/gkr1140]
  - 38 **Smith DB**, Simmonds P. Characteristics of nucleotide substitution in the hepatitis C virus genome: constraints on sequence change in coding regions at both ends of the genome. *J Mol Evol* 1997; **45**: 238-246 [PMID: 9302317 DOI: 10.1007/PL00006226]
  - 39 **Tuplin A**, Wood J, Evans DJ, Patel AH, Simmonds P. Thermodynamic and phylogenetic prediction of RNA secondary structures in the coding region of hepatitis C virus. *RNA* 2002; **8**: 824-841 [PMID: 12088154 DOI: 10.1017/S1355838202554066]
  - 40 **Kim YK**, Lee SH, Kim CS, Seol SK, Jang SK. Long-range RNA-RNA interaction between the 5' nontranslated region and the core-coding sequences of hepatitis C virus modulates the IRES-dependent translation. *RNA* 2003; **9**: 599-606 [PMID: 12702818 DOI: 10.1261/rna.2185603]
  - 41 **Beguiristain N**, Robertson HD, Gómez J. RNase III cleavage demonstrates a long range RNA: RNA duplex element



- flanking the hepatitis C virus internal ribosome entry site. *Nucleic Acids Res* 2005; **33**: 5250-5261 [PMID: 16170153 DOI: 10.1093/nar/gki822]
- 42 **McMullan LK**, Grakoui A, Evans MJ, Mihalik K, Puig M, Branch AD, Feinstone SM, Rice CM. Evidence for a functional RNA element in the hepatitis C virus core gene. *Proc Natl Acad Sci USA* 2007; **104**: 2879-2884 [PMID: 17299041 DOI: 10.1073/pnas.0611267104]
  - 43 **Vassilaki N**, Friebe P, Meuleman P, Kallis S, Kaul A, Paranhos-Baccalà G, Leroux-Roels G, Mavromara P, Bartenschlager R. Role of the hepatitis C virus core+1 open reading frame and core cis-acting RNA elements in viral RNA translation and replication. *J Virol* 2008; **82**: 11503-11515 [PMID: 18799568 DOI: 10.1128/JVI.01640-08]
  - 44 **Lee H**, Shin H, Wimmer E, Paul AV. cis-acting RNA signals in the NS5B C-terminal coding sequence of the hepatitis C virus genome. *J Virol* 2004; **78**: 10865-10877 [PMID: 15452207 DOI: 10.1128/JVI.78.20.10865-10877.2004]
  - 45 **You S**, Stump DD, Branch AD, Rice CM. A cis-acting replication element in the sequence encoding the NS5B RNA-dependent RNA polymerase is required for hepatitis C virus RNA replication. *J Virol* 2004; **78**: 1352-1366 [PMID: 14722290 DOI: 10.1128/JVI.78.3.1352-1366.2004]
  - 46 **Pestova TV**, Shatsky IN, Fletcher SP, Jackson RJ, Hellen CU. A prokaryotic-like mode of cytoplasmic eukaryotic ribosome binding to the initiation codon during internal translation initiation of hepatitis C and classical swine fever virus RNAs. *Genes Dev* 1998; **12**: 67-83 [PMID: 9420332 DOI: 10.1101/gad.12.1.67]
  - 47 **Kieft JS**, Zhou K, Jubin R, Doudna JA. Mechanism of ribosome recruitment by hepatitis C IRES RNA. *RNA* 2001; **7**: 194-206 [PMID: 11233977 DOI: 10.1017/S1355838201001790]
  - 48 **Otto GA**, Puglisi JD. The pathway of HCV IRES-mediated translation initiation. *Cell* 2004; **119**: 369-380 [PMID: 15507208 DOI: 10.1016/j.cell.2004.09.038]
  - 49 **Spahn CM**, Kieft JS, Grassucci RA, Penczek PA, Zhou K, Doudna JA, Frank J. Hepatitis C virus IRES RNA-induced changes in the conformation of the 40S ribosomal subunit. *Science* 2001; **291**: 1959-1962 [PMID: 11239155 DOI: 10.1126/science.1058409]
  - 50 **Sizova DV**, Kolupaeva VG, Pestova TV, Shatsky IN, Hellen CU. Specific interaction of eukaryotic translation initiation factor 3 with the 5' nontranslated regions of hepatitis C virus and classical swine fever virus RNAs. *J Virol* 1998; **72**: 4775-4782 [PMID: 9573242]
  - 51 **Locker N**, Easton LE, Lukavsky PJ. HCV and CSFV IRES domain II mediate eIF2 release during 80S ribosome assembly. *EMBO J* 2007; **26**: 795-805 [PMID: 17255934 DOI: 10.1038/sj.emboj.7601549]
  - 52 **Brown EA**, Zhang H, Ping LH, Lemon SM. Secondary structure of the 5' nontranslated regions of hepatitis C virus and pestivirus genomic RNAs. *Nucleic Acids Res* 1992; **20**: 5041-5045 [PMID: 1329037 DOI: 10.1093/nar/20.19.5041]
  - 53 **Kieft JS**, Zhou K, Jubin R, Murray MG, Lau JY, Doudna JA. The hepatitis C virus internal ribosome entry site adopts an ion-dependent tertiary fold. *J Mol Biol* 1999; **292**: 513-529 [PMID: 10497018 DOI: 10.1006/jmbi.1999.3095]
  - 54 **Honda M**, Brown EA, Lemon SM. Stability of a stem-loop involving the initiator AUG controls the efficiency of internal initiation of translation on hepatitis C virus RNA. *RNA* 1996; **2**: 955-968 [PMID: 8849773]
  - 55 **Berry KE**, Waghray S, Mortimer SA, Bai Y, Doudna JA. Crystal structure of the HCV IRES central domain reveals strategy for start-codon positioning. *Structure* 2011; **19**: 1456-1466 [PMID: 22000514 DOI: 10.1016/j.str.2011.08.002]
  - 56 **Honda M**, Beard MR, Ping LH, Lemon SM. A phylogenetically conserved stem-loop structure at the 5' border of the internal ribosome entry site of hepatitis C virus is required for cap-independent viral translation. *J Virol* 1999; **73**: 1165-1174 [PMID: 9882318]
  - 57 **Pestova TV**, Hellen CU. Internal initiation of translation of bovine viral diarrhea virus RNA. *Virology* 1999; **258**: 249-256 [PMID: 10366562 DOI: 10.1006/viro.1999.9741]
  - 58 **Lyons AJ**, Lytle JR, Gomez J, Robertson HD. Hepatitis C virus internal ribosome entry site RNA contains a tertiary structural element in a functional domain of stem-loop II. *Nucleic Acids Res* 2001; **29**: 2535-2541 [PMID: 11410661 DOI: 10.1093/nar/29.12.2535]
  - 59 **Lukavsky PJ**, Kim I, Otto GA, Puglisi JD. Structure of HCV IRES domain II determined by NMR. *Nat Struct Biol* 2003; **10**: 1033-1038 [PMID: 14578934 DOI: 10.1038/nsb1004]
  - 60 **Kolupaeva VG**, Pestova TV, Hellen CU. An enzymatic footprinting analysis of the interaction of 40S ribosomal subunits with the internal ribosomal entry site of hepatitis C virus. *J Virol* 2000; **74**: 6242-6250 [PMID: 10864633 DOI: 10.1128/JVI.74.14.6242-6250.2000]
  - 61 **Lytle JR**, Wu L, Robertson HD. Domains on the hepatitis C virus internal ribosome entry site for 40S subunit binding. *RNA* 2002; **8**: 1045-1055 [PMID: 12212848 DOI: 10.1017/S1355838202029965]
  - 62 **Kalliampakou KI**, Psaridi-Linardaki L, Mavromara P. Mutational analysis of the apical region of domain II of the HCV IRES. *FEBS Lett* 2002; **511**: 79-84 [PMID: 11821053 DOI: 10.1016/S0014-5793(01)03300-2]
  - 63 **Ji H**, Fraser CS, Yu Y, Leary J, Doudna JA. Coordinated assembly of human translation initiation complexes by the hepatitis C virus internal ribosome entry site RNA. *Proc Natl Acad Sci USA* 2004; **101**: 16990-16995 [PMID: 15563596 DOI: 10.1073/pnas.0407402101]
  - 64 **Spahn CM**, Jan E, Mulder A, Grassucci RA, Sarnow P, Frank J. Cryo-EM visualization of a viral internal ribosome entry site bound to human ribosomes: the IRES functions as an RNA-based translation factor. *Cell* 2004; **118**: 465-475 [PMID: 15315759 DOI: 10.1016/j.cell.2004.08.001]
  - 65 **Passmore LA**, Schmeing TM, Maag D, Applefield DJ, Acker MG, Algire MA, Lorsch JR, Ramakrishnan V. The eukaryotic translation initiation factors eIF1 and eIF1A induce an open conformation of the 40S ribosome. *Mol Cell* 2007; **26**: 41-50 [PMID: 17434125 DOI: 10.1016/j.molcel.2007.03.018]
  - 66 **Filbin ME**, Kieft JS. HCV IRES domain IIb affects the configuration of coding RNA in the 40S subunit's decoding groove. *RNA* 2011; **17**: 1258-1273 [PMID: 21606179 DOI: 10.1261/rna.2594011]
  - 67 **Lafuente E**, Ramos R, Martínez-Salas E. Long-range RNA-RNA interactions between distant regions of the hepatitis C virus internal ribosome entry site element. *J Gen Virol* 2002; **83**: 1113-1121 [PMID: 11961266]
  - 68 **Malygin AA**, Kossinova OA, Shatsky IN, Karpova GG. HCV IRES interacts with the 18S rRNA to activate the 40S ribosome for subsequent steps of translation initiation. *Nucleic Acids Res* 2013; **41**: 8706-8714 [PMID: 23873958 DOI: 10.1093/nar/gkt632]
  - 69 **Buratti E**, Tisminetzky S, Zotti M, Baralle FE. Functional analysis of the interaction between HCV 5'UTR and putative subunits of eukaryotic translation initiation factor eIF3. *Nucleic Acids Res* 1998; **26**: 3179-3187 [PMID: 9628916]
  - 70 **Hashem Y**, des Georges A, Dhote V, Langlois R, Liao HY, Grassucci RA, Pestova TV, Hellen CU, Frank J. Hepatitis-C-virus-like internal ribosome entry sites displace eIF3 to gain access to the 40S subunit. *Nature* 2013; **503**: 539-543 [PMID: 24185006 DOI: 10.1038/nature12658]
  - 71 **Jubin R**, Vantuno NE, Kieft JS, Murray MG, Doudna JA, Lau JY, Baroudy BM. Hepatitis C virus internal ribosome entry site (IRES) stem loop IIIId contains a phylogenetically conserved GGG triplet essential for translation and IRES folding. *J Virol* 2000; **74**: 10430-10437 [PMID: 11044087 DOI: 10.1128/JVI.74.22.10430-10437.2000]
  - 72 **Babaylova E**, Graifer D, Malygin A, Stahl J, Shatsky I, Karpova G. Positioning of subdomain IIIId and apical loop of domain II of the hepatitis C IRES on the human 40S ribosome.



- Nucleic Acids Res* 2009; **37**: 1141-1151 [PMID: 19129232 DOI: 10.1093/nar/gkn1026]
- 73 **Klinck R**, Westhof E, Walker S, Afshar M, Collier A, Aboul-Ela F. A potential RNA drug target in the hepatitis C virus internal ribosomal entry site. *RNA* 2000; **6**: 1423-1431 [PMID: 11073218 DOI: 10.1017/S1355838200000935]
  - 74 **Lukavsky PJ**, Otto GA, Lancaster AM, Sarnow P, Puglisi JD. Structures of two RNA domains essential for hepatitis C virus internal ribosome entry site function. *Nat Struct Biol* 2000; **7**: 1105-1110 [PMID: 11101890 DOI: 10.1038/81951]
  - 75 **Berry KE**, Waghray S, Doudna JA. The HCV IRES pseudoknot positions the initiation codon on the 40S ribosomal subunit. *RNA* 2010; **16**: 1559-1569 [PMID: 20584896 DOI: 10.1261/rna.2197210]
  - 76 **Lukavsky PJ**. Structure and function of HCV IRES domains. *Virus Res* 2009; **139**: 166-171 [PMID: 18638512 DOI: 10.1016/j.virusres.2008.06.004]
  - 77 **Kolykhalov AA**, Mihalik K, Feinstone SM, Rice CM. Hepatitis C virus-encoded enzymatic activities and conserved RNA elements in the 3' nontranslated region are essential for virus replication in vivo. *J Virol* 2000; **74**: 2046-2051 [PMID: 10644379 DOI: 10.1128/JVI.74.4.2046-2051.2000]
  - 78 **Yi M**, Lemon SM. 3' nontranslated RNA signals required for replication of hepatitis C virus RNA. *J Virol* 2003; **77**: 3557-3568 [PMID: 12610131 DOI: 10.1128/JVI.77.6.3557-3568.2003]
  - 79 **Kolykhalov AA**, Feinstone SM, Rice CM. Identification of a highly conserved sequence element at the 3' terminus of hepatitis C virus genome RNA. *J Virol* 1996; **70**: 3363-3371 [PMID: 8648666]
  - 80 **You S**, Rice CM. 3' RNA elements in hepatitis C virus replication: kissing partners and long poly(U). *J Virol* 2008; **82**: 184-195 [PMID: 17942554 DOI: 10.1128/JVI.01796-07]
  - 81 **Gontarek RR**, Gutshall LL, Herold KM, Tsai J, Sathe GM, Mao J, Prescott C, Del Vecchio AM. hnRNP C and polypyrimidine tract-binding protein specifically interact with the pyrimidine-rich region within the 3'NTR of the HCV RNA genome. *Nucleic Acids Res* 1999; **27**: 1457-1463 [PMID: 10037806 DOI: 10.1093/nar/27.6.1457]
  - 82 **Chang KS**, Luo G. The polypyrimidine tract-binding protein (PTB) is required for efficient replication of hepatitis C virus (HCV) RNA. *Virus Res* 2006; **115**: 1-8 [PMID: 16102869 DOI: 10.1016/j.virusres.2005.06.012]
  - 83 **Spångberg K**, Goobar-Larsson L, Wahren-Herlenius M, Schwartz S. The La protein from human liver cells interacts specifically with the U-rich region in the hepatitis C virus 3' untranslated region. *J Hum Virol* 1999; **2**: 296-307 [PMID: 10551736]
  - 84 **Petrik J**, Parker H, Alexander GJ. Human hepatic glyceraldehyde-3-phosphate dehydrogenase binds to the poly(U) tract of the 3' non-coding region of hepatitis C virus genomic RNA. *J Gen Virol* 1999; **80** (Pt 12): 3109-3113 [PMID: 10567641]
  - 85 **Luo G**. Cellular proteins bind to the poly(U) tract of the 3' untranslated region of hepatitis C virus RNA genome. *Virology* 1999; **256**: 105-118 [PMID: 10087231 DOI: 10.1006/viro.1999.9639]
  - 86 **Wang H**, Shen XT, Ye R, Lan SY, Xiang L, Yuan ZH. Roles of the polypyrimidine tract and 3' noncoding region of hepatitis C virus RNA in the internal ribosome entry site-mediated translation. *Arch Virol* 2005; **150**: 1085-1099 [PMID: 15747050 DOI: 10.1007/s00705-005-0491-3]
  - 87 **Kim M**, Kim H, Cho SP, Min MK. Template requirements for de novo RNA synthesis by hepatitis C virus nonstructural protein 5B polymerase on the viral X RNA. *J Virol* 2002; **76**: 6944-6956 [PMID: 12072495 DOI: 10.1128/JVI.76.14.6944-6956.2002]
  - 88 **Shim JH**, Larson G, Wu JZ, Hong Z. Selection of 3'-template bases and initiating nucleotides by hepatitis C virus NS5B RNA-dependent RNA polymerase. *J Virol* 2002; **76**: 7030-7039 [PMID: 12072503 DOI: 10.1128/JVI.76.14.7030-7039.2002]
  - 89 **Murayama A**, Weng L, Date T, Akazawa D, Tian X, Suzuki T, Kato T, Tanaka Y, Mizokami M, Wakita T, Toyoda T. RNA polymerase activity and specific RNA structure are required for efficient HCV replication in cultured cells. *PLoS Pathog* 2010; **6**: e1000885 [PMID: 20442786 DOI: 10.1371/journal.ppat.1000885]
  - 90 **Huang L**, Hwang J, Sharma SD, Hargittai MR, Chen Y, Arnold JJ, Raney KD, Cameron CE. Hepatitis C virus nonstructural protein 5A (NS5A) is an RNA-binding protein. *J Biol Chem* 2005; **280**: 36417-36428 [PMID: 16126720 DOI: 10.1074/jbc.M508175200]
  - 91 **Harris D**, Zhang Z, Chaubey B, Pandey VN. Identification of cellular factors associated with the 3'-nontranslated region of the hepatitis C virus genome. *Mol Cell Proteomics* 2006; **5**: 1006-1018 [PMID: 16500930 DOI: 10.1074/mcp.M500429-MCP200]
  - 92 **Tingting P**, Caiyun F, Zhigang Y, Pengyuan Y, Zhenghong Y. Subproteomic analysis of the cellular proteins associated with the 3' untranslated region of the hepatitis C virus genome in human liver cells. *Biochem Biophys Res Commun* 2006; **347**: 683-691 [PMID: 16842740 DOI: 10.1016/j.bbrc.2006.06.144]
  - 93 **Kim CS**, Seol SK, Song OK, Park JH, Jang SK. An RNA-binding protein, hnRNP A1, and a scaffold protein, septin 6, facilitate hepatitis C virus replication. *J Virol* 2007; **81**: 3852-3865 [PMID: 17229681 DOI: 10.1128/JVI.01311-06]
  - 94 **Scheller N**, Mina LB, Galão RP, Chari A, Giménez-Barcons M, Noueiry A, Fischer U, Meyerhans A, Díez J. Translation and replication of hepatitis C virus genomic RNA depends on ancient cellular proteins that control mRNA fates. *Proc Natl Acad Sci USA* 2009; **106**: 13517-13522 [PMID: 19628699 DOI: 10.1073/pnas.0906413106]
  - 95 **Weinlich S**, Hüttelmaier S, Schierhorn A, Behrens SE, Ostareck-Lederer A, Ostareck DH. IGF2BP1 enhances HCV IRES-mediated translation initiation via the 3'UTR. *RNA* 2009; **15**: 1528-1542 [PMID: 19541769 DOI: 10.1261/rna.1578409]
  - 96 **Yu KL**, Jang SI, You JC. Identification of in vivo interaction between Hepatitis C Virus core protein and 5' and 3' UTR RNA. *Virus Res* 2009; **145**: 285-292 [PMID: 19665505 DOI: 10.1016/j.virusres.2009.07.023]
  - 97 **Bai Y**, Zhou K, Doudna JA. Hepatitis C virus 3'UTR regulates viral translation through direct interactions with the host translation machinery. *Nucleic Acids Res* 2013; **41**: 7861-7874 [PMID: 23783572 DOI: 10.1093/nar/gkt543]
  - 98 **Friebe P**, Boudet J, Simorre JP, Bartenschlager R. Kissing-loop interaction in the 3' end of the hepatitis C virus genome essential for RNA replication. *J Virol* 2005; **79**: 380-392 [PMID: 15596831 DOI: 10.1128/JVI.79.1.380-392.2005]
  - 99 **Tuplin A**, Struthers M, Simmonds P, Evans DJ. A twist in the tail: SHAPE mapping of long-range interactions and structural rearrangements of RNA elements involved in HCV replication. *Nucleic Acids Res* 2012; **40**: 6908-6921 [PMID: 22561372 DOI: 10.1093/nar/gks370]
  - 100 **Ina Y**, Mizokami M, Ohba K, Gojobori T. Reduction of synonymous substitutions in the core protein gene of hepatitis C virus. *J Mol Evol* 1994; **38**: 50-56 [PMID: 8151715 DOI: 10.1007/BF00175495]
  - 101 **Xu Z**, Choi J, Yen TS, Lu W, Strohecker A, Govindarajan S, Chien D, Selby MJ, Ou J. Synthesis of a novel hepatitis C virus protein by ribosomal frameshift. *EMBO J* 2001; **20**: 3840-3848 [PMID: 11447125 DOI: 10.1093/emboj/20.14.3840]
  - 102 **Varaklioti A**, Vassilaki N, Georgopoulou U, Mavromara P. Alternate translation occurs within the core coding region of the hepatitis C viral genome. *J Biol Chem* 2002; **277**: 17713-17721 [PMID: 11884417 DOI: 10.1074/jbc.M201722200]
  - 103 **Vassilaki N**, Kalliampakou KI, Mavromara P. Differences in the expression of the hepatitis C virus core+1 open reading frame between a nuclear and a cytoplasmic expression system. *J Gen Virol* 2008; **89**: 222-231 [PMID: 18089746 DOI: 10.1099/vir/2007/01796-07]

- 10.1099/vir.0.83260-0]
- 104 **Díaz-Toledano R**, Ariza-Mateos A, Birk A, Martínez-García B, Gómez J. In vitro characterization of a miR-122-sensitive double-helical switch element in the 5' region of hepatitis C virus RNA. *Nucleic Acids Res* 2009; **37**: 5498-5510 [PMID: 19578061 DOI: 10.1093/nar/gkp553]
- 105 **Jopling CL**, Yi M, Lancaster AM, Lemon SM, Sarnow P. Modulation of hepatitis C virus RNA abundance by a liver-specific MicroRNA. *Science* 2005; **309**: 1577-1581 [PMID: 16141076 DOI: 10.1126/science.1113329]
- 106 **Henke JI**, Goergen D, Zheng J, Song Y, Schüttler CG, Fehr C, Jünemann C, Niepmann M. microRNA-122 stimulates translation of hepatitis C virus RNA. *EMBO J* 2008; **27**: 3300-3310 [PMID: 19020517 DOI: 10.1038/emboj.2008.244]
- 107 **Roberts AP**, Lewis AP, Jopling CL. miR-122 activates hepatitis C virus translation by a specialized mechanism requiring particular RNA components. *Nucleic Acids Res* 2011; **39**: 7716-7729 [PMID: 21653556 DOI: 10.1093/nar/gkr426]
- 108 **Hofacker IL**, Fekete M, Flamm C, Huynen MA, Rauscher S, Stolorz PE, Stadler PF. Automatic detection of conserved RNA structure elements in complete RNA virus genomes. *Nucleic Acids Res* 1998; **26**: 3825-3836 [PMID: 9685502 DOI: 10.1093/nar/26.16.3825]
- 109 **Diviney S**, Tuplin A, Struthers M, Armstrong V, Elliott RM, Simmonds P, Evans DJ. A hepatitis C virus cis-acting replication element forms a long-range RNA-RNA interaction with upstream RNA sequences in NS5B. *J Virol* 2008; **82**: 9008-9022 [PMID: 18614633 DOI: 10.1128/JVI.02326-07]
- 110 **Zhang J**, Yamada O, Sakamoto T, Yoshida H, Araki H, Murata T, Shimotohno K. Inhibition of hepatitis C virus replication by pol III-directed overexpression of RNA decoys corresponding to stem-loop structures in the NS5B coding region. *Virology* 2005; **342**: 276-285 [PMID: 16139319 DOI: 10.1016/j.virol.2005.08.003]
- 111 **Ito T**, Lai MM. Determination of the secondary structure of and cellular protein binding to the 3'-untranslated region of the hepatitis C virus RNA genome. *J Virol* 1997; **71**: 8698-8706 [PMID: 9343228]
- 112 **Shetty S**, Stefanovic S, Mihailescu MR. Hepatitis C virus RNA: molecular switches mediated by long-range RNA-RNA interactions? *Nucleic Acids Res* 2013; **41**: 2526-2540 [PMID: 23275555 DOI: 10.1093/nar/gks1318]
- 113 **Banerjee R**, Dasgupta A. Specific interaction of hepatitis C virus protease/helicase NS3 with the 3'-terminal sequences of viral positive- and negative-strand RNA. *J Virol* 2001; **75**: 1708-1721 [PMID: 11160669 DOI: 10.1128/JVI.75.4.1708-1721.2001]
- 114 **Yi M**, Lemon SM. Structure-function analysis of the 3' stem-loop of hepatitis C virus genomic RNA and its role in viral RNA replication. *RNA* 2003; **9**: 331-345 [PMID: 12592007 DOI: 10.1261/rna.2144203]
- 115 **Foster TL**, Belyaeva T, Stonehouse NJ, Pearson AR, Harris M. All three domains of the hepatitis C virus nonstructural NS5A protein contribute to RNA binding. *J Virol* 2010; **84**: 9267-9277 [PMID: 20592076 DOI: 10.1128/JVI.00616-10]
- 116 **Kanamori H**, Yuhashi K, Ohnishi S, Koike K, Kodama T. RNA-dependent RNA polymerase of hepatitis C virus binds to its coding region RNA stem-loop structure, 5BSL3.2, and its negative strand. *J Gen Virol* 2010; **91**: 1207-1212 [PMID: 20053824 DOI: 10.1099/vir.0.016907-0]
- 117 **Oakland TE**, Haselton KJ, Randall G. EWSR1 binds the hepatitis C virus cis-acting replication element and is required for efficient viral replication. *J Virol* 2013; **87**: 6625-6634 [PMID: 23552423 DOI: 10.1128/JVI.01006-12]
- 118 **Mathews MB**. Structure, function, and evolution of adenovirus virus-associated RNAs. *Curr Top Microbiol Immunol* 1995; **199** (Pt 2): 173-187 [PMID: 7555067 DOI: 10.1007/978-3-642-79499-5\_7]

**P- Reviewers:** Bock T, Lee sw **S- Editor:** Ma YJ

**L- Editor:** A **E- Editor:** Liu SQ



## Molecular epidemiology of hepatitis B virus in Asia

Takako Utsumi, Yoshihiko Yano, Hak Hotta

Takako Utsumi, Indonesia-Japan Collaborative Research Center for Emerging and Re-emerging Infectious Diseases, Institute of Tropical Disease, Airlangga University, Surabaya 60115, Indonesia

Takako Utsumi, Yoshihiko Yano, Hak Hotta, Center for Infectious Diseases, Kobe University Graduate School of Medicine, Kobe 650-0017, Japan

**Author contributions:** All the authors contributed substantially to the conception and design, acquisition of data described in the articles, drafting the article, designing the figures of the manuscript, revising it critically for important intellectual content, and gave approval for the final version to be published.

**Supported by** The Japan Initiative for Global Research Network on Infectious Diseases (J-GRID) Program from the Ministry of Education, Culture, Sports, Science and Technology, Japan; The Ministry of Health, Labour and Welfare, Japan, and a SATREPS Grant from Japan Science and Technology Agency and Japan International Cooperation Agency

**Correspondence to:** Hak Hotta, MD, PhD, Center for Infectious Diseases, Kobe University Graduate School of Medicine, 7-5-1 Kusunoki-cho, Chuo-ku, Kobe 650-0017, Japan. [hotta@kobe-u.ac.jp](mailto:hotta@kobe-u.ac.jp)

Telephone: +81-78-3825500 Fax: +81-78-3825519

Received: December 17, 2013 Revised: February 25, 2014

Accepted: March 13, 2014

Published online: May 27, 2014

### Abstract

Although safe and effective vaccines against hepatitis B virus (HBV) have been available for three decades, HBV infection remains the leading cause of chronic hepatitis, cirrhosis and hepatocellular carcinoma (HCC) worldwide, especially in Asian countries. HBV has been classified into at least 9 genotypes according to the molecular evolutionary analysis of the genomic DNA sequence and shown to have a distinct geographical distribution. Novel HBV genotypes/subgenotypes have been reported, especially from Southeast Asian countries. The clinical characteristics and therapeutic effectiveness of interferon (IFN) and nucleos(t)ide analogues vary among different HBV genotypes. Mutations at T1653C in subgenotype C2 from Japan and South

Korea, C/A1753T and C1858T in subgenotype C1 from Vietnam, and C1638T and T1753V in subgenotype B3 from Indonesia were reported to be associated with advanced liver diseases including HCC. Genotype distribution in Japan has been changed by an increasing ratio of subgenotype A2 in chronic hepatitis B. While a large number of epidemiological and clinical studies have been reported from Asian countries, most of the studies were conducted in developed countries such as Taiwan, China, South Korea and Japan. In this review, the most recent publications on the geographical distribution of genetic variants of HBV and related issues such as disease progression and therapy in Asia are updated and summarized.

© 2014 Baishideng Publishing Group Inc. All rights reserved.

**Key words:** Hepatitis B virus; Genotype; Subgenotype; Molecular epidemiology; Asia; Pathogenicity; Drug resistance

**Core tip:** Chronic hepatitis B virus (HBV) infection usually progresses to liver cirrhosis and hepatocellular carcinoma. The variation of the HBV genotype is related to the geographical distribution. Also, the clinical characteristics and therapeutic effectiveness of interferon and nucleos(t)ide analogue vary among different HBV genotypes. A large number of epidemiological and clinical studies have been reported from Asian countries. However, most of the studies were conducted in developed countries such as Taiwan, China, South Korea and Japan. In this review, epidemiologically and clinically important aspects of HBV genotypes/subgenotypes found in East and Southeast Asian countries are updated and summarized.

Utsumi T, Yano Y, Hotta H. Molecular epidemiology of hepatitis B virus in Asia. *World J Med Genet* 2014; 4(2): 19-26 Available from: URL: <http://www.wjgnet.com/2220-3184/full/v4/i2/19.htm>  
DOI: <http://dx.doi.org/10.5496/wjmg.v4.i2.19>

## INTRODUCTION

Although safe and effective vaccines against Hepatitis B virus (HBV) have been available for more than three decades, HBV infection remains a burden to global public health, resulting in 600000 to 1 million deaths per year worldwide<sup>[1]</sup>. Two billion people are estimated to be exposed to HBV infection once in their life and it causes a wide spectrum of liver disease, including acute or fulminant hepatitis, inactive carrier state, reactivation, chronic hepatitis, cirrhosis and hepatocellular carcinoma (HCC)<sup>[2]</sup>. More than 420 million individuals in the world are estimated to have chronic HBV infections; 15%-40% of them are at risk of death due to liver failure or HCC<sup>[3]</sup>. The prevalence of HBV infection varies markedly in different geographical areas of the world. Overall, approximately 45% of the global population live in areas of high HBV prevalence, such as sub-Saharan Africa, the Pacific and particularly Asia<sup>[4]</sup>.

HBV has been classified into at least 9 genotypes (A through H and J) and shown to have a distinct geographical distribution<sup>[5,6]</sup>. In Asia, HBV genotypes B and C are prevalent, with genotype C having been shown to cause more serious liver diseases than genotype B. High prevalence of HBV mutants with various forms, such as the pre-S mutants, basal core promoter (BCP) mutants, YMDD motif mutants and vaccine escape mutants<sup>[7,8]</sup>, were seen in Asia and these were found to be related to severe liver diseases and resistance to treatment and prevention. This article provides an overview of the molecular-based epidemiology of HBV in Asian countries.

## HBV GENOME

HBV contains a partially double-stranded DNA genome of approximately 3200 base pairs. HBV replicates *via* a RNA intermediate anti-genome sequence, which encodes a potentially error-prone polymerase without proof-reading activity. The error frequencies are similar to those of retroviruses and other RNA viruses. The HBV genome encodes viral proteins through four open and partially overlapping reading frames: surface (S), core (C), polymerase (P) and X genes. This unusual genomic structure can compress a large amount of information into short sequences but implies a constrained evolution for the virus. This constraint can be reflected on the calculated rate of substitution,  $10^{-5}$  per site per year, slower than the rate displayed by the retroviruses of around  $10^{-3}$  per site per year<sup>[9]</sup>.

## EPIDEMIOLOGY OF HBV INFECTION

The prevalence of chronic HBV infection varies greatly in different parts of the world and can be categorized as high ( $\geq 8\%$ ), intermediate (2%-7%) and low ( $< 2\%$ ) endemicity. Table 1 shows the prevalence of hepatitis B surface antigen (HBsAg)-positive individuals in the general population of Southeast Asia and East Asia. HBV infection is highly endemic in Myanmar<sup>[10]</sup>; has intermediate to

**Table 1** Prevalence of hepatitis B surface antigen in the general Asian population

Country	HBsAg positivity (%)	Ref.
Southeast Asia		
Brunei	4.7	Sebastian <i>et al</i> <sup>[30]</sup>
	6.0	Alexander <i>et al</i> <sup>[31]</sup>
Cambodia	7.7	OI <i>et al</i> <sup>[14]</sup>
	10.8	Sa-Nguanmoo <i>et al</i> <sup>[10]</sup>
Indonesia	3.5-9.1	Hasan <sup>[11]</sup>
	4.9	Achwan <i>et al</i> <sup>[12]</sup>
	2.1-10.5	Lusida <i>et al</i> <sup>[13]</sup>
Laos	6.9	Jutavijittum <i>et al</i> <sup>[24]</sup>
	8.7	Sa-Nguanmoo <i>et al</i> <sup>[10]</sup>
Malaysia	3.0-5.0	Merican <i>et al</i> <sup>[22]</sup>
	0.5-1.8	Yousuf <i>et al</i> <sup>[23]</sup>
Myanmar	9.7	Sa-Nguanmoo <i>et al</i> <sup>[10]</sup>
Philippines	10.0	Lingao <i>et al</i> <sup>[17]</sup>
	2.0-16.0	Lansang <i>et al</i> <sup>[18]</sup>
	16.7	Wong <i>et al</i> <sup>[19]</sup>
Singapore	3.6-4.0	James <i>et al</i> <sup>[28]</sup>
	2.7-4.0	Ang <i>et al</i> <sup>[29]</sup>
Thailand	4.0	Suwannakarn <i>et al</i> <sup>[15]</sup>
	13.8	Louisirirothanakul <i>et al</i> <sup>[16]</sup>
Vietnam	11.4	Viet <i>et al</i> <sup>[20]</sup>
	7.5	Reekie <i>et al</i> <sup>[21]</sup>
East Asia		
China	2.4	Ting-Lu <i>et al</i> <sup>[25]</sup>
	1.0	Liu <i>et al</i> <sup>[26]</sup>
	10.6	Chen <i>et al</i> <sup>[27]</sup>
Japan	0.8	Merican <i>et al</i> <sup>[22]</sup>
South Korea	3.0-4.0, 6.0	Kim <i>et al</i> <sup>[32]</sup>
	6.0	Hyun <i>et al</i> <sup>[33]</sup>

HBsAg: Hepatitis B surface antigen.

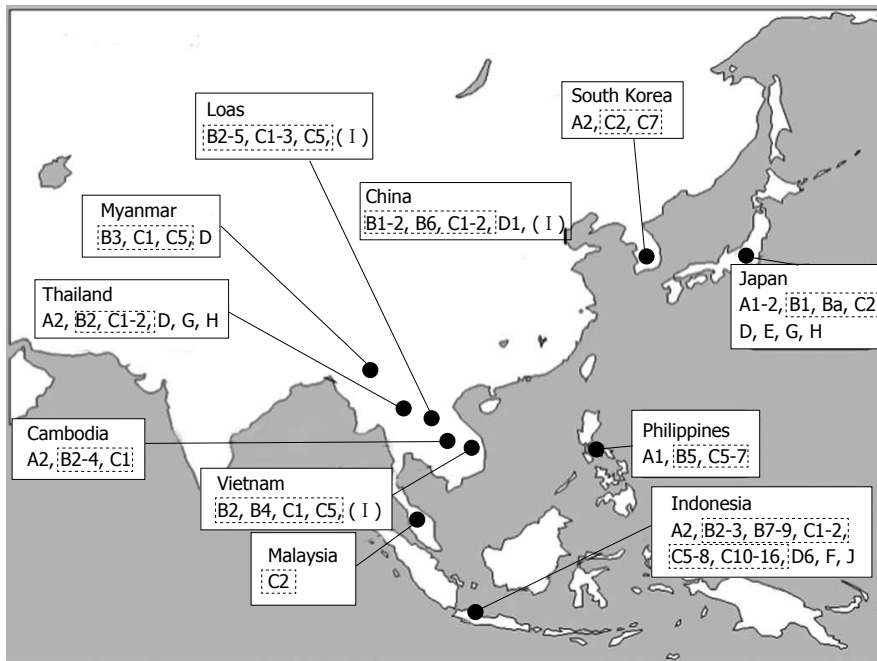
high endemicity in Indonesia<sup>[11-13]</sup>, Cambodia<sup>[10,14]</sup>, Thailand<sup>[15,16]</sup>, the Philippines<sup>[17-19]</sup>, Vietnam<sup>[20,21]</sup> and Laos<sup>[10,24]</sup>, low to high endemicity in Malaysia<sup>[22,23]</sup> and China<sup>[25-27]</sup>, and intermediate endemicity in Singapore<sup>[28,29]</sup>, Brunei<sup>[30,31]</sup> and South Korea<sup>[32,33]</sup>. Japan is the only country with low endemicity of HBV infection in Asia<sup>[22]</sup>.

HBV infection is highly endemic in developing regions with a large population such as Southeast Asia and China, where at least 8% of the population are HBV chronic carriers. For example, in Indonesia, which consists of thousands of islands with many ethnicities, the endemicity of HBV infection greatly varies even within the country. The wide range of the HBV prevalence is largely related to differences in age at the time of infection<sup>[3]</sup>.

## HBV GENOTYPES/SUBGENOTYPES AND THEIR GEOGRAPHICAL DISTRIBUTIONS

HBV is currently grouped into at least 9 genotypes (A through H and J, with I still being controversial)<sup>[6,34,35]</sup>, based on a full genome diversity of more than 8% at the nucleotide (nt) level, and phylogenetic analyses have shown that most of the genotypes can be further divided into subgenotypes differing by at least 4% of their full genome sequences. The prevalence of each HBV genotype and subgenotype varies in different geographical regions and is strongly associated with ethnicity<sup>[36]</sup>.





**Figure 1 Genotype/subgenotype distribution in East and Southeast Asia.** Subgenotypes of genotypes B and C commonly found in Asia are circled with dotted lines.

Genotype A is highly prevalent in Sub-Saharan Africa (A1 or Aa: a for Africa), Northern Europe (A2 or Ae; e stands for Europe) and Western Africa (A3). Genotypes B and C are the major HBV genotypes circulating in East and Southeast Asia<sup>[37]</sup> (Figure 1) and co-infection has led to a frequent occurrence of recombination between these two genotypes<sup>[38,39]</sup>. Subgenotype B1 (or Bj; j for Japan) is found almost exclusively in Japan and B2 (or Ba; a for Asia) is found in the rest of Asia<sup>[40,41]</sup>, but mainly in China and Vietnam. B1 is not a recombinant while B2 is considered to be B/C recombinants with the precore and core genes from genotype C. B3 is mostly found in Indonesia<sup>[42]</sup> while B4 is in Vietnam<sup>[5]</sup>. B5 was initially reported in 2006 from the Philippines<sup>[43]</sup>. B6 was identified in 2007 from the Arctic<sup>[39]</sup>. B7 to B9 were isolated in eastern Indonesia during the years 2007 to 2011<sup>[44-46]</sup>. C1 (or Cs: s for Southeast Asia) is the dominant strain in Southeast Asia and southern China, while C2 (or Ce: e for East Asia) is found mainly in East Asia (South Korea and Japan) and the northern part of China, C3 in Oceania<sup>[47]</sup> and C4 in the Aborigines from Australia<sup>[48]</sup>. C5 was initially reported in 2006 from the Philippines with B5<sup>[44]</sup>. C6 was identified from a Papuan population in Indonesia<sup>[13,49]</sup> and the Philippines<sup>[50]</sup> in 2008. Surprisingly, ten novel subgenotypes (C7 to C16) were isolated in Indonesia during 2009 to 2012<sup>[45,51-54]</sup>. Subgenotypes D1 to D4 of genotype D are widely distributed globally<sup>[5]</sup>, D5 in India<sup>[55]</sup> and D6 in Papua, Indonesia<sup>[13]</sup>. Genotype E is found mainly in sub-Saharan Africa. Genotypes F and H are found mainly in South and Central America, respectively. Genotype G has been found in Europe, United States and Japan. Genotype I was originally identified in Laos<sup>[56]</sup>, Vietnam and Southern China. However, this classification is still controversial as the sequence divergence hovers at but is slightly less than 8%, with a close relationship to genotype C<sup>[55]</sup>. Genotype J was found in a Japanese soldier

who was thought to have been infected in the forests in Kalimantan, Indonesia, during World War II<sup>[57]</sup>. Thus, novel HBV genotypes and novel subgenotypes have been found in Southeast Asia, especially in Laos, Vietnam, the Philippines and Indonesia, all consisting of many islands and ethnic groups. In addition to genotypes B and C which are common in Asia, an increasing rate of infection with rare HBV genotypes, such as genotypes A, D, E, G and H, has been recognized throughout Asia. Globalization may yield HBV strains of possible novel genotypes containing novel nucleotide sequences in the precore/core region<sup>[58]</sup>. The distribution of genotypes/subgenotypes varies even in different regions of a country, as observed in Indonesia, which may partly be related to the ethnic origin of the infected patients.

## HBV GENOTYPES AND DISEASE PROGRESSION

Chronic HBV infections usually progress to liver cirrhosis and HCC. Several studies revealed that the presence of hepatitis B e antigen (HBeAg) and high levels of HBV DNA were independent risk factors for the development of liver cirrhosis and HCC<sup>[59-62]</sup>. HBV genotypes are also related to the clinical characteristics<sup>[63]</sup>. In northeast Asian countries, where genotypes B and C are prevalent, the dominant mode of transmission is vertical (mother-to-child). A large number of studies have shown that genotype B is associated with HBeAg seroconversion at an earlier age, more sustained remission after HBeAg seroconversion, less active hepatic necroinflammation, a slower rate of progression to cirrhosis, and a lower rate of HCC development compared to genotype C<sup>[59,64-67]</sup>. On the other hand, genotypes D and A are prevalent in the southwest Asian countries, such as India and Pakistan<sup>[68]</sup>. The transmission route among Pakistanis, includ-

**Table 2 Summary of nucleos(t)ide analogues**

	Lamivudine	Adefovir	Entecavir	Telbivudine	Tenofovir	Ref.
Analogue type	Nucleoside	Nucleotide	Nucleoside	Nucleoside	Nucleotide	
Introduction (yr)	1999	2002	2005	2006	2008	
Product name (company)	Zefix (GSK)	Hepsera (Gilead)	Baraclude (BMS)	Sebivo (Novartis)	Viread (Gilead)	
Dose	100 mg Once daily	10 mg Once daily	0.5 mg Once daily	600 mg Once daily	300 mg Once daily	
Advantage	Low cost	Effective for HIV coinfection		Possible for pregnancy	Effective for HIV coinfection	[89]
Disadvantage	High rate of drug resistance	Renal dysfunction Fanconi anemia	Not recommend for pregnancy	Renal dysfunction	Renal dysfunction Fanconi anemia	
Undetectable HBV-DNA						
HBeAg positive	36%	21%	67%	60%	76%	[90]
HBeAg negative	89%	72%	90%	88%	93%	
HBeAg seroconversion	22%	12%	21%	23%	21%	[91]
Drug-resistance	24%	0%	0.2%	4%	0%	[92]
Drug-resistant mutation	V173I, L180M, A181T, M204V/I	A181V/T, N236T	I169T, L180M, T184A/F/L/S, S202G/I, M204V, M250V	M204V/I	A181V/T, N236T	

GSK: Glaxo Smith Kline; BMS: Bristol-Myers Squibb; HIV: Human immunodeficiency virus; HBV: Hepatitis B virus; HBeAg: Hepatitis B e antigen.

ing Afghan refugees, is not only vertical transmission but also through unsterilized materials and intravenous drug use<sup>[69,70]</sup>. Reports concerning the risk factors of advanced liver diseases are still limited in those countries.

Mutations in the viral genome, including the X region, are also important factors in association with disease progression. A study from Taiwan revealed that the precore G1896A wild-type and the BCP A1762T/G1764A mutation were strongly associated with HCC development among genotype C<sup>[71]</sup>. A study from north India also showed that the BCP A1762T/G1764A mutation was associated with progressive liver diseases among genotype D<sup>[72]</sup>. In Japan and South Korea, the T1653C mutation was reported as a predictive factor for the development of advanced liver diseases in HBV genotype C2 infection<sup>[73,74]</sup>. Whereas the C/A1753T and C1858T mutations were associated with advanced liver diseases in genotype C1 infection in Vietnam, C1638T and T1753V were independent risk factors for advanced liver diseases in genotype B3 infection in Indonesia<sup>[42,75]</sup>. In addition, several studies from Taiwan and Japan showed that the pre-S mutation also contributed to the progressive liver disease and HCC<sup>[76,77]</sup>. The progression from acute hepatitis to chronic infection occurs more frequently in genotype A (23%) compared with genotypes B (11%) and C (7%)<sup>[78]</sup>. This might change genotype distribution in the future. In Japan, indeed, the prevalence of genotype A in chronic hepatitis B increased from 1.7% to 3.5% during the period between 2000 and 2006<sup>[79]</sup>.

## HBV GENOTYPE AND ANTIVIRAL THERAPY

The purpose of antiviral therapy for chronic hepatitis B is the sustained suppression of HBV replication, biochemical remission, HBeAg seroconversion and ultimately HBsAg seroconversion. The annual rate of spontaneous HBsAg seroclearance is approximately 0.4%-2.3%, and the HBsAg seroclearance rates of genotypes A and

B are higher than that of genotypes C and D<sup>[80,81]</sup>.

Interferon (IFN) and nucleos(t)ide analogues (NA) are commonly used for the treatment of chronic hepatitis B. Antiviral regimens for chronic hepatitis B are decided based on the age, HBV-DNA viral load, alanine aminotransferase (ALT) levels and the degree of fibrosis. In general, younger patients with high ALT levels are recommended to be treated with IFN therapy and older and/or clinically advanced patients with NA. Due to the economic growth, the treatment of chronic hepatitis B has become universal in most developed and developing Asian countries. However, most of the clinical studies about antiviral therapy were reported from developed countries, with few studies being reported from developing countries. IFN has antiviral, antiproliferative and immunomodulatory effects. The response to IFN treatment is poorer in Asian patients compared with Caucasian patients, which may be due partly to the difference in the genotype distribution<sup>[82]</sup>. It was shown that patients infected with HBV genotypes A and B showed better response than those with genotypes C and D<sup>[83-87]</sup>. A meta-analysis also revealed that IFN therapy was more effective in patients infected with genotype A than in those with genotype D, and also more effective in genotype B than in genotype C infection<sup>[88]</sup>.

Currently, lamivudine, adefovir, entecavir, telbivudine and tenofovir have been approved for the treatment of chronic hepatitis B (Table 2). Lamivudine (Zeffix<sup>®</sup>) was first introduced in 1999 and the clinical efficacy was shown by a long-term follow-up study<sup>[93,94]</sup>. However, drug-resistant mutations, especially multidrug-resistant mutations, are the major concern with patients receiving long-term NA treatment. It was reported that the drug resistance against lamivudine monotherapy reached 70% after 4 years of treatment<sup>[95,96]</sup>. Entecavir (Baraclude<sup>®</sup>) is widely used and a first-line drug in many Asian countries, including China, South Korea, Thailand, Hong Kong and Japan. Entecavir is still expensive but the occurrence of drug resistance is very low for naïve patients.

However, the chemical structure of entecavir is similar to lamivudine, which resulted in the cross-resistance between lamivudine and entecavir. Recent long-term follow up studies conducted in South Korea and Hong Kong revealed that entecavir reduced liver-related death and HCC<sup>[97,98]</sup>. Adefovir (Hepsera®) is effective against lamivudine-resistant mutants and add-on therapy of adefovir and lamivudine is common for suppression of lamivudine-resistant mutants. Tenofovir (Viread®) and telbivudine (Sebivo®) are also safe and effective drugs but their introduction to clinical use is still limited. Telbivudine has recently been approved and is being used as a first-line drug in Indonesia. Unlike IFN therapy, meta-analysis revealed no significant difference between genotypes and response to NA<sup>[88]</sup>. However, as entecavir and telbivudine were introduced recently in developing countries, further studies will be needed to assess their efficacy against the different HBV genotypes/subgenotypes prevailing in those countries.

## CONCLUSION

HBV is widespread in Asian countries and contributes to the mortality from HCC. To reduce HBV infection and HCC mortality, appropriate national immunization programs are required in HBV-endemic countries, including Japan. Although HBV infection is predominant and a number of novel genotypes/subgenotypes have been discovered in Asian countries, studies have not been sufficient regarding disease prognosis and antiviral treatment. It is possible that certain genotypes or variants of HBV prevailing in these regions possess stronger pathogenicity and are associated with more severe outcomes of liver diseases. The studies on HBV genotypes related to their pathogenicity in chronic liver diseases, including liver cirrhosis and HCC, and their effects on treatment outcome are awaited with great interest, especially in Southeast Asia, which is the most endemic region of HBV in Asia with unique HBV genotypes/subgenotypes.

## REFERENCES

- 1 **Kao JH**, Chen DS. Global control of hepatitis B virus infection. *Lancet Infect Dis* 2002; **2**: 395-403 [PMID: 12127351]
- 2 **Mizokami M**. Changing concept of hepatitis B virus. *ISBT Science Series* 2009; **4**: 192 [DOI: 10.1111/j.1751-2824.2009.01281.x]
- 3 **Lok AS**. Chronic hepatitis B. *N Engl J Med* 2002; **346**: 1682-1683 [PMID: 12037146]
- 4 **Lavanchy D**. Hepatitis B virus epidemiology, disease burden, treatment, and current and emerging prevention and control measures. *J Viral Hepat* 2004; **11**: 97-107 [PMID: 14996343]
- 5 **Norder H**, Couroucé AM, Coursaget P, Echevarria JM, Lee SD, Mushahwar IK, Robertson BH, Locarnini S, Magnus LO. Genetic diversity of hepatitis B virus strains derived worldwide: genotypes, subgenotypes, and HBsAg subtypes. *Intervirology* 2004; **47**: 289-309 [PMID: 15564741]
- 6 **Locarnini S**, Littlejohn M, Aziz MN, Yuen L. Possible origins and evolution of the hepatitis B virus (HBV). *Semin Cancer Biol* 2013; **23**: 561-575 [PMID: 24013024 DOI: 10.1016/j.semcancer.2013.08.006]
- 7 **Utsumi T**, Yano Y, Lusida MI, Nasronudin M, Juniastuti H, Hayashi Y. Detection of highly prevalent hepatitis B virus co-infection with HIV in Indonesia. *Hepatol Res* 2013; **43**: 1032-1039 [PMID: 23336705 DOI: 10.1111/hepr.12053]
- 8 **Utsumi T**, Yano Y, Lusida MI, Amin M, Soetjipto H, Hayashi Y. Serologic and molecular characteristics of hepatitis B virus among school children in East Java, Indonesia. *Am J Trop Med Hyg* 2010; **83**: 189-193 [PMID: 20595500 DOI: 10.4269/ajtmh.2010.09-0589]
- 9 **Mizokami M**, Orito E, Ohba K, Ikeo K, Lau JY, Gojobori T. Constrained evolution with respect to gene overlap of hepatitis B virus. *J Mol Evol* 1997; **44** Suppl 1: S83-S90 [PMID: 9071016]
- 10 **Sa-Nguanmoo P**, Tangkijvanich P, Thawornsuk N, Vichaiwattana P, Prianantathavorn K, Theamboonlers A, Tanaka Y, Poovorawan Y. Molecular epidemiological study of hepatitis B virus among migrant workers from Cambodia, Laos, and Myanmar to Thailand. *J Med Virol* 2010; **82**: 1341-1349 [PMID: 20572086 DOI: 10.1002/jmv.21828]
- 11 **Hasan I**. Epidemiology of hepatitis B. *Acta Med Indones* 2005; **37**: 231-234 [PMID: 16317222]
- 12 **Achwan WA**, Muttaqin Z, Zakaria E, Depamede SA, Mulyanto S, Tsuda F, Takahashi K, Abe N, Mishihiro S. Epidemiology of hepatitis B, C, and E viruses and human immunodeficiency virus infections in Tahuna, Sangihe-Talaud Archipelago, Indonesia. *Intervirology* 2007; **50**: 408-411 [PMID: 18185013 DOI: 10.1159/000112915]
- 13 **Lusida MI**, Nugrahaputra VE, Soetjipto R, Nagano-Fujii M, Sasayama M, Utsumi T, Hotta H. Novel subgenotypes of hepatitis B virus genotypes C and D in Papua, Indonesia. *J Clin Microbiol* 2008; **46**: 2160-2166 [PMID: 18463220 DOI: 10.1128/JCM.01681-07]
- 14 **Ol HS**, Bjoerkvoll B, Sothy S, Van Heng Y, Hoel H, Husebekk A, Gutteberg T, Larsen S, Husum H. Prevalence of hepatitis B and hepatitis C virus infections in potential blood donors in rural Cambodia. *Southeast Asian J Trop Med Public Health* 2009; **40**: 963-971 [PMID: 19842380]
- 15 **Suwannakarn K**, Tangkijvanich P, Thawornsuk N, Theamboonlers A, Tharmaphornpilas P, Yoocharoen P, Chongsri-sawat V, Poovorawan Y. Molecular epidemiological study of hepatitis B virus in Thailand based on the analysis of pre-S and S genes. *Hepatol Res* 2008; **38**: 244-251 [PMID: 17711443]
- 16 **Louisiriratchanakul S**, Olinger CM, Arunkaewchaemsri P, Poovorawan Y, Kanoksinsombat C, Thongme C, Sa-Nguanmoo P, Krasae S, Theamboonlert A, Oota S, Fongsatitkul L, Puapairoj C, Promwong C, Weber B. The distribution of hepatitis B virus genotypes in Thailand. *J Med Virol* 2012; **84**: 1541-1547 [PMID: 22930500 DOI: 10.1002/jmv.23363]
- 17 **Lingao AL**, Torres NT, Muñoz N, Lansang MA, West SK, Bosch FX, Domingo EO. Mother to child transmission of hepatitis B virus in the Philippines. *Infection* 1989; **17**: 275-279 [PMID: 2599650]
- 18 **Lansang MA**. Epidemiology and control of hepatitis B infection: a perspective from the Philippines, Asia. *Gut* 1996; **38** Suppl 2: S43-S47 [PMID: 8786053]
- 19 **Wong SN**, Ong JP, Labio ME, Cabahug OT, Daez ML, Valdellon EV, Sollano JD, Arguillas MO. Hepatitis B infection among adults in the Philippines: A national seroprevalence study. *World J Hepatol* 2013; **5**: 214-219 [PMID: 23671726 DOI: 10.4254/wjh.v5.i4.214]
- 20 **Viet L**, Lan NT, Ty PX, Björkvoll B, Hoel H, Gutteberg T, Husebekk A, Larsen S, Skjerve E, Husum H. Prevalence of hepatitis B & hepatitis C virus infections in potential blood donors in rural Vietnam. *Indian J Med Res* 2012; **136**: 74-81 [PMID: 22885267]
- 21 **Reekie J**, Gidding HF, Kaldor JM, Liu B. Country of birth and other factors associated with hepatitis B prevalence in a population with high levels of immigration. *J Gastroenterol Hepatol* 2013; **28**: 1539-1544 [PMID: 23621437 DOI: 10.1111/jgh.12245]
- 22 **Merican I**, Guan R, Amarapuka D, Alexander MJ, Chuta-



- putti A, Chien RN, Hasnain SS, Leung N, Lesmana L, Phiet PH, Sjalfoellah Noer HM, Sollano J, Sun HS, Xu DZ. Chronic hepatitis B virus infection in Asian countries. *J Gastroenterol Hepatol* 2000; **15**: 1356-1361 [PMID: 11197043]
- 23 **Yousuf R**, Rapiaah M, Ahmed SA, Rosline H, Salam A, Selamah S, Roshan TM. Trends in hepatitis B virus infection among blood donors in Kelantan, Malaysia: a retrospective study. *Southeast Asian J Trop Med Public Health* 2007; **38**: 1070-1074 [PMID: 18613548]
- 24 **Jutavijittum P**, Yousukh A, Samountry B, Samountry K, Ounavong A, Thammavong T, Keokhamphue J, Toriyama K. Seroprevalence of hepatitis B and C virus infections among Lao blood donors. *Southeast Asian J Trop Med Public Health* 2007; **38**: 674-679 [PMID: 17883005]
- 25 **Ting-Lu Z**, Zhi-Ping X, Hong-Yu L, Chang-Hong G, Liang Y, Qiang D, Kai-Ling X, Yan-Ming M, Yue-He D, Ling-Yang Z. A community-based sero-epidemiological study of hepatitis B infection in Lianyungang, China, 2010. *Western Pac Surveill Response J* 2012; **3**: 69-75 [PMID: 23908927 DOI: 10.5365/WP-SAR.2011.2.1.008]
- 26 **Liu GC**, Sui GY, Liu GY, Zheng Y, Deng Y, Gao YY, Wang L. A Bayesian meta-analysis on prevalence of hepatitis B virus infection among Chinese volunteer blood donors. *PLoS One* 2013; **8**: e79203 [PMID: 24236110 DOI: 10.1371/journal.pone.0079203]
- 27 **Chen P**, Yu C, Ruan B, Yang S, Ren J, Xu W, Luo Z, Li L. Prevalence of hepatitis B in insular regions of southeast China: a community-based study. *PLoS One* 2013; **8**: e56444 [PMID: 23437134 DOI: 10.1371/journal.pone.0056444]
- 28 **James L**, Fong CW, Foong BH, Wee MK, Chow A, Shum E, Chew SK. Hepatitis B Seroprevalence Study 1999. *Singapore Med J* 2001; **42**: 420-424 [PMID: 11811609]
- 29 **Ang LW**, Cutter J, James L, Goh KT. Seroepidemiology of hepatitis B virus infection among adults in Singapore: a 12-year review. *Vaccine* 2013; **32**: 103-110 [PMID: 24200974 DOI: 10.1016/j.vaccine.2013.10.057]
- 30 **Sebastian VJ**, Bhattacharya S, Ray S, Ahmad MZ. Hepatitis-B surface antigen and VDRL in healthy blood donors of Brunei Darussalam. *Singapore Med J* 1989; **30**: 568-570 [PMID: 2635403]
- 31 **Alexander MJ**, Sinnatamby AS, Rohaimah MJ, Harun AH, Ng JS. Incidence of hepatitis B infection in Brunei Darussalam--analysis of racial distribution. *Ann Acad Med Singapore* 1990; **19**: 344-346 [PMID: 2393233]
- 32 **Kim H**, Shin AR, Chung HH, Kim MK, Lee JS, Shim JJ, Kim BH. Recent trends in hepatitis B virus infection in the general Korean population. *Korean J Intern Med* 2013; **28**: 413-419 [PMID: 23864799 DOI: 10.3904/kjim.2013.28.4.413]
- 33 **Hyun HJ**, Shim JJ, Kim JW, Lee JS, Lee CK, Jang JY, Kim BH. The Prevalence of Elevated Alanine Transaminase and its Possible Causes in the General Korean Population. *J Clin Gastroenterol* 2013; Epub ahead of print [PMID: 24275711]
- 34 **Roman S**, Panduro A. HBV endemicity in Mexico is associated with HBV genotypes H and G. *World J Gastroenterol* 2013; **19**: 5446-5453 [PMID: 24023487 DOI: 10.3748/wjg.v19.i33.5446]
- 35 **Seeger C**, Zoulim F, Mason WS. Hepadnaviruses. In: Knipe DM, Howley PM. *Fields Virology*. Philadelphia: Lippincott, Williams and Wilkins, 2013: 2185-2221
- 36 **Lindh M**, Andersson AS, Gusdal A. Genotypes, nt 1858 variants, and geographic origin of hepatitis B virus--large-scale analysis using a new genotyping method. *J Infect Dis* 1997; **175**: 1285-1293 [PMID: 9180165]
- 37 **Norder H**, Hammas B, Lee SD, Bile K, Couroucé AM, Mushahwar IK, Magnius LO. Genetic relatedness of hepatitis B viral strains of diverse geographical origin and natural variations in the primary structure of the surface antigen. *J Gen Virol* 1993; **74** (Pt 7): 1341-1348 [PMID: 8336122]
- 38 **Sugauchi F**, Orito E, Ichida T, Kato H, Sakugawa H, Kakumu S, Ishida T, Chutaputti A, Lai CL, Ueda R, Miyakawa Y, Mizokami M. Hepatitis B virus of genotype B with or without recombination with genotype C over the precore region plus the core gene. *J Virol* 2002; **76**: 5985-5992 [PMID: 12021331]
- 39 **Sakamoto T**, Tanaka Y, Simonetti J, Osiowy C, Borresen ML, Koch A, Kurbanov F, Sugiyama M, Minuk GY, McMahon BJ, Joh T, Mizokami M. Classification of hepatitis B virus genotype B into 2 major types based on characterization of a novel subgenotype in Arctic indigenous populations. *J Infect Dis* 2007; **196**: 1487-1492 [PMID: 18008228]
- 40 **Orito E**, Ichida T, Sakugawa H, Sata M, Horiike N, Hino K, Okita K, Okanoue T, Iino S, Tanaka E, Suzuki K, Watanabe H, Hige S, Mizokami M. Geographic distribution of hepatitis B virus (HBV) genotype in patients with chronic HBV infection in Japan. *Hepatology* 2001; **34**: 590-594 [PMID: 11526547]
- 41 **Chan HL**. JGH Foundation emerging leadership lecture. Significance of hepatitis B virus genotypes and mutations in the development of hepatocellular carcinoma in Asia. *J Gastroenterol Hepatol* 2011; **26**: 8-12 [PMID: 21175787 DOI: 10.1111/j.1440-1746.2010.06514.x]
- 42 **Heriyanto DS**, Yano Y, Utsumi T, Anggorowati N, Rinonce HT, Lusida MI, Soetjipto C, Ratnasari N, Maduseno S, Purnama PB, Nurdjanah S, Hayashi Y. Mutations within enhancer II and BCP regions of hepatitis B virus in relation to advanced liver diseases in patients infected with subgenotype B3 in Indonesia. *J Med Virol* 2012; **84**: 44-51 [PMID: 22095534 DOI: 10.1002/jmv.22266]
- 43 **Sakamoto T**, Tanaka Y, Orito E, Co J, Clavio J, Sugauchi F, Ito K, Ozasa A, Quino A, Ueda R, Sollano J, Mizokami M. Novel subtypes (subgenotypes) of hepatitis B virus genotypes B and C among chronic liver disease patients in the Philippines. *J Gen Virol* 2006; **87**: 1873-1882 [PMID: 16760389]
- 44 **Nurainy N**, Muljono DH, Sudoyo H, Marzuki S. Genetic study of hepatitis B virus in Indonesia reveals a new subgenotype of genotype B in east Nusa Tenggara. *Arch Virol* 2008; **153**: 1057-1065 [PMID: 18463783 DOI: 10.1007/s00705-008-0092-z]
- 45 **Mulyanto SN**, Surayah K, Tsuda F, Ichiyama K, Takahashi M, Okamoto H. A nationwide molecular epidemiological study on hepatitis B virus in Indonesia: identification of two novel subgenotypes, B8 and C7. *Arch Virol* 2009; **154**: 1047-1059 [PMID: 19499283 DOI: 10.1007/s00705-009-0406-9]
- 46 **Thedja MD**, Muljono DH, Nurainy N, Sukowati CH, Verhoef J, Marzuki S. Ethnogeographical structure of hepatitis B virus genotype distribution in Indonesia and discovery of a new subgenotype, B9. *Arch Virol* 2011; **156**: 855-868 [PMID: 21318309 DOI: 10.1007/s00705-011-0926-y]
- 47 **Utsumi T**, Yano Y, Truong BX, Tanaka Y, Mizokami M, Seo Y, Kasuga M, Kawabata M, Hayashi Y. Molecular epidemiological study of hepatitis B virus infection in two different ethnic populations from the Solomon Islands. *J Med Virol* 2007; **79**: 229-235 [PMID: 17245721]
- 48 **Sugauchi F**, Mizokami M, Orito E, Ohno T, Kato H, Suzuki S, Kimura Y, Ueda R, Butterworth LA, Cooksley WG. A novel variant genotype C of hepatitis B virus identified in isolates from Australian Aborigines: complete genome sequence and phylogenetic relatedness. *J Gen Virol* 2001; **82**: 883-892 [PMID: 11257194]
- 49 **Utsumi T**, Lusida MI, Yano Y, Nugrahaputra VE, Amin M, Juniastuti Y, Hotta H. Complete genome sequence and phylogenetic relatedness of hepatitis B virus isolates in Papua, Indonesia. *J Clin Microbiol* 2009; **47**: 1842-1847 [PMID: 19386834 DOI: 10.1128/JCM.02328-08]
- 50 **Cavinta L**, Sun J, May A, Yin J, von Meltzer M, Radtke M, Barzaga NG, Cao G, Schaefer S. A new isolate of hepatitis B virus from the Philippines possibly representing a new subgenotype C6. *J Med Virol* 2009; **81**: 983-987 [PMID: 19382274 DOI: 10.1002/jmv.21475]
- 51 **Mulyanto SN**, Surayah K, Tjahyono AA, Jirintai S, Takahashi M, Okamoto H. Identification and characterization of



- novel hepatitis B virus subgenotype C10 in Nusa Tenggara, Indonesia. *Arch Virol* 2010; **155**: 705-715 [PMID: 20306210 DOI: 10.1007/s00705-010-0628-x]
- 52 **Mulyanto SN**, Wahyono A, Jirintai S, Takahashi M, Okamoto H. Analysis of the full-length genomes of novel hepatitis B virus subgenotypes C11 and C12 in Papua, Indonesia. *J Med Virol* 2011; **83**: 54-64 [PMID: 21108339 DOI: 10.1002/jmv.21931]
- 53 **Utsumi T**, Nugrahaputra VE, Amin M, Hayashi Y, Hotta H, Lusida MI. Another novel subgenotype of hepatitis B virus genotype C from papuans of Highland origin. *J Med Virol* 2011; **83**: 225-234 [PMID: 21181916 DOI: 10.1002/jmv.21963]
- 54 **Mulyanto P**, Depamede SN, Wahyono A, Jirintai S, Nagashima S, Takahashi M, Nishizawa T, Okamoto H. Identification of four novel subgenotypes (C13-C16) and two inter-genotypic recombinants (C12/G and C13/B3) of hepatitis B virus in Papua province, Indonesia. *Virus Res* 2012; **163**: 129-140 [PMID: 21925554 DOI: 10.1016/j.virusres.2011.09.002]
- 55 **Banerjee A**, Kurbanov F, Datta S, Chandra PK, Tanaka Y, Mizokami M, Chakravarty R. Phylogenetic relatedness and genetic diversity of hepatitis B virus isolates in Eastern India. *J Med Virol* 2006; **78**: 1164-1174 [PMID: 16847957]
- 56 **Olinger CM**, Jutavijittum P, Hübschen JM, Yousukh A, Samounry B, Thammavong T, Toriyama K, Muller CP. Possible new hepatitis B virus genotype, southeast Asia. *Emerg Infect Dis* 2008; **14**: 1777-1780 [PMID: 18976569 DOI: 10.3201/eid1411.080437]
- 57 **Tatematsu K**, Tanaka Y, Kurbanov F, Sugauchi F, Mano S, Maeshiro T, Nakayoshi T, Wakuta M, Miyakawa Y, Mizokami M. A genetic variant of hepatitis B virus divergent from known human and ape genotypes isolated from a Japanese patient and provisionally assigned to new genotype J. *J Virol* 2009; **83**: 10538-10547 [PMID: 19640977 DOI: 10.1128/JVI.00462-09]
- 58 **Uchida Y**, Kouyama JI, Naiki K, Sugawara K, Inao M, Nakayama N, Mochida S. Novel hepatitis B virus strain developing due to recombination between genotypes H and B strains isolated from a Japanese patient. *Hepatol Res* 2013; Epub ahead of print [PMID: 24020990 DOI: 10.1111/hepr.12238]
- 59 **Yu MW**, Yeh SH, Chen PJ, Liaw YF, Lin CL, Liu CJ, Shih WL, Kao JH, Chen DS, Chen CJ. Hepatitis B virus genotype and DNA level and hepatocellular carcinoma: a prospective study in men. *J Natl Cancer Inst* 2005; **97**: 265-272 [PMID: 15713961]
- 60 **Yang HI**, Lu SN, Liaw YF, You SL, Sun CA, Wang LY, Hsiao CK, Chen PJ, Chen DS, Chen CJ, Taiwan Community-Based Cancer Screening Project Group. Hepatitis B e antigen and the risk of hepatocellular carcinoma. *N Engl J Med* 2002; **347**: 168-174 [PMID: 12124405]
- 61 **Iloeje UH**, Yang HI, Su J, Jen CL, You SL, Chen CJ, Risk Evaluation of Viral Load Elevation and Associated Liver Disease/Cancer-In HBV (the REVEAL-HBV) Study Group. Predicting cirrhosis risk based on the level of circulating hepatitis B viral load. *Gastroenterology* 2006; **130**: 678-686 [PMID: 16530509]
- 62 **Chen CJ**, Yang HI, Su J, Jen CL, You SL, Lu SN, Huang GT, Iloeje UH. Risk of hepatocellular carcinoma across a biological gradient of serum hepatitis B virus DNA level. *JAMA* 2006; **295**: 65-73 [PMID: 16391218]
- 63 **Ogawa M**, Hasegawa K, Naritomi T, Torii N, Hayashi N. Clinical features and viral sequences of various genotypes of hepatitis B virus compared among patients with acute hepatitis B. *Hepatol Res* 2002; **23**: 167-177 [PMID: 12076712]
- 64 **Chan HL**, Hui AY, Wong ML, Tse AM, Hung LC, Wong VW, Sung JJ. Genotype C hepatitis B virus infection is associated with an increased risk of hepatocellular carcinoma. *Gut* 2004; **53**: 1494-1498 [PMID: 15361502]
- 65 **Chu CJ**, Hussain M, Lok AS. Hepatitis B virus genotype B is associated with earlier HBeAg seroconversion compared with hepatitis B virus genotype C. *Gastroenterology* 2002; **122**: 1756-1762 [PMID: 12055581]
- 66 **Chu CM**, Liaw YF. Genotype C hepatitis B virus infection is associated with a higher risk of reactivation of hepatitis B and progression to cirrhosis than genotype B: a longitudinal study of hepatitis B e antigen-positive patients with normal aminotransferase levels at baseline. *J Hepatol* 2005; **43**: 411-417 [PMID: 16006001]
- 67 **Sumi H**, Yokosuka O, Seki N, Arai M, Imazeki F, Kurihara T, Kanda T, Fukai K, Kato M, Saisho H. Influence of hepatitis B virus genotypes on the progression of chronic type B liver disease. *Hepatology* 2003; **37**: 19-26 [PMID: 12500184]
- 68 **Idrees M**, Khan S, Riazuddin S. Common genotypes of hepatitis B virus. *J Coll Physicians Surg Pak* 2004; **14**: 344-347 [PMID: 15233887]
- 69 **Masood Z**, Jawaaid M, Khan RA, Rehman S. Screening for hepatitis B and C: A routine preoperative investigation. *Pak J Med Sci* 2005; **21**: 455-459
- 70 **Strathdee SA**, Zafar T, Brahmabhatt H, Baksh A, ul Hassan S. Rise in needle sharing among injection drug users in Pakistan during the Afghanistan war. *Drug Alcohol Depend* 2003; **71**: 17-24 [PMID: 12821202]
- 71 **Yang HI**, Yeh SH, Chen PJ, Iloeje UH, Jen CL, Su J, Wang LY, Lu SN, You SL, Chen DS, Liaw YF, Chen CJ, REVEAL-HBV Study Group. Associations between hepatitis B virus genotype and mutants and the risk of hepatocellular carcinoma. *J Natl Cancer Inst* 2008; **100**: 1134-1143 [PMID: 18695135 DOI: 10.1093/jnci/djn243]
- 72 **Sharma S**, Sharma B, Singla B, Chawla YK, Chakraborti A, Saini N, Duseja A, Das A, Dhiman RK. Clinical significance of genotypes and precore/basal core promoter mutations in HBV related chronic liver disease patients in North India. *Dig Dis Sci* 2010; **55**: 794-802 [PMID: 20043209 DOI: 10.1007/s10620-009-1083-y]
- 73 **Tanaka Y**, Mukaide M, Orito E, Yuen MF, Ito K, Kurbanov F, Sugauchi F, Asahina Y, Izumi N, Kato M, Lai CL, Ueda R, Mizokami M. Specific mutations in enhancer II/core promoter of hepatitis B virus subgenotypes C1/C2 increase the risk of hepatocellular carcinoma. *J Hepatol* 2006; **45**: 646-653 [PMID: 16935384]
- 74 **Kim JK**, Chang HY, Lee JM, Baatarkhuu O, Yoon YJ, Park JY, Kim do Y, Han KH, Chon CY, Ahn SH. Specific mutations in the enhancer II/core promoter/precore regions of hepatitis B virus subgenotype C2 in Korean patients with hepatocellular carcinoma. *J Med Virol* 2009; **81**: 1002-1008 [PMID: 19382267 DOI: 10.1002/jmv.21501]
- 75 **Truong BX**, Yano Y, Seo Y, Phuong TM, Tanaka Y, Kato H, Miki A, Utsumi T, Azuma T, Trach NK, Mizokami M, Hayashi Y, Kasuga M. Variations in the core promoter/precore region in HBV genotype C in Japanese and Northern Vietnamese patients. *J Med Virol* 2007; **79**: 1293-1304 [PMID: 17607788]
- 76 **Liu S**, Zhang H, Gu C, Yin J, He Y, Xie J, Cao G. Associations between hepatitis B virus mutations and the risk of hepatocellular carcinoma: a meta-analysis. *J Natl Cancer Inst* 2009; **101**: 1066-1082 [PMID: 19574418 DOI: 10.1093/jnci/djp180]
- 77 **Sugauchi F**, Ohno T, Orito E, Sakugawa H, Ichida T, Komatsu M, Kuramitsu T, Ueda R, Miyakawa Y, Mizokami M. Influence of hepatitis B virus genotypes on the development of preS deletions and advanced liver disease. *J Med Virol* 2003; **70**: 537-544 [PMID: 12794715]
- 78 **Suzuki Y**, Kobayashi M, Ikeda K, Suzuki F, Arfase Y, Akuta N, Hosaka T, Saitoh S, Kobayashi M, Someya T, Matsuda M, Sato J, Watabiki S, Miyakawa Y, Kumada H. Persistence of acute infection with hepatitis B virus genotype A and treatment in Japan. *J Med Virol* 2005; **76**: 33-39 [PMID: 15779048]
- 79 **Matsuura K**, Tanaka Y, Hige S, Yamada G, Murawaki Y, Komatsu M, Kuramitsu T, Kawata S, Tanaka S, Izumi N, Okue C, Kakumu S, Okanoue T, Hino K, Hiasa Y, Sata M, Maeshiro T, Sugauchi F, Nojiri S, Joh T, Miyakawa Y, Mizokami M. Distribution of hepatitis B virus genotypes among

- patients with chronic infection in Japan shifting toward an increase of genotype A. *J Clin Microbiol* 2009; **47**: 1476-1483 [PMID: 19297602 DOI: 10.1128/JCM.02081-08]
- 80 **Sánchez-Tapias JM**, Costa J, Mas A, Bruguera M, Rodés J. Influence of hepatitis B virus genotype on the long-term outcome of chronic hepatitis B in western patients. *Gastroenterology* 2002; **123**: 1848-1856 [PMID: 12454842]
  - 81 **Yuen MF**, Wong DK, Sablon E, Tse E, Ng IO, Yuan HJ, Siu CW, Sander TJ, Bourne EJ, Hall JG, Condearey LD, Lai CL. HBsAg seroclearance in chronic hepatitis B in the Chinese: virological, histological, and clinical aspects. *Hepatology* 2004; **39**: 1694-1701 [PMID: 15185311]
  - 82 **Lok AS**, Lai CL, Wu PC, Leung EK. Long-term follow-up in a randomised controlled trial of recombinant alpha 2-interferon in Chinese patients with chronic hepatitis B infection. *Lancet* 1988; **2**: 298-302 [PMID: 2899719]
  - 83 **Janssen HL**, van Zonneveld M, Senturk H, Zeuzem S, Akarca US, Cakaloglu Y, Simon C, So TM, Gerken G, de Man RA, Niesters HG, Zondervan P, Hansen B, Schalm SW, HBV 99-01 Study Group; Rotterdam Foundation for Liver Research. Pegylated interferon alfa-2b alone or in combination with lamivudine for HBeAg-positive chronic hepatitis B: a randomised trial. *Lancet* 2005; **365**: 123-129 [PMID: 15639293]
  - 84 **Perrillo RP**, Schiff ER, Davis GL, Bodenheimer HC, Lindsay K, Payne J, Dienstag JL, O'Brien C, Tamburro C, Jacobson IM. A randomized, controlled trial of interferon alfa-2b alone and after prednisone withdrawal for the treatment of chronic hepatitis B. The Hepatitis Interventional Therapy Group. *N Engl J Med* 1990; **323**: 295-301 [PMID: 2195346]
  - 85 **Lok AS**, Wu PC, Lai CL, Lau JY, Leung EK, Wong LS, Ma OC, Lauder IJ, Ng CP, Chung HT. A controlled trial of interferon with or without prednisone priming for chronic hepatitis B. *Gastroenterology* 1992; **102**: 2091-2097 [PMID: 1587429]
  - 86 **Wai CT**, Chu CJ, Hussain M, Lok AS. HBV genotype B is associated with better response to interferon therapy in HBeAg(+) chronic hepatitis than genotype C. *Hepatology* 2002; **36**: 1425-1430 [PMID: 12447868]
  - 87 **Erhardt A**, Blondin D, Hauck K, Sagir A, Kohnle T, Heintges T, Häussinger D. Response to interferon alfa is hepatitis B virus genotype dependent: genotype A is more sensitive to interferon than genotype D. *Gut* 2005; **54**: 1009-1013 [PMID: 15951551]
  - 88 **Wiegand J**, Hasenclever D, Tillmann HL. Should treatment of hepatitis B depend on hepatitis B virus genotypes? A hypothesis generated from an explorative analysis of published evidence. *Antivir Ther* 2008; **13**: 211-220 [PMID: 18505172]
  - 89 **Deng M**, Zhou X, Gao S, Yang SG, Wang B, Chen HZ, Ruan B. The effects of telbivudine in late pregnancy to prevent intra-uterine transmission of the hepatitis B virus: a systematic review and meta-analysis. *Virol J* 2012; **9**: 185 [PMID: 22947333 DOI: 10.1186/1743-422X-9-185]
  - 90 **Lai CL**, Yuen MF. Chronic hepatitis B—new goals, new treatment. *N Engl J Med* 2008; **359**: 2488-2491 [PMID: 19052131 DOI: 10.1056/NEJMe0808185]
  - 91 **Lau GK**. Current treatments for patients with HBeAg-positive chronic hepatitis B virus infection: a comparison focusing on HBeAg seroconversion. *Liver Int* 2010; **30**: 512-520 [PMID: 20102511 DOI: 10.1111/j.1478-3231.2009.02198.x]
  - 92 **European Association For The Study Of The Liver**. EASL Clinical Practice Guidelines: management of chronic hepatitis B. *J Hepatol* 2009; **50**: 227-242 [PMID: 19054588 DOI: 10.1016/j.jhep.2008.10.001]
  - 93 **Eun JR**, Lee HJ, Kim TN, Lee KS. Risk assessment for the development of hepatocellular carcinoma: according to on-treatment viral response during long-term lamivudine therapy in hepatitis B virus-related liver disease. *J Hepatol* 2010; **53**: 118-125 [PMID: 20471129 DOI: 10.1016/j.jhep.2010.02.026]
  - 94 **Kurokawa M**, Hiramatsu N, Oze T, Yakushijin T, Miyazaki M, Hosui A, Miyagi T, Yoshida Y, Ishida H, Tatsumi T, Kiso S, Kanto T, Kasahara A, Iio S, Doi Y, Yamada A, Oshita M, Kaneko A, Mochizuki K, Hagiwara H, Mita E, Ito T, Inui Y, Katayama K, Yoshihara H, Imai Y, Hayashi E, Hayashi N, Takehara T. Long-term effect of lamivudine treatment on the incidence of hepatocellular carcinoma in patients with hepatitis B virus infection. *J Gastroenterol* 2012; **47**: 577-585 [PMID: 22231575 DOI: 10.1007/s00535-011-0522-7]
  - 95 **Lok AS**, Zoulim F, Locarnini S, Bartholomeusz A, Ghany MG, Pawlotsky JM, Liaw YF, Mizokami M, Kuiken C, Hepatitis B Virus Drug Resistance Working Group. Antiviral drug-resistant HBV: standardization of nomenclature and assays and recommendations for management. *Hepatology* 2007; **46**: 254-265 [PMID: 17596850]
  - 96 **Chang TT**, Lai CL, Chien RN, Guan R, Lim SG, Lee CM, Ng KY, Nicholls GJ, Dent JC, Leung NW. Four years of lamivudine treatment in Chinese patients with chronic hepatitis B. *J Gastroenterol Hepatol* 2004; **19**: 1276-1282 [PMID: 15482535]
  - 97 **Jin YJ**, Shim JH, Lee HC, Yoo DJ, Kim KM, Lim YS, Suh DJ. Suppressive effects of entecavir on hepatitis B virus and hepatocellular carcinoma. *J Gastroenterol Hepatol* 2011; **26**: 1380-1388 [PMID: 21884247 DOI: 10.1111/j.1440-1746.2011.06776.x]
  - 98 **Wong GL**, Chan HL, Mak CW, Lee SK, Ip ZM, Lam AT, Iu HW, Leung JM, Lai JW, Lo AO, Chan HY, Wong VW. Entecavir treatment reduces hepatic events and deaths in chronic hepatitis B patients with liver cirrhosis. *Hepatology* 2013; **58**: 1537-1547 [PMID: 23389810 DOI: 10.1002/hep.26301]

**P- Reviewers:** Narciso-Schiavon JL, Sandblom G

**S- Editor:** Zhai HH **L- Editor:** Roemmele A **E- Editor:** Liu SQ



## Acquisition and dissemination mechanisms of CTX $\Phi$ in *Vibrio cholerae*: New paradigm for *dif* residents

Bhabatosh Das, G Balakrish Nair, Rupak K Bhadra

Bhabatosh Das, G Balakrish Nair, Centre for Human Microbial Ecology, Translational Health Science and Technology Institute, Gurgaon 122 016, Haryana, India

Rupak K Bhadra, Infectious Diseases and Immunology Division, CSIR-Indian Institute of Chemical Biology, Kolkata 700032, India

Supported by Research in the Laboratory of Das B and Nair GB is funded in part by Department of Science Technology, No. SB/FT/LS-309/2012; Government of India (GOI) and the Department of Biotechnology, No. BT/MB/THSTI/HMC-SFC/2011; Research in the Laboratory of Bhadra RK is partly financially supported by Council of Scientific and Industrial Research, GOI and Indian Council of Medical Research, GOI

Author contributions: Das B, Nair GB and Bhadra RK solely contributed to this article.

Correspondence to: Rupak K Bhadra, MSc, PhD, FAScT, Senior Principal Scientist, Infectious Diseases and Immunology Division, CSIR-Indian Institute of Chemical Biology, 4 Raja S. C. Mullick Road, Jadavpur, Kolkata 700032, India. [rupakbhadr@iicb.res.in](mailto:rupakbhadr@iicb.res.in)

Telephone: +91-33-24995840 Fax: +91-33-24735197

Received: December 26, 2013 Revised: April 10, 2014

Accepted: May 13, 2014

Published online: May 27, 2014

### Abstract

*Vibrio cholerae* (*V. cholerae*) genome is equipped with a number of integrative mobile genetic element (IMGE) like prophages, plasmids, transposons or genomic islands, which provides fitness factors that help the pathogen to survive in changing environmental conditions. Metagenomic analyses of clinical and environmental *V. cholerae* isolates revealed that dimer resolution sites (*dif*) harbor several structurally and functionally distinct IMGEs. All IMGEs present in the *dif* region exploit chromosomally encoded tyrosine recombinases, XerC and XerD, for integration. Integration takes place due to site-specific recombination between two specific DNA sequences; chromosomal sequence is called *attB* and IMGEs sequence is called *attP*. Different IMGEs present in the *attP* region have different

*attP* structure but all of them are recognized by XerC and XerD enzymes and mediate either reversible or irreversible integration. Cholera toxin phage (CTX $\Phi$ ), a lysogenic filamentous phage carrying the cholera toxin genes *ctxAB*, deserves special attention because it provides *V. cholerae* the crucial toxin and is always present in the *dif* region of all epidemic cholera isolates. Therefore, understanding the mechanisms of integration and dissemination of CTX $\Phi$ , genetic and ecological factors which support CTX $\Phi$  integration as well as production of virion from chromosomally integrated phage genome and interactions of CTX $\Phi$  with other genetic elements present in the genomes of *V. cholerae* is important for learning more about the biology of cholera pathogen.

© 2014 Baishideng Publishing Group Inc. All rights reserved.

**Key words:** *Vibrio cholerae*; Cholera toxin phage; VGJ $\Phi$ ; Plasmids, Integrative mobile genetic element; XerC; XerD; Dimer resolution sites; *attP*; *attB*

**Core tip:** Integrative mobile genetic element (IMGE) like prophage, plasmid, transposon or genomic island plays crucial roles in the evolution of bacterial pathogens. The *Vibrio cholerae* (*V. cholerae*) genome harbors several such IMGEs, which provides virulence, antibiotic resistance and other fitness traits to cholera pathogen and directly contributes in its evolution. Cholera toxin encoding phage (CTX $\Phi$ ) is a well characterized IMGE, found integrated in the dimer resolution sites of all epidemic *V. cholerae* strains and exploits host encoded XerC and XerD recombinases for its lysogenic conversion. In this review we discussed about integration and dissemination of CTX $\Phi$  and related IMGEs of *V. cholerae*.

Das B, Nair GB, Bhadra RK. Acquisition and dissemination mechanisms of CTX $\Phi$  in *Vibrio cholerae*: New paradigm for *dif* residents. *World J Med Genet* 2014; 4(2): 27-33 Available from: URL: <http://www.wjgnet.com/2220-3184/full/v4/i2/27.htm> DOI:

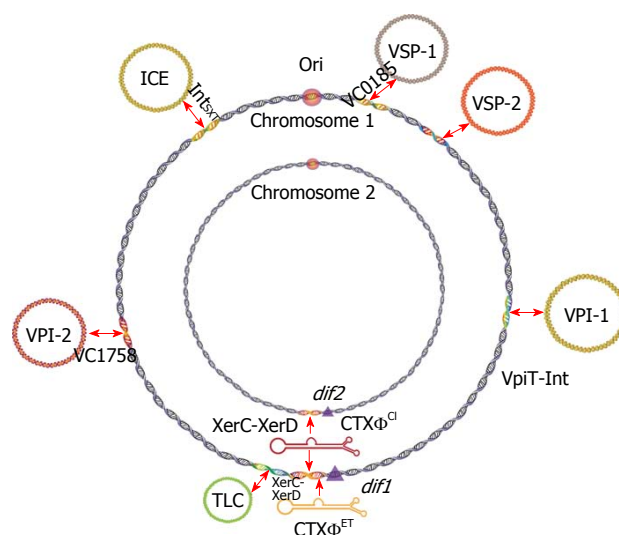


## INTRODUCTION

Bacterial genomes are highly dynamic and equipped with large number of integrative mobile genetic element (IMGE) like, prophages, plasmids, transposons or genomic islands (GIs), which allow bacteria to respond rapidly to changing environmental conditions and help them to survive in hostile environments<sup>[1]</sup>. The movement of genetic traits attributed to horizontal gene transfer systems is mediated by variety of homologous and non-homologous recombination processes that integrate, excise, and translocate genes into specific sites by DNA recombinases. The polynucleotidyl transferase that catalyzes the cutting and joining of phosphodiester bonds between DNA molecules necessary to move a segment of DNA is termed recombinase. DNA recombinases are generally needed to act at specific DNA sequences. Most IMGEs encode specific dedicated recombinase for their movement<sup>[2]</sup>. Among several IMGEs, phages are the best-characterized genetic elements, which play major role in real-time evolution of toxigenic bacterial pathogens<sup>[3]</sup>. Most pathogenic bacterial cells acquired their virulence traits either from phages<sup>[3]</sup> or other IMGEs like, plasmids<sup>[4]</sup>, transposons<sup>[5]</sup> or GIs<sup>[6]</sup>. Understanding the biology that supports emergence of pathogenic bacteria would be worthy to reduce disease burden and development of therapeutic agents. *Vibrio cholerae* (*V. cholerae*), the etiological agent of the acute secretory diarrheal disease cholera, represents a paradigm for toxin gene acquisition from cholera toxin phage (CTX $\Phi$ ) and evolution from non-pathogenic strains to toxigenic cholera pathogen<sup>[3]</sup>.

Both toxigenic and non-toxigenic *V. cholerae* strains are autochthonous inhabitant of estuaries and are introduced into the human intestine through contaminated water or food<sup>[7]</sup>. Only toxigenic *V. cholerae* isolates belonging to the serogroup O1 or O139 cause epidemic or pandemic cholera. Toxigenic *V. cholerae* harbor one or multiple copies of integrated CTX $\Phi$  either in large or in small or in both chromosomes. CTX $\Phi$  exploits host encoded tyrosine recombinases XerC and XerD to integrate into the chromosome dimer resolution site, called dimer resolution sites (*dif*), present in close proximity of the replication termination region of either chromosome of *V. cholerae*<sup>[8]</sup>.

In this review, we provided a brief view about the genome of cholera pathogen, the major IMGEs reported in the genome of cholera pathogen and molecular insights into integration mechanisms of CTX $\Phi$  in the *V. cholerae* chromosomes. The main emphasis is on the mechanistic part of CTX $\Phi$  integration, how (+)ssDNA genome of CTX $\Phi$  is recognized by the XerC and XerD enzymes and proceed for unusual irreversible integration. We also discussed about other IMGEs that follow CTX $\Phi$  like integration. Finally, we discussed how cooperative interactions between CTX $\Phi$  and other IMGEs could lead



**Figure 1** Schematic representation of the integrative mobile genetic elements present in the chromosome 1 (large) or 2 (small) of *Vibrio cholerae*. Except CTX $\Phi$ , all other IMGEs have unique *attB* site either in large or small chromosome. IMGEs present in the *dif* region exploit host-encoded recombinases for integration while the genome of other IMGEs encode their own recombinases for integration (see detail in the text). Except CTX $\Phi$ , integration of rest of the IMGEs relies on their dsDNA genome and the integration is reversible. IMGE: Integrative mobile genetic element; CTX $\Phi$ : Cholera toxin phage; ICE: Integrative conjugative elements; VPI: *Vibrio* pathogenicity island; VSP: *Vibrio* seventh pandemic; TLC: Plasmids; *dif*: Dimer resolution sites.

to dissemination of CTX $\Phi$  among closely or distantly related bacterial cells.

## CTX $\Phi$ AND OTHER IMGEs IN *V. CHOLERA* CHROMOSOMES

IMGEs are segments of DNA that encode or exploit host encoded recombinase(s) for intra- and inter-chromosomal movement within or between closely or distantly related bacterial cells. Comparative analysis of *V. cholerae* genomes revealed presence of IMGEs in both the chromosomes of toxigenic and non-toxigenic strains<sup>[9]</sup>. Three major classes of IMGEs are reported in cholera pathogens: prophages, GIs and integrative conjugative elements (ICEs) as shown in Figure 1. Both pathogenic determinants and antimicrobial resistance traits of cholera pathogen are encoded by these IMGEs<sup>[3]</sup>.

As mentioned, CTX $\Phi$  and several other phages always integrate in the *dif* site present one in each of the two circular chromosomes of *V. cholerae*<sup>[10,11]</sup>. The *dif* sites present in the large and small chromosomes are called *dif1* and *dif2*, respectively. *dif* consists of 28-bp DNA sequences, 11-bp binding sites for XerC and XerD separated by a 6-bp central region<sup>[12]</sup>. Strand exchange occurs immediately after binding sites of XerC and XerD. CTX $\Phi$ , RS1 and TLC element integrates as a single copy or in multiple tandemly arrayed copies in *dif1*<sup>[10,13]</sup>. Other IMGEs were found integrated at the *dif* sites as a single copy<sup>[9]</sup>. Prophages and other genomic elements present in the *dif* regions do not encode any recombinase but exploit host encoded XerC and XerD enzymes for their



inter- and intra-chromosomal movement.

Other than prophages, almost all the current 7<sup>th</sup> pandemic *V. cholerae* isolates harbor four other GIs, namely, *Vibrio* pathogenicity islands-1 (VPI-1) and VPI-2 and *Vibrio* seventh pandemic islands-1 (VSP-1) and VSP-2<sup>[14]</sup> as shown in Figure 1. All four GIs encode putative tyrosine recombinases for their integration and excision. Although no experiment has yet been conducted to demonstrate the integration of GIs in the *V. cholerae* chromosomes, their excision has been established by conventional genetic analysis<sup>[15]</sup>. Among the four pathogenicity islands, VPI-1 is crucial since it participates at different levels in the CTXΦ associated cholera development: (1) VPI-1 encodes toxin co-regulated type-IV pilus that acts as a receptor for CTXΦ as well as it is a major colonization factor of *V. cholerae* and is essential for the disease development; and (2) it encodes the transcriptional factor ToxT that positively regulates the expression of cholera toxin (CT) genes *ctxAB*<sup>[16,17]</sup>. VPI-1 is a 41-kb DNA segment physically linked to a tmRNA gene (*ssrA*), flanked by two nearly identical repeat sequences<sup>[18]</sup>. It is found preferentially in the toxigenic strains. VPI-1 carries two putative tyrosine recombinases, called Int<sub>vpi</sub> and VpiT<sup>[15]</sup>. Sequence analyses of these two putative recombinases indicate that they are quite different from each other. While Int<sub>vpi</sub> contains the conserved R1-H-R2-Y signature motif of the tyrosine recombinases, this motif is not clear in VpiT. VPI-1 appeared to be mobile since an extrachromosomal circular form was detected in *V. cholerae* cells. This is further supported by the fact that it is absent in the genomes of non-pathogenic *V. cholerae* strains<sup>[15]</sup>.

Presence or absence of VSPs could serve as distinct genetic signatures to differentiate the previous (6<sup>th</sup>) and the current (7<sup>th</sup>) pandemic *V. cholerae* isolates. VSP-1, a 16-kb DNA segment, harbors 13 open reading frame (ORF)s found in the sequenced seventh pandemic *V. cholerae* strain N16961 (Heidelberg *et al*<sup>[18]</sup> 2000). A XerCD like putative tyrosine recombinase (Int<sub>VSP-1</sub>) is present in the VSP-1 might participate in the integration and dissemination of VSP-1 island. Compared to other GIs, VSP-1 is highly conserved among O1 El Tor isolates. Recent study has demonstrated that the VSP-1 encoded enzyme DncV preferentially synthesizes hybrid c-AMP-GMP molecule, which directly contribute to the fitness of cholera pathogen<sup>[19]</sup>.

ICEs are usually large DNA fragments that can integrate in specific positions of bacterial chromosomes using their own tyrosine recombinase and able to excise and disseminate in closely or distantly related bacterial species by conjugation. Diverse ICEs have been detected in several gram-positive and gram-negative bacteria<sup>[20]</sup>. sulfamethoxazole and trimethoprim resistance traits (SXT), a 99.5 kb ICE, was first discovered in the chromosome of *V. cholerae* O139 MO10 strain from Southern India and it encodes resistances to several antibiotics like sulfamethoxazole (*sul2*), trimethoprim (*dfrA18* or *dfrA1*) chloramphenicol (*floR*), streptomycin (*strA* and *strB*), tetracycline (*tetA* and *tetR*), *etc*<sup>[21]</sup>. Wozniak *et al*<sup>[22]</sup> have sequenced and analyzed several SXT like elements present

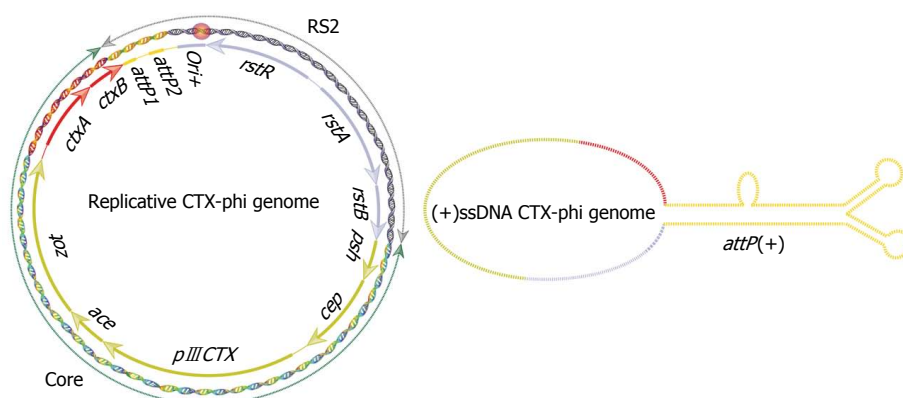
in diverse bacterial species including *V. cholerae* and their analysis indicated that similar organization and conservation of the core genes of these elements.

## CTXΦ AND ITS INTEGRATION IN THE CHROMOSOMES OF *V. CHOLERAE*

All epidemic *V. cholerae* isolates carry CTX prophage in their large and/or small chromosomes. CT, the toxin responsible for profuse diarrhoeal disease cholera, is encoded by the genome of CTXΦ. CTX prophages characterized from different toxigenic *V. cholerae* strains are fairly different at the DNA sequence level but their genomic organization and function of each of the protein encoded by the phage genome are identical. CTXΦ has a approximately 7000-nt ss (+) DNA genome arranged in structurally and functionally distinct two modular structures, RS2 (repeat sequence 2) and core (Figure 2). RS2 comprises three genes designated as *rstR*, *rstA* and *rstB*. RstA, initiator of rolling circle replication, is essential for phage replication<sup>[23,24]</sup>. It carries a conserved Y-X-X-X-Y motif for initiation and termination of phage replication at *ori* region. RstB, a single stranded DNA binding protein, plays a crucial role in phage integration<sup>[8]</sup>. It was proposed that RstB maintains proper (+) *attP* structure of CTXΦ that is recognized by XerC-XerD and prevents access of cytoplasmic single stranded DNA binding proteins interference in phage integration<sup>[25]</sup>. RstR acts as a transcriptional modulator and repressed transcription of *rstA* and *rstB* from *P<sub>rstA</sub>*, the only phage promoter required for CTXΦ replication and integration<sup>[26]</sup>. Core region comprises of seven genes responsible for phage morphogenesis and toxin production. Five genes, namely, *psh*, *cep*, *gIII<sup>CTX</sup>*, *ace* and *zot* encode proteins essential for phage morphogenesis and phage assembly. The Psh, Cep, OrfU (*pIII<sup>CTX</sup>*) and Ace proteins are phage structural proteins, which encapsulate single stranded phage genome whereas Zot protein play central role in phage assembly<sup>[3]</sup>. Interestingly, *ctxA* and *ctxB* genes coding for CT are not required for phage morphogenesis but is essential for disease development as discussed above.

Replicative genome of CTXΦ is detrimental to *V. cholerae* growth<sup>[27]</sup>. All toxigenic *V. cholerae* cells harbor integrated CTXΦ in large and/or small chromosomes. Although CTXΦ does not encode any recombinase for its integration, the replicative genome of CTXΦ carries two XerC-XerD binding sites in inverted orientations, called *attP1* and *attP2*<sup>[28]</sup>. In *attP1*, XerC and XerD binding sites are separated by 12-bp overlap region while in *attP2* the length of the overlap region is 5-6 bp. Both *attPs* are connected by a 90-bp DNA sequence. In the (+)ssDNA phage genome the 150-bp region encompassing *attP1* and *attP2* formed a hairpin structure by intra-strand complementary base pairing, which creates a phage integration site *attP*(+) (Figure 2).

The integration of CTXΦ is mediated by a complex DNA-protein reaction within a nucleoprotein complex consisting of one pair each of host encoded XerC and



**Figure 2 Replicative and integrative genomes of cholera toxin phage.** Replicative genome of CTX $\Phi$  arranged in structurally and functionally distinct two modular structures RS2 and core. RS2 encoded proteins are essential for phage replication, integration and transcriptional regulation of phage genes. Core encoded proteins are essential for phage morphogenesis and virion production. Folded (+)ssDNA phage genome is essential for phage integration. Functional XerCD binding site, *attP*(+), is formed by complementary base pairing between *attP1* and *attP2* of (+)ssDNA of phage genome. We have used similar color codes for replicative and integrative phage genomes for easy understanding of genetic attributes of *attP*(+) region. CTX $\Phi$ : Cholera toxin phage; RS2: Repeat sequence 2.

XerD recombinases, *attP*(+) and *dif* site. Within this nucleoprotein complex, first the XerC recombinase catalyzes the cleavage of the phosphodiester bond between last base of its binding site and first base of overlap region and creates transient XerC-DNA covalent phosphotyrosyl linkages with its binding site and leave free 5' hydroxyl extremities on the overlap side. After cleavage, few bases from the 5' end of both overlap regions melt from their complementary strand and attack the XerC-DNA phosphotyrosyl bond of their recombining partner. For the formation of covalent phosphodiester bond between overlap region base and XerC binding site of *attP* and *dif*, stabilization of exchanged strand by Watson-Crick or wobble base-pairing interactions is essential<sup>[29]</sup>. Strand stabilization between exchanged strands determines the tropism of phage integration. Although both Xer proteins are essential for integration reaction, CTX $\Phi$  integration needs only the XerC catalytic activity for one pair of strand exchange and final integration<sup>[28]</sup>. As a consequence, *dif* compatibility of CTX $\Phi$  solely determines by the homology between the overlap region bases next to XerC binding sites. The pseudo-Holliday junction form after XerC mediated strand exchanges is probably resolved during host chromosome replication. Once integrated, *attP*(+) is again converted to *attP1* and *attP2*, none of these sequences are suitable for Xer-mediated reactions. This process makes CTX $\Phi$  integration irreversible. The only possibility to form the *attP*(+) hairpins from integrated phage genome is by cruciform, *i.e.*, two opposite hairpins could extrude through intra-strand complementary base pairing. This phenomenon is very rare and *in vivo* the cruciform structure is very unstable. *V. cholerae* cells carry several ssDNA nucleases, which could easily destabilize *attP*(+). Probably, for this reason CTX $\Phi$  excision is not detected under standard laboratory condition. Recently, CTX $\Phi$  excision from toxigenic *V. cholerae* isolates has been reported<sup>[30]</sup>. Although, the mechanisms of excision and detail genotype of the reported *V. cholerae* strain have not been investigated, it might be due to

lack of some ssDNA nucleases, which stabilize the loop structure and form *attP*(+) for XerC-XerD mediated excision. Other possibilities might be some other IMGEs co-integrate with CTX $\Phi$  and provides functional XerC-XerD binding sites, hence clearance of CTX $\Phi$ .

## CTX $\Phi$ VARIANTS: CONVENTIONAL VS CURRENT GROUPING

Conventionally, sequence of *rstR* gene is used for classification of CTX $\Phi$ s. Based on the sequence similarity of *rstR*, CTX $\Phi$  has been grouped into four categories: CTX $\Phi$ <sup>ET</sup>, CTX $\Phi$ <sup>Cl</sup>, CTX $\Phi$ <sup>Cl<sub>a</sub></sup> and CTX $\Phi$ <sup>Env</sup>. Name of each of the phage classes has been chosen according to the host cells in which they were most frequently isolated. CTX $\Phi$  isolated from current pandemic isolates harbored a mosaic genome, few ORFs are identical to CTX $\Phi$ <sup>ET</sup> and few are similar to CTX $\Phi$ <sup>Cl</sup>. For example, CTX prophages present in the genome of current African (B33) and Asian (MJ-1236) isolates carry *ctxB* allele of CTX $\Phi$ <sup>Cl</sup> type whereas *rstR* genes in the RS2 region are identical to those of CTX $\Phi$ <sup>ET</sup> phage<sup>[31]</sup>. Nevertheless, recent Haitian *V. cholerae* isolates carries *ctxB* allele<sup>[32]</sup> neither identical to *ctxB*<sup>Cl</sup> or *ctxB*<sup>ET</sup>.

On the other side, CTX $\Phi$  can be classified into three broad classes based on their *dif* compatibility (Table 1). CTX $\Phi$  isolated from 6<sup>th</sup> pandemic *V. cholerae* isolates could integrate at both *dif1* and *dif2* sites. Bases immediate to XerC cleavage site of *attP*(+) could form Watson-Crick or Wobble base pair interaction with the similar bases of overlap region of both *dif1* and *dif2*<sup>[29]</sup>. In contrast, most well characterized CTX $\Phi$  from 7<sup>th</sup> pandemic El Tor cholera isolates could form such interaction only with the overlap region of *dif1* but not with *dif2*. Thus, the integration CTX $\Phi$  is specific for *dif1*. Recent *V. cholerae* isolates from Africa and India harbour CTX $\Phi$  in either chromosomes indicating their *attP* is similar to *attP* of CTX $\Phi$  isolated from 6<sup>th</sup> pandemic *V. cholerae*. In contrast, some environmental isolates carry *dif* sequenc-

**Table 1** Cholera toxin phage variants and their dimer resolution sites specificities

CTXΦ variants	Isolated from	Chromosomal integration site	Ref.
CTXΦ <sup>d1</sup>	O1 El Tor	<i>difA</i> , <i>dif1</i>	Das <i>et al</i> <sup>[29]</sup> Das <i>et al</i> <sup>[10]</sup>
CTXΦ <sup>d1/2</sup>	O1 classical	<i>difA</i> , <i>dif1</i> and <i>dif2</i>	Das <i>et al</i> <sup>[29]</sup> Das <i>et al</i> <sup>[10]</sup>
CTXΦ <sup>dG</sup>	Environmental	<i>difG</i>	Das <i>et al</i> <sup>[29]</sup> Das <i>et al</i> <sup>[10]</sup>

CTXΦ: Cholera toxin phage; *dif*: Dimer resolution sites.

es<sup>[33]</sup>, which support integration of only specific type of phage, CTXΦ<sup>dG</sup>. All these CTXΦ variants are different in terms of their integration specificity as well as integration efficiency. Highest level of integration efficiency was reported in between *dif1* and CTXΦ isolated from 7<sup>th</sup> pandemic El Tor cholera isolates while the lowest integration efficiency is detected in between *difG* and CTXΦ<sup>dG</sup> isolated from environmental *V. cholerae*<sup>[29]</sup>. Our extensive work on the integration mechanisms revealed that integration specificity and efficiency of all CTX phages including Haitian and altered variants rely only on their *attP* sequence but not on the type of *ctxB* allele present in the core genome.

## ECOLOGY OF *V. CHOLERA* *DIF* SITES

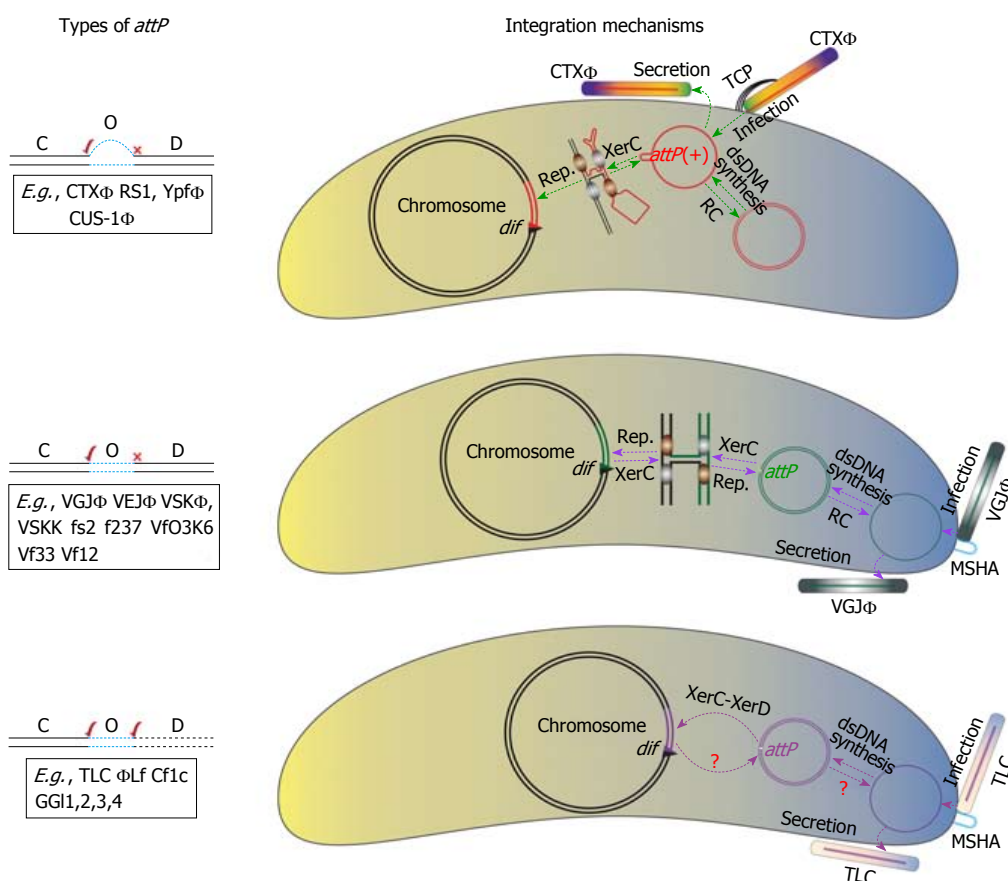
Each *V. cholerae* genome carries two *dif* sites, one for each of the two circular chromosomes of the bacterium. Till now, four different *dif* alleles have been identified in *V. cholerae* strains, *difA*, *dif1*, *dif2* and *difG*<sup>[10]</sup>. *difA* and *dif1* are identical except two bases in the XerC binding site<sup>[10,13]</sup>. All *dif* sites consist of 28 bp DNA sequence, 11 bp binding sites for XerC and XerD separated by a 6 bp central region. *dif1*, *dif2* and *difG* have different bases in the overlap region at the borders of which XerC and XerD mediate strand exchanges during chromosome dimer resolution or IMGEs integration<sup>[10]</sup>. Metagenomic analysis of clinical and environmental *V. cholerae* isolates revealed *dif1* could be the integration site of several IMGEs, including phages (CTXΦ, VGJΦ, VEJΦ, VSK, fs2, f237, Vf33), satellite phages (RS1), plasmids (TLC) and small GIs<sup>[9]</sup>. Several *V. cholerae* cells have different arrangements of IMGE at *dif1* site, *e.g.*, single or multiple IMGE(s) in multiple combinations<sup>[9]</sup>. Almost all integrated IMGEs are flanked by *dif*-like sequences. Several *in vivo* and *in vitro* experiments have been conducted to demonstrate the XerC and XerD mediated integration of CTXΦ, RS1, VGJΦ, and TLC<sup>[8,28,29,34]</sup>. Based on the sequence of *attP* site, IMGEs have been classified into three groups, CTXΦ-type, VGJΦ-type and TLC-type (Figure 3). Although, all of them rely on Xer recombinases for their integration, the integration mechanisms are very different (Figure 3). While the integration of CTXΦ-type IMGE rely on its folded single stranded genome, VGJΦ-type and TLC-type IMGEs use their replicative double stranded DNA

for integration. Integration of CTXΦ-type IMGE are irreversible, other two could excise from the chromosomes. Both CTXΦ-type and VGJΦ-type IMGEs use only XerC catalytic activity but this is not clear for TLC. Bases in the overlap region of CTXΦ-type and VGJΦ-type *attP* have homology to *dif1* next to the XerC binding site only, but the bases of overlap region of TLC-type *attP* has homology next to both XerC and XerD binding sites. At this point it is not clear whether TLC uses catalytic activity of XerC or XerD or both. Since the XerD binding site of *attP* of TLC is very degenerative it is not clear whether any other accessory protein(s) is needed for its integration. Currently, Barre and his colleagues at CGM-CNRS, France are working on TLC to address all these unanswered issues.

## CONCLUSION

Integration mechanism of CTXΦ and related elements like RS1, CUSΦ, YpfΦ are quite similar to the integration mechanism of integron cassette. In both the cases *attP* site is formed by intra-strand DNA base pairing which is recognized by tyrosine recombinases followed by catalysis of single pair of strand exchange with double stranded *attB*. The resulting pseudo-Holiday junction is resolved either by DNA replication or other DNA repairing system. Since integration of CTXΦ is irreversible, phage production from the integrated prophage genome relies on rolling circle replication. Virion production from integrated phage genome is detected only when multiple CTX prophage are present in tandem or CTX-prophages are flanked by pre-CTX or RS1 element. Recent metagenomic analysis revealed that several toxigenic *V. cholerae* isolates carry a single copy of the CTX prophage in the *dif* region. It is not clear whether these isolates are unable to produce virion or the special genotype of such toxigenic isolates could allow phage production due to excision of CTX prophage by forming cruciform structure under specific environmental conditions. Nevertheless, CTXΦ could replicate only in *V. cholerae* and few other vibrios but not in all species of *Vibrio* or other *γ-proteobacteria* like *Escherichia coli*. Further research is needed to determine what are the host factors involved in CTXΦ replication and what type of signal is recognised by the prophage to initiate rolling circle replication and virion production from toxigenic cholera pathogen. Clinical *V. cholerae* isolates carrying CTXΦ in the large chromosome also carry single or multiple copies of TLC element. Although, TLC is not essential for CTXΦ replication or integration, at least in the laboratory conditions, it is not clear why CTXΦ is always accompanied by the TLC element. More importantly, it is quite interesting to know the precise integration mechanism of TLC element, and whether the integration and excision of TLC could help CTXΦ dissemination in clinical and environmental *V. cholerae* strains. A recent study reported that TLC may promote CTXΦ integrations by contributing ideal attachment site (*dif1*) to the host chromosome<sup>[13]</sup>. Further work in this





**Figure 3** Schematic representation of the integration mechanisms of three classes of integrative mobile genetic elements present exclusively in the dimer resolution sites of bacterial genome. C-O-D denotes XerC-Overlap region-XerD. The XerC-XerD recombinases, attachment site of phage and bacterial genome are essential components of nucleoprotein complex. Compatibility of terminal bases *attP* with *attB* immediate after XerC or XerD cleavage is indicated by ✓. The nucleoprotein complex and the sequential strand exchanges are not yet characterized for TLC and related genetic elements. CTX $\Phi$  integration is irreversible. By contrast, VGJ $\Phi$  and TLC integration is reversible. Key steps in the life cycle of CTX $\Phi$ , VGJ $\Phi$  and TLC are also indicated. Host-encoded TCP and MSHA served as phage receptors for CTX $\Phi$ , VGJ $\Phi$  and TLC, respectively. Generally, TLC exploit fs2 encoded protein for encapsulation of its ssDNA genome and produces virion. Virion recognizes specific receptor present on its host cell surfaces and delivers its single-stranded DNA genome into the host cytoplasm. Once inside the host cell, the ssDNA phage genome is either converted into a double-stranded replicative phage genome by host machineries or directly integrated into the host chromosome. TLC: Plasmids; CTX $\Phi$ : Cholera toxin phage; TCP: Toxin co-regulated pilus; MSHA: Mannose sensitive hemagglutinin A.

direction is warranted to understand the intricate biology of TLC and CTX $\Phi$ .

## REFERENCES

- 1 Frost LS, Leplae R, Summers AO, Toussaint A. Mobile genetic elements: the agents of open source evolution. *Nat Rev Microbiol* 2005; **3**: 722-732 [PMID: 16138100 DOI: 10.1038/nrmicro1235]
- 2 Nash HA. Integration and excision of bacteriophage lambda: the mechanism of conservation site specific recombination. *Annu Rev Genet* 1981; **15**: 143-167 [PMID: 6461289 DOI: 10.1146/annurev.ge.15.120181.001043]
- 3 Waldor MK, Mekalanos JJ. Lysogenic conversion by a filamentous phage encoding cholera toxin. *Science* 1996; **272**: 1910-1914 [PMID: 8658163 DOI: 10.1126/science.272.5270.1910]
- 4 Brüggemann H. Insights in the pathogenic potential of *Propionibacterium acnes* from its complete genome. *Semin Cutan Med Surg* 2005; **24**: 67-72 [PMID: 16092793]
- 5 Veilleux S, Dubreuil JD. Presence of *Escherichia coli* carrying the EAST1 toxin gene in farm animals. *Vet Res* 2006; **37**: 3-13 [PMID: 16336921 DOI: 10.1051/vetres:2005045]
- 6 Dinges MM, Orwin PM, Schlievert PM. Exotoxins of *Staphylococcus aureus*. *Clin Microbiol Rev* 2000; **13**: 16-34, table of contents [PMID: 10627489 DOI: 10.1128/CMR.13.1.16-34.2000]
- 7 Sack DA, Sack RB, Nair GB, Siddique AK. Cholera. *Lancet* 2004; **363**: 223-233 [PMID: 14738797 DOI: 10.1016/S0140-6736(03)15328-7]
- 8 Huber KE, Waldor MK. Filamentous phage integration requires the host recombinases XerC and XerD. *Nature* 2002; **417**: 656-659 [PMID: 12050668 DOI: 10.1038/nature00782]
- 9 Chun J, Grim CJ, Hasan NA, Lee JH, Choi SY, Haley BJ, Taviani E, Jeon YS, Kim DW, Lee JH, Brettin TS, Bruce DC, Challacombe JF, Detter JC, Han CS, Munk AC, Chertkov O, Meincke L, Saunders E, Walters RA, Huq A, Nair GB, Colwell RR. Comparative genomics reveals mechanism for short-term and long-term clonal transitions in pandemic *Vibrio cholerae*. *Proc Natl Acad Sci USA* 2009; **106**: 15442-15447 [PMID: 19720995 DOI: 10.1073/pnas.0907787106]
- 10 Das B, Martínez E, Midonet C, Barre FX. Integrative mobile elements exploiting Xer recombination. *Trends Microbiol* 2013; **21**: 23-30 [PMID: 23127381 DOI: 10.1016/j.tim.2012.10.003]
- 11 Das B, Halder K, Pal P, Bhadra RK. Small chromosomal integration site of classical CTX prophage in Mozambique *Vibrio cholerae* O1 biotype El Tor strain. *Arch Microbiol* 2007; **188**: 677-683 [PMID: 17618421]
- 12 Carnoy C, Roten CA. The dif/Xer recombination systems in proteobacteria. *PLoS One* 2009; **4**: e6531 [PMID: 19727445 DOI: 10.1371/journal.pone.0006531]
- 13 Hassan F, Kamruzzaman M, Mekalanos JJ, Faruque SM. Satellite phage TLC $\Phi$  enables toxigenic conversion by CTX phage through dif site alteration. *Nature* 2010; **467**: 982-985



- [PMID: 20944629 DOI: 10.1038/nature09469]
- 14 **Murphy RA**, Boyd EF. Three pathogenicity islands of *Vibrio cholerae* can excise from the chromosome and form circular intermediates. *J Bacteriol* 2008; **190**: 636-647 [PMID: 17993521 DOI: 10.1128/JB.00562-07]
  - 15 **Rajanna C**, Wang J, Zhang D, Xu Z, Ali A, Hou YM, Karaolis DK. The vibrio pathogenicity island of epidemic *Vibrio cholerae* forms precise extrachromosomal circular excision products. *J Bacteriol* 2003; **185**: 6893-6901 [PMID: 14617653 DOI: 10.1128/JB.185.23.6893-6901.2003]
  - 16 **Karaolis DK**, Johnson JA, Bailey CC, Boedeker EC, Kaper JB, Reeves PR. A *Vibrio cholerae* pathogenicity island associated with epidemic and pandemic strains. *Proc Natl Acad Sci USA* 1998; **95**: 3134-3139 [PMID: 9501228]
  - 17 **Karaolis DK**, Somara S, Maneval DR, Johnson JA, Kaper JB. A bacteriophage encoding a pathogenicity island, a type-IV pilus and a phage receptor in cholera bacteria. *Nature* 1999; **399**: 375-379 [PMID: 10360577 DOI: 10.1038/20715]
  - 18 **Heidelberg JF**, Eisen JA, Nelson WC, Clayton RA, Gwinn ML, Dodson RJ, Haft DH, Hickey EK, Peterson JD, Umayam L, Gill SR, Nelson KE, Read TD, Tettelin H, Richardson D, Ermolaeva MD, Vamathevan J, Bass S, Qin H, Dragoi I, Sellers P, McDonald L, Utterback T, Fleischmann RD, Niernan WC, White O, Salzberg SL, Smith HO, Colwell RR, Mekalanos JJ, Venter JC, Fraser CM. DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. *Nature* 2000; **406**: 477-483 [PMID: 10952301 DOI: 10.1038/35020000]
  - 19 **Davies BW**, Bogard RW, Young TS, Mekalanos JJ. Coordinated regulation of accessory genetic elements produces cyclic di-nucleotides for *V. cholerae* virulence. *Cell* 2012; **149**: 358-370 [PMID: 22500802 DOI: 10.1016/j.cell.2012.01.053]
  - 20 **Burrus V**, Waldor MK. Shaping bacterial genomes with integrative and conjugative elements. *Res Microbiol* 2004; **155**: 376-386 [PMID: 15207870]
  - 21 **Hochhut B**, Lotfi Y, Mazel D, Faruque SM, Woodgate R, Waldor MK. Molecular analysis of antibiotic resistance gene clusters in *vibrio cholerae* O139 and O1 SXT constins. *Antimicrob Agents Chemother* 2001; **45**: 2991-3000 [PMID: 11600347 DOI: 10.1128/AAC.45.11.2991-3000.2001]
  - 22 **Wozniak RA**, Fouts DE, Spagnoletti M, Colombo MM, Ceccarelli D, Garriss G, Déry C, Burrus V, Waldor MK. Comparative ICE genomics: insights into the evolution of the SXT/R391 family of ICEs. *PLoS Genet* 2009; **5**: e1000786 [PMID: 20041216 DOI: 10.1371/journal.pgen.1000786]
  - 23 **Waldor MK**, Rubin EJ, Pearson GD, Kimsey H, Mekalanos JJ. Regulation, replication, and integration functions of the *Vibrio cholerae* CTXphi are encoded by region RS2. *Mol Microbiol* 1997; **24**: 917-926 [PMID: 9220000 DOI: 10.1046/j.1365-2958.1997.3911758.x]
  - 24 **Moyer KE**, Kimsey HH, Waldor MK. Evidence for a rolling-circle mechanism of phage DNA synthesis from both replicative and integrated forms of CTXphi. *Mol Microbiol* 2001; **41**: 311-323 [PMID: 11489120 DOI: 10.1046/j.1365-2958.2001.02517.x]
  - 25 **Falero A**, Caballero A, Ferrán B, Izquierdo Y, Fando R, Campos J. DNA binding proteins of the filamentous phages CTXphi and VGJphi of *Vibrio cholerae*. *J Bacteriol* 2009; **191**: 5873-5876 [PMID: 19617366 DOI: 10.1128/JB.01206-08]
  - 26 **Davis BM**, Kimsey HH, Kane AV, Waldor MK. A satellite phage-encoded antirepressor induces repressor aggregation and cholera toxin gene transfer. *EMBO J* 2002; **21**: 4240-4249 [PMID: 12169626 DOI: 10.1093/emboj/cdf427]
  - 27 **Faruque SM**, Rahman MM, Hasan AK, Nair GB, Mekalanos JJ, Sack DA. Diminished diarrheal response to *Vibrio cholerae* strains carrying the replicative form of the CTX(Phi) genome instead of CTX(Phi) lysogens in adult rabbits. *Infect Immun* 2001; **69**: 6084-6090 [PMID: 11553546 DOI: 10.1128/IAI.69.10.6084-6090.2001]
  - 28 **Val ME**, Bouvier M, Campos J, Sherratt D, Cornet F, Mazel D, Barre FX. The single-stranded genome of phage CTX is the form used for integration into the genome of *Vibrio cholerae*. *Mol Cell* 2005; **19**: 559-566 [PMID: 16109379 DOI: 10.1016/j.molcel.2005.07.002]
  - 29 **Das B**, Bischerour J, Val ME, Barre FX. Molecular keys of the tropism of integration of the cholera toxin phage. *Proc Natl Acad Sci USA* 2010; **107**: 4377-4382 [PMID: 20133778 DOI: 10.1073/pnas.0910212107]
  - 30 **Mantri CK**, Mohapatra SS, Singh DV. Effect of storage and sodium chloride on excision of CTXPhi or pre-CTXPhi and CTXPhi from *Vibrio cholerae* O139 strains. *Infect Genet Evol* 2010; **10**: 925-930 [PMID: 20621579 DOI: 10.1016/j.meegid.2010.05.015]
  - 31 **Safa A**, Nair GB, Kong RY. Evolution of new variants of *Vibrio cholerae* O1. *Trends Microbiol* 2010; **18**: 46-54 [PMID: 19942436 DOI: 10.1016/j.tim.2009.10.003]
  - 32 **Chin CS**, Sorenson J, Harris JB, Robins WP, Charles RC, Jean-Charles RR, Bullard J, Webster DR, Kasarskis A, Peluso P, Paxinos EE, Yamaichi Y, Calderwood SB, Mekalanos JJ, Schadt EE, Waldor MK. The origin of the Haitian cholera outbreak strain. *N Engl J Med* 2011; **364**: 33-42 [PMID: 21142692 DOI: 10.1056/NEJMoa1012928]
  - 33 **Das B**, Bischerour J, Barre FX. Molecular mechanism of acquisition of the cholera toxin genes. *Indian J Med Res* 2011; **133**: 195-200 [PMID: 21415494]
  - 34 **Das B**, Bischerour J, Barre FX. VGJphi integration and excision mechanisms contribute to the genetic diversity of *Vibrio cholerae* epidemic strains. *Proc Natl Acad Sci USA* 2011; **108**: 2516-2521 [PMID: 21262799 DOI: 10.1073/pnas.1017061108]

**P- Reviewers:** Gudlavalleti SK, Sarkar BL **S- Editor:** Ji FF

**L- Editor:** A **E- Editor:** Liu SQ



## Osteopathia striata with cranial sclerosis, Wilms' tumor and the *WTX* gene

Elisa Cattaneo, Sara Ciceri, Natascia Liberati, Paolo Radice, Luigi Tarani, Angelo Selicorni, Daniela Perotti

Elisa Cattaneo, Angelo Selicorni, UOS, Pediatric Genetic Unit, Pediatric Department, MBBM Foundation, S Gerardo Hospital, 20900 Monza, Italy

Sara Ciceri, Paolo Radice, Daniela Perotti, Molecular Bases of Genetic Risk and Genetic Testing Unit, Department of Preventive and Predictive Medicine, Fondazione IRCCS Istituto Nazionale dei Tumori, 20133 Milano, Italy

Natascia Liberati, Luigi Tarani, Department of Pediatrics, Policlinico Umberto I, Sapienza University, 00186 Roma, Italy

Author contributions: Cattaneo E, Ciceri S and Liberati N contributed equally to this work, they wrote the draft of the manuscript, developing the themes and the structure; Radice P, Tarani L, Selicorni A and Perotti D contributed equally to this work, they edited, wrote some sections and advised on the content.

Supported by Associazione Bianca Garavaglia, 21052 Busto Arsizio, Varese, Italy; Italian Association for Cancer Research (AIRC); Fondazione Pierfranco e Luisa Mariani, 20129 Milano, Italy

Correspondence to: Daniela Perotti, PhD, Molecular Bases of Genetic Risk and Genetic Testing Unit, Department of Preventive and Predictive Medicine, Fondazione IRCCS Istituto Nazionale dei Tumori, Via Venezian, 1, 20133 Milano, Italy. [daniela.perotti@istitutotumori.mi.it](mailto:daniela.perotti@istitutotumori.mi.it)

Telephone: +39-02-23902644 Fax: +39-02-23903073

Received: October 29, 2013 Revised: February 10, 2014

Accepted: February 16, 2014

Published online: May 27, 2014

patients and the involvement of *WTX* anomalies in OPCS and in WT.

© 2014 Baishideng Publishing Group Inc. All rights reserved.

**Key words:** Osteopathia striata with cranial sclerosis; Wilms' tumor; *WTX*; Mutation; Genetics

**Core tip:** Osteopathia striata with cranial sclerosis (OSCS), a condition often benign in females and severe and lethal in males, has a clinically heterogeneous presentation. Germline anomalies affecting the *WTX* gene, mapped to chromosome X, are causative of OSCS. Despite *WTX* mutations in Wilms' tumor (WT) that closely mirror those identified in OSCS patients, individuals with OSCS do not develop WT. This is in contrast with other syndromic conditions, in which germline mutations or epimutations, also found as somatic events in sporadic WTs, predispose to WT development.

Cattaneo E, Ciceri S, Liberati N, Radice P, Tarani L, Selicorni A, Perotti D. Osteopathia striata with cranial sclerosis, Wilms' tumor and the *WTX* gene. *World J Med Genet* 2014; 4(2): 34-38 Available from: URL: <http://www.wjgnet.com/2220-3184/full/v4/i2/34.htm> DOI: <http://dx.doi.org/10.5496/wjmg.v4.i2.34>

### Abstract

Osteopathia striata with cranial sclerosis (OSCS, OMIM#300373) is an X-linked dominant sclerosing bone dysplasia that shows a distinct phenotype in females and males. In 2009, Zandra Jenkins *et al* found that germline mutations in the *FAM123B/WTX/AMER1* gene, mapped to chromosome Xq11.2, cause both the familial and sporadic forms of OSCS. Intriguingly, the *WTX* gene was already known as a putative tumor suppressor gene, since in 2007 Rivera *et al* had reported inactivating *WTX* mutations in Wilms' tumor (WT), the most frequent renal tumor of childhood. Here we review the heterogeneous clinical presentation of OSCS

### OSTEOPATHIA STRIATA WITH CRANIAL SCLEROSIS

The X-linked inheritance pattern of osteopathia striata with cranial sclerosis (OSCS, OMIM#300373), previously predicted on clinical grounds, found confirmation when germline mutations involving the *WTX* gene, mapped to chromosome Xq11.2, were identified as the cause of OSCS<sup>[1]</sup>. All mutations either deleted the whole gene or resulted in the premature termination of translation<sup>[1]</sup>.

The *FAM123B/WTX/AMER1* gene encodes a 1135-amino acid protein with multiple protein-protein

interaction domains and N-terminal phosphatidylinositol 4,5-bisphosphate binding domains mediating its localization to the plasma membrane<sup>[1,3]</sup>. The *WTX* protein has been demonstrated to regulate the stability of  $\beta$ -catenin<sup>[2]</sup>, a key effector of the WNT/ $\beta$ -catenin signaling pathway (reviewed in Clevers and Nusse 2012<sup>[4]</sup>). The critical importance of this pathway during embryogenesis is clearly demonstrated by the pleiotropic clinical presentation of OPCS patients. The *WTX* gene has two splice forms, *WTXs1* full length and *WTXs2*, a shorter form encoding a 858-amino acid protein that lacks residues 50-326 and does not localize to the plasma membrane<sup>[1,5]</sup>. Only *WTXs1* is considered to be important in regulating the WNT signaling in the context of the development since disease-causing *WTX* mutations that do not affect the integrity of *WTXs2* have been reported<sup>[1,6-8]</sup>.

Females affected with OPCS present a great variability of the phenotype and, while previous studies<sup>[9-11]</sup> suggested, at least in some cases, a nonrandom X-inactivation<sup>[9-11]</sup> that could explain this phenomenon<sup>[6]</sup>, Jenkins *et al.*<sup>[11]</sup> demonstrated that in 19 *WTX* mutation-bearing heterozygous females X-inactivation ratios were not skewed.

Among the features that constitute the OPCS female phenotype, sclerosis of bone (especially the increased thickness and density mostly of the cranial base) and the fine, uniform, linear striations of the tubular bones are considered the hallmarks of the disease.

Cranial sclerosis is the most typical and early feature, being present since birth. It appears before the longitudinal striae that become evident in the first years of life. Fan-like striations of the iliac bones are present in more than 50% of cases. It is worth mentioning that longitudinal striations at the metaphyses and diaphyses of the tubular bones are seen only in females and in males that are mosaic for a *WTX* mutation<sup>[6,7,9,12]</sup>.

Other skeletal defects reported in the literature, although quite rare, are thoracic (pectus excavatum, broad flat ribs) and vertebral anomalies (2%), digital flexion contractures, phalangeal duplication, syndactyly, short or absent fibula and club feet (3%)<sup>[12]</sup>. Coronal craniostenosis has been described in one patient only<sup>[9]</sup>.

Facial dysmorphisms are rather frequent and sometimes the only pathological feature in addition to the sclerosis of the skull and the longitudinal striation of the long bones. Macrocephaly is documented in almost half of patients, followed by frontal and occipital bossing, prominent forehead, maxillary hypoplasia, mandible overgrowth with protuberance of the jaw and dental malocclusion. Female patients can also manifest ocular hypertelorism, downslanting palpebral fissures, broad and depressed nasal bridge, narrow high arched palate and low set dysplastic ears<sup>[13-22]</sup>. Dental anomalies have been reported in 30% of patients<sup>[20]</sup>. Regarding the neurological manifestations, intellectual disability has been described in a small percentage of patients, mainly associated with central nervous system defects (ventricular dilatation, abnormal gyration, corpus callosum hypoplasia or agenesis, hydrocephalus), as well as developmental and

speech delay.

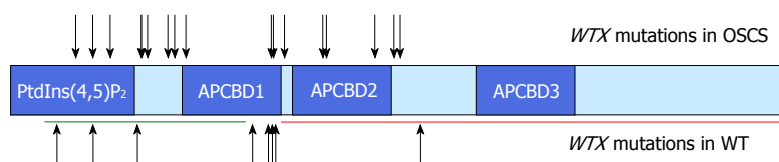
Conductive hearing loss can be considered a distinctive symptom of the disease, occurring in almost 50% of patients. In the remaining half, hearing loss is sensorineural or of mixed type<sup>[6,7,12,20,23]</sup>. Deafness is the result of bone sclerosis of mastoid cells, narrowing of the middle ear cavity, the mastoid antrum and the Eustachian canal, and of impaired mobility of the ossicles. High resolution computed tomography of the temporal bone has shown the presence in different patients of bilateral thickening and bone sclerosis of the skull base and mastoid cells, and the abnormal ossicular fixation to the bone surface of the middle ear cavity<sup>[23]</sup>. Sensorineural hearing loss could instead be due to the nerve encroachment.

Other cranial nerve deficiencies (oculomotor and hypoglossal, abducens and maxillary nerves) due to the narrowing of the nerve canals and foramina by the sclerosing process are reported. Unilateral peripheral facial palsy and congenital facial palsy were described in 4 patients<sup>[20,24]</sup>. The optic nerve may also be involved due to the narrowing of the optic foramina. Nerve palsy might be due to the sclerosing bone process<sup>[25]</sup> but it is also hypothesized that disruption of the nerve supporting vessels may lead to secondary cranial nerve deficiencies<sup>[20]</sup>. Lumbar spinal stenosis, defined as narrowing of the lumbar spinal canal, nerve root canal or intervertebral foramina, has been described in one patient only and could be thought of as a neurological complication of the disease<sup>[26]</sup>.

As already mentioned, OPCS manifests only with the hallmarks of the syndrome (cranial sclerosis and longitudinal striations of the long bones) accompanied by minor facial dysmorphisms or in association with internal organ anomalies, growth and mental retardation.

In female patients, the most frequently affected organs are the heart, with congenital defects including ventricular septal defects, patent ductus arteriosus, pulmonary atresia and valve stenosis, the lungs and the respiratory system in general, and the gastrointestinal and urogenital systems. The respiratory system may be affected in many patients. In particular, laryngotracheomalacia, nasal obstruction and recurrent bronchitis are reported in 13.5% of patients<sup>[12,20]</sup>. Cleft palate (Pierre Robin's triad) and bifid uvula can also be observed. Gastrointestinal anomalies, including omphalocele, intestinal malrotation and Hirschsprung's disease have been reported in 12% of patients<sup>[9,27,28]</sup>. Anal stenosis has been described in two girls only<sup>[29,30]</sup>.

OPCS in males is more severe than in females because it follows an X-linked dominant pattern of inheritance that determines hemizygoty of the mutation and, consequently, a wide spectrum of severe clinical manifestations, such as abortion, stillbirth and post-natal lethality. Despite this, cases with long survival are also described, allowing a clinical distinction in severe and mild forms in males. The male severe phenotype exhibits macrocephaly, facial dysmorphisms (frontal bossing, hypertelorism, low set ears, broad depressed nasal bridge and micrognathia) and bony sclerosis (more marked than in females), with



**Figure 1** *WTX* mutations in osteopathia striata with cranial sclerosis and Wilms' tumor. The full length *WTX* protein possesses two phosphatidylinositol(4,5)-bisphosphate [PtdIns(4,5)P<sub>2</sub>] binding domains, three adenomatous polyposis coli binding domains (APCBD) and a  $\beta$ -catenin binding region (red line)<sup>[2,3]</sup>. The smaller *WTX* isoform lacks amino acids 50-326 (green line)<sup>[1,5]</sup>. Functionally relevant mutations include whole gene deletions and truncating mutations. Arrows indicate the position of mutations introducing a stop codon or causing a frameshift of the reading frame and a premature stop codon<sup>[1,6-8,12,24,38,39,44,46,47]</sup>.

no metaphyseal striations, while other features are less frequent. The latter include genitourinary malformations (18%), bilateral absence of fibula (65%), cardiac defects (patent ductus arteriosus, atrial and ventricular septal defects, left ventricular non-compaction, tricuspid insufficiency and vascular ring) (31%), omphalocele and cleft of lips and palate (50%), ventriculomegaly and duplicated phalanges (30%)<sup>[8,31]</sup>. Prominent lumbar lordosis, joint luxation, camptodactyly and flexion contractures are also present with a lower incidence. Gastrointestinal anomalies such as omphalocele, duodenal web, malrotation of the gut, inguinal hernia and Hirschsprung's disease have also been reported<sup>[9,10,21,27,32]</sup>.

In the severe form, the prognosis is related to the severity of visceral malformations and a short survival is often present.

The mild phenotype is qualitatively different from the severe one, being characterized by the presence of mild neurodevelopmental delay (50% of patients), which might be attributable to the relative longevity of these patients and progressive neuromuscular disease, histologically similar to nemaline myopathy<sup>[8]</sup>, and by the absence of some anomalies such as fibular aplasia, duplicated phalanges, syndactyly<sup>[1,8]</sup>, and gastrointestinal<sup>[9,10,21,27,32]</sup>, cardiac<sup>[9,21,27,31,33,34]</sup> and genitourinary malformations<sup>[1,8,21,35,36]</sup>.

Characteristic features of the mild form are short stature, facial dysmorphisms and macrocephaly with cranial sclerosis, frontal bossing, hearing loss, high arched and cleft palate (75%), bifid uvula (25%), and extensive bony sclerosis with absent metaphyseal striations<sup>[8]</sup>. Milder bony sclerosis has been detected in males with mosaic mutation of *WTX*<sup>[37]</sup>. Striations of the long bones have also been observed in molecularly confirmed or suspected mosaics for *WTX* mutations<sup>[8,24,26,34,37,38]</sup>.

A possible genotype-phenotype correlation between the position of the *WTX* mutation and survival in males had been initially proposed<sup>[1]</sup> but further studies showed that this correlation is not absolute<sup>[6,7]</sup>.

## WILMS' TUMOR

Intriguingly, *WTX* has been also identified as a putative tumor suppressor gene in Wilms' tumor (WT). Since this gene, as already mentioned, resides on the X chromosome, it has been speculated that its anti-oncogenic activity can be inactivated by a single "hit" both in hemizygous males and in heterozygous females if the mutation affects the only functional allele on the active X chromosome<sup>[39]</sup>.

WT, the most common renal tumor of childhood, is an embryonal malignancy of the kidney that is thought to arise from metanephric mesenchyme. Histologically, it resembles fetal kidney, with varying proportions of blastemal, epithelial and stromal elements<sup>[40]</sup>. Approximately 40% of WTs occur in association with nephrogenic rests (NRs), embryonal remnants in the kidney which are known precursor lesions for WT<sup>[41]</sup>. The genetics of WT is heterogeneous and the *WT1* gene at 11p13 and the *WT2* locus at 11p15.5 have been associated with WT pathogenesis (reviewed in Huff<sup>[42]</sup> 2011, Royer-Pokora<sup>[43]</sup> 2013). Further genes involved in WT development include, in addition to *WTX*, *CTNNB1* and *TP53*<sup>[42,43]</sup>. *WTX* anomalies have been described in approximately 20% of WTs<sup>[39,44-47]</sup>. While *WTX* deletions and truncating mutations are somatically acquired, missense mutations of unknown functional relevance can be present in the germline (reviewed in Huff<sup>[42]</sup> 2011).

## OSCS, WT AND *WTX* MUTATIONS: IS THERE AN ASSOCIATION?

Whole gene deletions represent the majority of *WTX* mutations in WTs, whereas truncating mutations are more common in OSCS (reviewed in Huff<sup>[42]</sup> 2011). However, the spectrum of *WTX* truncating mutations in OSCS patients and in WTs is very similar<sup>[1,42]</sup> (Figure 1).

Different syndromic conditions associated with susceptibility to WT, such as the WAGR, the Denys-Drash and the Beckwith-Wiedemann syndromes, are due to germline mutations or epimutations affecting genes/loci also found involved in somatic events in sporadic WTs (reviewed in Scott *et al.*<sup>[48]</sup> 2006). In contrast, individuals with OSCS do not seem to have any predisposition to develop either WT or other malignancies<sup>[1,22,35]</sup>, although in a few of these patients the presence of bilateral multifocal NRs has been reported<sup>[8,35]</sup>. However, it has to be noted that NRs are found in approximately 1% of infant autopsies and that most of them do not form WT but spontaneously undergo regression or involution<sup>[40]</sup>. Thus, the detection of NRs in OSCS patients does not allow establishing a link with WT. Consistently, the *WTX*-knockout mice, despite exhibiting somatic overgrowth and malformation of several organs including kidney, do not appear to be tumor prone<sup>[49]</sup>. The lack of association between OSCS and WT could be explained assuming that *WTX* is mainly involved in WT progression rather than



in its early phase. This possibility is supported by a study detecting various levels of *WTX* mutation in different microdissected areas of the same tumor<sup>[46]</sup>.

Overall, current evidence suggests a possible involvement of the *WTX* gene in kidney development but is not consistent with its role in WT predisposition.

## REFERENCES

- Jenkins ZA, van Kogelenberg M, Morgan T, Jeffs A, Fukuzawa R, Pearl E, Thaller C, Hing AV, Porteous ME, Garcia-Miñaur S, Bohring A, Lacombe D, Stewart F, Fiskerstrand T, Bindoff L, Berland S, Adès LC, Tchan M, David A, Wilson LC, Hennekam RC, Donnai D, Mansour S, Cormier-Daire V, Robertson SP. Germline mutations in *WTX* cause a sclerosing skeletal dysplasia but do not predispose to tumorigenesis. *Nat Genet* 2009; **41**: 95-100 [PMID: 19079258 DOI: 10.1038/ng.270]
- Major MB, Camp ND, Berndt JD, Yi X, Goldenberg SJ, Hubbert C, Biechele TL, Gingras AC, Zheng N, Maccoss MJ, Angers S, Moon RT. Wilms tumor suppressor *WTX* negatively regulates WNT/beta-catenin signaling. *Science* 2007; **316**: 1043-1046 [PMID: 17510365 DOI: 10.1126/science.1141515]
- Grohmann A, Tanneberger K, Alzner A, Schneikert J, Behrens J. AMER1 regulates the distribution of the tumor suppressor APC between microtubules and the plasma membrane. *J Cell Sci* 2007; **120**: 3738-3747 [PMID: 17925383 DOI: 10.1242/jcs.011320]
- Clevers H, Nusse R. Wnt/ $\beta$ -catenin signaling and disease. *Cell* 2012; **149**: 1192-1205 [PMID: 22682243 DOI: 10.1016/j.cell.2012.05.012]
- Rivera MN, Kim WJ, Wells J, Stone A, Burger A, Coffman EJ, Zhang J, Haber DA. The tumor suppressor *WTX* shuttles to the nucleus and modulates WT1 activity. *Proc Natl Acad Sci USA* 2009; **106**: 8338-8343 [PMID: 19416806 DOI: 10.1073/pnas.0811349106]
- Perdu B, de Freitas F, Frints SG, Schouten M, Schrandt-Stumpel C, Barbosa M, Pinto-Basto J, Reis-Lima M, de Vernejoul MC, Becker K, Freckmann ML, Keymolen K, Haan E, Savarirayan R, Koenig R, Zabel B, Vanhoenacker FM, Van Hul W. Osteopathia striata with cranial sclerosis owing to *WTX* gene defect. *J Bone Miner Res* 2010; **25**: 82-90 [PMID: 20209645 DOI: 10.1359/jbmr.090707]
- Perdu B, Lakeman P, Mortier G, Koenig R, Lachmeijer AM, Van Hul W. Two novel *WTX* mutations underscore the unpredictability of male survival in osteopathia striata with cranial sclerosis. *Clin Genet* 2011; **80**: 383-388 [PMID: 20950377 DOI: 10.1111/j.1399-0004.2010.01553.x]
- Holman SK, Daniel P, Jenkins ZA, Herron RL, Morgan T, Savarirayan R, Chow CW, Bohring A, Mosel A, Lacombe D, Steiner B, Schmitt-Mechelke T, Schroter B, Raas-Rothschild A, Miñaur SG, Porteous M, Parker M, Quarrell O, Tapon D, Cormier-Daire V, Mansour S, Nash R, Bindoff LA, Fiskerstrand T, Robertson SP. The male phenotype in osteopathia striata congenita with cranial sclerosis. *Am J Med Genet A* 2011; **155A**: 2397-2408 [PMID: 22043478]
- Viot G, Lacombe D, David A, Mathieu M, de Broca A, Faivre L, Gigarel N, Munnich A, Lyonnet S, Le Merrer M, Cormier-Daire V. Osteopathia striata cranial sclerosis: non-random X-inactivation suggestive of X-linked dominant inheritance. *Am J Med Genet* 2002; **107**: 1-4 [PMID: 11807859 DOI: 10.1002/ajmg.10028]
- Rott HD, Krieg P, Rüttschle H, Kraus C. Multiple malformations in a male and maternal osteopathia striata with cranial sclerosis (OSCS). *Genet Couns* 2003; **14**: 281-288 [PMID: 14577672]
- Kraus C, König R, Rott HD. Comments on „osteopathia striata cranial sclerosis: non-random X-inactivation suggestive of X-linked dominant inheritance“. *Am J Med Genet A* 2003; **119A**: 400 [PMID: 12784316 DOI: 10.1002/ajmg.a.20178]
- Zicari AM, Tarani L, Perotti D, Papetti L, Nicita F, Liberati N, Spalice A, Salvatori G, Guaraldi F, Duse M. *WTX* R353X mutation in a family with osteopathia striata and cranial sclerosis (OS-CS): case report and literature review of the disease clinical, genetic and radiological features. *Ital J Pediatr* 2012; **38**: 27 [PMID: 22716240 DOI: 10.1186/1824-7288-38-27]
- Jones MD, Mulcahy ND. Osteopathia striata, osteopetrosis, and impaired hearing. A case report. *Arch Otolaryngol* 1968; **87**: 116-118 [PMID: 5635349 DOI: 10.1001/archotol.1968.00760060118004]
- Franklyn PP, Wilkinson D. Two cases of osteopathia striata, deafness and cranial osteopetrosis. *Ann Radiol (Paris)* 1978; **21**: 91-93 [PMID: 677756]
- Cortina H, Vallcanera A, Vidal J. Familial osteopathia striata with cranial condensation. *Pediatr Radiol* 1981; **11**: 87-90 [PMID: 7301452 DOI: 10.1007/BF00971786]
- Clément A, Garrigues C, Coursault-Durand R, Ledoux-Lebard G, Bonnin A. [A rare bone disease, but one not to be ignored: osteopathia striata]. *J Radiol* 1982; **63**: 673-676 [PMID: 7153962]
- Robinow M, Unger F. Syndrome of osteopathia striata, macrocephaly, and cranial sclerosis. *Am J Dis Child* 1984; **138**: 821-823 [PMID: 6475869]
- Piechowiak H, Goebel FD, Hirche U, Tyrell R. Cranial sclerosis with striated bone disease (osteopathia striata). *Klin Padiatr* 1986; **198**: 418-424 [PMID: 3784439]
- Kornreich L, Grunebaum M, Ziv N, Shuper A, Mimouni M. Osteopathia striata, cranial sclerosis with cleft palate and facial nerve palsy. *Eur J Pediatr* 1988; **147**: 101-103 [PMID: 3276523 DOI: 10.1007/BF00442625]
- König R, Dukiet C, Dörries A, Zabel B, Fuchs S. Osteopathia striata with cranial sclerosis: variable expressivity in a four generation pedigree. *Am J Med Genet* 1996; **63**: 68-73 [PMID: 8723089 DOI: 10.1002/(SICI)1096-8628(19960503)63:1<68::AID-AJMG14>3.0.CO;2-S]
- Pellegrino JE, McDonald-McGinn DM, Schneider A, Markowitz RI, Zackai EH. Further clinical delineation and increased morbidity in males with osteopathia striata with cranial sclerosis: an X-linked disorder? *Am J Med Genet* 1997; **70**: 159-165 [PMID: 9128936 DOI: 10.1002/(SICI)1096-8628(19970516)70:2<159::AID-AJMG11>3.3.CO;2-C]
- Behninger C, Rott HD. Osteopathia striata with cranial sclerosis: literature reappraisal argues for X-linked inheritance. *Genet Couns* 2000; **11**: 157-167 [PMID: 10893667]
- Magliulo G, Parrotto D, Zicari AM, Zappala D, Lo Mele L, Primicerio P, Marini M. Osteopathia striata-cranial sclerosis: otorhinolaryngologic clinical presentation and radiologic findings. *Am J Otolaryngol* 2007; **28**: 59-63 [PMID: 17162136 DOI: 10.1016/j.amjoto.2006.05.003]
- Ciceri S, Cattaneo E, Fossati C, Radice P, Selicorni A, Perotti D. First evidence of vertical paternal transmission of osteopathia striata with cranial sclerosis. *Am J Med Genet A* 2013; **161A**: 1173-1176 [PMID: 23494899 DOI: 10.1002/ajmg.a.35813]
- De Keyser J, Bruyland M, De Greve J, Leemans J, Potvlieghe R, Six R, Ebinger G. Osteopathia striata with cranial sclerosis. Report of a case and review of the literature. *Clin Neurol Neurosurg* 1983; **85**: 41-48 [PMID: 6303664 DOI: 10.1016/0304-8467(83)90022-7]
- Deniz FE, Köseoğlu RD. Osteopathia striata with cranial sclerosis and lumbar spinal stenosis. *Acta Neurochir (Wien)* 2007; **149**: 811-815; discussion 815 [PMID: 17558457 DOI: 10.1007/s00701-007-1182-5]
- Bar-Oz B, Mogle P, Ben-Neriah Z, Sheffer R, Arad I. Duodenal web in the syndrome of osteopathia striata with cranial sclerosis. *Clin Genet* 1996; **49**: 329-330 [PMID: 8884088 DOI: 10.1111/j.1399-0004.1996.tb03802.x]
- Koudstaal MJ, Wolvius EB, Ongkosuwo EM, van der Wal

- KG. Surgically assisted rapid maxillary expansion in two cases of osteopathia striata with cranial sclerosis. *Cleft Palate Craniofac J* 2008; **45**: 337-342 [PMID: 18452358 DOI: 10.1597/07-016]
- 29 **Savarirayan R**, Nance J, Morris L, Haan E, Couper R. Osteopathia striata with cranial sclerosis: highly variable phenotypic expression within a family. *Clin Genet* 1997; **52**: 199-205 [PMID: 9383023 DOI: 10.1111/j.1399-0004.1997.tb02547.x]
  - 30 **Ward LM**, Rauch F, Travers R, Roy M, Montes J, Chabot G, Glorieux FH. Osteopathia striata with cranial sclerosis: clinical, radiological, and bone histological findings in an adolescent girl. *Am J Med Genet A* 2004; **129A**: 8-12 [PMID: 15266607 DOI: 10.1002/ajmg.a.30107]
  - 31 **Curarino G**, Friedman JM. Severe craniofacial sclerosis with multiple anomalies in a boy and his mother. *Pediatr Radiol* 1986; **16**: 441-447 [PMID: 3774390 DOI: 10.1007/BF02387953]
  - 32 **Keymolen K**, Bonduelle M, De Maeseneer M, Liebaers I. How to counsel in osteopathia striata with cranial sclerosis. *Genet Couns* 1997; **8**: 207-211 [PMID: 9327263]
  - 33 **Bueno AL**, Ramos FJ, Bueno O, Olivares JL, Bello ML, Bueno M. Severe malformations in males from families with osteopathia striata with cranial sclerosis. *Clin Genet* 1998; **54**: 400-405 [PMID: 9842992]
  - 34 **Lüerssen K**, Ptok M. Osteopathia striata with cranial sclerosis and hearing loss. *Eur Arch Otorhinolaryngol* 2006; **263**: 123-126 [PMID: 16010569 DOI: 10.1007/s00405-005-0972-8]
  - 35 **Fukuzawa R**, Holman SK, Chow CW, Savarirayan R, Reeve AE, Robertson SP. WTX mutations can occur both early and late in the pathogenesis of Wilms tumour. *J Med Genet* 2010; **47**: 791-794 [PMID: 20679664 DOI: 10.1136/jmg.2010.080663]
  - 36 **Murphy-Ryan M**, Kirmani S, Thompson DM, Binkovitz LA, Thomas KB, Babovic-Vuksanovic D. A novel sclerosing skeletal dysplasia with mixed sclerosing bone dysplasia, characteristic syndromic features, and clinical and radiographic evidence of male-male transmission. *Am J Med Genet A* 2012; **158A**: 2292-2296 [PMID: 22821701 DOI: 10.1002/ajmg.a.35497]
  - 37 **Chénier S**, Noor A, Dupuis L, Stavropoulos DJ, Mendoza-Londono R. Osteopathia striata with cranial sclerosis and developmental delay in a male with a mosaic deletion in chromosome region Xq11.2. *Am J Med Genet A* 2012; **158A**: 2946-2952 [PMID: 22987541 DOI: 10.1002/ajmg.a.35619]
  - 38 **Joseph DJ**, Ichikawa S, Econs MJ. Mosaicism in osteopathia striata with cranial sclerosis. *J Clin Endocrinol Metab* 2010; **95**: 1506-1507 [PMID: 20150574 DOI: 10.1210/jc.2009-2343]
  - 39 **Rivera MN**, Kim WJ, Wells J, Driscoll DR, Brannigan BW, Han M, Kim JC, Feinberg AP, Gerald WL, Vargas SO, Chin L, Iafrate AJ, Bell DW, Haber DA. An X chromosome gene, WTX, is commonly inactivated in Wilms tumor. *Science* 2007; **315**: 642-645 [PMID: 17204608 DOI: 10.1055/s-2008-1033900]
  - 40 **Beckwith JB**. Nephrogenic rests and the pathogenesis of Wilms tumor: developmental and clinical considerations. *Am J Med Genet* 1998; **79**: 268-273 [PMID: 9781906 DOI: 10.1002/(SICI)1096-8628]
  - 41 **Beckwith JB**, Kiviat NB, Bonadio JF. Nephrogenic rests, nephroblastomatosis, and the pathogenesis of Wilms' tumor. *Pediatr Pathol* 1990; **10**: 1-36 [PMID: 2156243 DOI: 10.3109/15513819009067094]
  - 42 **Huff V**. Wilms' tumours: about tumour suppressor genes, an oncogene and a chameleon gene. *Nat Rev Cancer* 2011; **11**: 111-121 [PMID: 21248786 DOI: 10.1038/nrc3002]
  - 43 **Royer-Pokora B**. Genetics of pediatric renal tumors. *Pediatr Nephrol* 2013; **28**: 13-23 [PMID: 22461142 DOI: 10.1007/s00467-012-2146-4]
  - 44 **Ruteshouser EC**, Robinson SM, Huff V. Wilms tumor genetics: mutations in WT1, WTX, and CTNNB1 account for only about one-third of tumors. *Genes Chromosomes Cancer* 2008; **47**: 461-470 [PMID: 18311776 DOI: 10.1002/gcc.20553]
  - 45 **Perotti D**, Gamba B, Sardella M, Spreafico F, Terenziani M, Collini P, Pession A, Nantron M, Fossati-Bellani F, Radice P. Functional inactivation of the WTX gene is not a frequent event in Wilms' tumors. *Oncogene* 2008; **27**: 4625-4632 [PMID: 18391980 DOI: 10.1038/onc.2008.93]
  - 46 **Wegert J**, Wittmann S, Leuschner I, Geissinger E, Graf N, Gessler M. WTX inactivation is a frequent, but late event in Wilms tumors without apparent clinical impact. *Genes Chromosomes Cancer* 2009; **48**: 1102-1111 [PMID: 19760609 DOI: 10.1002/gcc.20712]
  - 47 **Fukuzawa R**, Anaka MR, Weeks RJ, Morison IM, Reeve AE. Canonical WNT signalling determines lineage specificity in Wilms tumour. *Oncogene* 2009; **28**: 1063-1075 [PMID: 19137020 DOI: 10.1038/onc.2008.455]
  - 48 **Scott RH**, Stiller CA, Walker L, Rahman N. Syndromes and constitutional chromosomal abnormalities associated with Wilms tumour. *J Med Genet* 2006; **43**: 705-715 [PMID: 16690728]
  - 49 **Moisan A**, Rivera MN, Lotinun S, Akhavanfard S, Coffman EJ, Cook EB, Stoykova S, Mukherjee S, Schoonmaker JA, Burger A, Kim WJ, Kronenberg HM, Baron R, Haber DA, Bardeesy N. The WTX tumor suppressor regulates mesenchymal progenitor cell fate specification. *Dev Cell* 2011; **20**: 583-596 [PMID: 21571217 DOI: 10.1016/j.devcel.2011.03.013]

**P- Reviewers:** Fukuzawa R, Sangkhathat S    **S- Editor:** Gou SX  
**L- Editor:** Roemmele A    **E- Editor:** Liu SQ



## Earlier onset and multiple primaries in familial as opposed to sporadic esophageal cancer

Xiao-Duo Wen, Deng-Gui Wen, Yi Yang, Bao-En Shan, Shi-Jie Wang

Xiao-Duo Wen, Yi Yang, Department of Diagnosis and Image, 4<sup>th</sup> Hospital of Hebei Medical University, Shijiazhuang 050011, Hebei Province, China

Deng-Gui Wen, Bao-En Shan, Shi-Jie Wang, Cancer Center, 4<sup>th</sup> Hospital of Hebei Medical University, Shijiazhuang 050011, Hebei Province, China

Author contributions: Wen XD and Wen DG performed data collection, analysis, and drafting the manuscript; Yang Y, Shan BE and Wang SJ established the cohort and designed the study.

Correspondence to: Dr. Shi-Jie Wang, MD, Cancer Center, 4<sup>th</sup> Hospital of Hebei Medical University, Jiankanglu 12, Shijiazhuang 050011, Hebei Province, China. [wshjw136@aliyun.com](mailto:wshjw136@aliyun.com)

Telephone: +86-311-86095337 Fax: +86-311-86077634

Received: July 30, 2013 Revised: October 17, 2013

Accepted: November 15, 2013

Published online: May 27, 2014

### Abstract

**AIM:** To study the differences in onset age and multiple primary cancers between familial and sporadic esophageal squamous cell carcinoma (ESCC).

**METHODS:** The differences in onset age and multiple primary cancers were analyzed between ESCC patients with ( $n = 766$ ) and without ( $n = 1776$ ) a family history of the cancer. The cases analyzed constituted all consecutive patients who had undergone cure-intent surgery at the Department of Thoracic Surgery of the 4<sup>th</sup> Hospital of Hebei Medical University from January 1 1975 to December 31 1989. Because we also originally aimed to examine the difference in survival time, only older subjects with a long follow-up period were selected.

**RESULTS:** Overall, patients with ESCC and a positive family history of the cancer had a significantly younger age at onset and more multiple primary cancers than those without a positive family history ( $51.83 \pm 8.39$  vs  $53.49 \pm 8.23$  years old,  $P = 0.000$ ;  $5.50\%$  vs  $1.70\%$ ,  $P = 0.000$ ). Both of these differences were evident in subgroup analyses, however, no correlations were ob-

served. While age at onset differed significantly by family history in males, smokers, and drinkers, the difference in multiple primary cancers by family history was significant in nonsmoking, nondrinking, and younger onset patients. In multivariate analysis, age over 50 years, tobacco smoking, and multiple primary cancers were found to be significant predictors of familial cancer: the corresponding OR (95%CI) and  $P$ -value were 0.974 (0.963-0.985) and 0.000; 1.271 (1.053-1.535) and 0.012; and 4.265 (2.535-7.176) and 0.000, respectively.

**CONCLUSION:** Patients with ESCC and a positive family history of the cancer had a significantly younger onset age and more multiple primary cancers than those without a positive family history. Sub-group analyses indicated that younger onset age may be due to the interaction of genetic predisposition and environmental hazards, and multiple primary cancers may only be due to genetic predisposition.

© 2014 Baishideng Publishing Group Inc. All rights reserved.

**Key words:** Esophageal squamous cell carcinoma; Familial cancer; Sporadic cancer; Age at onset; Synchronous primary carcinoma

**Core tip:** Patients with esophageal squamous cell carcinoma (ESCC) and a positive family history had a significantly younger onset age and more multiple primary cancers than those without a positive family history ( $51.83 \pm 8.39$  vs  $53.49 \pm 8.23$  years old,  $P = 0.000$ ;  $5.50\%$  vs  $1.70\%$ ,  $P = 0.000$ ). Both of these differences were evident in sub-group analyses, however, no correlations were observed. While age at onset differed significantly by family history in males, smokers, and drinkers, the difference in multiple primary cancers by family history was significant in nonsmoking, nondrinking, and younger onset patients. These results suggest a genetic component in ESCC. Furthermore, a younger onset age may be due to the interaction of genetic predisposition and environmental hazards, and

multiple primary cancers may only be due to genetic predisposition.

Wen XD, Wen DG, Yang Y, Shan BE, Wang SJ. Earlier onset and multiple primaries in familial as opposed to sporadic esophageal cancer. *World J Med Genet* 2014; 4(2): 39-45 Available from: URL: <http://www.wjgnet.com/2220-3184/full/v4/i2/39.htm> DOI: <http://dx.doi.org/10.5496/wjmg.v4.i2.39>

## INTRODUCTION

The incidence rates of most cancers increase with age, suggesting that cancers develop due to the accumulation of somatic mutations. If a germline mutation exists, however, fewer later life somatic mutations will be needed, and the cancer will develop at a younger age<sup>[1]</sup>. The “two-hit” hypothesis has been widely accepted to explain the occurrence of both familial and sporadic cancers by the inactivation of tumor-suppressor genes mechanism. In familial cancer, because the function of one allele of a crucial tumor suppressor gene has already been lost due to inheritance of a germline mutation (the first-hit, which theoretically is present in every cell), only inactivation of the remaining allele by a second-hit is necessary. For a sporadic cancer, however, both alleles have to be inactivated by two fateful somatic mutations: the chance is so small that under most circumstances only one cell is likely to be implicated<sup>[2,3]</sup>. Therefore, differences in age at onset and multiple primary cancers between cancer cases with or without a family history may suggest a genetic component in the etiology of the cancer<sup>[4-7]</sup>.

In the Taihang mountain region in northern central China, the incidence rates of esophageal squamous cell carcinoma (ESCC) are the highest in the world, however, the risk factors for ESCC are not yet clear<sup>[8-16]</sup>. On analyzing a large surgical cohort at a field cancer center, we compared age of onset and multiple primary cancers in ESCC patients with and without a family history of the cancer. Our aim was to identify a genetic component using reliable clinicopathological data in a hospital-based surgical cohort.

## MATERIALS AND METHODS

### Ethics

The study was approved by the Institutional Ethics Review Board of the 4<sup>th</sup> Hospital of Hebei Medical University.

### Subject selection

The high risk region and the Upper Gastrointestinal Cancer Center were previously described<sup>[4,5,7,9]</sup>. The data source was a hospital-based surgical registry of ESCC patients who had been operated on as early as October 1965. The subjects analyzed comprised all ESCC cases who had undergone cure-intent surgery in the Department of Thoracic Surgery of the 4<sup>th</sup> Affiliated Hospital

of Hebei Medical University (also the Hebei Province Cancer Center) from January 1 1975 to December 31 1989. Because the 4<sup>th</sup> Hospital of Hebei Medical University used to be the only local cancer center capable of performing thoracic surgery and all surgical resections of esophageal cancer during that period were performed at the hospital, the cases analyzed comprised almost all incident resectable ESCC cases in the population.

### Definition of family history and onset age

As ESCC and gastric cardia adenocarcinoma (GCA) have a common etiology in China<sup>[17-19]</sup> and have similar symptoms such as swallowing disturbance and substantial pain, a recalled family history was unable to distinguish between these two types of cancer when crucial documentation was missing<sup>[20,21]</sup>. Therefore, a positive family history was defined as at least one first- or two second-degree relatives diagnosed with ESCC/GCA. The family history was usually obtained by the surgeon on the first day of hospitalization. The onset age was the age at which the symptoms first appeared. Information on a positive family history of cancer included the site of cancer, blood relationship, when and where diagnosed, and vital status of the relative with cancer. In patients with a negative family history recalled at the time of hospitalization, if a first- or second-degree relative was subsequently diagnosed with ESCC or GCA, this question was routinely asked in the follow-up interview, and the information on family history was updated. Blood relationships were categorized as first, second, and third degree relatives.

### Definition of synchronous multiple primary cancers

As ESCC and GCA in China share a susceptible genetic locus<sup>[17-19]</sup>, we used not only multiple primary ESCCs, but also the coincidence of ESCC and GCA as evidence of genetic predisposition (a further explanation can be found in the discussion section). The presence of multiple primary cancers was investigated by reviewing the pathology report and slides of serial histological examinations. Evidence of a second primary cancer did not include high grade dysplasia or intraepithelial tumors as the diagnostic criteria were not consistent for these lesions during 1975-1989. With regard to the definition of second primaries, we used the following criteria: (1) both lesions exhibited definite malignant morphologic features and were not connected through the lymphatic system; and (2) both tumors were surrounded by intraepithelial tumors or dysplastic tissue. All 56 second primaries were discovered by endoscopy or barium X-ray examination during the out-clinic stage and were successfully resected.

### Statistical analyses

Differences in demographic and clinicopathological distributions by family history were examined using the  $\chi^2$  test. A comparison of onset age between familial and sporadic cases was made using the Student's *t* test. A two-sided *P* value of less than 0.05 was considered statistically significant. To identify significant independent predic-



**Table 1** Difference in age at onset in patients with esophageal squamous cell carcinomas with and without a family history of the cancer

	Positive family history		Negative family history		Differences (1)-(2)	P t-test (1)-(2)
	n (%)	Mean age (yr) $\pm$ SD (1)	n (%)	Mean age (yr) $\pm$ SD (2)		
Total	766	51.83 $\pm$ 8.39	1776	53.49 $\pm$ 8.23	-1.66	0.000 <sup>b</sup>
Sex						
Female	229 (33.80)	52.46 $\pm$ 7.91	600 (29.90)	53.48 $\pm$ 7.81	-1.02	0.100
Male	537 (66.20)	51.56 $\pm$ 8.58	1176 (70.10)	53.49 $\pm$ 8.44	-1.93	0.000 <sup>b</sup>
Tobacco						
Nonsmoker	282 (36.80)	52.66 $\pm$ 8.03	739 (41.60)	53.63 $\pm$ 8.27	-0.98	0.090
Smoker	484 (63.20)	51.35 $\pm$ 8.56	1037 (58.40)	53.38 $\pm$ 8.13	-2.03	0.000 <sup>b</sup>
Alcohol						
Nondrinker	480 (62.70)	52.10 $\pm$ 8.37	1200 (67.60)	53.38 $\pm$ 8.21	-1.29	0.004 <sup>a</sup>
Drinker	286 (37.30)	51.38 $\pm$ 8.42	578 (32.40)	53.70 $\pm$ 8.28	-2.32	0.000 <sup>b</sup>
Surgery year						
1975-1979	224 (29.20)	50.90 $\pm$ 8.21	376 (21.20)	52.36 $\pm$ 8.00	-1.46	0.034 <sup>a</sup>
1980-1984	216 (28.20)	51.69 $\pm$ 8.58	545 (30.70)	52.95 $\pm$ 8.18	-1.26	0.050 <sup>a</sup>
1985-1989	326 (42.60)	52.56 $\pm$ 8.34	855 (48.10)	54.32 $\pm$ 8.28	-1.76	0.001 <sup>b</sup>
TNM						
T <sub>1-3</sub> N <sub>0</sub> M <sub>0</sub>	34 (4.40)	50.94 $\pm$ 9.78	66 (3.70)	55.52 $\pm$ 7.24	-4.57	0.010 <sup>a</sup>
T <sub>2-3</sub> N <sub>0</sub> M <sub>0</sub>	422 (55.10)	51.64 $\pm$ 8.30	983 (55.30)	53.30 $\pm$ 8.37	-1.67	0.001 <sup>b</sup>
T <sub>2-3,4</sub> N <sub>1</sub> M <sub>0</sub>	310 (40.50)	52.19 $\pm$ 8.37	727 (40.90)	53.55 $\pm$ 8.10	-1.36	0.010 <sup>a</sup>
Resection						
Exploratory	16 (2.10)	49.88 $\pm$ 8.16	28 (1.60)	52.57 $\pm$ 9.18	-2.70	0.340
R1 or R2	70 (9.10)	49.33 $\pm$ 10.00	152 (8.60)	54.73 $\pm$ 8.49	-5.40	0.000 <sup>b</sup>
R0	680 (88.80)	52.13 $\pm$ 8.17	1596 (89.90)	53.38 $\pm$ 8.18	-1.25	0.001 <sup>b</sup>

Positive family history *vs* negative family history, <sup>a</sup>*P* < 0.05, <sup>b</sup>*P* < 0.01.

tors of familial cancer, multivariate logistic analyses were performed in a backward manner to study the association between sex, age, smoking, drinking, presence of multiple primary cancers and the occurrence of multiple ESCC/GCA cases in the family. All calculations were performed using SPSS software version 13.0<sup>[22]</sup>.

## RESULTS

### General demographic characteristics

The male:female ratio was 2.07:1 (1713/829) and the average age at onset was 52.89 years in males and 53.20 years in females. All 2542 ESCCs diagnosed in the 2524 patients were considered resectable before surgery, however, the actual resection rate was 89.5% (2276/2542). In general, no significant differences in demographic and clinicopathological characteristics were observed between the positive and negative family history groups (Table 1).

Of the 2542 ESCCs analyzed, 69.87% (1776/2542) developed in patients with no family history of the cancer and were regarded as sporadic; 30.13% (766/2542) were diagnosed in patients with a positive family history and were regarded as familial.

On reviewing the pathology reports and slides of serial histological examinations of the 2524 primary ESCC cases, 16 cases (0.63%) developed two primary ESCCs, 36 cases developed one ESCC and one GCA, two cases each developed two primary ESCCs as well as another primary GCA, and the remaining 2470 cases each developed a single primary ESCC. The 38 GCAs were taken as evidence of second primary cancer of ESCC, but were not

included in the analyzed dataset of 2542 primary ESCCs.

### Differences in age at onset in patients with ESCCs with and without a positive family history of the cancer

As shown in Table 1, the average age at onset of ESCCs (*n* = 766) with a positive family history was 51.83 years, significantly younger than that of 53.49 years for ESCCs with no family history (*n* = 1776, *P* = 0.000). Overall, the difference was 1.66 years, and the difference in sub-group analyses was significant for males, smokers, drinkers, non-drinkers, surgery during 1975-1979, 1980-1984, 1985-1989, tumor stage of T<sub>1-3</sub>N<sub>0</sub>M<sub>0</sub>, T<sub>2-3</sub>N<sub>0</sub>M<sub>0</sub>, T<sub>2-3,4</sub>N<sub>1</sub>M<sub>0</sub><sup>[23]</sup>, complete resection (R<sub>0</sub>) and partial resection (R<sub>1</sub>/R<sub>2</sub>) sub-groups; but not significant for females, nonsmokers, and the exploratory surgery subgroup.

### Patients with ESCCs and a positive family history are more likely to have multiple primary cancers

As shown in Table 2, 72 ESCCs had multiple primary cancers as they were associated with one or more primary ESCC/GCA in a single patient. The overall prevalence of multiple primary cancers was 2.8% (72/2542). The prevalence was 5.5% in the positive family history group compared with 1.7% in the negative family history group. The difference was significant ( $\chi^2 = 27.80$ , *P* = 0.000).

When the sub-group analyses were performed, the difference in multiple primary cancer by family history persisted in all subgroups and was significant for males, females, younger onset age (under 55 years old), non-smoking, non-drinking, T<sub>2-3</sub>N<sub>0</sub>M<sub>0</sub>, and T<sub>2-3,4</sub>N<sub>1</sub>M<sub>0</sub> groups; but not significant for the older onset age (over 55 years old), smoking, drinking and T<sub>1-3</sub>N<sub>0</sub>M<sub>0</sub> groups.

**Table 2** Difference in the prevalence of multiple primary esophageal squamous cell carcinoma/gastric cardia adenocarcinoma in patients with esophageal squamous cell carcinoma with and without a family history of the cancer

Classification	Family history	n (% of multiple primary cancer)	Total	$\chi^2$	<sup>a</sup> P
Overall		72 (2.8)	2542		
	Negative family history	30 (1.7)	1776		
	Positive family history	42 (5.5)	766	27.80	0.000 <sup>b</sup>
Sex					
Male	Negative family history	20 (1.7)	1176		
	Positive family history	27 (5.0)	537	15.30	0.000 <sup>b</sup>
Female	Negative family history	10 (1.7)	600		
	Positive family history	15 (6.6)	229	13.52	0.000 <sup>b</sup>
Age					
≥ 55 yr	Negative family history	23 (2.6)	877		
	Positive family history	15 (5.0)	303	3.29	0.070
< 55 yr	Negative family history	7 (0.8)	899		
	Positive family history	27 (5.8)	463	32.06	0.000 <sup>b</sup>
Tobacco					
Smoker	Negative family history	20 (1.9)	1037		
	Positive family history	1.8 (3.7)	484	3.63	0.060
Nonsmoker	Negative family history	10 (1.4)	739		
	Positive family history	24 (8.5)	282	32.48	0.000 <sup>b</sup>
Alcohol					
Drinker	Negative family history	16 (2.8)	576		
	Positive family history	13 (4.5)	286	1.84	0.230
Nondrinker	Negative family history	14 (1.2)	1200		
	Positive family history	29 (6.0)	480	32.67	0.000 <sup>b</sup>
TNM					
T <sub>1-3</sub> N <sub>0</sub> M <sub>0</sub>	Negative family history	2 (3.0)	66		
	Positive family history	2 (5.9)	34	0.48	0.480
T <sub>2-3</sub> N <sub>0</sub> M <sub>0</sub>	Negative family history	8 (0.8)	983		
	Positive family history	20 (4.7)	422	23.29	0.000 <sup>b</sup>
T <sub>2-3-4</sub> N <sub>1</sub> M <sub>0</sub>	Negative family history	20 (2.8)	727		
	Positive family history	20 (6.5)	310	8.03	0.010 <sup>b</sup>

Positive family *vs* negative family, <sup>a</sup>P < 0.05, <sup>b</sup>P < 0.01.

### Multivariate logistic regression analysis of factors associated with the occurrence of multiple ESCC/GCA cases in a family

When significant factors in univariate analysis as well as other related factors were entered into a binary logistic model, backward stepwise analysis removed insignificant variables step by step according to their contribution to the model. The order in which the insignificant variables were eliminated from the model was as follows: sex, primary tumor site, pre-surgical radiotherapy, and alcohol drinking. In the final model, age over 50 years, tobacco smoking, and multiple primary cancers were found to be significant independent predictors of familial cancer: the corresponding OR (95%CI) and *P*-value were 0.974 (0.963-0.985) and 0.000; 1.271 (1.053-1.535) and 0.012; and 4.265 (2.535-7.176) and 0.000, respectively (Table 3).

## DISCUSSION

In the present analysis, we found that cases with a positive family history developed ESCC at a significantly younger age than those without such a family history. This was not due to earlier diagnosis as the stage distribution did not vary significantly between the two groups. We also found that a positive family history was signifi-

cantly associated with a higher rate of multiple primary cancers. These observations suggest that this is a genetic component in ESCC. In Table 1, age at onset differed significantly by family history for males and smokers, but not for females and nonsmokers. Conversely, the age at diagnosis in males, smoking and alcoholic drinking cases was younger than that in females, non-smoking and non-drinking cases, respectively, when the family history was positive (51.56 *vs* 52.46, 51.35 *vs* 52.66, and 51.38 *vs* 52.10, respectively), but not observed when the family history was negative (53.49 *vs* 53.48, 53.38 *vs* 53.63, and 53.70 *vs* 53.38, respectively). These findings suggest a younger onset age of ESCC is due to the interaction between genetic and environmental risk factors (*i.e.*, only when both hazards are present).

Table 2 illustrates the different prevalences of multiple primary cancers by family history. What is interesting about this difference is that it was significant in the < 55 years, nonsmoking and nondrinking groups, but was non-significant in the > 55 years, smoking and drinking groups. The underlying reasons for this may be because age, smoking and drinking are established environmental risk factors for ESCC<sup>[10]</sup>. When these environmental risk factors do not exist, genetic predisposition must play a dominant role, therefore, significant differences in mul-

**Table 3** Significant and independent predictors of familial esophageal squamous cell cancer by multivariate logistic analysis

Variables entered at each step	Regression coefficient	Standard error	Wald test value	P-value	OR	95%CI	
						Lower	Upper
Step 1							
Sex	-0.057	0.146	0.153	0.696	0.944	0.709	1.258
Age	-0.026	0.006	21.06	0.000 <sup>b</sup>	0.975	0.964	0.985
Primary tumor site	0.235	0.349	0.451	0.502	0.791	0.399	1.568
Tobacco	0.153	0.140	1.195	0.274	1.165	0.886	1.532
Alcohol	0.113	0.112	1.019	0.313	1.120	0.899	1.396
Multiple cancer	1.427	0.267	28.49	0.000 <sup>b</sup>	4.166	2.467	7.036
Presurgical radiotherapy	0.161	0.179	0.813	0.367	1.175	0.828	1.669
Constant	0.592	0.827	0.512	0.474	1.808		
Step 2							
Age	-0.026	0.006	21.068	0.000 <sup>b</sup>	0.975	0.964	0.985
Primary tumor site	-0.241	0.349	0.479	0.489	0.786	0.397	1.556
Tobacco	0.187	0.108	3.001	0.083	1.206	0.967	1.491
Alcohol	0.123	0.110	1.250	0.264	1.131	0.912	1.402
Multiple cancer	1.427	0.267	28.497	0.000 <sup>b</sup>	4.166	2.467	7.035
Presurgical radiotherapy	0.160	0.179	0.802	0.371	1.174	0.827	1.667
Constant	0.462	0.758	0.372	0.542	1.588		
Step 3							
Age	-0.026	0.006	21.311	0.000 <sup>b</sup>	0.974	0.964	0.985
Tobacco	0.185	0.108	2.926	0.087	1.203	0.973	1.487
Alcohol	0.122	0.110	1.236	0.266	1.130	0.911	1.401
Multiple cancer	1.446	0.266	29.56	0.000 <sup>b</sup>	4.246	2.521	7.152
Presurgical radiotherapy	0.157	0.179	0.773	0.379	1.170	0.824	1.662
Constant	-0.004	0.346	0.000	0.990	0.996		
Step 4							
Age	-0.026	0.006	22.395	0.000 <sup>b</sup>	0.974	0.963	0.985
Tobacco	1.187	0.108	3.000	0.083	1.206	0.976	1.491
Alcohol	0.117	0.110	1.146	0.284	1.125	0.907	1.394
Multiple cancer	1.438	0.266	29.282	0.000 <sup>b</sup>	4.213	2.502	7.092
Constant	0.037	0.343	0.011	0.915	1.037		
Step 5							
Age	-0.026	0.006	22.364	0.000 <sup>b</sup>	0.974	0.963	0.985
Tobacco	0.240	0.096	6.245	0.012 <sup>a</sup>	1.271	1.053	1.535
Multiple cancer	1.451	0.265	29.862	0.000 <sup>b</sup>	4.265	2.535	7.176
Constant	0.108	0.337	0.104	0.747	1.114		

<sup>a</sup>P < 0.05, <sup>b</sup>P < 0.01 vs standard error.

multiple primary cancers by family history was only observed in the < 55 years, nonsmoking and nondrinking groups, but not in the > 55 years, smoking and drinking groups. This suggests multiple primary cancers, unlike age at diagnosis, reflects only genetic predisposition.

We used the coincidence of GCA as evidence of genetic predisposition for ESCC for the following reason: In China, GCA contrasts with esophagogastric junctional adenocarcinoma in the Western world, in that gastro-esophageal reflux disease or Barrett's esophagus is not a precursor<sup>[8]</sup>. Instead, GCA and ESCC have identical epidemiological distributions<sup>[17-21]</sup>. Molecularly, gene polymorphisms associated with elevated risks for both cancers were found in Hebei Province<sup>[13,14]</sup>, and identical DNA alterations in the two cancers were reported in Henan Province<sup>[15,16]</sup>. These two provinces are located in the central area or the Taihang Mountain high-risk region. Recently, two genome-wide association studies of ESCC/GCA cases in China reported two susceptible loci, one at PLCE1 and the other (C20orf54) at 20p13, to be significantly associated with the risk of both cancers<sup>[17,18]</sup>.

During surgery, the two cancers are often found to coincide in one patient: among the 2524 ESCC cases

undergoing curative-intent surgery in the 4<sup>th</sup> Hospital of Hebei Medical University, the coincidence rate of ESCC and GCA was 1.5% (38/2524), twice as high as that of two primary ESCCs [0.7% (16/2524)].

In cancer genetics, a patient developing both an ESCC and a GCA may be less familial than a patient developing two primary ESCCs, but is certainly more familial than a single primary ESCC case.

Although epidemiological studies showed that a positive family history of UGIC increases the risk of both ESCC and GCA<sup>[20,21]</sup>, the increased risk may be attributed to either common childhood household exposure or to genetic predisposition. As our results showed that younger onset age and multiple primary cancers are associated with familial as opposed to sporadic ESCC, we consider that the high risk associated with a positive family history is due to an inherited predisposition (for instance, the first-hit on a critical tumor suppressor gene) rather than childhood exposure to common household risk factors. This is because a background of inherited "first-hit", which theoretically exists in every cell, would make it more likely for independent tumors to develop in a tissue; in a sporadic case, however, the coincidence of

two mutations in one cell to inactivate both alleles of one tumor suppressor gene is so rare that under most circumstances only a single cell is likely to be involved, and thus the chance of two primary tumors is much less than it would be for a familial cancer case<sup>[23-25]</sup>.

Although the difference in onset age between familial and sporadic cases was just 1.66 years, it is not merely a statistical advantage inherited in large sample sizes, because it was widely observed in all sub-classifications, significant in stage-specific subgroups according to the International Staging System<sup>[26]</sup>, and more notably, in ESCCs with onset age < 55 years, family history was significantly associated with multiple primary cancers ( $P = 0.000$ ), however, this association was not significant in the  $\geq 55$  years group ( $P = 0.07$ ). This variation in strength by onset age was indicative of the effect of genetic predisposition.

Because all the double primary ESCC/GCA were clinically evident and surgically resected, these tumors only represent synchronous carcinomas detected clinically. If metachronous or intraepithelial tumors were included, much higher rates would have been observed.

In conclusion, we found significant differences in age at onset and multiple primary cancers between familial and sporadic ESCCs. Younger onset age results from genetic and environmental interaction, but multiple primary cancers may be more related to genetic predisposition.

## ACKNOWLEDGMENTS

They authors are grateful to colleagues in the Departments of Thoracic Surgery, Anesthesia, and Surgical Operation at the Fourth Hospital of Hebei Medical University for their great contribution to surgical treatment of esophageal cancer.

## COMMENTS

### Background

Although previous epidemiological studies have reported that a positive family history of esophageal cancer increases the risk of the cancer by 2- to 3-fold, the increased risk may be due to either common childhood household exposure or to genetic predisposition. If a younger onset age and more synchronous primary cancers were found to be associated with familial as opposed to sporadic esophageal squamous cell carcinoma (ESCC), the authors would be able to attribute the elevated risk associated with a positive family history to an inherited predisposition rather than childhood exposure to common household risk factors.

### Research frontiers

According to the "two-hit story of tumor-suppressor gene in carcinogenesis", a background of inherited "first-hit" on a critical tumor suppressor gene, which theoretically exists in every cell of a tissue, would make it more likely for independent tumors to develop; in a sporadic case, however, the coincidence of two mutations to inactivate both alleles of a tumor suppressor gene is so rare that under most circumstances, only a single cell is likely to be involved, and thus the chance of two primary tumors is much less than it would be for a familial cancer case. Although many molecular findings have been suggested for the mechanism of genetic predisposition in ESCC, clinicopathological evidence is lacking.

### Innovations and breakthroughs

For the first time, this study showed that ESCCs with a positive family history of the cancer have a significantly younger age at onset and more multiple primary cancers than those without a positive family history. This study clearly indicates that a genetic component exists in ESCC. It also demonstrates that clinico-

pathological characteristics of a cancer may be analyzed to identify evidence of genetic predisposition.

### Applications

The molecular mechanism of genetic predisposition has not yet been determined for ESCC. By focusing attention on the onset age and number of primary cancers, as well as family history, it may be possible to pinpoint familial esophageal cancer cases and obtain specimens to study the exact molecular mechanism involved. In clinical practice, attention should be paid to familial cases to identify synchronous or metachronous second primary cancers.

### Terminology

Like most cancers, ESCC exhibits considerable heterogeneity in etiology. Some are caused by the accumulation of somatic mutations brought about by environmental hazards, and others develop due to genetic predisposition. The former is known as sporadic cancer, while the latter is known as familial cancer. A family history of the cancer may be used to distinguish between the two cancer forms, but is not always reliable. Some familial cancer cases may lack a definite family history, under such circumstances, the onset age and second primary cancer may help to pinpoint a familial cancer.

### Peer review

The authors have analyzed family history of cancer, as well as age at diagnosis and number of tumors, in a cohort of 2542 patients from Hebei Province with history of ESCC and gastric cardia adenocarcinoma. This topic is current, methods used and the results obtained in the paper are of good scientific value.

## REFERENCES

- 1 Armitage P, Doll R. The age distribution of cancer and a multi-stage theory of carcinogenesis. 1954. *Int J Epidemiol* 2004; **33**: 1174-1179 [PMID: 15319408 DOI: 10.1093/ije/dyh216]
- 2 Berger AH, Knudson AG, Pandolfi PP. A continuum model for tumour suppression. *Nature* 2011; **476**: 163-169 [PMID: 21833082 DOI: 10.1038/nature10275]
- 3 Hino O. Hereditary renal carcinogenesis fitting Knudson's two-hit model: genotype, environment, and phenotype. *Genes Chromosomes Cancer* 2003; **38**: 357-367 [PMID: 14566856 DOI: 10.1002/gcc.10265]
- 4 Wen D, Wang S, Zhang L, Wei L, Zhou W, Peng Q. Early onset, multiple primary malignancies, and poor prognosis are indicative of an inherited predisposition to esophageal squamous cell carcinoma for the familial as opposed to the sporadic cases--an update on over 14-year survival. *Eur J Med Genet* 2009; **52**: 381-385 [PMID: 19632364 DOI: 10.1016/j.ejmg.2009.07.003]
- 5 Wen D, Shan B, Wang S, Zhang L, Wei L, Zhou W, Peng Q. A positive family history of esophageal/gastric cardia cancer with gastric cardia adenocarcinoma is associated with a younger age at onset and more likely with another synchronous esophageal/gastric cardia cancer in a Chinese high-risk area. *Eur J Med Genet* 2010; **53**: 250-255 [PMID: 20603233 DOI: 10.1016/j.ejmg.2010.06.011]
- 6 Chuang SC, Hashibe M, Scelo G, Brewster DH, Pukkala E, Friis S, Tracey E, Weiderpass E, Hemminki K, Tamara S, Chia KS, Pompe-Kirn V, Kliever EV, Tonita JM, Martos C, Jonasson JG, Dresler CM, Boffetta P, Brennan P. Risk of second primary cancer among esophageal cancer patients: a pooled analysis of 13 cancer registries. *Cancer Epidemiol Biomarkers Prev* 2008; **17**: 1543-1549 [PMID: 18559572 DOI: 10.1158/1055-9965.EPI-07-2876]
- 7 Wen D, Wang S, Zhang L, Zhang J, Wei L, Zhao X. Differences of onset age and survival rates in esophageal squamous cell carcinoma cases with and without family history of upper gastrointestinal cancer from a high-incidence area in North China. *Fam Cancer* 2006; **5**: 343-352 [PMID: 16724246 DOI: 10.1007/s10689-006-0004-x]
- 8 Wang JB, Fan JH, Liang H, Li J, Xiao HJ, Wei WQ, Dawsey SM, Qiao YL, Boffetta P. Attributable causes of esophageal cancer incidence and mortality in China. *PLoS One* 2012; **7**: e42281 [PMID: 22876312 DOI: 10.1371/journal.pone.0042281]



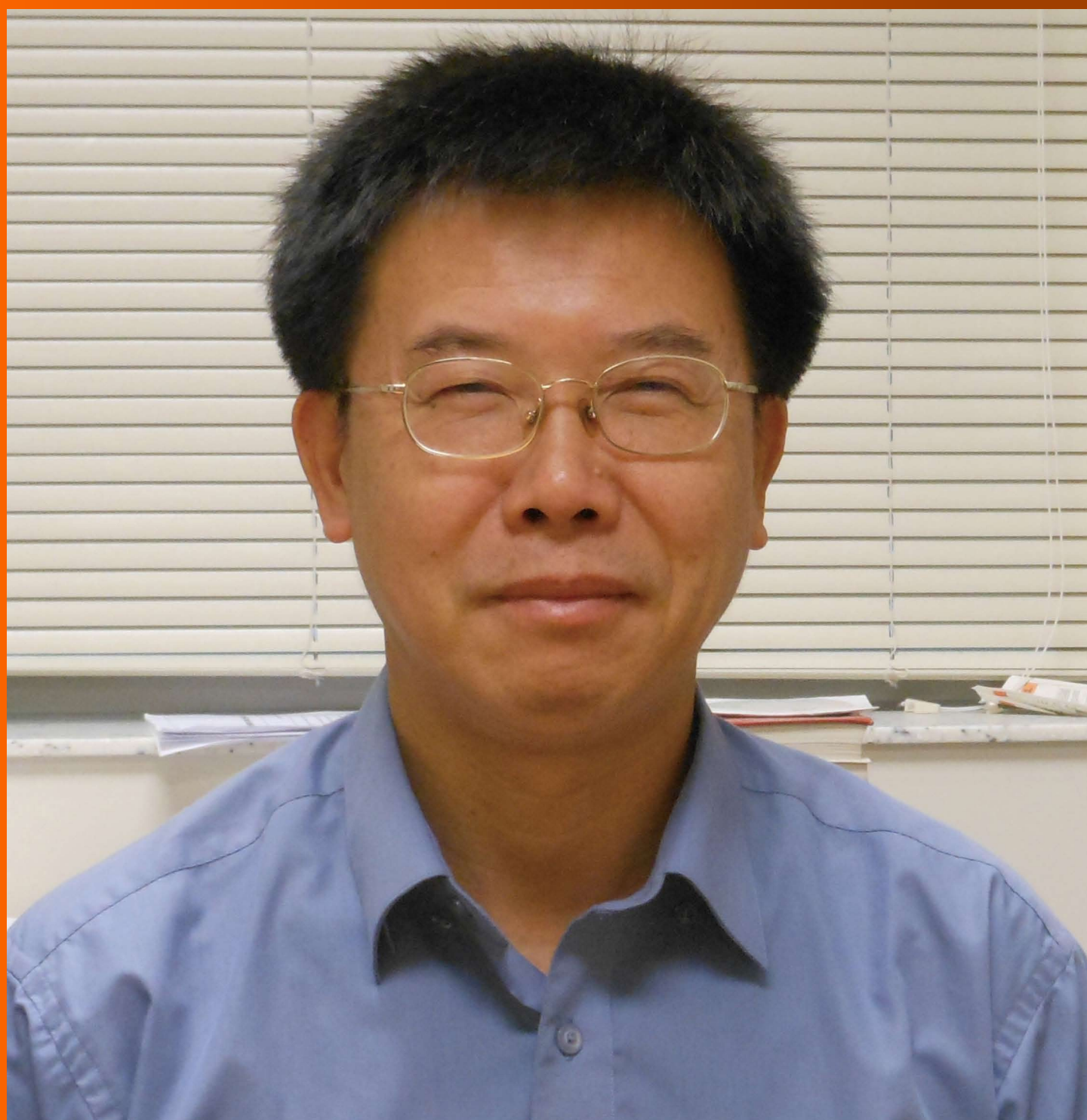
- 9 **Wen D**, Zhang N, Shan B, Wang S. Helicobacter pylori infection may be implicated in the topography and geographic variation of upper gastrointestinal cancers in the Taihang Mountain high-risk region in northern China. *Helicobacter* 2010; **15**: 416-421 [PMID: 21083747 DOI: 10.1111/j.1523-5378.2010.00790.x]
- 10 **Tran GD**, Sun XD, Abnet CC, Fan JH, Dawsey SM, Dong ZW, Mark SD, Qiao YL, Taylor PR. Prospective study of risk factors for esophageal and gastric cancers in the Linxian general population trial cohort in China. *Int J Cancer* 2005; **113**: 456-463 [PMID: 15455378 DOI: 10.1002/ijc.20616]
- 11 **Zhang HZ**, Jin GF, Shen HB. Epidemiologic differences in esophageal cancer between Asian and Western populations. *Chin J Cancer* 2012; **31**: 281-286 [PMID: 22507220 DOI: 10.5732/cjc.011.10390]
- 12 **Huang H**, Su M, Li X, Li H, Tian D, Gao Y, Guo Y. Y-chromosome evidence for common ancestry of three Chinese populations with a high risk of esophageal cancer. *PLoS One* 2010; **5**: e11118 [PMID: 20559544 DOI: 10.1371/journal.pone.0011118]
- 13 **Zhang J**, Li Y, Wang R, Wen D, Sarbia M, Kuang G, Wu M, Wei L, He M, Zhang L, Wang S. Association of cyclin D1 (G870A) polymorphism with susceptibility to esophageal and gastric cardiac carcinoma in a northern Chinese population. *Int J Cancer* 2003; **105**: 281-284 [PMID: 12673692 DOI: 10.1002/ijc.11067]
- 14 **Zhang J**, Cui Y, Kuang G, Li Y, Wang N, Wang R, Guo W, Wen D, Wei L, Yu F, Wang S. Association of the thymidylate synthase polymorphisms with esophageal squamous cell carcinoma and gastric cardiac adenocarcinoma. *Carcinogenesis* 2004; **25**: 2479-2485 [PMID: 15284183 DOI: 10.1093/carcin/bgh247]
- 15 **Wang LD**, Qin YR, Fan ZM, Kwong D, Guan XY, Tsao GS, Sham J, Li JL, Feng XS. Comparative genomic hybridization: comparison between esophageal squamous cell carcinoma and gastric cardia adenocarcinoma from a high-incidence area for both cancers in Henan, northern China. *Dis Esophagus* 2006; **19**: 459-467 [PMID: 17069589 DOI: 10.1111/j.1442-2050.2006.00620.x]
- 16 **Wang LD**, Shi ST, Zhou Q, Goldstein S, Hong JY, Shao P, Qiu SL, Yang CS. Changes in p53 and cyclin D1 protein levels and cell proliferation in different stages of human esophageal and gastric-cardia carcinogenesis. *Int J Cancer* 1994; **59**: 514-519 [PMID: 7960222 DOI: 10.1002/ijc.2910590414]
- 17 **Abnet CC**, Freedman ND, Hu N, Wang Z, Yu K, Shu XO, Yuan JM, Zheng W, Dawsey SM, Dong LM, Lee MP, Ding T, Qiao YL, Gao YT, Koh WP, Xiang YB, Tang ZZ, Fan JH, Wang C, Wheeler W, Gail MH, Yeager M, Yuenger J, Hutchinson A, Jacobs KB, Giffen CA, Burdett L, Fraumeni JF, Tucker MA, Chow WH, Goldstein AM, Chanock SJ, Taylor PR. A shared susceptibility locus in PLCE1 at 10q23 for gastric adenocarcinoma and esophageal squamous cell carcinoma. *Nat Genet* 2010; **42**: 764-767 [PMID: 20729852 DOI: 10.1038/ng.649]
- 18 **Wang LD**, Zhou FY, Li XM, Sun LD, Song X, Jin Y, Li JM, Kong GQ, Qi H, Cui J, Zhang LQ, Yang JZ, Li JL, Li XC, Ren JL, Liu ZC, Gao WJ, Yuan L, Wei W, Zhang YR, Wang WP, Sheyhidin I, Li F, Chen BP, Ren SW, Liu B, Li D, Ku JW, Fan ZM, Zhou SL, Guo ZG, Zhao XK, Liu N, Ai YH, Shen FF, Cui WY, Song S, Guo T, Huang J, Yuan C, Huang J, Wu Y, Yue WB, Feng CW, Li HL, Wang Y, Tian JY, Lu Y, Yuan Y, Zhu WL, Liu M, Fu WJ, Yang X, Wang HJ, Han SL, Chen J, Han M, Wang HY, Zhang P, Li XM, Dong JC, Xing GL, Wang R, Guo M, Chang ZW, Liu HL, Guo L, Yuan ZQ, Liu H, Lu Q, Yang LQ, Zhu FG, Yang XF, Feng XS, Wang Z, Li Y, Gao SG, Qige Q, Bai LT, Yang WJ, Lei GY, Shen ZY, Chen LQ, Li EM, Xu LY, Wu ZY, Cao WK, Wang JP, Bao ZQ, Chen JL, Ding GC, Zhuang X, Zhou YF, Zheng HF, Zhang Z, Zuo XB, Dong ZM, Fan DM, He X, Wang J, Zhou Q, Zhang QX, Jiao XY, Lian SY, Ji AF, Lu XM, Wang JS, Chang FB, Lu CD, Chen ZG, Miao JJ, Fan ZL, Lin RB, Liu TJ, Wei JC, Kong QP, Lan Y, Fan YJ, Gao FS, Wang TY, Xie D, Chen SQ, Yang WC, Hong JY, Wang L, Qiu SL, Cai ZM, Zhang XJ. Genome-wide association study of esophageal squamous cell carcinoma in Chinese subjects identifies susceptibility loci at PLCE1 and C20orf54. *Nat Genet* 2010; **42**: 759-763 [PMID: 20729853 DOI: 10.1038/ng.648]
- 19 **Ma H**, Wang LE, Liu Z, Sturgis EM, Wei Q. Association between novel PLCE1 variants identified in published esophageal cancer genome-wide association studies and risk of squamous cell carcinoma of the head and neck. *BMC Cancer* 2011; **11**: 258 [PMID: 21689432 DOI: 10.1186/1471-2407-11-258]
- 20 **Hu N**, Dawsey SM, Wu M, Bonney GE, He LJ, Han XY, Fu M, Taylor PR. Familial aggregation of oesophageal cancer in Yangcheng County, Shanxi Province, China. *Int J Epidemiol* 1992; **21**: 877-882 [PMID: 1468848 DOI: 10.1093/ije/21.5.877]
- 21 **Chang-Claude J**, Becher H, Blettner M, Qiu S, Yang G, Wahrendorf J. Familial aggregation of oesophageal cancer in a high incidence area in China. *Int J Epidemiol* 1997; **26**: 1159-1165 [PMID: 9447394 DOI: 10.1093/ije/26.6.1159]
- 22 **SPSS Incorporation**. SPSS 13.0 for the Windows. Chicago (IL): SPSS Inc, 2006
- 23 **Bigbold LP**. The cell-type-specificity of inherited predispositions to tumours: review and hypothesis. *Cancer Lett* 2004; **216**: 127-146 [PMID: 15533589 DOI: 10.1016/j.canlet.2004.07.037]
- 24 **Cheng CW**, Wu PE, Yu JC, Huang CS, Yue CT, Wu CW, Shen CY. Mechanisms of inactivation of E-cadherin in breast carcinoma: modification of the two-hit hypothesis of tumor suppressor gene. *Oncogene* 2001; **20**: 3814-3823 [PMID: 11439345 DOI: 10.1038/sj.onc.1204505]
- 25 **Haeno H**, Iwasa Y, Michor F. The evolution of two mutations during clonal expansion. *Genetics* 2007; **177**: 2209-2221 [PMID: 18073428 DOI: 10.1534/genetics.107.078915]
- 26 **Rice TW**, Blackstone EH, Rusch VW. 7th edition of the AJCC Cancer Staging Manual: esophagus and esophagogastric junction. *Ann Surg Oncol* 2010; **17**: 1721-1724 [PMID: 20369299 DOI: 10.1245/s10434-010-1024-1]

**P- Reviewers:** Stanilova S, Vargas FR **S- Editor:** Cui XM  
**L- Editor:** Webster JR **E- Editor:** Liu SQ



# World Journal of *Medical Genetics*

*World J Med Genet* 2014 August 27; 4(3): 46-76





**REVIEW**

- 46 Genetics of canine behavior: A review  
*Rigterink A, Houpt K*
- 58 Molecular genetics of gastric adenocarcinoma in clinical practice  
*Cho M, Eze O, Xu R*
- 69 Genome engineering using the CRISPR/Cas system  
*Horii T, Hatada I*

## Contents

*World Journal of Medical Genetics*  
Volume 4 Number 3 August 27, 2014

**APPENDIX** I-V Instructions to authors

**ABOUT COVER** Editorial Board Member of *World Journal of Medical Genetics*, Zong Sheng Guo, PhD, Assistant Professor, Division of Surgical Oncology, University of Pittsburgh Cancer Institute, Pittsburgh, PA 15213, United States

**AIM AND SCOPE** *World Journal of Medical Genetics* (*World J Med Genet*, *WJMG*, online ISSN 2220-3184, DOI: 10.5496) is a peer-reviewed open access academic journal that aims to guide clinical practice and improve diagnostic and therapeutic skills of clinicians.

*WJMG* covers topics concerning genes and the pathology of human disease, molecular analysis of simple and complex genetic traits, cancer genetics, epigenetics, gene therapy, developmental genetics, regulation of gene expression, strategies and technologies for extracting function from genomic data, pharmacological genomics, genome evolution. The current columns of *WJMG* include editorial, frontier, diagnostic advances, therapeutics advances, field of vision, mini-reviews, review, topic highlight, medical ethics, original articles, case report, clinical case conference (Clinicopathological conference), and autobiography.

We encourage authors to submit their manuscripts to *WJMG*. We will give priority to manuscripts that are supported by major national and international foundations and those that are of great basic and clinical significance.

**INDEXING/ABSTRACTING** *World Journal of Medical Genetics* is now indexed in Digital Object Identifier.

**FLYLEAF** I-IV Editorial Board

## EDITORS FOR THIS ISSUE

Responsible Assistant Editor: *Xiang Li*  
Responsible Electronic Editor: *Su-Qing Liu*  
Proofing Editor-in-Chief: *Lian-Sheng Ma*

Responsible Science Editor: *Fang-Fang Ji*  
Proofing Editorial Office Director: *Xiu-Xia Song*

**NAME OF JOURNAL**  
*World Journal of Medical Genetics*

**ISSN**  
ISSN 2220-3184 (online)

**LAUNCH DATE**  
December 27, 2011

**FREQUENCY**  
Quarterly

**EDITOR-IN-CHIEF**  
**Hans van Bokhoven, Professor, PhD**, Department of Human Genetics and Cognitive Neurosciences, Radboud university Nijmegen Medical centre, PO Box 9101, 6500 HB Nijmegen, The Netherlands

**EDITORIAL OFFICE**  
Jin-Lei Wang, Director  
Xiu-Xia Song, Vice Director  
*World Journal of Medical Genetics*

Room 903, Building D, Ocean International Center, No. 62 Dongsihuan Zhonglu, Chaoyang District, Beijing 100025, China  
Telephone: +86-10-59080039  
Fax: +86-10-85381893  
E-mail: [editorialoffice@wjgnet.com](mailto:editorialoffice@wjgnet.com)  
Help Desk: <http://www.wjgnet.com/esp/helpdesk.aspx>  
<http://www.wjgnet.com>

**PUBLISHER**  
Baishideng Publishing Group Inc  
8226 Regency Drive,  
Pleasanton, CA 94588, USA  
Telephone: +1-925-223-8242  
Fax: +1-925-223-8243  
E-mail: [bpgoffice@wjgnet.com](mailto:bpgoffice@wjgnet.com)  
Help Desk: <http://www.wjgnet.com/esp/helpdesk.aspx>  
<http://www.wjgnet.com>

**PUBLICATION DATE**  
August 27, 2014

## COPYRIGHT

© 2014 Baishideng Publishing Group Inc. Articles published by this Open Access journal are distributed under the terms of the Creative Commons Attribution Non-commercial License, which permits use, distribution, and reproduction in any medium, provided the original work is properly cited, the use is non commercial and is otherwise in compliance with the license.

## SPECIAL STATEMENT

All articles published in journals owned by the Baishideng Publishing Group (BPG) represent the views and opinions of their authors, and not the views, opinions or policies of the BPG, except where otherwise explicitly indicated.

## INSTRUCTIONS TO AUTHORS

Full instructions are available online at [http://www.wjgnet.com/2220-3184/g\\_info\\_20100722180909.htm](http://www.wjgnet.com/2220-3184/g_info_20100722180909.htm)

## ONLINE SUBMISSION

<http://www.wjgnet.com/esp/>



## Genetics of canine behavior: A review

Amanda Rigterink, Katherine Houpt

Amanda Rigterink, Indiana Veterinary Behavior Services, LLC,  
West Lafayette, IN 47906, United States

Katherine Houpt, College of Veterinary Medicine, Department  
of Clinical Sciences Cornell University, CVM Cornell University,  
Ithaca, NY 14853-6401, United States

Author contributions: Rigterink A reviewed domestication and  
breed differences in behavior; Houpt K reviewed aggression and  
other behavior abnormalities.

Correspondence to: Katherine Houpt, Professor, College of  
Veterinary Medicine, Department of Clinical Sciences Cornell  
University, CVM Cornell University, 3065 E Dixon Lake Road,  
Ithaca, NY 14853-6401, United States. [kah3@cornell.edu](mailto:kah3@cornell.edu)

Telephone: +1-989-4488083 Fax: +1-989-4488083

Received: December 26, 2013 Revised: May 20, 2014

Accepted: June 10, 2014

Published online: August 27, 2014

**Key words:** Dog; Wolf; Fox; Canine aggression; Dopamine; Serotonin; Canid

**Core tip:** This review incorporates the latest findings in the rapidly moving field of canine behavioral genetics. The genes involved in tameness of foxes and in domestication of dogs from wolves are discussed. The genes involved in several obsessive compulsive behaviors such as flank sucking and circling are mentioned. The genetic and physiological differences between aggressive and non-aggressive dogs of various breeds are emphasized.

Rigterink A, Houpt K. Genetics of canine behavior: A review. *World J Med Genet* 2014; 4(3): 46-57 Available from: URL: <http://www.wjgnet.com/2220-3184/full/v4/i3/46.htm> DOI: <http://dx.doi.org/10.5496/wjmg.v4.i3.46>

### Abstract

The past decade has seen rapid progress in the field of canid behavioral genetics. The recent advances are summarized in this review. The identification of the genes responsible for tameness in silver foxes is the culmination of a half century of behavioral testing and, more recently, genomic investigation. There is agreement that domestic dogs evolved from wolves, but when and from which population remains controversial. The genetic differences between wolves and dogs identified include those for neurotransmitters and digestion. Breed differences in behavior are well known, but only recently have the genetics underlying these differences been investigated. The genes responsible for flank sucking in Doberman Pinschers and for several other obsessive compulsive problems in other breeds have been identified. Aggression is the least desirable canine trait, and several laboratories have detected differences in neurotransmitters and their receptors between aggressive and non-aggressive dogs. In English Cocker Spaniels, the genes linked to aggressive behavior code for dopamine, serotonin, and glutamate receptors. A dopamine transporter gene has been associated with impulsive behavior in Malinois.

© 2014 Baishideng Publishing Group Inc. All rights reserved.

### INTRODUCTION

This review is a summary of recent research focusing on the current knowledge of the genetic contribution to behavior in the Canidae family. We first review the farm fox experiment and how this long-term study has led to greater understanding of the process of canine domestication at the phenotypic and molecular levels. We then turn our attention to the relationship between dogs and wolves and canine breed differences in behavior. Finally, we review the current knowledge of the genetic basis of aggressive behavior in dogs.

### TAME FOX EXPERIMENT

The farm-fox experiment constitutes a major milestone in canid behavioral genetics, clearly demonstrating the genetic basis of behavior. No discussion of the genetics of canine behavior would be complete without summarizing some of the groundbreaking research performed at the Institute of Cytology and Genetics (ICG) of the Russian Academy of Sciences<sup>[1-3]</sup>. For more than 50 years,

scientists at the ICG in Novosibirsk, Russia have been reconstructing experimentally the domestication process in farm-bred silver foxes (a variant form of the red fox, *Vulpes vulpes*) as a novel model for studying the genetic basis of canine domestication and behavior. In 1959, scientist Dmitry Belyaev and his team began an intensive selective breeding program of silver farm foxes to isolate the trait of tame behavior towards humans<sup>[2,4]</sup>. After several generations of selective breeding in a controlled environment, Belyaev succeeded in attaining a subset of tame foxes. During fifty years of continuous selective breeding, the farm-fox experiment has tested over 52000 foxes for tameness, with the resultant tame population of foxes showing friendly dog-like responses to humans as early as one month of age<sup>[2,3,5]</sup>.

The goal of selective breeding of the farm foxes was limited strictly to behavioral criteria related to tameness. However, physical, developmental, physiological, and other behavioral differences also emerged in the tame foxes compared to the original farm-bred foxes. The selection for tameness led to numerous physical changes in the foxes, including piebald coats, floppy ears, and curly tails, despite no selection criteria for these traits<sup>[2]</sup>. In addition, the socialization period elongated from approximately 45 to 60 d in the selected tame foxes, similar to the socialization period in the domestic dog<sup>[1]</sup>. Tame foxes also developed a novel repertoire of affiliative vocalizations towards humans to promote interaction<sup>[6]</sup>. Hare *et al.*<sup>[7]</sup> found that tame fox kits are as skillful as puppies in using human point and gaze gestures for finding hidden food, demonstrating that domestication has led to improved social cognitive ability. Physiological differences also were found with hormonal assays showing that tame foxes do not experience stress when in contact with humans. A comparative study of hypothalamic-pituitary-adrenal axis (HPA) function in tame *vs* unselected foxes showed that in tame foxes, basal and stress-induced blood cortisol levels were respectively three- and five-fold lower than in the unselected foxes<sup>[2,8]</sup>.

In the 1970s, a second parallel strain of farm foxes began to be bred selectively at the Institute for Cytology and Genetics—those with aggressive behaviors towards humans. Fifty farm-bred silver foxes with the most aggressive responses towards humans were selected and used as the basis of the aggressive population<sup>[9]</sup>. Criteria for measuring behavior in the aggressive population were the critical distances between the experimenter and the caged animals at which the animals first demonstrated aggression and the intensity of the aggressive responses<sup>[10]</sup>.

From the evolution of these tame and aggressive populations of foxes, much information has been learned about the changes that can occur with intensive behavior selection pressures. Because the fox-farm domesticated foxes were created in only a few decades through intense selection and by focusing exclusively on certain behavioral traits, it seemed reasonable to assume that a small number of genetic loci determined the behavioral traits<sup>[11]</sup>. A rudimentary map of the fox genome with karyotype

and some linkage groups was available by the late 1990s; however, a meiotic linkage map of the fox was needed to determine which loci were implicated in tame behavior<sup>[9]</sup>. Fortunately, the fox and the dog share a close evolutionary and genetic relationship, and since the dog genome was sequenced by 2005<sup>[12]</sup>, available canine genomic information then could be utilized to develop the necessary fox meiotic map<sup>[3,11]</sup>. The availability of high resolution canine genome maps and sequence data aided in the creation of the fox meiotic linkage map, with the high genomic sequence identity between dog and fox permitting the adaptation of canine microsatellites for genotyping and meiotic mapping in foxes. Using 320 such markers, Kukekova *et al.*<sup>[3]</sup> constructed the first meiotic linkage map of the fox genome. This first mapping covers 16 fox autosomes and the X chromosome. After alignment with a canine genome sequence of similar length, high conservation of marker order between homologous regions of the two species was apparent<sup>[11]</sup>. Utilizing and adapting scoring systems (for tameness and aggression phenotypes) developed by the fox-farm experiment over the years for the selective breeding process, Kukekova *et al.*<sup>[3]</sup> created a new principal-component analysis of fox behavior with selected traits. This new scoring system effectively reduced 311 binary scoring behaviors to fifty of the most important traits that would serve as quantitative phenotypes (and continuous variables) to represent heritable differences in behavior among individual foxes and the fox populations and permit quantitative genetic analysis<sup>[10]</sup>. By interval mapping using fox and canine meiotic maps, a locus for tame behavior on fox chromosome VVU12 was identified. This locus is orthologous to a genomic region implicated in canine domestication<sup>[13]</sup>. Tameness as the defining trait of domestication is a complex “phenotype” consisting of many behavioral variables. In fact, when genome-wide association studies were performed by Kukekova *et al.*<sup>[3]</sup>, the resulting data suggested that at least two VVU12 loci are associated with tame *vs* aggressive behavior and active *vs* passive behavior. Moreover, differing mapping characteristics of specific behavioral traits were found, suggesting different genotype/phenotype relationships; for example, floppy *vs* erect ears are associated with different regions of VVU12 and vary between tame and aggressive foxes. Expression of the VVU12 loci thus appears to depend on interaction with other parts of the genome and on individual fox parents<sup>[13]</sup>.

At the molecular level, the development of transcriptome sequencing significantly enhances genetic study without the need for a fully sequenced genome. The comparison of transcriptome sequencing from the prefrontal cortices of a tame and an aggressive fox is in the preliminary stages at this time<sup>[9]</sup>. Thus far, preliminary analysis of “comparison of transcriptome sequences of the same genes between the tame and aggressive fox samples has identified a large set of informative single nucleotide polymorphism (SNP) markers and begun a catalogue of gene-specific sequence variants between the

two strains<sup>113]</sup>.

The farm-fox experiment demonstrates that over generations, intensive selection for tame behavior in foxes can serve as a reliable model for studying the genetic basis of canine domestication. The identification of genetic loci that both influence tame behavior in foxes and are homologous to regions in the dog genome supports the hypothesis that domesticated behavior in dogs and foxes may have similar genetic bases. These recent advances will help identify more genes implicated in fox behavior that can be correlated to dog domestication.

## WOLVES TO DOGS

The complete sequencing of the dog genome has greatly expanded general knowledge of the processes of genome evolution and the genetic basis of phenotypic traits in dogs and other animals. However, the evolutionary path leading from wild ancestor to domesticated dog continues to remain elusive. Comparative genomics utilizing the completed dog genome has confirmed the close relationship of dogs to such other canidae as foxes, coyotes, and wolves. It appears that modern canids share a common ancestor dating back approximately ten million years; the closest relatives to the dog such as the gray wolf and coyote share a common ancestor dating to approximately three to four million years ago<sup>114]</sup>. Like the dog, all wolf-like canids have 78 chromosomes and can mate with one another to produce fertile offspring. Thus, wolf-like canid species are among the strongest candidates for the ancestors of today's dog. Moreover, molecular genetic data from the past two decades<sup>112,115,116]</sup> strongly support the origin of the dog from the gray wolf in particular<sup>117]</sup>. Molecular evidence also suggests that divergence of dog from wolf and the beginning of the dog's relationship with humans occurred as recently as 15000 years ago<sup>114]</sup>. Other studies looking at genomic variation in wolves, Chinese indigenous dogs, and modern breeds point to an even earlier beginning to domestication, possibly about 30000 years ago, prior to the development of an agricultural human society<sup>118,119]</sup>. Very early domestication may have involved the intentional taming of small groups of wolves who, less fearful of humans and motivated by hunger, scavenged the camps of Mesolithic human hunters-gatherers<sup>120]</sup>.

Where canine domestication originated also is debatable. While DNA genomic data suggest a Middle Eastern origin, analyses of mitochondrial DNA and Y-chromosome markers from various dog breeds and from geographically-dispersed wolf populations suggest that canine domestication originated in East Asia<sup>11]</sup>. Wang *et al.*<sup>118]</sup> used whole-genome sequencing to compare gray wolves, Chinese indigenous dogs, and modern breeds. They found that the genetic variation between the three canid groups generally decreased step-wise from wolf to Chinese dog to modern dog breed. Based on these findings, they speculate that the Chinese indigenous dog may represent the link between wolf and dog and the pro-

genitor of today's diverse modern dog breeds. They identified 311 genes that appear to have been selected in dogs compared to wolves and that have functions affecting sexual reproduction, digestion/metabolism, neurological processes, and cancer. The fact that these particular genes overlap to a great extent with those also selected in humans suggests a parallel evolutionary process in dogs and humans, especially in the realm of neurological processes. They note that: As domestication is often associated with large increases in population density and crowded living conditions, these "unfavourable" environments might be the selective pressure that drove the rewiring of both species. Positive selection in neurological pathways, in particular the serotonin system, could be associated with constant need for reduced aggression stemming from the crowded living environment<sup>118]</sup>.

Another study employed mitochondrial DNA sequencing, showing a closer relationship of dogs to gray wolves from East Asia<sup>121]</sup>. VonHoldt *et al.*<sup>122]</sup> sought to identify the primary source of genetic diversity for domestic dogs and conducted an extensive genome-wide survey of over 48000 SNPs in dogs and gray wolves. Their data, however, showed that dogs share a greater percentage of multi-locus haplotypes unique to gray wolves from the Middle East rather than from East Asia<sup>122]</sup>.

Although genetic data support the theory that the process of canine domestication began in East Asia over 15000 years ago, a recent study compared the complete mitochondrial genome sequences of 18 European prehistoric canids to a comprehensive panel of modern dogs and wolves. The researchers found phylogenetic relatedness between the modern dogs and the ancient canids of Europe dating back to more than 30000 years ago, thus suggesting that canine domestication first may have occurred in Europe rather than in Asia<sup>123]</sup>.

Behavior differences between dogs and wolves are the most striking result of the domestication process, even more than the marked differences in physical size and shape. In fact, the canine breeds in existence today have diverse physical characteristics that distinguish them from one another just as much as from wolves. However, the fact that all breeds of domestic dog as a group are more similar in behavior when compared to one another than when compared to the wolf suggests that genetic selection for behavior drove the domestication process. It is logical to hypothesize that ancestral wolves initially may have experienced natural selection for tame behavior, permitting coexistence with humans. Based on the findings of the fox-farm experiment where genetic loci influencing tame behavior in foxes are homologous to regions in the dog genome and also related to selection differences between dogs and wolves, it is plausible to suggest that domesticated behavior in dogs and foxes share a similar genomic basis<sup>11]</sup>.

Several studies before and after the complete sequencing of the dog genome in 2005 have attempted to target, at the molecular level, the genetic basis of behavioral differences between the domestic dog and its wolf progeni-

tor. A study by Sætre *et al.*<sup>[24]</sup> used microarray technology to evaluate mRNA expression levels of 7762 genes in the post-mortem brains of dogs, wolves, and coyotes. They found markedly altered gene expression of two neuropeptides, CALCB and NPY, in the dogs as compared to the wolves and coyotes. These neuropeptides, present in all mammalian brains, are implicated in energy control and feeding behavior, neuroendocrine stress response *via* the HPA axis, and possibly play a role in anxiety and depression. The findings of species-specific differences in the elaboration of the neuropeptides suggest that selection for behavior during domestication may have resulted in modification of mRNA expression patterns in genes located in the hypothalamus of the dog<sup>[24]</sup>. Björnerfeldt *et al.*<sup>[25]</sup> postulate that domestication of dogs created a new lifestyle that changed selective forces acting on the species, in turn affecting the dog's genome. Using mitochondrial DNA sequencing in 14 dogs, 6 wolves, and 3 coyotes, they showed that dogs have accumulated into their genome non-synonymous changes in mitochondrial genes at rates faster than in wolves. In turn, this results in elevated levels of protein variations in the dog as compared to the wolf. Björnerfeldt *et al.*<sup>[25]</sup> conclude that an important consequence of domestication is a "relaxation of selective constraint on dog mitochondrial DNA" that also could have affected other parts of the dog genome to facilitate "the generation of novel functional genetic diversity"<sup>[25]</sup>. Cruz *et al.*<sup>[26]</sup> compared the genome of the dog to that of the gray wolf to examine the effect of domestication. Using whole-genome SNP data, they compared the variation in dog and wolf genes. They also found increased frequency in the trend for non-synonymous mutations in dogs as compared to their wild canid counterparts. They concluded that the increase in mutation rate could have myriad effects, some deleterious, and may indicate that the process of domestication in the dog led to an increase in functional genetic variation that has contributed to the markedly diverse physical and behavioral phenotypes characteristic of dog breeds, as well as to the prevalence of pathology in modern breeds<sup>[26]</sup>.

Li *et al.*<sup>[27]</sup> studied the expression profiles of a specific subset of developmental genes believed to be implicated in the evolution of dog domestication. They ran comparative genomic analyses by assaying the SNP genotypes in Chinese native dogs (believed to have the genetic structure most similar to that of ancient dog), German Shepherd (purebred) dogs, and gray wolves to detect a genetic basis for the behavior transformation from wolf to primitive dog to modern purebred dog<sup>[27]</sup>. Genomic regions that have undergone strong selection in the recent past should show extended haplotype homozygosity<sup>[28]</sup>. Following this line of reasoning, Li *et al.*<sup>[27]</sup> detected four regions of high extended haplotype homozygosity that contained only a single highly differentiated SNP located within a single gene. Comparison of candidate genes between the Chinese native dogs and wolves showed a high bias for expression localized in the brain's prefrontal cortex, the center for complex cognitive-type behaviors.

However, candidate genes showing large population differentiation between the Chinese dogs and German Shepherds did not demonstrate significant expression bias. Thus, the finding that wolves and dogs have highly differentiated brain-based genes suggests that behavioral transformation most likely was key to the onset of domestication and that "this rapid evolution likely was driven by artificial selection during the primary transition from wolves to ancient dogs, and was consistent with the evolution of dog-specific characteristics, such as behavior transformation, for thousands of years"<sup>[27,28]</sup>.

Other recent studies have taken a closer look at the genetic processes underlying physiological and behavior differences resulting from dog domestication. Utilizing whole-genome resequencing of wolves and dogs, Axelson *et al.*<sup>[29]</sup> identified 36 genomic regions that likely are implicated in selection during the domestication of the dog. It is of interest that more than half of the regions play roles in brain function with 8 regions in particular involved in neurophysiologic pathways that may underlie behavioral changes characteristic of dog domestication. Moreover, they identified 10 genes with selection signals that play key roles in starch digestion and fat metabolism. In terms of starch digestion, three genes (*AMY2B*, *MGAM*, and *SGLT1*) that facilitate the digestion of starches show evidence of being selected for during the process of dog domestication. These findings may indicate that, unlike in carnivorous wolves, genetic mutations found in modern dog facilitate the adaptation to and even thriving on a diet available in cohabitation with humans<sup>[29]</sup>.

## BREED DIFFERENCES

Over the past hundreds of years, the selective breeding of domestic dogs has given rise to more than 400 modern dog breeds with many unique differences in both physical appearance and behavior characteristics<sup>[30]</sup>. The physical differences among the dog breeds mostly are obvious to the naked eye, and the behavior differences between breeds also are distinctive and diverse<sup>[31]</sup>. Humans have exerted genetic pressure on dogs by selecting various traits to create breeds better adapted to utilitarian purposes such as herding, guarding, or hunting. The modern dog's extraordinary diversity in phenotype, behavior, and ability to perform tasks is unmatched by any other species on earth<sup>[32]</sup>. A study by McGreevy *et al.*<sup>[33]</sup> investigated the relationship between height, bodyweight, and canine cephalic index (CI: the ratio of skull width to skull length) and how these values correlated with certain behavior traits using the Canine Behavioral Assessment and Research Questionnaire (C-BARQ). It is of interest that certain canine morphotypes were associated reliably with particular behavior profiles. For example, brachycephalic skull shape (high CI) may be a by-product for human selection of neotenuous behavioral characteristics, and dolichocephalic skull shape is a product of human selection for hunting and chasing ability. The authors note that it is unclear if these associations between mor-



phology and behavior represent functional co-adaptations or accidental by-products of allometric change. Therefore, the relationships noted in this study could be either genetically or environmentally driven or both<sup>[33]</sup>.

With its wealth of phenotypic diversity, the dog clearly is a valuable genetic model for studying both breed-specific behaviors and abnormal behaviors. The persistence of such breed-specific behaviors as herding, pointing, tracking, and hunting in the absence of training or motivation suggests that these behaviors are, at least in part, controlled at a genetic level<sup>[34,35]</sup>.

Prior to the completion of the dog genome in 2005, genetic studies used mitochondrial sequencing to reveal a large amount of variation in relatively short sequences. Although some breed clustering could be demonstrated, researchers found that mitochondrial sequences were more successful at distinguishing between species than between breeds<sup>[36]</sup>. Early genetic studies also utilized microsatellite-based marker sets to study the genomes of a small number of breeds. Differences in allele frequencies occurred in different breeds supporting the hypothesis that there was less variation within breeds than across the species<sup>[36]</sup>. Parker *et al.*<sup>[37]</sup> investigated the relationships among 85 breeds using 96 microsatellite markers, demonstrating marked population stratification within the dog species and establishing that the breeds were indeed genetically separate. Once the whole genome of the dog became available, use of SNPs became favored over microsatellites due to the ease of genotyping bialleles and analyzing thousands of markers in a single assay<sup>[36]</sup>. SNP genotyping chips were derived from the over 2 million SNPs in the dog genome<sup>[38]</sup>. Both analytic techniques are useful; clustering analysis using mitochondrial DNA demonstrates hybridization among groups, while SNP analysis results in a phylogenetic tree that show the unique placement of a breed within a group<sup>[36]</sup>. Moreover, SNP analysis corroborates earlier research showing that genetic variation among breeds is greater than that among individuals. A study by Vonholdt *et al.*<sup>[22]</sup> demonstrated a 4% overall variation between breed clusters.

In the 1950s through the 1960s, Scott and Fuller pioneered research on identifying heritable differences in behavior and cognition in the dog using five different breeds in a laboratory model setting<sup>[39]</sup>. More recent studies have assessed heritability of behavior in working and/or pet dog populations outside of the laboratory setting<sup>[40]</sup>.

In 1989, the Swedish Dog Mentality Assessment (DMA) was initiated as a tool for selective breeding in working dogs. The test originally was developed as a tool for selective breeding of working dogs, but it is used today as a general behavioral test by many breeding clubs in Sweden. The DMA has been applied to over 24000 dogs representing more than 180 breeds. Using this data set and the pedigrees of German Shepherds and Rottweiler dogs, Sætre *et al.*<sup>[41]</sup> noted that the genetic correlation of the score on one test was not independent of the score on another test. In fact, their analysis provides evidence that

there may be substantial shared genetics underlying most of the behavioral response in all of the test situations except for aggression which tended to be distinct. Sætre *et al.*<sup>[41]</sup> identified “shyness-boldness” as a generalized trait underlying many behavioral scores with a heritability of 0.25-0.27.

Recent research has shown that the genetic similarities among different breeds may not correlate well to characteristic behavior traits attributed to historical functional breed groups such as herders and hunters<sup>[30,42]</sup>. Turcsan *et al.*<sup>[30]</sup> investigated whether or not behavioral traits historically believed to characterize certain breed categories actually correlated with genetic relatedness. Using online questionnaires submitted by 5733 dog owners of 98 breeds, they looked at trainability, boldness, calmness, and dog sociability. They found that the breeds differed to a great extent in the four traits and that breed-specific behavior in trainability and boldness appeared to be determined partly by genetics. However, breeds that were similar in behavioral characteristics per report of the owners did not correspond well to recognized functional/conventional breed classification nor to genetic breed clusters. The authors state that this lack of correlation between the questionnaire results and commonly acknowledged breed or functional group traits could be associated with cross-breeding with breeds of dissimilar behavioral traits or could represent differences in socialization and/or relationship with owners. The authors conclude “...the behavioural breed clusters showed poor correspondence to both the functional and genetic categorization, which may reflect the effect of recent selective processes. Behavioural breed clusters can provide a more reliable characterization of the breeds’ current typical behaviour”<sup>[30]</sup>.

Meyer *et al.*<sup>[43]</sup> estimated the heritability and correlation of 7 behavioral traits in German Shepherd Dogs in Switzerland using data from 4855 animals that underwent the standardized behavior test of the German Shepherd Dog Club of Switzerland between 1978 and 2010. The traits tested were self-confidence, nerve stability, hardness, temperament, sharpness, defense drive, and reaction to gunfire. Sex, year of testing, judge, place of testing, and age at testing were found to have significant effects on the outcome of the test. Overall, estimated heritability of the traits was low, ranging from 0.05 (5%) to 0.21 (21%). It also is of interest that some traits were highly correlated; self-confidence and nerve stability had a genetic correlation of 0.98 and sharpness and defensive drive, 0.93. Meyer *et al.*<sup>[43]</sup> suggest that while the heritability of behavioral traits is generally low, genetic evaluation of behavior can be helpful as a basis for selection of a given trait, with the caveat that precise definition of the desired traits along with accurate scoring of the dog’s behavior are requisite<sup>[43]</sup>. Mehrkam *et al.*<sup>[44]</sup> recently reviewed the current state of knowledge regarding canine breed differences in behavior, finding scientific evidence for differences both between breeds as well as within-breed differences<sup>[44]</sup>.

The genetics underlying racing performance has been

studied in sled dogs<sup>[45]</sup>. The Alaskan sled dog is considered genetically distinct in that the population has been shaped to create a group of high-performance athletes through selective interbreeding with purebred dogs based on working ability rather than breed physical appearance. New breeds have been introduced gradually into the lines of racing dogs to improve racing performance. Therefore, Alaskan sled dogs provide a unique opportunity to research the impact of trait selection and breed composition and their influence on genomic structure. Huson *et al*<sup>[45]</sup> genotyped 199 Alaskan sled dogs using 96 microsatellite markers and compared the data to that from 141 genotyped purebred breeds. The breed composition of each sled dog was compared to its performance phenotype, including speed, endurance, and work ethic. It is of interest that the sled dogs separated into two groups that aligned with their racing style-sprint *vs* distance<sup>[46]</sup>. Huson *et al*<sup>[46]</sup> then used a set of 7644 ancestry informative marker SNPs to model ancestry in the sprint and distance sled dog populations with four known reference breeds, the Alaskan Malamute, Siberian Husky, German Shorthaired Pointer, and Borzoi. It was found that the distance sled dogs had, on average, highest Alaskan Malamute allele patterns compared to the sprint dogs who had the highest German Shorthaired Pointer allele patterns. In addition, genetic comparison between sprint *vs* distance racing Alaskan sled dogs identified several genomic regions associated with differences in racing style and pinpointed a variant of *MYH9* gene that is associated with increased heat tolerance in sprint dogs<sup>[46]</sup>. Although variants responsible for improved muscle function are important, those responsible for the motivation to perform are also involved.

There are many genetic differences in behavior, but few of the genes are known. The laboratory of Veterinary Ethology of Tokyo University has located putative genes affecting canine behavior. The researchers have identified polymorphisms in five breeds of dogs (Golden Retriever, Labrador Retriever, Maltese, Miniature Schnauzer, and Shiba) that pinpoint differences in SNPs in genes regulating neurotransmitters, the enzymes that synthesize or destroy the neurotransmitters, and the receptors<sup>[47]</sup>. SNP (T199C) is located on the putative third exon of the canine monoamine oxidase B gene that causes an amino acid substitution from cysteine to arginine. Takeuchi *et al*<sup>[47]</sup> also found 4 SNPs in the tyrosine hydroxylase and dopamine beta hydroxylase genes. Ogata *et al*<sup>[48]</sup> found 2 SNPs in the glutamine transporter gene. The Tokyo University researchers have related the polymorphisms with breed behaviors as identified by Hart *et al*<sup>[49]</sup>, although there is no direct evidence that these could explain interbreed differences<sup>[51]</sup>.

Due to the great diversity of dog breeds, the dog is a valuable genetic model for studying both breed-specific behaviors and abnormal behaviors. At a molecular level, analytic techniques to study breed differences include using mitochondrial DNA to perform cluster analysis that shows hybridization among groups and SNP analysis that develops a phylogenetic tree and places a breed

within a group on that tree. Recent studies also have assessed heritability of behavior in both working dog and pet dog populations. The genetic similarities among different breeds may not correlate well to characteristic behavior traits attributed to historical function of the breed groups. However, behavioral breed clusters may provide a more reliable characterization of the breeds' current typical behavior. Currently, only a few genes that underlie inheritable behavior characteristics are known. Polymorphisms have been identified in five breeds of dogs, pinpointing differences in SNPs in genes regulating neurotransmitters, enzymes acting on neurotransmitter enzymes, and receptors.

## GENETICS OF ABNORMAL BEHAVIOR

### Flank sucking

Yokoyama *et al*<sup>[50]</sup> pointed out that genome wide association testing is more profitable than the candidate gene approach to determining the genetics of behavior. Using this approach, the first gene for a specific behavior was found. Flank sucking, a very specific and easily recognized compulsive problem, is a behavior seen almost exclusively in Doberman Pinschers. *CDH2* is the gene associated with this compulsive behavior. Occasionally a blanket or another material can serve as the substrate for sucking. It is not a serious behavior problem because the irritation to the skin is mild. More owners complain about fabric sucking because the material must be replaced. The sucking behavior occurs mostly as the dog is resting prior to sleeping. Using genome-wide analysis, Dodman *et al*<sup>[51]</sup> found an association of SNPs peak on canine chromosome 7. The most significantly associated SNP is located within the *CDH2* gene. *CDH2* is widely expressed, mediating synaptic activity-regulated neuronal adhesion. Dogs showing multiple compulsive behaviors have a higher frequency of the risk allele than do dogs with a less severe phenotype (60% and 43%, respectively) compared with 22% in unaffected dogs<sup>[51]</sup>.

In an interesting follow-up to the genetic basis of this abnormal behavior, Ogata *et al*<sup>[52]</sup> found that the brains of flank sucking Dobermans differed from those of unaffected Dobermans. Magnetic resonance imaging revealed higher total brain and gray matter volumes and lower dorsal anterior cingulate cortex and right anterior insula gray matter densities in the affected dogs. The affected Dobermans also had higher fractional anisotropy in the splenium of the corpus callosum, the degree of which correlated with the severity of the behavioral phenotype<sup>[52]</sup>.

Another behavior abnormality, tail chasing, can have multiple etiologies including neuropathic pain, so it is not surprising that there is no association with the *CDH2* gene<sup>[53,54]</sup>. Single photon emission computed tomography (SPECT) was used with <sup>123</sup>I-R91150 and <sup>123</sup>I-FP-CIT, in combination with <sup>99m</sup>Tc-ECD brain perfusion co-registration, to measure the serotonin (5-HT) 2A receptor, dopamine transporter (DAT), and serotonin transporter (SERT) availability. There was significantly less 5-HT2A

receptor binding in the frontal and temporal cortex of obsessive compulsive dogs. The midbrain SERT also was lower. The DAT differences between normal and compulsive dogs were mixed<sup>[53]</sup>.

More recently the original data from Dodman *et al.*<sup>[51]</sup> was reanalyzed using a new calling algorithm called MAGIC was used to identify genes, in addition to cadherin, that are involved in flank sucking and other obsessive compulsive behavior (OCD). The genome wide association revealed 119 variants in evolutionarily conserved sites that are specific to dogs with OCD. Using small numbers of dogs, (< 16 of each breed), case dogs (exhibiting OCDs), control dogs, and unphenotyped dogs were compared. Four genes have an excess of case-only variation in evolutionarily constrained elements, even after correcting for gene size: ataxin-1 (ATXN1), neuronal cadherin (CDH2), catenin alpha2 (CTNNA2), and plasma glutamate carboxypeptidase (PGCP). CDH2, a neural cadherin, encodes a calcium dependent cell-cell adhesion glycoprotein important for synapse assembly, where it mediates presynaptic to postsynaptic adhesions.

CTNNA2 encodes a neuronal-specific catenin protein that links cadherins to the cytoskeleton. ATXN1 encodes a chromatin binding protein that regulates the Notch pathway<sup>[42]</sup>, a developmental pathway also active in the adult brain, where it mediates neuronal migration, morphology and synaptic plasticity<sup>[55]</sup>. All three of these genes are involved in synaptic formation. The fourth gene PGCP, encodes a poorly characterized plasma glutamate carboxypeptidase. It may be involved in the hydrolysis of N-acetylaspartylglutamate. One might consider glutamate targeting drugs for treatment of OCD's.

## NEUROTRANSMITTERS AND AGGRESSION

Canine aggression has been the subject of many genetics studies because it is the most common behavior presented as a problem and the only one responsible for human injury or even death<sup>[56]</sup>. Hyperactivity and impulsive (unpredictable) aggression by dogs are problems frequently presented to veterinarians. Since behavior is the consequence of central nervous activity, it is not surprising that differences in neurotransmitters are associated with differences in behavior. These differences can be at any stage in the production and function of the neurotransmitter. The levels of neurotransmitter or their metabolites in brain, blood or cerebral spinal fluid have been investigated, and transporters and receptors of neurotransmitters have been associated genetically with aggression and other behaviors.

Dopamine and serotonin are the neurotransmitters examined most frequently in studies of aggression. Serotonin is produced from tryptophan and is widely believed to be important in the etiology and treatment of mood disorders, including aggression in dogs<sup>[57]</sup>. It is logical to conclude that serotonin levels in the body fluids or number of serotonin receptors should be measured in normal

and abnormal dogs with the prediction that serotonin levels would be lower in aggressive dogs. The results of these studies are summarized below.

Dopamine (D1 and D2) is formed from tyrosine and catalyzed by the enzyme tyrosine kinase. Dopamine has multiple receptors and is inactivated by another enzyme, monoamine oxidase (MAO). Dopamine is transported back into the pre-synaptic neuron *via* a transporter. Studies in dogs exhibiting aggression have examined blood and cerebrospinal fluid levels of dopamine and its expression in the brain. In genetic studies, alleles regulating dopamine transporters, receptors, and dopamine deactivating enzymes have been compared in non-aggressive dogs and dogs exhibiting aggression. The results of these studies are summarized below.

### Blood and body fluids

Cakiroğlu *et al.*<sup>[58]</sup> found that serotonin in blood varied with canine disposition. Serum serotonin was 33 ng/mL in non-aggressive dogs and 12 ng/mL in aggressive dogs<sup>[58]</sup>. In a later study, Leon *et al.*<sup>[59]</sup> found lower levels of serotonin in plasma, serum and platelets in aggressive dogs of various breeds that presented to a behavior clinic than in the control group of Beagles. However, the differing serotonin levels might represent breed differences in serotonin rather than differences between aggressive and non-aggressive dogs.

It is probably more fruitful to look for genetic differences between dogs within the same breed. For that reason, English Cocker Spaniels were studied because dogs of that breed frequently exhibit unpredictable or impulsive aggression towards their owners<sup>[60]</sup>. Moreover, the prevalence of aggression varies with coat color; red (blonde or buff) spaniels are more aggressive than black ones and solid color spaniels are more likely to be aggressive than parti-colored ones. It is not clear how the production of pheomelanin (yellow pigment) rather than melanin (black pigment) leads to or is related to aggression although melanin and dopamine share a common precursor-tyrosine. This area bears investigation<sup>[61]</sup>. Amat *et al.*<sup>[62]</sup> compared serum serotonin levels in aggressive English Cocker Spaniels with those of aggressive dogs of a variety of other breeds and found the serotonin levels were significantly lower in the cockers.

MAO-A is an enzyme that catalyzes monoaminergic neurotransmitters such as dopamine and serotonin. A mutation that lowers the amount of MAO-A is associated with incarcerated humans, if they had bad childhood environments<sup>[63]</sup>. There is evidence in dogs that aggressive individuals have lower cerebrospinal levels of 5-hydroxyindole acetic acid and homovanillic acid, the major metabolites of serotonin and dopamine respectively<sup>[64]</sup>.

Based on current studies, dopamine is the neurotransmitter most involved in aggression. Different breeds appear to have genes that are active at different points in the pharmacodynamics of the catecholamine. For example, compared to their non-aggressive counterparts, aggressive English Cocker Spaniels have significantly different



alleles for a dopamine receptor as well as a serotonin receptor. The gene for a dopamine receptor also appears to affect impulsive behavior in working German Shepherds, and the dopamine transporter appears to be involved in aggression, at least in the Malinois. In addition, the short form of the tyrosine hydroxylase gene appears to be involved in dopamine synthesis in German Shepherds and Siberian Huskies with particular behaviors. These studies will be discussed in detail below.

### Brain receptors for neurotransmitters

The amygdala is a structure in the brain that is associated with fear. The basolateral nuclear group of the amygdala is involved directly in the modulation of aggressive behavior in dogs. This structure has an increased volume and a higher number of neurons in aggressive dogs<sup>[65]</sup>. Serotonin 1B receptors act as auto-receptors regulating serotonin release. Indirect immunohistochemistry revealed that aggressive dogs had a higher number of serotonin 1B receptors than non-aggressive dogs. One might have expected the number to be lower in aggressive dogs, but one possible explanation is that a lower serotonergic activity is present in aggressive dogs because stimulation of presynaptic serotonin-1 autoreceptors causes a reduction of the serotonin release<sup>[65]</sup>.

Substance P is a neuropeptide that stimulates defensive aggression in cats<sup>[66]</sup> and mice<sup>[67]</sup>. It binds preferentially to neurokinin receptors. Using immunohistochemistry, Jacobs *et al.*<sup>[65]</sup> found that although the brains of aggressive dogs had more neurokinin reactivity in the amygdala than did normal dogs, the numerical densities and fractions of receptor-positive neurons did not differ significantly between the two groups. As noted above aggressive dogs have 27% more neurons in the amygdala than do normal dogs<sup>[65]</sup>.

Vermeire *et al.*<sup>[68]</sup> found differences in serotonin 2A receptors in the brains of impulsively aggressive dogs compared to normal dogs. Aggressive dogs had higher binding indexes for serotonin 2A receptors in the frontal and temporal cortex as revealed by SPECT following a 5-hydroxytryptophan (5-HT) antagonist radioligand injection. Although expensive and technically difficult, SPECT could be used to confirm a diagnosis of impulsive aggression<sup>[68]</sup>.

The brains of aggressive German Shepherds were compared with those of non-aggressive dogs of the same breed for beta adrenergic and serotonergic receptors using radioligand binding assays<sup>[69]</sup>. More binding of low affinity 5-HT (serotonergic) receptors were found in the whole brains of aggressive dogs. High affinity 5-HT was greater only in the hypothalamus and thalamus of the aggressive dogs. One might have expected 5-HT receptors to be decreased in aggressive dogs however, the increase in the number of 5-HT receptors may be due to a decrease in physiological serotonin levels at synaptic clefts or to an altered turnover of the neurotransmitter<sup>[69]</sup>.

It is not surprising that the adrenergic neurotransmitter norepinephrine might be involved in aggression.

Badino *et al.*<sup>[69]</sup> found that beta adrenergic binding was decreased in the frontal cortex, hippocampus, and thalamus of aggressive dogs. The decrease in beta adrenergic concentrations observed in these brain regions of aggressive dogs may be explained by a prolonged stimulation exerted by the high catecholamine levels resulting in beta adrenergic receptor down-regulation<sup>[69]</sup>.

In summary, there are differences in the brain, blood and cerebrospinal fluid between aggressive and non-aggressive dogs. Serotonin and its metabolites have been investigated most thoroughly. In general, blood serotonin levels are low and its metabolites are lower in aggressive dogs. The studies of receptors in the brain present a more complicated picture with serotonin receptors higher in aggressive dogs.

## HUMAN-DIRECTED IMPULSIVE AGGRESSION

### Heritability

Pérez-Guisado *et al.*<sup>[70]</sup> investigated the heritability (the percent variability due to genetics) of aggression in English Cocker Spaniels. They found that in addition to sex and coat color, nurture also influenced whether or not a dog was aggressive. The variance due to the sire heritability of aggression was only 0.2 (20%) whereas that due to the dam was 0.46 (46%) indicating a maternal-environmental effect<sup>[70]</sup>.

Although commonly perceived as gentle, non-aggressive dogs, Golden Retrievers can be aggressive, especially in European populations. Linamo *et al.*<sup>[71]</sup> used the Restricted Maximum Likelihood method to determine heritability of aggression based on the dog owner's impression of the animal's human and dog-directed aggression or the responses on C-BARQ<sup>[71,72]</sup>. They found heritability of 0.77 for human-directed aggression and 0.81 for dog-directed aggression. There is little correlation between the two types of aggression indicating separate genetic causes of the traits. There were high heritability estimates on several C-BARQ items such as strange dog approaching leashed dog (0.85), family member grooming dog (0.83), family member removing food (0.95), and stranger trying to touch dog (0.99)<sup>[71]</sup>. The next step in researching the etiology of aggression is to determine which mutations in the neurotransmitter, its receptor, or its transporters might be involved in aggression or other behavior abnormality.

### Genes

van den Berg *et al.*<sup>[73]</sup> did an extensive study of the genetic differences in four candidate genes affecting serotonin in aggressive and non-aggressive Golden Retrievers. They used mutation screens, linkage analysis, an association study, and a quantitative genetic analysis. There were no systematic differences in the coding DNA sequence of the candidate genes in aggressive and non-aggressive Golden Retrievers. An affected-only parametric linkage analysis revealed no strong major locus effect on human-



directed aggression related to the candidate genes. An analysis of 41 SNPs in the 1 Mb regions flanking the genes in 49 unrelated human-directed aggressive and in 49 unrelated non-aggressive dogs did not show association of SNP alleles, genotypes, or haplotypes with aggression at the candidate loci. They completed their analyses with a study of the effect of variation in the candidate genes on a collection of aggression-related phenotypic measures. The effects of the candidate gene haplotypes were estimated using the Restricted Maximum Likelihood method, with the haplotypes included as fixed effects in a linear animal model. They found no effect of the candidate gene haplotypes on a range of aggression-related phenotypes<sup>[73]</sup>.

Hejjas *et al.*<sup>[74]</sup> genotyped police and pet German Shepherd Dogs and diagnosed hyperactivity and impulsivity based on questionnaires. They compared the dopamine D4 receptors subtypes 2/2 with 2a/3a and 3a/3a (combined because 3a/3a is rare) with the behaviors. There was no difference in the activity-impulsivity scores between dogs with 2/2 genotype *vs* the 2/3a and 3a/3a combined genotype group either in the total sample or in the pet dog group. In contrast, police dogs with 2/2 genotype showed significantly lower activity-impulsivity scores compared with police dogs with 2/3a or 3a/3a genotype<sup>[74]</sup>.

Kubinyi *et al.*<sup>[75]</sup> found that German Shepherds with the short form of the tyrosine hydroxylase (*TH*, the enzyme involved in dopamine formation) gene were more active and impulsive. Wan *et al.*<sup>[76]</sup> also found that Siberian Huskies with the short form of the *TH* gene were more impulsive. They also reported that Siberian Huskies possessing at least one short dopamine D4 allele displayed greater activity-impulsivity in the behavioral tests than did those with two long alleles; dogs with the short allele tended to receive higher ratings on the activity-impulsivity scale of the questionnaire<sup>[76]</sup>.

Våge *et al.*<sup>[77]</sup> have used English Cocker Spaniels, a breed in which aggressive behavior has been noted for the past forty years. By using one breed, breed differences in the genotype can be eliminated so that any differences found should reflect differences in temperament. In a study comparing non-aggressive English Cocker Spaniels with English Cocker Spaniels that had bitten and broken skin, there were significant associations between aggression and four SNPs in the region of the dopamine D1 receptor (*DRD1*), two SNPs in the serotonin 1D receptor (*HTR1D*), and five SNPs in a glutamate receptor (*SLC6A1*)<sup>[77]</sup>.

The same laboratory later identified 62 SNPs occurring in or in the close vicinity of 16 neurotransmitter-related genes. Allelic associations with aggression were identified for *DRD1*, *HTR1D*, *HTR2C* (5-HT receptors D1 and 2C) and *SLC6A1* (solute carrier family 6 neurotransmitter transporter gamma amino acid member). Risk or protective haplotypes for aggressive behavior based on 2-5 SNPs were identified. The frequency of aggressive dogs varied significantly between the haplotypes within loci, and the odds ratios of aggression

in dogs with risk haplotypes compared with protective haplotypes varied from 4.4 (*HTR2C*) to 9.0 (*SLC6A1*). No haplotypes in complete association with the recorded phenotypes were identified, supporting a complex inheritance of aggression. Gene *SLC6A1* on chromosome 20 should be investigated in association with aggression in other breeds, and use of benzodiazepines which bind with gamma amino acid receptors should be investigated further as treatments for aggression<sup>[78]</sup>.

Most dogs are homozygous for the dopamine transporter-variable number tandem repeat two-tandem-repeat allele (2/2). The one-tandem-repeat allele is over-represented in American Malinois, both as heterozygotes and homozygotes (1/2 or 1/1). All American Malinois with reported seizures were 1/1 genotype. Those with at least one “1” allele (1/1 or 1/2 genotype), were more likely display hypervigilance and exhibit episodic aggression as well as more fearful postures<sup>[78]</sup>.

### Methylation

Although the genome acts as a blueprint for the production of observable morphological, physiological, and behavioral characteristics (*i.e.*, the phenotype), the expression of these traits may vary in different social or ecological contexts and in generations. Environmentally-induced phenotypic variation resulting from differential gene expression may be regulated by processes that do not include the DNA sequence itself (*i.e.*, “epigenetic mechanisms”). DNA methylation is one such epigenetic mechanism that allows organisms to respond to environmental change *via* changes in gene expression that alter the phenotype. DNA methylation during development and early life can have long-term consequences for gene expression, physiology, and behavior in many vertebrates. This is a completely uninvestigated subject in canine behavior.

## CONCLUSION

In the last ten years, the field of canine behavioral genetics has experienced rapid and exciting scientific advances, especially after completion of the sequencing of the dog genome. Although the history of dog domestication in terms of time and location is still debated, the divergence of dogs from wolves based on friendliness towards humans clearly has been outlined and experimentally repeated in the tame fox experiment. Genetic research also has focused on the great diversity of dog breeds, the genetic differences between breeds, and normal and abnormal behavioral traits. While much progress has been made in elucidating the genetics underlying aggression in dogs, future scientific studies will continue to examine this most serious problem threatening the human-canine bond and expand our knowledge about the genetic basis of canine behavior.

## REFERENCES

- 1 Belyaev DK, Plyusnina IZ, Trut LN. Domestication in the silver fox (*Vulpes fulvus* Desm): changes in physiological

- boundaries of the sensitive period of primary socialization. *Appl Anim Behav Sci* 1985; **13**: 359-370 [DOI: 10.1016/0168-1591(85)90015-2]
- 2 **Trut LN**. Early canid domestication: the farm-fox experiment. *Am Sci* 1999; **87**: 160-169
- 3 **Kukekova AV**, Trut LN, Oskina IN, Kharlamova AV, Shikhevich SG, Kirkness EF, Aguirre GD, Acland GM. A marker set for construction of a genetic map of the silver fox (*Vulpes vulpes*). *J Hered* 2002; **95**: 185-194 [PMID: 15220384 DOI: 10.1093/jhered/esh033]
- 4 **Statham MJ**, Trut LN, Sacks BN, Kharlamova AV, Oskina IN, Gulevich RG, Johnson JL, Temnykh SV, Acland GM, Kukekova AV. On the origin of a domesticated species: Identifying the parent population of Russian silver foxes (*Vulpes vulpes*). *Biol J Linn Soc Lond* 2011; **103**: 168-175 [PMID: 21625363 DOI: 10.1111/j.1095-8312.2011.01629.x]
- 5 **Trut LN**. Experimental studies of early canid domestication. In: Ruvinsky A, Sampson J, Eds. The genetics of the dog. New York: CABI, 2001: 15-43
- 6 **Gogoleva SS**, Volodin IA, Volodina EV, Kharlamova AV, Trut LN. Explosive vocal activity for attracting human attention is related to domestication in silver fox. *Behav Processes* 2011; **86**: 216-221 [PMID: 21145949 DOI: 10.1016/j.beproc.2010.12.001]
- 7 **Hare B**, Plyusnina I, Ignacio N, Schepina O, Stepika A, Wrangham R, Trut L. Social cognitive evolution in captive foxes is a correlated by-product of experimental domestication. *Curr Biol* 2005; **15**: 226-230 [PMID: 15694305 DOI: 10.1016/j.cub.2005.01.040]
- 8 **Oskina IN**, Herbeck, YE, Shikhevich SG, Plyusnina, IZ, Gulevich, RG, 2008. Alterations in the hypothalamus-pituitary-adrenal and immune systems during selection of animals for tame behavior. *VOGis Herald* 2008; **12**: 39-49 (in Russian)
- 9 **Kukekova AV**, Trut LN, Acland GM. Genetics of domesticated behavior in dogs and foxes. In: Grandin T, Deesing MJ, Eds. Genetics and the behavior of domestic animals 2nd ed. London Elsevier 2014: 361-396 [DOI: 10.1016/B978-0-12-394586-0.00010-X]
- 10 **Kukekova AV**, Trut LN, Chase K, Shepeleva DV, Vladimirova AV, Kharlamova AV, Oskina IN, Stepika A, Klebanov S, Erb HN, Acland GM. Measurement of segregating behaviors in experimental silver fox pedigrees. *Behav Genet* 2008; **38**: 185-194 [PMID: 18030612 DOI: 10.1007/s10519-007-9180-1]
- 11 **Kukekova AV**, Trut LN, Oskina IN, Johnson JL, Temnykh SV, Kharlamova AV, Shepeleva DV, Gulevich RG, Shikhevich SG, Graphodatsky AS, Aguirre GD, Acland GM. A meiotic linkage map of the silver fox, aligned and compared to the canine genome. *Genome Res* 2007; **17**: 387-399 [PMID: 17284676 DOI: 10.1101/gr.5893307]
- 12 **Lindblad-Toh K**, Wade CM, Mikkelsen TS, Karlsson EK, Jaffe DB, Kamal M, Clamp M, Chang JL, Kulbokas EJ, Zody MC, Mauceli E, Xie X, Breen M, Wayne RK, Ostrander EA, Ponting CP, Galibert F, Smith DR, DeJong PJ, Kirkness E, Alvarez P, Biagi T, Brockman W, Butler J, Chin CW, Cook A, Cuff J, Daly MJ, DeCaprio D, Gnerre S, Grabherr M, Kellis M, Kleber M, Bardeleben C, Goodstadt L, Heger A, Hitte C, Kim L, Koepfli KP, Parker HG, Pollinger JP, Searle SM, Sutter NB, Thomas R, Webber C, Baldwin J, Abebe A, Abouelleil A, Aftuck L, Ait-Zahra M, Aldredge T, Allen N, An P, Anderson S, Antoine C, Arachchi H, Aslam A, Ayotte L, Bachantsang P, Barry A, Bayul T, Benamara M, Berlin A, Bessette D, Blihshteyn B, Bloom T, Blye J, Boguslavskiy L, Bonnet C, Boukhgalter B, Brown A, Cahill P, Calixte N, Camarata J, Cheshatsang Y, Chu J, Citroen M, Collymore A, Cooke P, Dawoe T, Daza R, Decktor K, DeGray S, Dhargay N, Dooley K, Dooley K, Dorje P, Dorjee K, Dorris L, Duffey N, Dupes A, Egbiremolen O, Elong R, Falk J, Farina A, Faro S, Ferguson D, Ferreira P, Fisher S, FitzGerald M, Foley K, Foley C, Franke A, Friedrich D, Gage D, Garber M, Gearin G, Giannoukos G, Goode T, Goyette A, Graham J, Grandbois E, Gyaltsen K, Hafez N, Hagopian D, Hagos B, Hall J, Healy C, Hegarty R, Honan T, Horn A, Houde N, Hughes L, Hunnicutt L, Husby M, Jester B, Jones C, Kamat A, Kanga B, Kells C, Khazanovich D, Kieu AC, Kisner P, Kumar M, Lance K, Landers T, Lara M, Lee W, Leger JP, Lennon N, Leuper L, LeVine S, Liu J, Liu X, Lokyitsang Y, Lokyitsang T, Lui A, Macdonald J, Major J, Marabella R, Maru K, Matthews C, McDonough S, Mehta T, Meldrim J, Melnikov A, Meneus L, Mihalev A, Mihova T, Miller K, Mittelman R, Mlenga V, Mulrain L, Munson G, Navidi A, Naylor J, Nguyen T, Nguyen N, Nguyen C, Nguyen T, Nicol R, Norbu N, Norbu C, Novod N, Nyima T, Olandt P, O'Neill B, O'Neill K, Osman S, Oyono L, Patti C, Perrin D, Phunkhang P, Pierre F, Priest M, Rachupka A, Raghuraman S, Rameau R, Ray V, Raymond C, Rege F, Rise C, Rogers J, Rogov P, Sahalie J, Settipalli S, Sharpe T, Shea T, Sheehan M, Sherpa N, Shi J, Shih D, Sloan J, Smith C, Sparrow T, Stalker J, Stange-Thomann N, Stavropoulos S, Stone C, Stone S, Sykes S, Tchuinga P, Tenzing P, Tesfaye S, Thoulutsang D, Thoulutsang Y, Topham K, Topping I, Tsamla T, Vassiliev H, Venkataraman V, Vo A, Wangchuk T, Wangdi T, Weiland M, Wilkinson J, Wilson A, Yadav S, Yang S, Yang X, Young G, Yu Q, Zainoun J, Zembek L, Zimmer A, Lander ES. Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature* 2005; **438**: 803-819 [PMID: 16341006]
- 13 **Kukekova AV**, Trut LN, Chase K, Kharlamova AV, Johnson JL, Temnykh SV, Oskina IN, Gulevich RG, Vladimirova AV, Klebanov S, Shepeleva DV, Shikhevich SG, Acland GM, Lark KG. Mapping Loci for fox domestication: deconstruction/reconstruction of a behavioral phenotype. *Behav Genet* 2011; **41**: 593-606 [PMID: 21153916 DOI: 10.1007/s10519-010-9418-1]
- 14 **Wayne RK**, Ostrander EA. Lessons learned from the dog genome. *Trends Genet* 2007; **23**: 557-567 [DOI: 10.1016/j.tig.2007.08.013]
- 15 **Wayne RK**, Ostrander EA. Origin, genetic diversity, and genome structure of the domestic dog. *Bioessays* 1999; **21**: 247-57 [DOI: 10.1002/(SICI)1521-1878(199903)21:3<247::AID-BIES9>3.0.CO;2-Z]
- 16 **Vilà C**, Maldonado JE, Wayne RK. Phylogenetic relationships, evolution, and genetic diversity of the domestic dog. *J Hered* 1999; **90**: 71-77 [DOI: 10.1093/jhered/90.1.71]
- 17 **Trut LN**, Oskina IN, Kharlamova AV. Experimental studies of early canid domestication. In Ostrander EA, Ruvinsky A. The genetics of the dog, 2nd Edition. Oxfordshire, UK: CABI, 2012: 12-37
- 18 **Wang GD**, Zhai W, Yang HC, Fan RX, Cao X, Zhong L, Wang L, Liu F, Wu H, Cheng LG, Poyarkov AD, Poyarkov NA, Tang SS, Zhao WM, Gao Y, Lv XM, Irwin DM, Savolainen P, Wu CI, Zhang YP. The genomics of selection in dogs and the parallel evolution between dogs and humans. *Nat Commun* 2013; **4**: 1860 [PMID: 23673645]
- 19 **Skoglund P**, Götherström A, Jakobsson M. Estimation of population divergence times from non-overlapping genomic sequences: examples from dogs and wolves. *Mol Biol Evol* 2011; **28**: 1505-1517 [PMID: 21177316 DOI: 10.1093/molbev/msq342]
- 20 **Driscoll CA**, Macdonald DW, O'Brien SJ. From wild animals to domestic pets, an evolutionary view of domestication. *Proceedings of the National Academy of Sciences*, 2009: 9971-9978 [DOI: 10.1073/pnas.0901586106]
- 21 **Savolainen P**, Zhang YP, Luo J, Lundeberg J, Leitner T. Genetic evidence for an East Asian origin of domestic dogs. *Science* 2002; **298**: 1610-1613 [DOI: 10.1126/science.1073906]
- 22 **Vonholdt BM**, Pollinger JP, Lohmueller KE, Han E, Parker HG, Quignon P, Degenhardt JD, Boyko AR, Earl DA, Auton A, Reynolds A, Bryc K, Brisbin A, Knowles JC, Mosher DS, Spady TC, Elkhouloun A, Geffen E, Pilot M, Jedrzejewski W, Greco C, Randi E, Bannasch D, Wilton A, Shearman J, Musiani M, Cargill M, Jones PG, Qian Z, Huang W, Ding ZL, Zhang YP, Bustamante CD, Ostrander EA, Novembre J, Wayne RK. Genome-wide SNP and haplotype analyses

- reveal a rich history underlying dog domestication. *Nature* 2010; **464**: 898-902 [PMID: 20237475 DOI: 10.1038/nature08837]
- 23 **Thalmann O**, Shapiro B, Cui P, Schuenemann VJ, Sawyer SK, Greenfield DL, Germonpré MB, Sablin MV, López-Giráldez F, Domingo-Roura X, Napierala H, Uerpmann HP, Loponte DM, Acosta AA, Gianschi L, Schmitz RW, Worthington B, Buikstra JE, Druzhkova A, Graphodatsky AS, Ovodov ND, Wahlberg N, Freedman AH, Schweizer RM, Koepfli KP, Leonard JA, Meyer M, Krause J, Pääbo S, Green RE, Wayne RK. 2013 Complete mitochondrial genomes of ancient canids suggest a European origin of domestic dogs. *Science* 2013; **342**: 871-874 [DOI: 10.1126/science.1243650]
- 24 **Saetre P**, Lindberg J, Leonard JA, Olsson K, Pettersson U, Ellegren H, Bergström TF, Vilà C, Jazin E. From wild wolf to domestic dog: gene expression changes in the brain. *Brain Res Mol Brain Res* 2004; **126**: 198-206 [PMID: 15249144]
- 25 **Björnerfeldt S**, Webster MT, Vilà C. Relaxation of selective constraint on dog mitochondrial DNA. *Genome Res* 2006; **16**: 990-994 [DOI: 10.1101/gr.5117706]
- 26 **Cruz F**, Vilà C, Webster MT. The legacy of domestication: accumulation of deleterious mutations in the dog genome. *Mol Biol Evol* 2008; **25**: 2331-2336 [DOI: 10.1093/molbev/msn177]
- 27 **Li Y**, Vonholdt BM, Reynolds A, Boyko AR, Wayne RK, Wu DD, Zhang YP. Artificial selection on brain-expressed genes during the domestication of dog. *Mol Biol Evol* 2013; **30**: 1867-1876 [DOI: 10.1093/molbev/mst088]
- 28 **Sabeti PC**, Reich DE, Higgins JM, Levine HZ, Richter DJ, Schaffner SF, Gabriel SB, Platko JV, Patterson NJ, McDonald GJ, Ackerman HC, Campbell SJ, Altshuler D, Cooper R, Kwiatkowski D, Ward R, Lander ES. Detecting recent positive selection in the human genome from haplotype structure. *Nature* 2002; **419**: 832-837 [PMID: 12397357 DOI: 10.1038/nature01140]
- 29 **Axelsson E**, Ratnakumar A, Arendt ML, Maqbool K, Webster MT, Perloski M, Liberg O, Arnemo JM, Hedhammar A, Lindblad-Toh K. The genomic signature of dog domestication reveals adaptation to a starch-rich diet. *Nature* 2013; **495**: 360-364 [PMID: 23354050 DOI: 10.1038/nature11837]
- 30 **Turcsan B**, Kubinyi E, Miklosi A. Trainability and boldness traits differ between dog breed clusters based on conventional breed categories and genetic relatedness. *Appl Anim Behav Sci* 2011; **(132)**: 61-70 [DOI: 10.1016/j.applanim.2011.03.006]
- 31 **Houpt KA**. Genetics of canine behavior. *Acta Vet Brno* 2007; **76**: 431-444 [DOI: 10.2754/avb200776030431]
- 32 **Galibert F**, André C. The dog: A powerful model for studying genotype-phenotype relationships. *Comp Biochem Physiol Part D Genomics Proteomics* 2008; **3**: 67-77 [PMID: 20483208 DOI: 10.1016/j.cbpd.2007.06.001]
- 33 **McGreevy PD**, Georgievsky D, Carrasco J, Valenzuela M, Duffy DL, Serpell JA. Dog behavior co-varies with height, bodyweight and skull shape. *PLoS One* 2013; **8**: e80529 [PMID: 24358107 DOI: 10.1371/journal.pone.0080529]
- 34 **Spady TC**, Ostrander EA. Canine behavioral genetics: pointing out the phenotypes and herding up the genes. *Am J Hum Genet* 2008; **82**: 10-18 [DOI: 10.1016/j.ajhg.2007.12.001]
- 35 **Ostrander EA**, Ruvinsky A. The genetics of the Dog 2nd ed. Wallingford UK: CABI, 2012: 275-320
- 36 **Parker HG**. Genomic analyses of modern dog breeds. *Mamm Genome* 2012; **23**: 19-27 [PMID: 22231497 DOI: 10.1007/s00335-011-9387-6]
- 37 **Parker HG**, Kim LV, Sutter NB, Carlson S, Lorentzen TD, Malek TB, Johnson GS, DeFrance HB, Ostrander EA, Kruglyak L. Genetic structure of the purebred domestic dog. *Science* 2004; **304**: 1160-1164 [PMID: 15155949 DOI: 10.1126/science.1097406]
- 38 **Boyko AR**, Quignon P, Li L, Schoenebeck JJ, Degenhardt JD, Lohmueller KE, Zhao K, Brisbin A, Parker HG, vonHoldt BM, Cargill M, Auton A, Reynolds A, Elkhouloun AG, Castel-hano M, Mosher DS, Sutter NB, Johnson GS, Novembre J, Hubisz MJ, Siepel A, Wayne RK, Bustamante CD, Ostrander EA. A simple genetic architecture underlies morphological variation in dogs. *PLoS Biol* 2010; **8**: e1000451 [PMID: 20711490 DOI: 10.1371/journal.pbio.1000451]
- 39 **Scott JP**, Fuller JL, 1965. Genetics and the social behavior of the dog. Chicago: The University of Chicago Press, 1965
- 40 **Hall NJ**, Wynne CD. The canid genome: behavioral geneticists' best friend? *Genes Brain Behav* 2012; Epub ahead of print [PMID: 22979960 DOI: 10.1111/j.1601-183X.2012.00851]
- 41 **Saetre P**, Strandberg E, Sundgren PE, Pettersson U, Jazin E, Bergström TF. The genetic contribution to canine personality. *Genes Brain Behav* 2006; **5**: 240-248 [DOI: 10.1111/j.1601-183X.2005.001]
- 42 **Svartberg K**. Breed-typical behaviour in dogs- Historical remnants or recent constructs? *Appl Ani Behav Sci* 2006; **96**: 293-313 [DOI: 10.1016/j.applanim.2005.06.014]
- 43 **Meyer F**, Schawalder P, Gaillard C, Dolf G. Estimation of genetic parameters for behavior based on results of German Shepherd Dogs in Switzerland. *Appl Ani Behav Sci* 2012; **140**: 53- 61 [DOI: 10.1016/j.applanim.2012.05.007]
- 44 **Mehrkam LR**, Wynne CDL. Behavioral differences among breeds of domestic dogs (*Canis lupus familiaris*): Current status of the science. *Appl Anim Behav Sci* 2014; **155**: 12-27 [DOI: 10.1016/j.applanim.2014.03.005]
- 45 **Huson HJ**, Parker HG, Runstadler J, Ostrander EA. A genetic dissection of breed composition and performance enhancement in the Alaskan sled dog. *BMC Genet* 2010; **11**: 71 [PMID: 20649949]
- 46 **Huson HJ**, vonHoldt BM, Rimbault M, Byers AM, Runstadler JA, Parker HG, Ostrander EA. Breed-specific ancestry studies and genome-wide association analysis highlight an association between the MYH9 gene and heat tolerance in Alaskan sprint racing sled dogs. *Mamm Genome* 2012; **23**: 178-94 [DOI: 10.1007/s00335-011-9374-y]
- 47 **Takeuchi Y**, Hashizume C, Chon EM, Momozawa Y, Masuda K, Kikusui T, Mori Y. Canine tyrosine hydroxylase (TH) gene and dopamine beta-hydroxylase (DBH) gene: their sequences, genetic polymorphisms, and diversities among five different dog breeds. *J Vet Med Sci* 2005; **67**: 861-867 [PMID: 16210796 DOI: 10.1292/jvms.67.861]
- 48 **Ogata N**, Hashizume C, Momozawa Y, Masuda K, Kikusui T, Takeuchi Y, Mori Y. Polymorphisms in the canine glutamate transporter-1 gene: identification and variation among five dog breeds. *J Vet Med Sci* 2006; **68**: 157-159 [PMID: 16520538 DOI: 10.1292/jvms.68.157]
- 49 **Hart BL**, Hart LA. Selecting pet dogs on the basis of cluster analysis of breed behavior profiles and gender. *J Am Vet Med Assoc* 1985; **186**: 1181-1185 [PMID: 4008297]
- 50 **Yokoyama JS**, Hamilton SP. Genetics of Canine Behavioral disorders In: Ostrander E Ruvinsky A. The Genetics of the Dog 2nd ed. Wallingford UK: CABI, 2012: 275-320 [DOI: 10.1079/9781845939403.0275]
- 51 **Dodman NH**, Karlsson EK, Moon-Fanelli A, Galdzicka M, Perloski M, Shuster L, Lindblad-Toh K, Ginns EI. A canine chromosome 7 locus confers compulsive disorder susceptibility. *Mol Psychiatry* 2010; **15**: 8-10 [PMID: 20029408 DOI: 10.1038/mp.2009.111]
- 52 **Ogata N**, Gillis TE, Liu X, Cunningham SM, Lowen SB, Adams BL, Sutherland-Smith J, Mintzopoulos D, Janes AC, Dodman NH, Kaufman MJ. Brain structural abnormalities in Doberman pinschers with canine compulsive disorder. *Prog Neuropsychopharmacol Biol Psychiatry* 2013; **45**: 1-6 [PMID: 23590875 DOI: 10.1016/j.pnpbp.2013.04.002]
- 53 **Vermeire S**, Audenaert K, De Meester R, Vandermeulen E, Waelbers T, De Spiegeleer B, Eersels J, Dobbeleir A, Peremans K. Serotonin 2A receptor, serotonin transporter and dopamine transporter alterations in dogs with compulsive behaviour as a promising model for human obsessive-compulsive disorder. *Psychiatry Res* 2012; **201**: 78-87 [PMID: 22285716 DOI: 10.1016/j.psychres.2011.06.006]



- 54 **Horwitz DE**, Neilson JC. Blackwell's five-minute veterinary consult clinical companion-canine and feline behavior. Ames Iowa Blackwell Publ, 2007: 482-483
- 55 **Tang R**, Noh HJ, Wang D, Sigurdsson S, Swofford R, Perloski M, Karlsson EK. Candidate genes and functional non-coding variants identified in a canine model of obsessive-compulsive disorder. *Genome Biology* 2014; **15**: R25
- 56 **Bamberger M**, Houpt KA. Signalment factors, comorbidity, and trends in behavior diagnoses in dogs: 1,644 cases (1991-2001). *J Am Vet Med Assoc* 2006; **229**: 1591-1601 [PMID: 17107314 DOI: 10.2460/javma.229.10.1591]
- 57 **Dodman NH**, Donnelly R, Shuster L, Mertens P, Rand W, Miczek K. Use of fluoxetine to treat dominance aggression in dogs. *J Am Vet Med Assoc* 1996; **209**: 1585-1587 [PMID: 8899022]
- 58 **Cakiroglu D**, Meral Y, Sancak AA, Cifti G. Relationship between the serum concentrations of serotonin and lipids and aggression in dogs. *Vet Rec* 2007; **161**: 59-61 [PMID: 17630419 DOI: 10.1136/vr.161.2.59]
- 59 **Leon M**, Rosado B, Garcia-Belenguer S, Chacon C, Villegas A, Palacio J. Assessment of serotonin in serum, plasma, and platelets of aggressive dogs. *J Vet Behav* 2012; **7**: 348-352 [DOI: 10.1016/j.jveb.2012.01.005]
- 60 **Podberscek AL**, Serpell JA. The English Cocker Spaniel: preliminary findings on aggressive behaviour. *Appl Anim Behav Sci* 1996; **47**: 75-89 [DOI: 10.1016/0168-1591(95)01012-2]
- 61 **Kaelin CB**, Barsh GS. Molecular Genetics of coat colour, texture and length in the dog. In Ostrander EA, Ruvinsky A. The genetics of the dog, 2nd Edition. Oxfordshire, UK: CABI, 2012: 57-7737
- 62 **Amat M**, Le Brech S, Camps T, Torrente C, Mariotti VM, José L. Ruiz JL, Manteca X. Differences in serotonin levels between aggressive English cocker spaniels and aggressive dogs of other breeds. *Jour Vet Behav* 2010; **5**: 46 [DOI: 10.1016/j.jveb.2009.08.006]
- 63 **Lesch KP**, Bengel D, Heils A, Sabol SZ, Greenberg BD, Petri S, Benjamin J, Müller CR, Hamer DH, Murphy DL. Association of anxiety-related traits with a polymorphism in the serotonin transporter gene regulatory region. *Science* 1996; **274**: 1527-1531 [PMID: 8929413 DOI: 10.1126/science.274.5292.1527]
- 64 **Reisner IR**, Mann JJ, Stanley M, Huang YY, Houpt KA. Comparison of cerebrospinal fluid monoamine metabolite levels in dominant-aggressive and non-aggressive dogs. *Brain Res* 1996; **714**: 57-64 [PMID: 8861609 DOI: 10.1016/0006-8993(95)014640]
- 65 **Jacobs C**, Van Den Broeck W, Simoens P. Neurons expressing serotonin-1B receptor in the basolateral nuclear group of the amygdala in normally behaving and aggressive dogs. *Brain Res* 2007; **1136**: 102-109 [PMID: 17223093 DOI: 10.1016/j.brainres.2006.04.099]
- 66 **Shaikh MB**, Steinberg A, Siegel A. Evidence that substance P is utilized in medial amygdaloid facilitation of defensive rage behavior in the cat. *Brain Res* 1993; **625**: 283-294 [PMID: 7506110 DOI: 10.1016/0006-8993(93)91070-9]
- 67 **De Felipe C**, Herrero JF, O'Brien JA, Palmer JA, Doyle CA, Smith AJ, Laird JM, Belmonte C, Cervero F, Hunt SP. Altered nociception, analgesia and aggression in mice lacking the receptor for substance P. *Nature* 1998; **392**: 394-397 [PMID: 9537323 DOI: 10.1038/32904]
- 68 **Vermeire S**, Audenaert K, De Meester R, Vandermeulen E, Waelbers T, De Spiegeleer B, Eersels J, Dobbeleir A, Peremans K. Neuro-imaging the serotonin 2A receptor as a valid biomarker for canine behavioural disorders. *Res Vet Sci* 2011; **91**: 465-472 [PMID: 21051063 DOI: 10.1016/j.rvsc.2010.09.021]
- 69 **Badino P**, Odore R, Osella MC, Bergamasco L, Francione P, Girardi C, Re G. Modifications of serotonergic and adrenergic receptor concentrations in the brain of aggressive Canis familiaris. *Comp Biochem Physiol A Mol Integr Physiol* 2004; **139**: 343-350 [PMID: 15556390 DOI: 10.1016/j.chpb]
- 70 **Pérez-Guisado J**, Lopez-Rodriguez R, Munoz-Serrano A. Heritability of dominant-aggressive behaviour in English Cocker Spaniels. *Appl Anim Behav Sci* 2006; **100**: 219-227 [DOI: 10.1016/j.applanim.2005.11.005]
- 71 **Linamo AE**, van den Berg L, Leegwater PAF, Schiklder MBH, von Arendonk JAM, van Oost BA. Genetic variation in genetic related traits in Golden Retriever dogs. *Appl Anim Behav Sci* 2007; **104**: 95-106 [DOI: 10.1016/j.applanim.2006.04.025]
- 72 **Hsu Y**, Serpell JA. Development and validation of a questionnaire for measuring behavior and temperament traits in pet dogs. *J Am Vet Med Assoc* 2003; **223**: 1293-1300 [PMID: 14621216 DOI: 10.2460/javma.2003.223.1293]
- 73 **van den Berg L**, Vos-Loohuis M, Schilder MBH, van Oost BA, Hazewinkel HAW, Wade CM, Karlsson K, Lindblad-Toh Liinamo AE, Leegwater PAG. Evaluation of the serotonergic genes htr1A, htr1B, htr2A, and slc6A4 in aggressive behavior of Golden Retriever dogs. *Behav Genet* 2008; **38**: 55-66 [PMID: 18066658 DOI: 10.1007/s10519-007-9179-7]
- 74 **Hejjas K**, Vas J, Topal J, Szantai E, Ronai Z, Szekely A, Kubinyi E, Horvath Z, Sasvari-Szekely M, Miklosi A. Association of polymorphisms in the dopamine D4 receptor gene and the activity-impulsivity endophenotype in dogs. *Anim Genet* 2007; **38**: 629-633 [PMID: 17986156 DOI: 10.1111/j.1365-2052.2007.01657.x]
- 75 **Kubinyi E**, Vas J, Hejjas K, Ronai Z, Brúder I, Turcsán B, Sasvari-Szekely M, Miklósi A. Polymorphism in the tyrosine hydroxylase (TH) gene is associated with activity-impulsivity in German Shepherd Dogs. *PLoS One* 2012; **7**: e30271 [PMID: 22272320 DOI: 10.1371/journal.pone.0030271]
- 76 **Wan M**, Hejjas K, Ronai Z, Elek Z, Sasvari-Szekely M, Champagne FA, Miklósi A, Kubinyi E. DRD4 and TH gene polymorphisms are associated with activity, impulsivity and inattention in Siberian Husky dogs. *Anim Genet* 2013; **44**: 717-727 [PMID: 23713429 DOI: 10.1111/age.12058]
- 77 **Våge J**, Bønsdorff TB, Arnet E, Tverdal A, Lingaas F. Differential gene expression in brain tissues of aggressive and non-aggressive dogs. *BMC Vet Res* 2010; **6**: 34 [PMID: 20553618]
- 78 **Lit L**, Belanger JM, Boehm D, Lybarger N, Haverbeke A, Diederich C, Oberbauer AM. Characterization of a dopamine transporter polymorphism and behavior in Belgian Malinois. *BMC Genet* 2013; **14**: 45 [PMID: 23718893 DOI: 10.1186/1471-2156-14-45]

**P- Reviewer:** Clark LA, Ekenstedt KJ **S- Editor:** Ji FF

**L- Editor:** A **E- Editor:** Liu SQ





## Molecular genetics of gastric adenocarcinoma in clinical practice

Margaret Cho, Ogechukwu Eze, Ruliang Xu

Margaret Cho, Ogechukwu Eze, Ruliang Xu, Department of Pathology, New York University Langone Medical Center, New York, NY 10016, United States

Author contributions: Cho M and Eze O searched for literatures and wrote portions of manuscript; Xu R wrote and edited the majority of manuscript.

Correspondence to: Ruliang Xu, MD, PhD, Department of Pathology, New York University Langone Medical Center, 550 First Avenue, New York, NY 10016, United States. [ruliang.xu@nyumc.org](mailto:ruliang.xu@nyumc.org)

Telephone: +1-212-2630728 Fax: +1-212-2637916

Received: December 24, 2013 Revised: April 16, 2014

Accepted: May 15, 2014

Published online: August 27, 2014

### Abstract

The molecular genetics of gastric carcinoma (GC) dictates their biology and clinical behavior. The two morphologically distinct types of gastric carcinoma by Lauren classification, *i.e.*, intestinal and diffuse cell types, have a significant difference in clinical outcome. These two types of GC have different molecular pathogenetic pathways with unique genetic alterations. In addition to environmental and other etiologies, intestinal type GC is associated with *Helicobacter pylori* (*H. pylori*) infection and involves a multistep molecular pathway driving the normal epithelium to intestinal metaplasia, dysplasia, and malignant transformation by chromosomal and/or microsatellite instability (MSI), mutation of tumor suppressor genes, and loss of heterozygosity among others. Diffuse type shows no clear causal relationship with *H. pylori* infection, but is commonly associated with deficiency of cell-cell adhesion due to mutation of the E-cadherin gene (*CDH1*), and a manifestation of the hereditary gastric cancer syndrome. Thus, detection of *CDH1* mutation or loss of expression of E-cadherin may aid in early diagnosis or screening of diffuse type GC. Detection of certain genetic markers, for example, MSI and matrix metalloproteinases, may

provide prognostic information, particularly for intestinal type. The common genetic alterations may offer therapeutic targets for treatment of GC. Polymorphisms in Thymidylate synthase to metabolize 5-fluorouracil, glutathione S-transferase for degradation of Cisplatin, and amplification/overexpression of human epidermal growth factor receptor 2 targeted by monoclonal antibody Trastuzumab, are a few examples. P13K/Akt/mTOR pathway, c-Met pathways, epidermal growth factor receptor, insulin-like growth factor receptor, vascular endothelial growth factor receptor fibroblast growth factor receptor, and micro RNAs are several potential therapeutic biomarkers for GC under investigation.

© 2014 Baishideng Publishing Group Inc. All rights reserved.

**Key words:** Molecular genetics; Lauren classification; Intestinal type gastric cancer; Diffuse type gastric cancer; Molecular Biomarker

**Core tip:** Intestinal and diffuse cell types of gastric carcinoma have a significant difference in clinical outcome with different molecular pathogenetic pathways. Intestinal type gastric carcinoma (GC) is associated with chromosomal and/or microsatellite instability, mutation of tumor suppressor genes, and loss of heterozygosity. Diffuse type GC is commonly associated with mutation of the E-cadherin gene, and a manifestation of the hereditary gastric cancer syndrome. Detection of certain mutations may aid in early diagnosis, screening, and prognostication of GC, and common genetic alterations may offer therapeutic targets for treatment. Furthermore, potential therapeutic biomarkers for GC are under investigation and may hold future promise.

Cho M, Eze O, Xu R. Molecular genetics of gastric adenocarcinoma in clinical practice. *World J Med Genet* 2014; 4(3): 58-68 Available from: URL: <http://www.wjgnet.com/2220-3184/full/v4/i3/58.htm> DOI: <http://dx.doi.org/10.5496/wjmg.v4.i3.58>

## INTRODUCTION

Gastric carcinoma (GC) is the second major leading cause of cancer-related death and fourth most common cancer worldwide<sup>[1]</sup>. The relatively unfavorable outcome is largely attributable to complex biology and marginal effectiveness of treatment options, including surgical resection, chemotherapy, and multidisciplinary approach. Even with use of multimodality approaches, overall survival continues to be poor with 30%-36% 5-year survival rates<sup>[2,3]</sup>. Chemotherapy is the main treatment in cases of metastatic disease, and the median survival time is only 9 to 14 mo<sup>[3,4]</sup>. However, not all of GCs have the same outcomes. The biological behavior and clinical presentation of GCs differ with their histological and molecular features.

The current World Health Organization classification (2010 edition), classifies GC into many different types based upon the histology combined with molecular genetic information. However, the traditional Lauren classification, purely on histologic basis, is the most commonly used system. It classifies GC into three different groups: (1) intestinal type with glandular differentiation or pattern; (2) diffuse, or poorly differentiated type, including a signet ring cell histology; and (3) mixed or indeterminate type<sup>[5]</sup>. The first two types of GCs (intestinal and diffuse types) have distinct histogenesis as well as clinical characteristics. Recent data suggest that these two groups largely differ in their molecular genetics. This paper will mainly review the molecular characteristics of non-hereditary intestinal and diffuse types of GC to understand the molecular pathways involved in GC development and to identify molecular targets for diagnosis, therapy, and prognostication.

## MOLECULAR BASIS OF PATHOGENESIS

### Genetic predisposition of GC

Certain genetic polymorphisms are predisposed to an increased risk for gastric cancer. These polymorphisms were found in genes involved in the inflammatory response to *Helicobacter pylori* (*H. pylori*) infection<sup>[6-8]</sup>, prevention of DNA to oxidative damage<sup>[9]</sup>, and mucosal protection against *H. pylori* infection<sup>[10-12]</sup>, and detoxification<sup>[13,14]</sup>. Polymorphisms of the interleukin 1 (*IL-1β*) gene consistently show strong association with GC<sup>[15]</sup>. The association is also seen with other genes, including IL-1 receptor antagonist genes<sup>[15-19]</sup>, tumor necrosis factor-α gene<sup>[18,20,21]</sup>, rs11556218 T/G polymorphism of the *IL-16* gene<sup>[8]</sup>, and genes encoding glutathione-S-transferase (GST) (GSTT1 and GSTM1)<sup>[22,23]</sup>. Many hereditary tumor syndromes increase the risk to develop GC. The high risk association is well-documented in Hereditary diffuse-type gastric cancer syndrome, familial adenomatous polyposis, hereditary nonpolyposis colorectal cancer, Peutz-Jeghers syndrome, Juvenile polyposis and Li-Fraumeni syndrome<sup>[24,25]</sup>.

### Molecular pathogenesis of intestinal cell type of GC

The major etiology of the intestinal type includes dietary,

environmental factors, and *H. pylori* infection<sup>[26]</sup>. There are many good reviews that have discussed the role and significance of dietary habits and environmental factors in gastric carcinogenesis, but *H. pylori* infection has also been shown to emerge as an important carcinogen in the stomach. The bacterial virulence factors of *H. pylori* contributing to GC risk include vacuolating cytotoxin A (*vacA*), blood group antigen binding adhesion 2, outer inflammatory protein, and cytotoxin-associated gene product (*cagA*) genes<sup>[27,28]</sup>. Infection with a *cagA*-positive *H. pylori* strain in comparison with a *cagA*-negative strain increases the risk for development of GC<sup>[29]</sup>. *CagA* is translocated into host cells and induces a growth factor-like response in gastric epithelial cells by forming a physical complex with the Src homology 2 domain-containing tyrosine phosphatase in a phosphorylation dependent manner<sup>[30,31]</sup>. In addition, aberrant expression of activation-induced cytidine deaminase, a gene originally linked to immunoglobulin class switching and B lymphocyte hypermutation, results in accumulation of mutations in the p53 tumor suppressor gene<sup>[31,32]</sup>. A second virulence gene, the *vacA*, induces gastric epithelial cell apoptosis and interferes with T cell activation which suppresses local immune response<sup>[33]</sup>. Chronic inflammation also causes genetic instability through the generation of reactive oxygen and nitrogen species which can directly damage the genomic and mitochondrial DNA<sup>[31,34]</sup>.

The development of GC is a multi-step process. Chronic or atrophic gastritis may lead to intestinal metaplasia, subsequently dysplasia, and eventually carcinoma in some patients<sup>[26,35]</sup>. During the above progression, a series of genetic alterations occur. The inactivation of tumor suppressor gene p53 is involved in early carcinogenesis, and is found in 38% of intestinal metaplasia, 58% of dysplasia, and 38%-71% of GC<sup>[36-38]</sup>. Mutations occur more commonly in CpG sites of p53 and transition of G: C to A:T at these sites is the most common type of mutation irrespective of the histologic type of GC<sup>[39]</sup>. Mutation of p73, a member of the p53 family, is indicated in the carcinogenesis of GC associated with *H. pylori* infection in the mouse model<sup>[40]</sup>.

Chromosomal instability in intestinal-type GC includes gains at 8q, 17q, 20q and losses at 3p and 5q<sup>[28,41,42]</sup>. Microsatellite instability-high is often seen in intestinal type of GC, largely due to epigenetic effect, (*i.e.*, hypermethylation of the promoter regions of mismatch repair genes, most commonly mutL homolog 1 and mutS homolog 2, and in small percentage of cases, gene mutations<sup>[43-45]</sup>). The CpG island methylator phenotype was found in 24%-47% of GC, similar to colorectal cancer<sup>[43,46-48]</sup>. Recent studies have conducted a genome-wide search to identify novel methylation-silenced genes in GC<sup>[49,50]</sup>. GC cell lines were treated with a demethylating and/or deacetylating agent and were screened for epigenetically silenced genes using oligonucleotide microarrays<sup>[49]</sup>. The gene encoding serine proteases inhibitor Tissue Factor Pathway Inhibitor 2 was found to be highly methylated (81%) in GC, and its methylation was a significant and independent prognostic indicator in GC<sup>[51,49]</sup>.

Loss of heterozygosity (LOH) or mutation of Adenomatous polyposis coli (*APC*) gene may be found in approximately 25% of the cancer precursors, adenomas, and in up to 60% of intestinal-type GC<sup>[51-53]</sup>. Mutation of *CTNNB1*, which encodes  $\beta$ -catenin, appears to be exclusive to the mutations that inactivate APC protein<sup>[54]</sup>.  $\beta$ -catenin accumulates in the cytoplasm, binds to members of the Tcf/Lef family of transcription factors, and is translocated to the nucleus where the Tcf/ $\beta$ -catenin complex activates target genes such as MYC and cyclin D1 gene<sup>[51,54]</sup>. The incidence of *CTNNB1* mutations in intestinal- vs diffuse-type GC remains unclear. One study reported no mutations in diffuse-type GC but 27% incidence in intestinal-type GC<sup>[55]</sup>. Clements *et al*<sup>[56]</sup> 2002 found that 26% of tumors with  $\beta$ -catenin nuclear staining contained *CTNNB1* mutations, with no difference between diffuse- and intestinal-type GC.

LOH at the bcl-2 locus and amplification of cyclin D1 and E genes are also associated with intestinal-type GCs<sup>[57,58]</sup>. The oncogene ErbB2 (Her-2/neu) is amplified in approximately 20% of intestinal type GCs<sup>[59]</sup>. E-cadherin gene (*CDH1*) mutations have an insignificant association with the development of intestinal-type GC, in contrast to diffuse type GC<sup>[60,61]</sup>.

RUNX3 is now accepted as a tumor suppressor gene and first reported in gastric epithelial cells of RUNX3 knockout mice in 2002<sup>[62]</sup>. Approximately 45%-60% of human GCs show loss of RUNX3 expression due to hemizygous deletion and hypermethylation of the promoter region, and RUNX3 hypermethylation is seen in *H. pylori* infection, intestinal metaplasia, and gastric adenoma<sup>[31,62,63]</sup>. In response to transforming growth factor (TGF)- $\beta$ , RUNX3 inhibits gastric epithelial proliferation by inducing the CDKN1A (*p21*) gene<sup>[64]</sup> and also upregulates the expression of proapoptotic gene BCL2L11 (Bim) in gastric cancer cells treated with TGF- $\beta$ <sup>[65]</sup>. Restoration of RUNX3 also strongly inhibited peritoneal metastases of GC in an animal model<sup>[66]</sup>. RUNX3 inhibited the expression of vascular endothelial growth factor A (VEGF-A) and suppressed angiogenesis and metastasis of GCs<sup>[67]</sup>.

### Molecular pathogenesis of diffuse type GC

Little is known about the etiology of diffuse type GC. Epidemiological studies did not link *H. pylori* infection to diffuse type of GC. Its association with hereditary gastric cancer predisposition syndrome is well-documented<sup>[68]</sup>. The unique molecular genetics of this type of GC, in contrast to intestinal-type, is deficiency of the cell-cell adhesion due to genetic or epigenetic inactivation/down regulation of E-cadherin gene (*CDH1*). Approximately 50% of diffuse-type GC harbor this mutation or inactivation<sup>[24,69]</sup>. The abnormality of *CDH1* gene can be found in early stage of diffuse GC development and loss of E-cadherin expression is seen in invasive and *in situ* carcinomas<sup>[70]</sup>. In a model proposed by Carneiro *et al*<sup>[70]</sup>, the development of diffuse GC in E-cadherin mutation carriers encompasses *in situ* signet ring carcinoma with pagetoid spread of signet ring cells as pre-invasive lesions.

In early hereditary GC, the wild-type *CDH1* allele is suppressed or lost in tumor cells with a second hit caused by promoter hypermethylation of *CDH1* in at least 50% of cases<sup>[71]</sup>. Promoter methylation is also part of the major mechanism underlying E-cadherin downregulation in sporadic diffuse gastric cancers<sup>[60]</sup>. Chromosomal instability in diffuse-type GC include gains at 12q, 13q and losses at 4q, 15q, 16q, and 17p<sup>[28,42,72,73]</sup>. The diffuse type GC is also associated with the alterations or mutations in other genes or gene products, including the met proto-oncogene encoding the hepatocyte growth factor receptor and the SC-1 antigen (an apoptosis receptor)<sup>[74-76]</sup>.

### Molecular pathogenesis of Epstein-Barr virus-associated GC

Five percent of GC is associated with monoclonal proliferation of Epstein-Barr virus (EBV)-infected epithelial cells and is a specific clinicopathologic subset with characteristics of younger age, male predominance, proximal location, lower rate of lymph node involvement, marked lymphocytic infiltration, and lace pattern within the mucosa<sup>[31,77]</sup>. EBV maintains its latent infection and expresses viral latent genes which include EBV-determined nuclear antigen 1, EBV-encoded small RNA, latent membrane protein 2A (LMP2A) and Bam H1-A rightward transcripts (BARTs)<sup>[31,78]</sup>. Frequent loss of p16 (CDKN2A), smad4, Fhit, and CD82 (KAI-1) are seen<sup>[79]</sup>. Global CpG island methylation in the PTEN promoter region is considered as a characteristic abnormality in EBV-associated GC<sup>[80]</sup> with viral LMP2A responsible for aberrant hypermethylation by activation of host DNA methyltransferase 1<sup>[81]</sup>. LMP2A also upregulates Birc5 (survivin) expression through the activation of nuclear factor- $\kappa$ B, activates extracellular signal regulated kinases (ERK/MAPK1), and inhibits TGF- $\beta$ -induced apoptosis through activation of the Ras/PI3K/Akt pathway<sup>[31,82-84]</sup>.

## GENETIC CHANGES ASSOCIATED WITH MUCIN PHENOTYPIC EXPRESSION IN GC

GC can be classified into four phenotypes according to mucin (MUC1, MUC2 and CD10) expression: gastric or foveolar phenotype (G-type), intestinal phenotype (I-type), intestinal and gastric mixed phenotype and neither gastric nor intestinal phenotype<sup>[31,85]</sup>. Genetic changes associated with mucin phenotypic expression in GC include *TP53* mutations in I-type GC and microsatellite instability in G-type. Specific epigenetic alterations include methylation of hMLH1 occurring more frequently in MUC2-negative GC and more frequently methylated MGMT in MUC2-positive GC than in MUC2-negative GC<sup>[31,85]</sup>.

## MOLECULAR DIAGNOSIS OF GC

Genetic markers associated with the development of GC are numerous. However, very few have diagnostic utility. *CDH1* probably is the best candidate marker for such purpose. About 50 of diffuse type of GC have *CDH1*



mutation, either complete or partial deletions of exons, in more than 70% of somatic E-cadherin mutations<sup>[24,68]</sup>. This unique gene alteration may have a diagnostic potential. *CDH1* mutations can be detected by polymerase chain reaction on paraffin-embedded tissue. Detection of the germline mutation in *CDH1* may help identify asymptomatic mutations carriers of hereditary gastric cancer syndrome and provide molecular basis for prophylactic total gastrectomy<sup>[86]</sup>.

### Molecular prognostication

Currently clinical stage is considered to be the gold standard to predict clinical behavior and the most valuable prognostic factor for all GC types. However, clinical stage does not address the issue of tumor heterogeneity. Many studies have investigated molecular biomarkers as alternatives or supplements to the current staging system. There is some success in identifying biomarkers potentially useful in predicting prognosis and therapeutic response.

Microsatellite instability-high is commonly seen in GC located in the distal stomach or antrum, usually intestinal type. It is less frequently associated with metastasis to local lymph nodes<sup>[45,87,88]</sup>. However, it is still controversial whether patients with MSI-H GC have a favorable long term survival than those with microsatellite instability-low or microsatellite stability GCs<sup>[89,90]</sup>.

Overexpression of matrix metalloproteinases (MMPs) is shown to be related to tumor invasiveness and metastasis<sup>[91]</sup>. MMP-1-overexpression in GCs has a worse prognosis than tumors without MMP overexpression. VEGF overexpression is associated with shorter survival time attributable to its enhancement of tumor angiogenesis. Amplification or overexpression of cyclin E is correlated with aggressiveness<sup>[92]</sup>. Amplification/overexpression of the *ERBB2* (Her-2/neu) oncogene in general is considered to be an independent, poor prognostic factor<sup>[93,94]</sup>. Overexpression of EGF-R and abnormal expression of E-cadherin and  $\beta$ -catenin decrease survival or have poor prognosis<sup>[95-97]</sup>. Abnormal gene expression of *IGF2*, *KIAA1093*, *OCT2*, *PCOLCE*, *PFN2*, *RBP4*, and three genes (*BIK*, *Aurora kinase B* and *EIF5A2*) identified in the primary tumor are related to node metastasis<sup>[98,99]</sup>. Expression of caudal type homeobox transcription factor 2 (*CDX2*) and combination of normal expression of E-cadherin and negative expression of the transmembrane protein *MUC1* predict a better prognosis for patients with GC<sup>[100,101]</sup>. Down-regulation of a cyclin dependent kinase inhibitor, *P27/Kip1*, is a negative prognostic factor<sup>[102,103]</sup>. Loss of expression of tumor suppressor gene *Rb* is related to worse overall survival or inversely correlates with tumor invasion<sup>[104,105]</sup>. Mutation or abnormal expression of *p53* may have a reduced cumulative survival, lymph node metastasis, and lower chemosensitivity<sup>[100,106,107]</sup>, but its overall prognostic significance is controversial<sup>[108]</sup>. Protection of telomere expression levels are also higher in advanced GC<sup>[28]</sup>.

### Molecular therapeutic predictors

The management of patients with GC, particularly those

in late stage of tumors, usually requires chemotherapy or target therapy as single or one of the components of combined modality. The chemotherapy or target therapy is toxic, and the effectiveness is variable with patients. Molecular biomarkers have been proven to be a useful tool to predicate therapeutic response and may be used clinically to select patients or chemotherapeutic regimen for optimal result.

**Predictors for Fluorouracil treatment:** (1) Thymidylate synthase (TYMS) is a catabolizing enzyme for fluorouracil (5-FU). Polymorphisms in the gene encoding TYMS affect expression and appear to be associated with poorer response with 5-FU (47 marker) levels<sup>[109,110]</sup>. A specific polymorphism in the 5'-untranslated regions is correlated with low sensitivity to 5-FU based chemotherapy and decreased survival in a retrospective study<sup>[111]</sup>; (2) Dihydropyrimidine dehydrogenase (DPD) and thymidine phosphorylase (TP) are two regulatory enzymes involved in the degradation of 5-FU. Low levels of DPD and TP have been shown to be associated with better response<sup>[24,112,113]</sup>; and (3) The role of other genes or products for predicting 5-FU response has been also investigated, but results are inconclusive. These molecular markers include glutathione S-transferase (GST), vascular endothelial growth factor, and apoptosis-related genes and gene products including *Bcl-2*, *Bax* and *p53*<sup>[43,114,115]</sup>.

**Molecular predictors for Cisplatin treatment:** Unlike for 5-FU, molecular predicting markers for chemosensitivity of Cisplatin are not well established. GST, an enzyme that degrades Cisplatin, is one of the potential markers. Its activity is affected by polymorphisms in the *GSTM1*, *GSTT1*, and *GSTP1* genes, which may in turn cause variable catabolism of Cisplatin and prognosis<sup>[111,116,117]</sup>. GC with a high LOH rate or MSI-high show a better response to a Cisplatin-based chemotherapy<sup>[118]</sup>.

**Molecular targeted therapy:** Trastuzumab is a monoclonal antibody targeting HER-2 that has shown an overall survival benefit when combined with palliative chemotherapy in patients with HER-2 amplified GC<sup>[119]</sup>. HER-2 is currently the only validated therapeutic target in GC with guidelines for HER-2 testing established by the ToGA trial<sup>[119]</sup>.

HER-2 expression may be assessed by immunohistochemistry (IHC), with scoring ranging from 0 to 3+, by gene amplification using fluorescence in situ hybridization (FISH) or by silver in situ hybridization<sup>[28,119]</sup>. The survival benefit associated with trastuzumab is seen greatest in IHC 3+ or IHC 2+ and FISH-positive patients. Complete membranous staining is not a prerequisite for IHC 2+ or IHC 3+ scores in GC as it is for breast cancer since gastric tumor cells may only show HER-2 staining at the basolateral or lateral membrane regions<sup>[28,120]</sup>.

Other potential candidates for targeted GC therapy include *P13K/Akt/mTOR* pathway, *c-Met* pathways, epidermal growth factor receptor (EGFR), VEGF receptor, insulin-like growth factor receptor, and fibroblast growth



factor receptor<sup>[121]</sup>.

Lapatinib is a dual kinase inhibitor of EGFR and HER-2 under investigation in two ongoing phase III clinical trials in a select group of patients positive HER-2<sup>[121]</sup>. These include the Lapatinib Optimization Study in HER-2 Positive Gastric Cancer study with capecitabine and cisplatin in the first-line setting and the TYTAN study in second-line therapy using paclitaxel<sup>[121-123]</sup>.

In terms of other agents, targeting human EGFR in GC remains controversial. Cetuximab is targeted against EGFR and is a recombinant human, chimeric IgG1 monoclonal antibody<sup>[121,124]</sup>. With combined chemotherapy and cetuximab, promising results have been shown in a phase II trial, but when compared to chemotherapy, the EXPAND study (phase III) failed in prolonging the progression free survival (PFS) and overall survival<sup>[125]</sup>. The REAL-III trial did not show any advantage of adding panitumumab to a combination of chemotherapy and also showed a worse overall survival and PFS<sup>[126-128]</sup>. The combination with matuzumab and chemotherapy seems more promising but was evaluated only in phase II trials<sup>[129]</sup>. Thus, additional studies are necessary.

Antiangiogenic therapy has shown minimal effectiveness when compared to existing treatments for GC<sup>[130,131]</sup>. Biomarkers such as serum VEGF-A and microvessel density still remain unconfirmed as potentially useful predictive markers by phase III trials<sup>[121]</sup>. However, in advanced cases of GC treated with the VEGF inhibitor bevacizumab, plasma VEGF-A and tumor neuropilin-1 are strong biomarker candidates for predicting clinical outcome<sup>[132]</sup>.

Various ongoing trials are testing potential targeting agents addressed to the downstream components of VEGF-R/EGFR, such as inhibitors of mTOR, c-Met, and Histone deacetylase<sup>[121]</sup>. The phase III trial (GRANITE-1) of everolimus, an inhibitor of the P13K/Akt/mTOR pathway, has reported prolonged PFS with a 34% reduction of the risk of progression<sup>[121,133]</sup>.

## MICRORNAS AS THERAPEUTIC TARGETS

In recent years, microRNAs (miRNAs) have been investigated as potential markers in treatment of GC. MiRNAs are important regulators of genes with critical roles in cell proliferation, differentiation, and survival<sup>[134]</sup>. MiRNAs play important roles in the pathogenesis of a variety of malignancies<sup>[135-139]</sup>. Different miRNA methylation profiles are seen in various cancers. MiR-155 is down-regulated and methylated in GC<sup>[140]</sup>. MiR-155 is up-regulated in breast cancer<sup>[141]</sup>, colorectal cancer<sup>[142]</sup> and pancreatic ductal adenocarcinoma<sup>[143]</sup>. Upregulated miRNAs might act as oncogenes and target tumor suppressors, while down-regulated miRNAs might act as tumor suppressors and target oncogenes<sup>[144]</sup>. Several miRNAs have been found to be deregulated in GC but the specific molecular mechanisms are unknown<sup>[144]</sup>. DNA hypermethylation in the miRNA 50 regulatory region accounts

for the downregulation of miRNA in tumors<sup>[145,146]</sup>, and many miRNAs have been reported to be down-regulated due to hypermethylation of the CpG islands in GC. MiR-124a-1, miR-124a-2 and miR-124a-3 have been found to be methylation-silenced in GC cell lines<sup>[144,147]</sup>. Such epigenetic changes are reversible, and make them a potential therapeutic target. Silenced miRNAs in GC could be restored by treating with demethylating agents, such as decitabine (5-aza-20-deoxycytidine), which leads to inhibition of growth, invasion, and metastasis of GC cells<sup>[148]</sup>.

Interestingly, studies have shown that the miRNA methylation levels are positively associated with the clinical stage of GC patients<sup>[144]</sup>. Low expression of miR-34b and miR-129-3p are associated with a poor clinical outcome of GC patients, and hypermethylation of miR-129-2 and miR-34b CpG islands had a tendency to show poor clinicopathological features<sup>[144,149]</sup>. Thus, specific miRNA methylation levels may be used in the prognosis of GC patients. However, limitations exist as several factors besides methylation can affect miRNA expression levels. As reported by Tsukamoto *et al.*<sup>[148]</sup>, the expression of miR-375 in NUGC3 cells can be significantly increased with either 5-aza-2-deoxycytidine and markedly up-regulated by greater than 20-fold when treated with both 5-aza-2-deoxycytidine and trichostatin A<sup>[148]</sup>. In addition, *H. pylori* infection can induce aberrant DNA methylation in gastric epithelial cells<sup>[150]</sup>. Individuals with *H. pylori* had 7.8-13.1-fold higher methylation levels than those without *H. pylori* infection<sup>[147,151]</sup>. Another limitation is the serious side effects of demethylating drugs. The use of demethylating agents may induce the expression of many otherwise normally silenced genes and cause a variety of diseases. Thus, the use of demethylating agents in restoring the expression of epigenetically silenced miRNA in GC still requires further investigation.

## CONCLUSION

The current research has provided some insights to the genetics of gastric cancer. Clinical trials based upon the genetic information have generated promising results. However, up to date, we still do not have an optimal solution for prevention, early diagnosis, and treatment of this disease. The advanced molecular technology, particularly next generation sequencing, may offer hope in deciphering the myth behind the molecular genetics of gastric cancer. Equipped with the advanced technology, together with efforts from clinical oncology and bioinformatics, we have gradually gained much more understanding about the genetic basis of the host-environmental interaction and will have a greater opportunity to identify diagnostic and therapeutic markers for gastric cancer. These advancements have shed light in finding a cure for gastric cancer in the near future.

## REFERENCES

- 1 Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statis-

- tics, 2002. *CA Cancer J Clin* 2005; **55**: 74-108 [PMID: 15761078 DOI: 10.3322/canjclin.55.2.74]
- 2 **Macdonald JS**, Smalley SR, Benedetti J, Hundahl SA, Estes NC, Stemmermann GN, Haller DG, Ajani JA, Gunderson LL, Jessup JM, Martenson JA. Chemoradiotherapy after surgery compared with surgery alone for adenocarcinoma of the stomach or gastroesophageal junction. *N Engl J Med* 2001; **345**: 725-730 [PMID: 11547741 DOI: 10.1056/NEJMoa010187]
- 3 **Cunningham D**, Starling N, Rao S, Iveson T, Nicolson M, Coxon F, Middleton G, Daniel F, Oates J, Norman AR. Capecitabine and oxaliplatin for advanced esophagogastric cancer. *N Engl J Med* 2008; **358**: 36-46 [PMID: 18172173 DOI: 10.1056/NEJMoa073149]
- 4 **Wagner AD**, Grothe W, Haerting J, Kleber G, Grothey A, Fleig WE. Chemotherapy in advanced gastric cancer: a systematic review and meta-analysis based on aggregate data. *J Clin Oncol* 2006; **24**: 2903-2909 [PMID: 16782930 DOI: 10.1200/JCO.2005.05.0245]
- 5 **Lauren P**. The two histological main types of gastric carcinoma: diffuse and so-called intestinal-type carcinoma. an attempt at a histo-clinical classification. *Acta Pathol Microbiol Scand* 1965; **64**: 31-49 [PMID: 14320675]
- 6 **Crusius JB**, Canzian F, Capellá G, Peña AS, Pera G, Sala N, Agudo A, Rico F, Del Giudice G, Palli D, Plebani M, Boeing H, Bueno-de-Mesquita HB, Carneiro F, Pala V, Save VE, Vineis P, Tumino R, Panico S, Berglund G, Manjer J, Stenling R, Hallmans G, Martínez C, Dorronsoro M, Barricarte A, Navarro C, Quirós JR, Allen N, Key TJ, Bingham S, Caldas C, Linseisen J, Kaaks R, Overvad K, Tjønneland A, Büchner FC, Peeters PH, Numans ME, Clavel-Chapelon F, Trichopoulou A, Lund E, Jenab M, Rinaldi S, Ferrari P, Riboli E, González CA. Cytokine gene polymorphisms and the risk of adenocarcinoma of the stomach in the European prospective investigation into cancer and nutrition (EPIC-EURGAST). *Ann Oncol* 2008; **19**: 1894-1902 [PMID: 18628242 DOI: 10.1093/annonc/mdn400]
- 7 **Canedo P**, Durães C, Pereira F, Regalo G, Lunet N, Barros H, Carneiro F, Seruca R, Rocha J, Machado JC. Tumor necrosis factor alpha extended haplotypes and risk of gastric carcinoma. *Cancer Epidemiol Biomarkers Prev* 2008; **17**: 2416-2420 [PMID: 18768512 DOI: 10.1158/1055-9965]
- 8 **Gao LB**, Rao L, Wang YY, Liang WB, Li C, Xue H, Zhou B, Sun H, Li Y, Lv ML, Du XJ, Zhang L. The association of interleukin-16 polymorphisms with IL-16 serum levels and risk of colorectal and gastric cancer. *Carcinogenesis* 2009; **30**: 295-299 [PMID: 19073878 DOI: 10.1093/carcin/bgn281]
- 9 **Capellá G**, Pera G, Sala N, Agudo A, Rico F, Del Giudice G, Plebani M, Palli D, Boeing H, Bueno-de-Mesquita HB, Carneiro F, Berrino F, Vineis P, Tumino R, Panico S, Berglund G, Simán H, Nyrén O, Hallmans G, Martínez C, Dorronsoro M, Barricarte A, Navarro C, Quirós JR, Allen N, Key T, Bingham S, Caldas C, Linseisen J, Nagel G, Overvad K, Tjønneland A, Boshuizen HC, Peeters PH, Numans ME, Clavel-Chapelon F, Trichopoulou A, Lund E, Jenab M, Kaaks R, Riboli E, González CA. DNA repair polymorphisms and the risk of stomach adenocarcinoma and severe chronic gastritis in the EPIC-EURGAST study. *Int J Epidemiol* 2008; **37**: 1316-1325 [PMID: 18641418 DOI: 10.1093/ije/dyn145]
- 10 **Hishida A**, Matsuo K, Goto Y, Mitsuda Y, Hiraki A, Naito M, Wakai K, Tajima K, Hamajima N. Toll-like receptor 4 +3725 G/C polymorphism, Helicobacter pylori seropositivity, and the risk of gastric atrophy and gastric cancer in Japanese. *Helicobacter* 2009; **14**: 47-53 [PMID: 19191896 DOI: 10.1111/j.1523-5378.2009.00659.x]
- 11 **Zhang Y**, Jin M, Liu B, Ma X, Yao K, Li Q, Chen K. Association between H-RAS T81C genetic polymorphism and gastrointestinal cancer risk: a population based case-control study in China. *BMC Cancer* 2008; **8**: 256 [PMID: 18783589 DOI: 10.1186/1471-2407-8-256]
- 12 **Sawa T**, Mounawar M, Tatemichi M, Gilbert I, Katoh T, Ohshima H. Increased risk of gastric cancer in Japanese subjects is associated with microsatellite polymorphisms in the heme oxygenase-1 and the inducible nitric oxide synthase gene promoters. *Cancer Lett* 2008; **269**: 78-84 [PMID: 18502573 DOI: 10.1016/j.canlet.2008.04.015]
- 13 **Shibata T**, Arisawa T, Tahara T, Ohkubo M, Yoshioka D, Maruyama N, Fujita H, Kamiya Y, Nakamura M, Nagasaka M, Iwata M, Takahama K, Watanabe M, Hirata I. Selenoprotein S (SEPS1) gene -105G>A promoter polymorphism influences the susceptibility to gastric cancer in the Japanese population. *BMC Gastroenterol* 2009; **9**: 2 [PMID: 19144102 DOI: 10.1186/1471-230X-9-2]
- 14 **Tripathi S**, Ghoshal U, Ghoshal UC, Mittal B, Krishnani N, Chourasia D, Agarwal AK, Singh K. Gastric carcinogenesis: Possible role of polymorphisms of GSTM1, GSTT1, and GSTP1 genes. *Scand J Gastroenterol* 2008; **43**: 431-439 [PMID: 18365908 DOI: 10.1080/00365520701742930]
- 15 **McNamara D**, El-Omar E. Helicobacter pylori infection and the pathogenesis of gastric cancer: a paradigm for host-bacterial interactions. *Dig Liver Dis* 2008; **40**: 504-509 [PMID: 18486572 DOI: 10.1016/j.dld.2008.02.031]
- 16 **Gehmert S**, Velapattino B, Herrera P, Balqui J, Santivañez L, Cok J, Vargas G, Combe J, Passaro DJ, Wen S, Meyer F, Berg DE, Gilman RH. Interleukin-1 beta single-nucleotide polymorphism's C allele is associated with elevated risk of gastric cancer in Helicobacter pylori-infected Peruvians. *Am J Trop Med Hyg* 2009; **81**: 804-810 [PMID: 19861615 DOI: 10.4269/ajtmh.2009.08-0494]
- 17 **Kumar S**, Kumar A, Dixit VK. Evidences showing association of interleukin-1B polymorphisms with increased risk of gastric cancer in an Indian population. *Biochem Biophys Res Commun* 2009; **387**: 456-460 [PMID: 19607807 DOI: 10.1016/j.bbrc.2009.07.033]
- 18 **Melo Barbosa HP**, Martins LC, Dos Santos SE, Demacki S, Assumpção MB, Aragão CD, de Oliveira Corvelo TC. Interleukin-1 and TNF-alpha polymorphisms and Helicobacter pylori in a Brazilian Amazon population. *World J Gastroenterol* 2009; **15**: 1465-1471 [PMID: 19322919 DOI: 10.3748/wjg.15.1465]
- 19 **Persson C**, Engstrand L, Nyrén O, Hansson LE, Enroth H, Ekström AM, Ye W. Interleukin 1-beta gene polymorphisms and risk of gastric cancer in Sweden. *Scand J Gastroenterol* 2009; **44**: 339-345 [PMID: 19031173 DOI: 10.1080/00365520802556015]
- 20 **Garza-González E**, Bosques-Padilla FJ, El-Omar E, Hold G, Tijerina-Menchaca R, Maldonado-Garza HJ, Pérez-Pérez GI. Role of the polymorphic IL-1B, IL-1RN and TNF-A genes in distal gastric cancer in Mexico. *Int J Cancer* 2005; **114**: 237-241 [PMID: 15540224 DOI: 10.1002/ijc.20718]
- 21 **El-Omar EM**, Carrington M, Chow WH, McColl KE, Bream JH, Young HA, Herrera J, Lissowska J, Yuan CC, Rothman N, Lanyon G, Martin M, Fraumeni JF, Rabkin CS. Interleukin-1 polymorphisms associated with increased risk of gastric cancer. *Nature* 2000; **404**: 398-402 [PMID: 10746728 DOI: 10.1038/35006081]
- 22 **Zendejdel K**, Bahmanyar S, McCarthy S, Nyren O, Andersson B, Ye W. Genetic polymorphisms of glutathione S-transferase genes GSTP1, GSTM1, and GSTT1 and risk of esophageal and gastric cardia cancers. *Cancer Causes Control* 2009; **20**: 2031-2038 [PMID: 19618282 DOI: 10.1007/s10552-009-9399-7]
- 23 **Moy KA**, Yuan JM, Chung FL, Wang XL, Van Den Berg D, Wang R, Gao YT, Yu MC. Isothiocyanates, glutathione S-transferase M1 and T1 polymorphisms and gastric cancer risk: a prospective study of men in Shanghai, China. *Int J Cancer* 2009; **125**: 2652-2659 [PMID: 19610060 DOI: 10.1002/ijc.24583]
- 24 **Napieralski R**, Ott K, Kremer M, Specht K, Vogelsang H, Becker K, Müller M, Lordick F, Fink U, Rüdiger Siewert J, Höfler H, Keller G. Combined GADD45A and thymidine phosphorylase expression levels predict response and sur-

- vival of neoadjuvant-treated gastric cancer patients. *Clin Cancer Res* 2005; **11**: 3025-3031 [PMID: 15837757]
- 25 **Hamilton JP**, Meltzer SJ. A review of the genomics of gastric cancer. *Clin Gastroenterol Hepatol* 2006; **4**: 416-425 [PMID: 16616344 DOI: 10.1016/j.cgh.2006.01.019]
- 26 **Kelley JR**, Duggan JM. Gastric cancer epidemiology and risk factors. *J Clin Epidemiol* 2003; **56**: 1-9 [PMID: 12589864 DOI: 10.1016/S0895-4356(02)00534-6]
- 27 **Tan IB**, Ng I, Tai WM, Tan P. Understanding the genetic basis of gastric cancer: recent advances. *Expert Rev Gastroenterol Hepatol* 2012; **6**: 335-341 [PMID: 22646255 DOI: 10.1586/egh.12.7]
- 28 **Zulfiqar M**, Bhalla A, Weindel M, Shidham VB. Molecular diagnostics in esophageal and gastric neoplasms. *Clin Lab Med* 2013; **33**: 867-873 [PMID: 24267191 DOI: 10.1016/j.cl.2013.08.006]
- 29 **Parsonnet J**, Friedman GD, Orentreich N, Vogelmann H. Risk for gastric cancer in people with CagA positive or CagA negative *Helicobacter pylori* infection. *Gut* 1997; **40**: 297-301 [PMID: 9135515]
- 30 **Higashi H**, Tsutsumi R, Muto S, Sugiyama T, Azuma T, Asaka M, Hatakeyama M. SHP-2 tyrosine phosphatase as an intracellular target of *Helicobacter pylori* CagA protein. *Science* 2002; **295**: 683-686 [PMID: 11743164]
- 31 **Jang BG**, Kim WH. Molecular pathology of gastric carcinoma. *Pathobiology* 2011; **78**: 302-310 [PMID: 22104201 DOI: 10.1159/000321703]
- 32 **Matsumoto Y**, Marusawa H, Kinoshita K, Endo Y, Kou T, Morisawa T, Azuma T, Okazaki IM, Honjo T, Chiba T. *Helicobacter pylori* infection triggers aberrant expression of activation-induced cytidine deaminase in gastric epithelium. *Nat Med* 2007; **13**: 470-476 [PMID: 17401375 DOI: 10.1038/nm1566]
- 33 **Yamasaki E**, Wada A, Kumatori A, Nakagawa I, Funao J, Nakayama M, Hisatsune J, Kimura M, Moss J, Hirayama T. *Helicobacter pylori* vacuolating cytotoxin induces activation of the proapoptotic proteins Bax and Bak, leading to cytochrome c release and cell death, independent of vacuolation. *J Biol Chem* 2006; **281**: 11250-11259 [PMID: 16436379 DOI: 10.1074/jbc.M509404200]
- 34 **Farinati F**, Cardin R, Cassaro M, Bortolami M, Nitti D, Tieppo C, Zaninotto G, Rugge M. *Helicobacter pylori*, inflammation, oxidative damage and gastric cancer: a morphological, biological and molecular pathway. *Eur J Cancer Prev* 2008; **17**: 195-200 [PMID: 18414189 DOI: 10.1097/CEJ.0b013e3282f0bfff]
- 35 **Oberhuber G**, Stolte M. Gastric polyps: an update of their pathology and biological significance. *Virchows Arch* 2000; **437**: 581-590 [PMID: 11193468 DOI: 10.1007/s004280000330]
- 36 **Shiao YH**, Rugge M, Correa P, Lehmann HP, Scheer WD. p53 alteration in gastric precancerous lesions. *Am J Pathol* 1994; **144**: 511-517 [PMID: 8129036]
- 37 **Xiangming C**, Hokita S, Natsugoe S, Tanabe G, Baba M, Takao S, Kuroshima K, Aikou T. p21 expression is a prognostic factor in patients with p53-negative gastric cancer. *Cancer Lett* 2000; **148**: 181-188 [PMID: 10695995 DOI: 10.1016/S0304-3835(99)00335-3]
- 38 **Joyppaul BV**, Newman EL, Hopwood D, Grant A, Qureshi S, Lane DP, Cuschieri A. Expression of p53 protein in normal, dysplastic, and malignant gastric mucosa: an immunohistochemical study. *J Pathol* 1993; **170**: 279-283 [PMID: 8133401 DOI: 10.1002/path.1711700310]
- 39 **Fenoglio-Preiser CM**, Wang J, Stemmermann GN, Noffsinger A. TP53 and gastric carcinoma: a review. *Hum Mutat* 2003; **21**: 258-270 [PMID: 12619111 DOI: 10.1002/humu.10180]
- 40 **Wei J**, O'Brien D, Vilgelm A, Piazuelo MB, Correa P, Washington MK, El-Rifai W, Peek RM, Zaika A. Interaction of *Helicobacter pylori* with gastric epithelial cells is mediated by the p53 protein family. *Gastroenterology* 2008; **134**: 1412-1423 [PMID: 18343378 DOI: 10.1053/j.gastro.2008.01.072]
- 41 **Kokkola A**, Monni O, Puolakkainen P, Larramendy ML, Victorzon M, Nordling S, Haapiainen R, Kivilaakso E, Knuutila S. 17q12-21 amplicon, a novel recurrent genetic change in intestinal type of gastric carcinoma: a comparative genomic hybridization study. *Genes Chromosomes Cancer* 1997; **20**: 38-43 [PMID: 9290952 DOI: 10.1002/(SICI)1098-2264(199709)20:1<38::AID-GCC6>3.0.CO;2-A]
- 42 **Wu MS**, Chang MC, Huang SP, Tseng CC, Sheu JC, Lin YW, Shun CT, Lin MT, Lin JT. Correlation of histologic subtypes and replication error phenotype with comparative genomic hybridization in gastric cancer. *Genes Chromosomes Cancer* 2001; **30**: 80-86 [PMID: 11107179 DOI: 10.1002/1098-2264(2000)9999:9999<::AID-GCC1062>3.0.CO;2-R]
- 43 **Yeh KH**, Shun CT, Chen CL, Lin JT, Lee WJ, Lee PH, Chen YC, Cheng AL. High expression of thymidylate synthase is associated with the drug resistance of gastric carcinoma to high dose 5-fluorouracil-based systemic chemotherapy. *Cancer* 1998; **82**: 1626-1631 [PMID: 9576280 DOI: 10.1002/(SICI)1097-0142(19980501)82:9<1626::AID-CNCR5>3.0.CO;2-8]
- 44 **Gu M**, Kim D, Bae Y, Choi J, Kim S, Song S. Analysis of microsatellite instability, protein expression and methylation status of hMLH1 and hMSH2 genes in gastric carcinomas. *Hepatogastroenterology* 2009; **56**: 899-904 [PMID: 19621725]
- 45 **Seo HM**, Chang YS, Joo SH, Kim YW, Park YK, Hong SW, Lee SH. Clinicopathologic characteristics and outcomes of gastric cancers with the MSI-H phenotype. *J Surg Oncol* 2009; **99**: 143-147 [PMID: 19117018 DOI: 10.1002/jso.21220]
- 46 **Oue N**, Oshimo Y, Nakayama H, Ito R, Yoshida K, Matsusaki K, Yasui W. DNA methylation of multiple genes in gastric carcinoma: association with histological type and CpG island methylator phenotype. *Cancer Sci* 2003; **94**: 901-905 [PMID: 14556664 DOI: 10.1111/j.1349-7006.2003.tb01373.x]
- 47 **Toyota M**, Ahuja N, Suzuki H, Itoh F, Ohe-Toyota M, Imai K, Baylton SB, Issa JP. Aberrant methylation in gastric cancer associated with the CpG island methylator phenotype. *Cancer Res* 1999; **59**: 5438-5442 [PMID: 10554013]
- 48 **Zhang KL**, Sun Y, Li Y, Liu M, Qu B, Cui SH, Kong QY, Chen XY, Li H, Liu J. Increased frequency of CpG island methylator phenotype and CDH1 methylation in a gastric cancer high-risk region of china. *Transl Oncol* 2008; **1**: 28-35 [PMID: 18607505 DOI: 10.1593/tlo.07103]
- 49 **Jee CD**, Kim MA, Jung EJ, Kim J, Kim WH. Identification of genes epigenetically silenced by CpG methylation in human gastric carcinoma. *Eur J Cancer* 2009; **45**: 1282-1293 [PMID: 19195878 DOI: 10.1016/j.ejca.2008.12.027]
- 50 **Yamashita S**, Tsujino Y, Moriguchi K, Tatamatsu M, Ushijima T. Chemical genomic screening for methylation-silenced genes in gastric cancer cell lines using 5-aza-2'-deoxycytidine treatment and oligonucleotide microarray. *Cancer Sci* 2006; **97**: 64-71 [PMID: 16367923 DOI: 10.1111/j.1349-7006.2006.00136.x]
- 51 **Nakatsuru S**, Yanagisawa A, Furukawa Y, Ichii S, Kato Y, Nakamura Y, Horii A. Somatic mutations of the APC gene in precancerous lesion of the stomach. *Hum Mol Genet* 1993; **2**: 1463-1465 [PMID: 8242071 DOI: 10.1093/hmg/2.9.1463]
- 52 **Nakatsuru S**, Yanagisawa A, Ichii S, Tahara E, Kato Y, Nakamura Y, Horii A. Somatic mutation of the APC gene in gastric cancer: frequent mutations in very well differentiated adenocarcinoma and signet-ring cell carcinoma. *Hum Mol Genet* 1992; **1**: 559-563 [PMID: 1338691 DOI: 10.1093/hmg/1.8.559]
- 53 **Horii A**, Nakatsuru S, Miyoshi Y, Ichii S, Nagase H, Kato Y, Yanagisawa A, Nakamura Y. The APC gene, responsible for familial adenomatous polyposis, is mutated in human gastric cancer. *Cancer Res* 1992; **52**: 3231-3233 [PMID: 1317264]
- 54 **Polakis P**. Wnt signaling and cancer. *Genes Dev* 2000; **14**: 1837-1851 [PMID: 10921899]
- 55 **Park WS**, Oh RR, Park JY, Lee SH, Shin MS, Kim YS, Kim SY, Lee HK, Kim PJ, Oh ST, Yoo NJ, Lee JY. Frequent somatic mutations of the beta-catenin gene in intestinal-type gastric cancer. *Cancer Res* 1999; **59**: 4257-4260 [PMID: 10485468]
- 56 **Clements WM**, Wang J, Sarnaik A, Kim OJ, MacDonald J, Fenoglio-Preiser C, Groden J, Lowy AM. beta-Catenin muta-



- tion is a frequent cause of Wnt pathway activation in gastric cancer. *Cancer Res* 2002; **62**: 3503-3506 [PMID: 12067995]
- 57 **Ayhan A**, Yasui W, Yokozaki H, Seto M, Ueda R, Tahara E. Loss of heterozygosity at the bcl-2 gene locus and expression of bcl-2 in human gastric and colorectal carcinomas. *Jpn J Cancer Res* 1994; **85**: 584-591 [PMID: 8063611 DOI: 10.1111/j.1349-7006.1994.tb02400.x]
- 58 **Müller W**, Noguchi T, Wirtz HC, Hommel G, Gabbert HE. Expression of cell-cycle regulatory proteins cyclin D1, cyclin E, and their inhibitor p21 WAF1/CIP1 in gastric cancer. *J Pathol* 1999; **189**: 186-193 [PMID: 10547573 DOI: 10.1002/(SICI)1096-9896(199910)189:2<186::AID-PATH418>3.0.CO;2-L]
- 59 **Wang YL**, Sheu BS, Yang HB, Lin PW, Chang YC. Overexpression of c-erb-B2 proteins in tumor and non-tumor parts of gastric adenocarcinoma--emphasis on its relation to H. pylori infection and clinicohistological characteristics. *Hepato-gastroenterology* 2002; **49**: 1172-1176 [PMID: 12143229]
- 60 **Machado JC**, Oliveira C, Carvalho R, Soares P, Berx G, Caldas C, Seruca R, Carneiro F, Sobrinho-Simões M. E-cadherin gene (CDH1) promoter methylation as the second hit in sporadic diffuse gastric carcinoma. *Oncogene* 2001; **20**: 1525-1528 [PMID: 11313896 DOI: 10.1038/sj.onc.1204234]
- 61 **Mingchao TR**, Stockton P, Sun K, Sills RC, Clayton N, Portier M, Flake G. Loss of E-cadherin expression in gastric intestinal metaplasia and later stage p53 altered expression in gastric carcinogenesis. *Exp Toxicol Pathol* 2001; **53**: 237-246 [PMID: 11665847 DOI: 10.1078/0940-2993-00190]
- 62 **Li QL**, Ito K, Sakakura C, Fukamachi H, Inoue Ki, Chi XZ, Lee KY, Nomura S, Lee CW, Han SB, Kim HM, Kim WJ, Yamamoto H, Yamashita N, Yano T, Ikeda T, Itoharu S, Inazawa J, Abe T, Hagiwara A, Yamagishi H, Ooe A, Kaneda A, Sugimura T, Ushijima T, Bae SC, Ito Y. Causal relationship between the loss of RUNX3 expression and gastric cancer. *Cell* 2002; **109**: 113-124 [PMID: 11955451 DOI: 10.1016/S0092-8674(02)00690-6]
- 63 **Kim TY**, Lee HJ, Hwang KS, Lee M, Kim JW, Bang YJ, Kang GH. Methylation of RUNX3 in various types of human cancers and premalignant stages of gastric carcinoma. *Lab Invest* 2004; **84**: 479-484 [PMID: 14968123 DOI: 10.1038/labinvest.3700060]
- 64 **Chi XZ**, Yang JO, Lee KY, Ito K, Sakakura C, Li QL, Kim HR, Cha EJ, Lee YH, Kaneda A, Ushijima T, Kim WJ, Ito Y, Bae SC. RUNX3 suppresses gastric epithelial cell growth by inducing p21(WAF1/Cip1) expression in cooperation with transforming growth factor {beta}-activated SMAD. *Mol Cell Biol* 2005; **25**: 8097-8107 [PMID: 16135801 DOI: 10.1128/MCB.25.18.8097-8107.2005]
- 65 **Yano T**, Ito K, Fukamachi H, Chi XZ, Wee HJ, Inoue K, Ida H, Bouillet P, Strasser A, Bae SC, Ito Y. The RUNX3 tumor suppressor upregulates Bim in gastric epithelial cells undergoing transforming growth factor beta-induced apoptosis. *Mol Cell Biol* 2006; **26**: 4474-4488 [PMID: 16738314 DOI: 10.1128/MCB.01926-05]
- 66 **Sakakura C**, Hasegawa K, Miyagawa K, Nakashima S, Yoshikawa T, Kin S, Nakase Y, Yazumi S, Yamagishi H, Okanoue T, Chiba T, Hagiwara A. Possible involvement of RUNX3 silencing in the peritoneal metastases of gastric cancers. *Clin Cancer Res* 2005; **11**: 6479-6488 [PMID: 16166423 DOI: 10.1158/1078-0432.CCR-05-0729]
- 67 **Peng Z**, Wei D, Wang L, Tang H, Zhang J, Le X, Jia Z, Li Q, Xie K. RUNX3 inhibits the expression of vascular endothelial growth factor and reduces the angiogenesis, growth, and metastasis of human gastric cancer. *Clin Cancer Res* 2006; **12**: 6386-6394 [PMID: 17085650 DOI: 10.1158/1078-0432.CCR-05-2359]
- 68 **Guilford P**, Hopkins J, Harraway J, McLeod M, McLeod N, Harawira P, Taite H, Scoular R, Miller A, Reeve AE. E-cadherin germline mutations in familial gastric cancer. *Nature* 1998; **392**: 402-405 [PMID: 9537325 DOI: 10.1038/32918]
- 69 **Berx G**, Becker KF, Höfler H, van Roy F. Mutations of the human E-cadherin (CDH1) gene. *Hum Mutat* 1998; **12**: 226-237 [PMID: 9744472 DOI: 10.1002/(SICI)1098-1004(1998)12:4<226::AID-HUMU2>3.0.CO;2-D]
- 70 **Carneiro F**, Huntsman DG, Smyrk TC, Owen DA, Seruca R, Pharoah P, Caldas C, Sobrinho-Simões M. Model of the early development of diffuse gastric cancer in E-cadherin mutation carriers and its implications for patient screening. *J Pathol* 2004; **203**: 681-687 [PMID: 15141383 DOI: 10.1002/path.1564]
- 71 **Humar B**, Guilford P. Hereditary diffuse gastric cancer: a manifestation of lost cell polarity. *Cancer Sci* 2009; **100**: 1151-1157 [PMID: 19432899 DOI: 10.1111/j.1349-7006.2009.01163.x]
- 72 **Weiss MM**, Kuipers EJ, Postma C, Snijders AM, Pinkel D, Meuwissen SG, Albertson D, Meijer GA. Genomic alterations in primary gastric adenocarcinomas correlate with clinicopathological characteristics and survival. *Cell Oncol* 2004; **26**: 307-317 [PMID: 15623941]
- 73 **Tsukamoto Y**, Uchida T, Karnan S, Noguchi T, Nguyen LT, Tanigawa M, Takeuchi I, Matsuura K, Hijiya N, Nakada C, Kishida T, Kawahara K, Ito H, Murakami K, Fujioka T, Seto M, Moriyama M. Genome-wide analysis of DNA copy number alterations and gene expression in gastric cancer. *J Pathol* 2008; **216**: 471-482 [PMID: 18798223 DOI: 10.1002/path.2424]
- 74 **Kuniyasu H**, Yasui W, Kitadai Y, Yokozaki H, Ito H, Tahara E. Frequent amplification of the c-met gene in scirrhous type stomach cancer. *Biochem Biophys Res Commun* 1992; **189**: 227-232 [PMID: 1333188]
- 75 **Vollmers HP**, Zimmermann U, Krenn V, Timmermann W, Illert B, Hensel F, Hermann R, Thiede A, Wilhelm M, Rückel-Lanz H, Reindl L, Müller-Hermelink HK. Adjuvant therapy for gastric adenocarcinoma with the apoptosis-inducing human monoclonal antibody SC-1: first clinical and histopathological results. *Oncol Rep* 1998; **5**: 549-552 [PMID: 9538151]
- 76 **Vollmers HP**, Hensel F, Hermann R, Dämmrich J, Wozniak E, Gessner P, Herrmann B, Zimmermann U, Müller-Hermelink HK. Tumor-specific apoptosis induced by the human monoclonal antibody SC-1: a new therapeutic approach for stomach cancer. *Oncol Rep* 1998; **5**: 35-40 [PMID: 9458378]
- 77 **Uozaki H**, Fukayama M. Epstein-Barr virus and gastric carcinoma--viral carcinogenesis through epigenetic mechanisms. *Int J Clin Exp Pathol* 2008; **1**: 198-216 [PMID: 18784828]
- 78 **Fukayama M**. Epstein-Barr virus and gastric carcinoma. *Pathol Int* 2010; **60**: 337-350 [PMID: 20518883 DOI: 10.1111/j.1440-1827.2010.02533.x]
- 79 **Lee HS**, Chang MS, Yang HK, Lee BL, Kim WH. Epstein-barr virus-positive gastric carcinoma has a distinct protein expression profile in comparison with epstein-barr virus-negative carcinoma. *Clin Cancer Res* 2004; **10**: 1698-1705 [PMID: 15014022]
- 80 **Kang GH**, Lee S, Kim WH, Lee HW, Kim JC, Rhyu MG, Ro JY. Epstein-barr virus-positive gastric carcinoma demonstrates frequent aberrant methylation of multiple genes and constitutes CpG island methylator phenotype-positive gastric carcinoma. *Am J Pathol* 2002; **160**: 787-794 [PMID: 11891177]
- 81 **Hino R**, Uozaki H, Murakami N, Ushiku T, Shinozaki A, Ishikawa S, Morikawa T, Nakaya T, Sakatani T, Takada K, Fukayama M. Activation of DNA methyltransferase 1 by EBV latent membrane protein 2A leads to promoter hypermethylation of PTEN gene in gastric carcinoma. *Cancer Res* 2009; **69**: 2766-2774 [PMID: 19339266 DOI: 10.1158/0008-5472.CAN-08-3070]
- 82 **Hino R**, Uozaki H, Inoue Y, Shintani Y, Ushiku T, Sakatani T, Takada K, Fukayama M. Survival advantage of EBV-associated gastric carcinoma: survivin up-regulation by viral latent membrane protein 2A. *Cancer Res* 2008; **68**: 1427-1435 [PMID: 18316606 DOI: 10.1158/0008-5472.CAN-07-3027]
- 83 **Chen SY**, Lu J, Shih YC, Tsai CH. Epstein-Barr virus latent membrane protein 2A regulates c-Jun protein through extracellular signal-regulated kinase. *J Virol* 2002; **76**: 9556-9561



- [PMID: 12186939 DOI: 10.1128/JVI.76.18.9556-9561.2002]
- 84 **Fukuda M**, Longnecker R. Latent membrane protein 2A inhibits transforming growth factor-beta 1-induced apoptosis through the phosphatidylinositol 3-kinase/Akt pathway. *J Virol* 2004; **78**: 1697-1705 [PMID: 14747535 DOI: 10.1128/JVI.78.4.1697-1705.2004]
  - 85 **Motoshita J**, Oue N, Nakayama H, Kuraoka K, Aung PP, Taniyama K, Matsusaki K, Yasui W. DNA methylation profiles of differentiated-type gastric carcinomas with distinct mucin phenotypes. *Cancer Sci* 2005; **96**: 474-479 [PMID: 16108828 DOI: 10.1111/j.1349-7006.2005.00074.x]
  - 86 **Fitzgerald RC**, Caldas C. E-cadherin mutations and hereditary gastric cancer: prevention by resection? *Dig Dis* 2002; **20**: 23-31 [PMID: 12145417 DOI: 10.1159/000063162]
  - 87 **Simpson AJ**, Caballero OL, Pena SD. Microsatellite instability as a tool for the classification of gastric cancer. *Trends Mol Med* 2001; **7**: 76-80 [PMID: 11286759 DOI: 10.1016/S1471-4914(01)01916-5]
  - 88 **Schneider BG**, Bravo JC, Roa JC, Roa I, Kim MC, Lee KM, Plaisance KT, McBride CM, Mera R. Microsatellite instability, prognosis and metastasis in gastric cancers from a low-risk population. *Int J Cancer* 2000; **89**: 444-452 [PMID: 11008207 DOI: 10.1002/1097-0215(20000920)89:5<444::AID-IJC8>3.0.CO;2-F]
  - 89 **Wirtz HC**, Müller W, Noguchi T, Scheven M, Rüschoff J, Hommel G, Gabbert HE. Prognostic value and clinicopathological profile of microsatellite instability in gastric cancer. *Clin Cancer Res* 1998; **4**: 1749-1754 [PMID: 9676851]
  - 90 **Hayden JD**, Cawkwell L, Quirke P, Dixon MF, Goldstone AR, Sue-Ling H, Johnston D, Martin IG. Prognostic significance of microsatellite instability in patients with gastric carcinoma. *Eur J Cancer* 1997; **33**: 2342-2346 [PMID: 9616279 DOI: 10.1016/S0959-8049(97)00343-2]
  - 91 **Noël A**, Jost M, Maquoui E. Matrix metalloproteinases at cancer tumor-host interface. *Semin Cell Dev Biol* 2008; **19**: 52-60 [PMID: 17625931 DOI: 10.1016/j.semcdb.2007.05.011]
  - 92 **Bani-Hani KE**, Almasri NM, Khader YS, Sheyab FM, Karam HN. Combined evaluation of expressions of cyclin E and p53 proteins as prognostic factors for patients with gastric cancer. *Clin Cancer Res* 2005; **11**: 1447-1453 [PMID: 15746045 DOI: 10.1158/1078-0432.CCR-04-1730]
  - 93 **Allgayer H**, Babic R, Gruetzner KU, Tarabichi A, Schildberg FW, Heiss MM. c-erbB-2 is of independent prognostic relevance in gastric cancer and is associated with the expression of tumor-associated protease systems. *J Clin Oncol* 2000; **18**: 2201-2209 [PMID: 10829039]
  - 94 **Parker WB**, Cheng YC. Metabolism and mechanism of action of 5-fluorouracil. *Pharmacol Ther* 1990; **48**: 381-395 [PMID: 1707544 DOI: 10.1016/0163-7258(90)90056-8]
  - 95 **Jonjić N**, Kovac K, Krasević M, Valković T, Ernjak N, Sasso F, Melato M. Epidermal growth factor-receptor expression correlates with tumor cell proliferation and prognosis in gastric cancer. *Anticancer Res* 1997; **17**: 3883-3888 [PMID: 9427797]
  - 96 **Zhou YN**, Xu CP, Han B, Li M, Qiao L, Fang DC, Yang JM. Expression of E-cadherin and beta-catenin in gastric carcinoma and its correlation with the clinicopathological features and patient survival. *World J Gastroenterol* 2002; **8**: 987-993 [PMID: 12439911]
  - 97 **Gamboa-Dominguez A**, Dominguez-Fonseca C, Quintanilla-Martinez L, Reyes-Gutierrez E, Green D, Angeles-Angeles A, Busch R, Hermannstädter C, Nährig J, Becker KF, Becker I, Höfler H, Fend F, Lubert B. Epidermal growth factor receptor expression correlates with poor survival in gastric adenocarcinoma from Mexican patients: a multivariate analysis using a standardized immunohistochemical detection system. *Mod Pathol* 2004; **17**: 579-587 [PMID: 15073595 DOI: 10.1038/modpathol.3800085]
  - 98 **Hippo Y**, Taniguchi H, Tsutsumi S, Machida N, Chong JM, Fukayama M, Kodama T, Aburatani H. Global gene expression analysis of gastric cancer by oligonucleotide microarrays. *Cancer Res* 2002; **62**: 233-240 [PMID: 11782383]
  - 99 **Marchet A**, Mocellin S, Belluco C, Ambrosi A, DeMarchi F, Mammano E, Digito M, Leon A, D'Arrigo A, Lise M, Nitti D. Gene expression profile of primary gastric cancer: towards the prediction of lymph node status. *Ann Surg Oncol* 2007; **14**: 1058-1064 [PMID: 17106627 DOI: 10.1245/s10434-006-9090-0]
  - 100 **Fondevila C**, Metges JP, Fuster J, Grau JJ, Palacín A, Castells A, Volant A, Pera M. p53 and VEGF expression are independent predictors of tumour recurrence and survival following curative resection of gastric cancer. *Br J Cancer* 2004; **90**: 206-215 [PMID: 14710231 DOI: 10.1038/sj.bjc.6601455]
  - 101 **Tanaka M**, Kitajima Y, Sato S, Miyazaki K. Combined evaluation of mucin antigen and E-cadherin expression may help select patients with gastric cancer suitable for minimally invasive therapy. *Br J Surg* 2003; **90**: 95-101 [PMID: 12520583 DOI: 10.1002/bjs.4014]
  - 102 **Chetty R**. p27 Protein and cancers of the gastrointestinal tract and liver: an overview. *J Clin Gastroenterol* 2003; **37**: 23-27 [PMID: 12811204 DOI: 10.1097/00004836-200307000-00008]
  - 103 **Catzavelos C**, Tsao MS, DeBoer G, Bhattacharya N, Shepherd FA, Slingerland JM. Reduced expression of the cell cycle inhibitor p27Kip1 in non-small cell lung carcinoma: a prognostic factor independent of Ras. *Cancer Res* 1999; **59**: 684-688 [PMID: 9973218]
  - 104 **Chou NH**, Chen HC, Chou NS, Hsu PI, Tseng HH. Expression of altered retinoblastoma protein inversely correlates with tumor invasion in gastric carcinoma. *World J Gastroenterol* 2006; **12**: 7188-7191 [PMID: 17131485]
  - 105 **Feakins RM**, Nickols CD, Bidd H, Walton SJ. Abnormal expression of pRb, p16, and cyclin D1 in gastric adenocarcinoma and its lymph node metastases: relationship with pathological features and survival. *Hum Pathol* 2003; **34**: 1276-1282 [PMID: 14691913 DOI: 10.1016/j.humpath.2003.07.005]
  - 106 **Oh SY**, Kwon HC, Kim SH, Jang JS, Kim MC, Kim KH, Han JY, Kim CO, Kim SJ, Jeong JS, Kim HJ. Clinicopathologic significance of HIF-1alpha, p53, and VEGF expression and preoperative serum VEGF level in gastric cancer. *BMC Cancer* 2008; **8**: 123 [PMID: 18452596 DOI: 10.1186/1471-2407-8-123]
  - 107 **Pinto-de-Sousa J**, Silva F, David L, Leitão D, Seixas M, Pimenta A, Cardoso-de-Oliveira M. Clinicopathological significance and survival influence of p53 protein expression in gastric carcinoma. *Histopathology* 2004; **44**: 323-331 [PMID: 15049897]
  - 108 **Anderson C**, Nijagal A, Kim J. Molecular markers for gastric adenocarcinoma: an update. *Mol Diagn Ther* 2006; **10**: 345-352 [PMID: 17154651]
  - 109 **Marsh S**. Thymidylate synthase pharmacogenetics. *Invest New Drugs* 2005; **23**: 533-537 [PMID: 16267625]
  - 110 **Kawakami K**, Watanabe G. Identification and functional analysis of single nucleotide polymorphism in the tandem repeat sequence of thymidylate synthase gene. *Cancer Res* 2003; **63**: 6004-6007 [PMID: 14522928]
  - 111 **Ruzzo A**, Graziano F, Kawakami K, Watanabe G, Santini D, Catalano V, Bissonni R, Canestrari E, Ficarella R, Menichetti ET, Mari D, Testa E, Silva R, Vincenzi B, Giordani P, Cascinu S, Giustini L, Tonini G, Magnani M. Pharmacogenetic profiling and clinical outcome of patients with advanced gastric cancer treated with palliative chemotherapy. *J Clin Oncol* 2006; **24**: 1883-1891 [PMID: 16622263]
  - 112 **Terashima M**, Irinoda T, Fujiwara H, Nakaya T, Takagane A, Abe K, Yonezawa H, Oyama K, Inaba T, Saito K, Takechi T, Fukushima M. Roles of thymidylate synthase and dihydropyrimidine dehydrogenase in tumor progression and sensitivity to 5-fluorouracil in human gastric cancer. *Anticancer Res* 2002; **22**: 761-768 [PMID: 12014648]
  - 113 **Hua D**, Huang ZH, Mao Y, Deng JZ. Thymidylate synthase and thymidine phosphorylase gene expression as predictive parameters for the efficacy of 5-fluorouracil-based adjuvant chemotherapy for gastric cancer. *World J Gastroenterol* 2007;

- 13: 5030-5034 [PMID: 17854149]
- 114 **Giatromanolaki A**, Stathopoulos GP, Koukourakis MI, Rigatos S, Vrettou E, Kittas C, Fountzilas G, Sivridis E. Angiogenesis and apoptosis-related protein (p53, bcl-2, and bax) expression versus response of gastric adenocarcinomas to paclitaxel and carboplatin chemotherapy. *Am J Clin Oncol* 2001; **24**: 222-226 [PMID: 11404489]
- 115 **Kikuyama S**, Inada T, Shimizu K, Miyakita M, Ogata Y. p53, bcl-2 and thymidine phosphorylase as predictive markers of chemotherapy in patients with advanced and recurrent gastric cancer. *Anticancer Res* 2001; **21**: 2149-2153 [PMID: 11501839]
- 116 **Mannervik B**. The isoenzymes of glutathione transferase. *Adv Enzymol Relat Areas Mol Biol* 1985; **57**: 357-417 [PMID: 3898742]
- 117 **Ott K**, Lordick F, Becker K, Ulm K, Siewert J, Höfler H, Keller G. Glutathione-S-transferase P1, T1 and M1 genetic polymorphisms in neoadjuvant-treated locally advanced gastric cancer: GSTM1-present genotype is associated with better prognosis in completely resected patients. *Int J Colorectal Dis* 2008; **23**: 773-782 [PMID: 18443805 DOI: 10.1007/s00384-008-0490-4]
- 118 **Ott K**, Vogelsang H, Mueller J, Becker K, Müller M, Fink U, Siewert JR, Höfler H, Keller G. Chromosomal instability rather than p53 mutation is associated with response to neoadjuvant cisplatin-based chemotherapy in gastric carcinoma. *Clin Cancer Res* 2003; **9**: 2307-2315 [PMID: 12796400]
- 119 **Bang YJ**, Van Cutsem E, Feyereislova A, Chung HC, Shen L, Sawaki A, Lordick F, Ohtsu A, Omuro Y, Satoh T, Aprile G, Kulikov E, Hill J, Lehle M, Rüschoff J, Kang YK. Trastuzumab in combination with chemotherapy versus chemotherapy alone for treatment of HER2-positive advanced gastric or gastro-oesophageal junction cancer (ToGA): a phase 3, open-label, randomised controlled trial. *Lancet* 2010; **376**: 687-697 [PMID: 20728210 DOI: 10.1016/S0140-6736(10)61121-X]
- 120 **Bang YJ**. Advances in the management of HER2-positive advanced gastric and gastroesophageal junction cancer. *J Clin Gastroenterol* 2012; **46**: 637-648 [PMID: 22751336 DOI: 10.1097/MCG.0b013e3182557307]
- 121 **Cidon EU**, Ellis SG, Inam Y, Adeleke S, Zarif S, Geldart T. Molecular targeted agents for gastric cancer: a step forward towards personalized therapy. *Cancers (Basel)* 2013; **5**: 64-91 [PMID: 24216699 DOI: 10.3390/cancers5010064]
- 122 **Satoh T**, Bang Y, Wang J, Xu J, Chung HC, Yeh K, Chen J, Mukaiyama A, Yoshida P, Ohtsu A. Interim safety analysis from TYTAN: A phase III Asian study of lapatinib in combination with paclitaxel as second-line therapy in gastric cancer. *J Clin Oncol* 2010; **28**: Abstract 4057
- 123 **LOGiC-Lapatinib Optimization Study in ErbB2 (HER2) Positive Gastric Cancer: A Phase III Global, Blinded Study Designed to Evaluate Clinical Endpoints and Safety of Chemotherapy Plus Lapatinib**. Available from: URL: <http://clinicaltrials.gov/ct2/show/NCT00680901?term=LOGiC&rank=8/>
- 124 **Martinelli E**, De Palma R, Orditura M, De Vita F, Ciardiello F. Anti-epidermal growth factor receptor monoclonal antibodies in cancer therapy. *Clin Exp Immunol* 2009; **158**: 1-9 [PMID: 19737224 DOI: 10.1111/j.1365-2249.2009.03992.x]
- 125 **Leach B**. Cetuximab Fails to Extend PFS in Advanced Gastric Cancer. Available from: URL: <http://www.onclive.com/web-exclusives/Cetuximab-Fails-to-Extend-PFS-in-Advanced-Gastric-Cancer/>
- 126 **Okines AF**, Ashley SE, Cunningham D, Oates J, Turner A, Webb J, Saffery C, Chua YJ, Chau I. Epirubicin, oxaliplatin, and capecitabine with or without panitumumab for advanced esophagogastric cancer: dose-finding study for the prospective multicenter, randomized, phase II/III REAL-3 trial. *J Clin Oncol* 2010; **28**: 3945-3950 [PMID: 20679619 DOI: 10.1200/JCO.2010.29.2847]
- 127 **Chau I**, Okines AFC, Gonzalez de Castro D, Saffery C, Barbachano Y, Wotherspoon A, Puckey L, Hulkki Wilson S, Coxon FY, Middleton GW, Ferry DR, Crosby TDL, Madhusudan S, Wadsley J, Waters JS, Hall M, Swinson D, Robinson A, Smith D, Cunningham D. REAL3: A multicenter randomized phase II/III trial of epirubicin, oxaliplatin, and capecitabine (EOC) versus modified (m) EOC plus panitumumab (P) in advanced oesophagogastric (OG) cancer. Response rate (RR), toxicity, and molecular analysis from phase II. *J Clin Oncol* 2011; **29**: Abstract 4131
- 128 **Waddell TS**, Chau I, Barbachano Y, de Castro DG, Wotherspoon A, Saffery C, Middleton GW, Wadsley J, Ferry DR, Mansoor W, Crosby TDL, Coxon FY, Smith D, Waters JS, Iveson T, Falk S, Slater S, Okines AFC, Cunningham D. A randomized multicenter trial of epirubicin, oxaliplatin, and capecitabine (EOC) plus panitumumab in advanced esophagogastric cancer (REAL3). *J Clin Oncol* 2012; **30**: Abstract 4000
- 129 **Rao S**, Starling N, Cunningham D, Sumpter K, Gilligan D, Ruhstaller T, Valladares-Ayerbes M, Wilke H, Archer C, Kurek R, Beadman C, Oates J. Matuzumab plus epirubicin, cisplatin and capecitabine (ECX) compared with epirubicin, cisplatin and capecitabine alone as first-line treatment in patients with advanced oesophago-gastric cancer: a randomised, multicentre open-label phase II study. *Ann Oncol* 2010; **21**: 2213-2219 [PMID: 20497967 DOI: 10.1093/annonc/mdq247]
- 130 **Kang Y**, Ohtsu A, Van Cutsem E, Rha SY, Sawaki A, Park S, Lim H, Wu J, Langer B, Shah MA. AVAGAST: A randomized, double-blind, placebo-controlled, phase III study of first-line capecitabine and cis plus bevacizumab or placebo in patients with advanced gastric cancer (AGC). *J Clin Oncol* 2010; **28**: Abstract LBA4007
- 131 **Kang Y**, Ohtsu A, Van Cutsem E, Roman L, Nunes J, Li C, Otero D, Rivera F, Aprile G, Pimentel Alvarez PR, Moehler MH, Wu J, Langer B, Shah MA. Survival analysis by pooling risk factors in AVAGAST: First-line capecitabine plus bevacizumab or placebo in patients with advanced gastric cancer. *J Clin Oncol* 2011; **29**: Abstract 4119
- 132 **Van Cutsem E**, de Haas S, Kang YK, Ohtsu A, Tebbutt NC, Ming Xu J, Peng Yong W, Langer B, Delmar P, Scherer SJ, Shah MA. Bevacizumab in combination with chemotherapy as first-line therapy in advanced gastric cancer: a biomarker evaluation from the AVAGAST randomized phase III trial. *J Clin Oncol* 2012; **30**: 2119-2127 [PMID: 22565005 DOI: 10.1200/JCO.2011.39.9824]
- 133 **Van Cutsem E**, Yeh KH, Bang YJ. Phase III trial of everolimus (EVE) in previously treated patients with advanced gastric cancer (AGC): GRANITE-1. *J Clin Oncol* 2012; **4**: LBA3
- 134 **Ambros V**. The functions of animal microRNAs. *Nature* 2004; **431**: 350-355 [PMID: 15372042]
- 135 **Wang H**, Li M, Zhang R, Wang Y, Zang W, Ma Y, Zhao G, Zhang G. Effect of miR-335 upregulation on the apoptosis and invasion of lung cancer cell A549 and H1299. *Tumour Biol* 2013; **34**: 3101-3109 [PMID: 23740614 DOI: 10.1007/s13277-013-0878-9]
- 136 **He H**, Di Y, Liang M, Yang F, Yao L, Hao S, Li J, Jiang Y, Jin C, Fu D. The microRNA-218 and ROBO-1 signaling axis correlates with the lymphatic metastasis of pancreatic cancer. *Oncol Rep* 2013; **30**: 651-658 [PMID: 23733161 DOI: 10.3892/or.2013.2516]
- 137 **Xu YY**, Wu HJ, Ma HD, Xu LP, Huo Y, Yin LR. MicroRNA-503 suppresses proliferation and cell-cycle progression of endometrioid endometrial cancer by negatively regulating cyclin D1. *FEBS J* 2013; **280**: 3768-3779 [PMID: 23731275 DOI: 10.1111/febs.12365]
- 138 **Bae HJ**, Noh JH, Kim JK, Eun JW, Jung KH, Kim MG, Chang YG, Shen Q, Kim SJ, Park WS, Lee JY, Nam SW. MicroRNA-29c functions as a tumor suppressor by direct targeting oncogenic SIRT1 in hepatocellular carcinoma. *Oncogene* 2014; **33**: 2557-2567 [PMID: 23728341 DOI: 10.1038/ncr.2013.216]

- 139 **Zhao S**, Yao DS, Chen JY, Ding N. Aberrant expression of miR-20a and miR-203 in cervical cancer. *Asian Pac J Cancer Prev* 2013; **14**: 2289-2293 [PMID: 23725129]
- 140 **Corral-Fernández NE**, Salgado-Bustamante M, Martínez-Leija ME, Cortez-Espinosa N, García-Hernández MH, Reynaga-Hernández E, Quezada-Calvillo R, Portales-Pérez DP. Dysregulated miR-155 expression in peripheral blood mononuclear cells from patients with type 2 diabetes. *Exp Clin Endocrinol Diabetes* 2013; **121**: 347-353 [PMID: 23616185 DOI: 10.1055/s-0033-1341516]
- 141 **Kong W**, He L, Richards EJ, Challa S, Xu CX, Permuth-Wey J, Lancaster JM, Coppola D, Sellers TA, Djeu JY, Cheng JQ. Upregulation of miRNA-155 promotes tumour angiogenesis by targeting VHL and is associated with poor prognosis and triple-negative breast cancer. *Oncogene* 2014; **33**: 679-689 [PMID: 23353819 DOI: 10.1038/onc.2012.636]
- 142 **Zhang GJ**, Xiao HX, Tian HP, Liu ZL, Xia SS, Zhou T. Up-regulation of microRNA-155 promotes the migration and invasion of colorectal cancer cells through the regulation of claudin-1 expression. *Int J Mol Med* 2013; **31**: 1375-1380 [PMID: 23588589 DOI: 10.3892/ijmm.2013.1348]
- 143 **Liu Q**, Chen J, Wang J, Amos C, Killary AM, Sen S, Wei C, Frazier ML. Putative tumor suppressor gene SEL1L was downregulated by aberrantly upregulated hsa-mir-155 in human pancreatic ductal adenocarcinoma. *Mol Carcinog* 2014; **53**: 711-721 [PMID: 23661430 DOI: 10.1002/mc.22023]
- 144 **Ma J**, Hong L, Chen Z, Nie Y, Fan D. Epigenetic regulation of microRNAs in gastric cancer. *Dig Dis Sci* 2014; **59**: 716-723 [PMID: 24248419 DOI: 10.1007/s10620-013-2939-8]
- 145 **Krek A**, Grün D, Poy MN, Wolf R, Rosenberg L, Epstein EJ, MacMenamin P, da Piedade I, Gunsalus KC, Stoffel M, Rajewsky N. Combinatorial microRNA target predictions. *Nat Genet* 2005; **37**: 495-500 [PMID: 15806104]
- 146 **Saito Y**, Liang G, Egger G, Friedman JM, Chuang JC, Coetzee GA, Jones PA. Specific activation of microRNA-127 with downregulation of the proto-oncogene BCL6 by chromatin-modifying drugs in human cancer cells. *Cancer Cell* 2006; **9**: 435-443 [PMID: 16766263]
- 147 **Ando T**, Yoshida T, Enomoto S, Asada K, Tatematsu M, Ichinose M, Sugiyama T, Ushijima T. DNA methylation of microRNA genes in gastric mucosae of gastric cancer patients: its possible involvement in the formation of epigenetic field defect. *Int J Cancer* 2009; **124**: 2367-2374 [PMID: 19165869 DOI: 10.1002/ijc.24219]
- 148 **Tsukamoto Y**, Nakada C, Noguchi T, Tanigawa M, Nguyen LT, Uchida T, Hijiya N, Matsuura K, Fujioka T, Seto M, Moriyama M. MicroRNA-375 is downregulated in gastric carcinomas and regulates cell survival by targeting PDK1 and 14-3-3zeta. *Cancer Res* 2010; **70**: 2339-2349 [PMID: 20215506 DOI: 10.1158/0008-5472.CAN-09-2777]
- 149 **Tsai KW**, Wu CW, Hu LY, Li SC, Liao YL, Lai CH, Kao HW, Fang WL, Huang KH, Chan WC, Lin WC. Epigenetic regulation of miR-34b and miR-129 expression in gastric cancer. *Int J Cancer* 2011; **129**: 2600-2610 [PMID: 21960261 DOI: 10.1002/ijc.25919]
- 150 **Maekita T**, Nakazawa K, Mihara M, Nakajima T, Yanaoka K, Iguchi M, Arii K, Kaneda A, Tsukamoto T, Tatematsu M, Tamura G, Saito D, Sugimura T, Ichinose M, Ushijima T. High levels of aberrant DNA methylation in Helicobacter pylori-infected gastric mucosae and its possible association with gastric cancer risk. *Clin Cancer Res* 2006; **12**: 989-995 [PMID: 16467114]
- 151 **Kalimutho M**, Di Cecilia S, Del Vecchio Blanco G, Roviello F, Sileri P, Cretella M, Formosa A, Corso G, Marrelli D, Pallone F, Federici G, Bernardini S. Epigenetically silenced miR-34b/c as a novel faecal-based screening marker for colorectal cancer. *Br J Cancer* 2011; **104**: 1770-1778 [PMID: 21610744 DOI: 10.1038/bjc.2011.82]

**P- Reviewer:** Declich P, Kumar R, Montecucco A

**S- Editor:** Ji FF **L- Editor:** A **E- Editor:** Liu SQ



## Genome engineering using the CRISPR/Cas system

Takuro Horii, Izuho Hatada

Takuro Horii, Izuho Hatada, Laboratory of Genome Science, Biosignal Genome Resource Center, Institute for Molecular and Cellular Regulation, Gunma University, Gunma 371-8512, Japan  
Author contributions: Horii T and Hatada I contributed to this paper.

Supported by The Grants from the Ministry of Education, Culture, Sports, Science and Technology of Japan; the Ministry of Health, Labour and Welfare of Japan; the National Institute of Biomedical Innovation; the Asahi Glass Foundation; the Ichiro Kanehara Foundation; and the Program for Cultivating Global Leaders in Heavy Ion Therapeutics and Engineering

Correspondence to: Izuho Hatada, PhD, Laboratory of Genome Science, Biosignal Genome Resource Center, Institute for Molecular and Cellular Regulation, Gunma University, 3-39-15 Showa-machi, Maebashi, Gunma 371-8512, Japan. [hatada@gunma-u.ac.jp](mailto:hatada@gunma-u.ac.jp)

Telephone: +81-27-2208057 Fax: +81-27-2208110

Received: December 28, 2013 Revised: May 13, 2014

Accepted: May 16, 2014

Published online: August 27, 2014

**Key words:** Clustered regularly at interspaced short palindromic repeats; Clustered regularly at interspaced short palindromic repeats associated 9; Genome engineering; Double-strand breaks; Non-homologous end joining; Homology-directed repair

**Core tip:** This review introduces the latest information about the genome manipulation technology of the clustered regularly at interspaced short palindromic repeats (CRISPR)/CRISPR associated (Cas) system to readers. We focus particularly on the application of CRISPR/Cas in mammalian cultured cells and mice. The problems of off-target effects and the prospects for therapeutic applications of CRISPR/Cas in the future are also discussed.

Horii T, Hatada I. Genome engineering using the CRISPR/Cas system. *World J Med Genet* 2014; 4(3): 69-76 Available from: URL: <http://www.wjgnet.com/2220-3184/full/v4/i3/69.htm> DOI: <http://dx.doi.org/10.5496/wjmg.v4.i3.69>

### Abstract

Recently, an epoch-making genome engineering technology using clustered regularly at interspaced short palindromic repeats (CRISPR) and CRISPR associated (Cas) nucleases, was developed. Previous technologies for genome manipulation require the time-consuming design and construction of genome-engineered nucleases for each target and have, therefore, not been widely used in mouse research where standard techniques based on homologous recombination are commonly used. The CRISPR/Cas system only requires the design of sequences complementary to a target locus, making this technology fast and straightforward. In addition, CRISPR/Cas can be used to generate mice carrying mutations in multiple genes in a single step, an achievement not possible using other methods. Here, we review the uses of this technology in genetic analysis and manipulation, including achievements made possible to date and the prospects for future therapeutic applications.

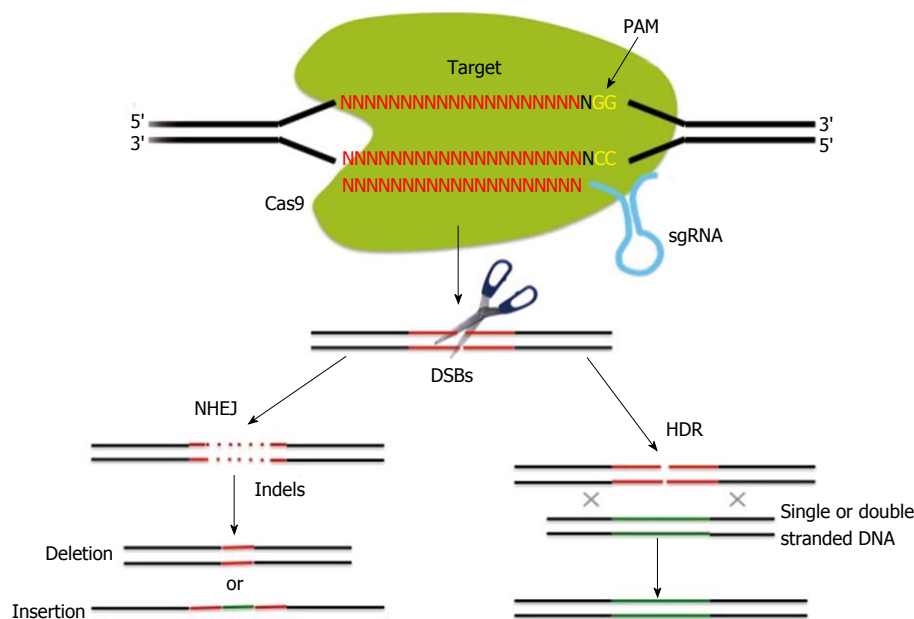
© 2014 Baishideng Publishing Group Inc. All rights reserved.

### INTRODUCTION

The recent development of site-specific endonuclease technologies for selective genome cleavage has been an important advance in mammalian genome engineering. Zinc-finger nucleases (ZFNs) consist of specific DNA-binding zinc-finger proteins and a nuclease domain of the *FokI* endonuclease<sup>[1-3]</sup>. Cleavage with *FokI* requires dimerization of the protein; therefore, fusion to a pair of zinc-finger proteins provides target specificity, and allows cleavage of the target DNA locus, generating double-strand breaks (DSBs).

On the other hand, transcription activator-like effector (TALE) nucleases (TALENs) are fusions of DNA-binding domain TALE repeats with the cleavage domain of the *FokI* restriction enzyme. TALE repeats are highly conserved 33-35 amino acid sequences found in naturally occurring TALEs encoded by *Xanthomonas* bacteria<sup>[4]</sup>. Each TALE repeat binds to a single base pair of DNA and the identities of the amino acids at two posi-





**Figure 1** Schematic of the principles of clustered regularly at interspaced short palindromic repeats (CRISPR)/CRISPR associated (Cas)-mediated genome editing. The CRISPR associated 9 (Cas9) endonuclease can generate sequence-specific double strand breaks (DSBs) of target DNAs bound to small guide RNAs (sgRNAs). The binding site of a target DNA requires a protospacer-adjacent motif (PAM) (with the sequence NGG). DSBs generated by the Cas9 endonuclease are repaired by non-homologous end joining (NHEJ) or homology-directed repair (HDR).

tions have been associated with specificities for different nucleotides<sup>[5,6]</sup>.

These chimeric nucleases enable genome editing by inducing targeted DNA DSBs that are repaired by error-prone, non-homologous end joining (NHEJ) or homology-directed repair (HDR)<sup>[7-10]</sup>. NHEJ-mediated repair induces small insertions or deletions (indels) at the cleavage site, and results in disruption of gene function by frame-shift mutations. In the presence of a single- or double-stranded DNA template containing homology to the sequences flanking the DSB, mutant alleles with precise-point mutations or DNA inserts can be produced by HDR. However, both ZFNs and TALENs require the design of DNA-binding proteins and the construction of complicated plasmids for expression of these, making these methods time-consuming and laborious.

Recently, a new efficient genome manipulation technology, clustered regularly at interspaced short palindromic repeats (CRISPR)/CRISPR associated (Cas) system, which uses the RNA-guided nuclease, Cas9, and is derived from the immune system of bacteria and archaea, has been developed. CRISPR/Cas technology has the advantages of a highly efficient mutation rate and simple-to-design target-specific RNA molecules, compared to the complex ZFN and TALEN systems. Therefore, CRISPR/Cas has been rapidly adopted and applied to many species in a short period of time<sup>[11-40]</sup>.

Several reviews about CRISPR/Cas have already been published<sup>[41-44]</sup>; however, this technology is progressing rapidly, with new reports published weekly. Here, we introduce recent research made possible by CRISPR/Cas technologies and discuss the application of these reagents for genetic analysis and manipulation. We also show the therapeutic potential of CRISPR/Cas and make discus-

sion of future prospects for the field.

## THE CRISPR/CAS SYSTEM

CRISPR/Cas is the RNA-based acquired immunity system in bacteria and archaea<sup>[45,46]</sup>. CRISPR RNA-guided Cas9 nucleases use short RNAs to target and cleave DNA elements captured from foreign invaders (termed “spacers”) in a sequence-specific manner. In the type II CRISPR/Cas system, a single gene encoding the Cas9 protein and two RNAs, a mature CRISPR RNA (crRNA) which is transcribed from spacers, and a partially complementary transacting RNA (tracrRNA) are sufficient for RNA-guided cleavage of foreign DNAs. For maturation of crRNA, RNase III and tracrRNA are necessary<sup>[47]</sup>; however, this process can be simplified by an engineered small guide RNA (sgRNA) containing a hairpin that mimics the tracrRNA-crRNA complex and short guide sequence<sup>[48]</sup> with a protospacer-adjacent motif (with the sequence NGG, Figure 1)<sup>[49]</sup>. Thus, the Cas9 endonuclease can generate sequence-specific DSBs of target DNAs bound to sgRNAs (Figure 1). DSBs generated by the Cas9 endonuclease are repaired by NHEJ or HDR<sup>[7-10]</sup>. NHEJ-mediated repair leads to the generation of small indels at the targeted site, which results in disruption of gene function *via* frame-shift mutations. In the presence of a single- or double-stranded DNA template with homology to the sequences flanking the DSB, mutant alleles with precise-point mutations or DNA inserts can be produced by HDR.

## GENOME EDITING IN CULTURED CELLS

When the CRISPR/Cas system in bacteria and archaea

was elucidated, many researchers expected that it functions in the cells of eukaryotic organisms such as yeast, plants, and even mammals. In January 2013, several papers using the CRISPR/Cas system in human cells were published in succession<sup>[50-53]</sup>. Cho *et al.*<sup>[52]</sup> showed that combination of Cas9 protein and artificial sgRNAs efficiently cleaved two genomic sites and induced indels with approximately 33% frequencies using human embryonic kidney (HEK) 293T-cells. Two papers published in Science used other cell types or targeting loci<sup>[50,51]</sup>. For the endogenous AAVS1 safe harbor genomic locus, Mali *et al.*<sup>[50]</sup> succeeded in gene targeting using 293T-cells (10%-25%), human chronic myelogenous leukemia K562 cells (8%-13%), and human induced pluripotent stem (iPS) cells (2%-4%). In addition, they also used HDR to integrate either a double-stranded DNA donor construct (SA-2A-Puro-pA + CAG-GFP-pA) or an oligo donor into the native AAVS1 locus, and obtained 293T or iPS clones showing HDR-mediated integration.

CRISPR/Cas also enables NHEJ- and HDR-mediated genome editing in mouse ES cells<sup>[54,55]</sup>. The high efficiency of the CRISPR/Cas system coupled with the ability to easily create synthetic sgRNAs make it possible to target multiple genes simultaneously, which is not possible using previous methods<sup>[54]</sup>. Wang *et al.*<sup>[54]</sup> transformed embryonic stem cells using CRISPR/Cas system for three different genes (*Tet1*, *Tet2*, and *Tet3*), and found that > 20% (20/96) of ES cell clones had mutations in all six alleles. To further test the potential of multiplexed gene targeting using the CRISPR/Cas system, sgRNAs targeting five genes (*Tet1*, *Tet2*, *Tet3*, *Sry*, and *Uty*) were mixed and co-transfected with a Cas9-expressing vector into ES cells; of 96 clones screened using an restriction fragment length polymorphism assay, 10% carried mutations at all five loci.

The use of the CRISPR/Cas system in combination with haploid ES cells<sup>[56-58]</sup> provides a powerful platform to manipulate the mammalian genome, because disruption of only one allele can cause loss-of-function phenotypes in haploid ES cells. We have recently reported that co-transfection of mouse haploid ES cells with vectors expressing Cas9 nuclease and sgRNAs targeting *Tet1*, *Tet2*, and *Tet3* results in the complete disruption of all three genes, causing a loss-of-function phenotype with higher efficiency (50%)<sup>[59]</sup> than that previously reported using diploid ES cells<sup>[54]</sup>. Thus, the CRISPR/Cas system used in the context of haploid cells will be useful for the efficient disruption of multiple genes.

## ONE-STEP GENERATION OF GENOME-EDITED ANIMALS

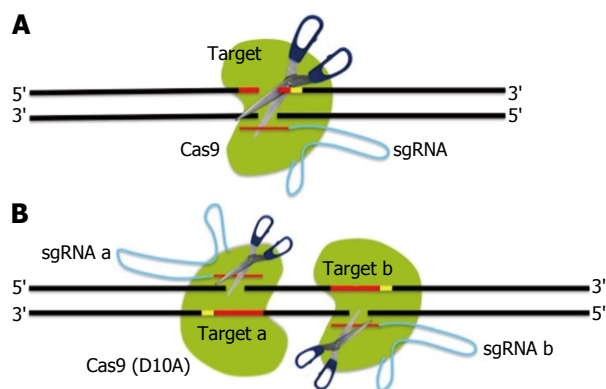
Homologous recombination in mouse ES cells is the most popular method for targeted modifications of the mouse genome; however, generating gene-modified mice through germline chimeras is both time consuming and expensive. Therefore, alternative methods have been developed to accelerate the process of genome modification by the introduction of site-specific nucleases into

fertilized embryos to generate DNA DSBs at a target locus in various species. ZFNs and TALENs have been used to produce several gene-modified rodents<sup>[60-62]</sup>. Although these technologies are widely used in other animals, their use in mice has been limited, principally because the ZFN and TALEN systems are labor-intensive and expensive techniques that do not perform substantially better than ordinary gene knockout technology. On the other hand, CRISPR/Cas-mediated genome editing has successfully demonstrated one-step generation of gene-modified mice, and this technology became widely used within only one year<sup>[54,55,63-65]</sup>. To understand the functions of genes in families of two or more members, animals carrying multiple mutated genes are required; however, ZFNs or TALENs cannot be multiplexed to generate animals with several targeted loci. In contrast, the CRISPR/Cas system can be used to generate mice carrying mutations in multiple genes in one step<sup>[54]</sup>. Co-injection of Cas9 mRNA and sgRNAs for *Tet1* and *Tet2* into fertilized embryos led to the generation of mice with biallelic mutations in both genes with an efficiency of 78% (22/28). Wang *et al.*<sup>[54]</sup> also showed that co-injection of Cas9 mRNA and sgRNAs with mutant oligos generated precise-point mutations simultaneously in two target genes with an efficiency of 20% (2/10). Using this “one-step” procedure, Yang *et al.*<sup>[55]</sup> produced mice carrying a tag or a fluorescent reporter construct in the *Oat4*, *Sax2*, and *Nanog* genes. In addition, *Mecp2* conditional mutant mice with two loxP sites were generated<sup>[55]</sup>. These results show that a single step by CRISPR/Cas-mediated genome editing can generate mice having NHEJ- or HDR-mediated mutations in multiple genes.

## OFF-TARGET MUTATIONS

Compared to ZFNs and TALENs, CRISPR/Cas technology has the advantages of a highly efficient mutation rate and the simplicity of the design of target-specific sgRNAs. It is difficult to compare the off-target effect risk among ZFN, TALEN, and CRISPR/Cas. Although the cleavage of off-target sites has also been observed in ZFN and TALEN systems<sup>[66,67]</sup>, it appears to be less likely because they require two adjacent recognition sites, while the CRISPR/Cas system requires only one. Therefore, it is important to pay careful attention to the specificity of CRISPR/Cas target sequences, because off-target mutations are detrimental to experimental results.

When genome-edited mice are produced using the CRISPR/Cas system, they are rarely influenced by off-target effects. For example, of seven double-mutant mice produced by injection with high RNA concentrations, none showed effects at potential off-target loci using the Surveyor assay<sup>[54]</sup>. Mashiko *et al.*<sup>[65]</sup> found only one off-target mutation in a total of 144 sites examined. In addition, Fujii *et al.*<sup>[64]</sup> proposed that off-target effects are mostly avoided by the careful control of Cas9 mRNA concentration. Surprisingly, the optimized CRISPR/Cas system has a higher gene targeting rate and a lower occurrence of



**Figure 2 Improvement of site-specificity by double nicking.** A: Double strand breaks (DSBs) using wild-type interspaced short palindromic repeats associated 9 (Cas9) endonuclease; B: DSBs using a pair of small guide RNAs (sgRNAs) guiding Cas9 D10A nickases. Using paired nicking can reduce off-target activity because individual nicks that unexpectedly occurred at off-target sites are predominantly repaired by the high-fidelity base excision repair pathway.

off-target effects compared to ZFN<sup>[64]</sup>. Mutant mouse ES cells generated by the CRISPR/Cas system also showed a very low Cas9-mediated cleavage rate in off-target loci<sup>[55]</sup>. These reports suggest that the CRISPR/Cas system is highly specific in the “one-step generation” of mutant mice and mouse ES cells.

By contrast, study of the CRISPR/Cas system in human cancer cell lines indicated a widespread occurrence of off-target mutations<sup>[68,69]</sup>. Cas9-mediated cleavage can be abolished by single mismatches at the sgRNA-target site interface, particularly in the last 10-12 nucleotides located at the 3' end of the 20-nt sgRNA-targeting sequence<sup>[48,51]</sup>. Using human cell lines (U2OS.EGFP, HEK293, and K562), Fu *et al.*<sup>[68]</sup> found that one or two mismatches are tolerated to varying degrees, depending on their position along the sgRNA-DNA interface. In addition, they easily detected off-target alterations induced by 66% (4/6) of CRISPR/Cas experiments targeting endogenous loci by examination of partially mismatched sites. However, these mismatches were mainly located in the 5' region, with only one base mismatch detected in the last 12 nucleotides at the 3' end of one off-target locus.

Yang *et al.*<sup>[55]</sup> considered several possibilities to explain the lower off-target cleavage rate observed in animals derived from manipulated zygotes compared to the results reported for CRISPR/Cas-treated human cell lines including the following: (1) the cells analyzed in mice and humans are clonal and heterogenous populations, respectively; (2) the transformed human cell lines may have different DNA damage responses, resulting in a different mutagenesis rate compared to normal one-cell embryos; and (3) introduced nucleotides are short-lived RNA or long-lived DNA plasmids in mouse and human systems, respectively, which lead to more extensive cleavage in human cells; however, a definitive explanation has not yet been found.

Several measures to improve the specificity of Cas9-mediated genome editing have been assessed. Firstly, it was hypothesized that cleavage specificity may be im-

proved by increasing the length of the region of base pairing between the sgRNA and its target locus. To test this, Ran *et al.*<sup>[70]</sup> generated sgRNAs with 20 or 30 nucleotides guide sequences; however, they found that extension of the guide sequence did not improve Cas9 targeting specificity. Next, Ran *et al.*<sup>[70]</sup> developed a strategy that combines the D10A mutant nickase version of Cas9<sup>[48,51,71]</sup> with a pair of offset sgRNAs complementary to opposite strands of the target site (Figure 2A: DSBs using wild-type Cas9 endonuclease; B: DSBs using a pair of sgRNAs guiding Cas9 D10A). Whereas nicking of both DNA strands by a pair of Cas9 nickases leads to site-specific DSBs and NHEJ, individual nicks are predominantly repaired by the high-fidelity base excision repair pathway<sup>[72]</sup>. As a result, this double nicking method can reduce off-target activity by 50- to 1500-fold and assisted gene knockout without reduction of on-target cleavage efficiency<sup>[70,73]</sup>. Double nicking allows not only NHEJ-mediated indels but also insertion into the genome *via* HDR in human cells.

In the case of mutant animals produced by CRISPR/Cas, off-target mutations will be eliminated by backcrossing to wild-type animals. Therefore, if researchers do not use F0 pups obtained by CRISPR/Cas for experiments, off-target effects should not be a concern. RNA interference (RNAi) experiment to induce sequence-specific gene silencing is now a standard method for the functional analysis of genes. However, designed small RNA frequently repress translation from unexpected loci<sup>[74-76]</sup>. To remove this off-target effect, two or more independent small RNAs are generally used in RNAi experiments. In CRISPR/Cas experiment, use of two or more independent sgRNAs for a gene will be also an effective control to remove off-target noise and improve the reliability of the obtained phenotype. Nevertheless, more detailed work will be necessary to determine the frequency of off-target mutations, and improve the specificity in CRISPR/Cas systems.

## PROSPECTIVE APPLICATIONS OF THE CRISPR/CAS SYSTEM

Precise genome modifications by CRISPR/Cas system excite the interest of scientists working in both basic science and applied fields, including gene therapy. Undoubtedly, the CRISPR/Cas system is a strong candidate for application in human gene therapy. Several human iPS cell lines have been generated from patients for stem cell-based gene therapy by correction of gene mutations. But, gene targeting in human pluripotent stem cells including ES and iPS cells has been very difficult historically<sup>[77]</sup>. Nevertheless, ZFNs and TALENs are capable of correcting gene mutations mediated by HDR repair mechanisms in human iPS cells<sup>[78-82]</sup> and the CRISPR/Cas system has also recently been applied to the gene therapy model<sup>[50, 83]</sup>. Of course, this application will require a highly efficient gene editing rate and no off-target mutations.

CRISPR/Cas is thought to be applicable for genome



editing based only on NHEJ or HDR; however, nuclease-null Cas9 (Cas9<sub>N</sub>) can work as a transcriptional activator or silencer without changing DNA sequences<sup>[84,85]</sup>. Mali *et al.*<sup>[84]</sup> produced a Cas9<sub>N</sub> directly fused with the VP64 activation domain to generate a Cas9<sub>N</sub>-fusion protein capable of transcriptional activation. This Cas9<sub>N</sub>-VP64 protein robustly activated transcription of reporter constructs and endogenous *REX1*, *OCT4*, *SOX2*, and *NANOG* genes when this fusion protein is combined with sgRNA-targeting sequences near the promoter<sup>[84]</sup>. This is the example of RNA-guided transcriptional activation. By contrast, a Cas9<sub>N</sub>-sgRNA complex is specifically able to interfere with transcriptional elongation, transcription factor binding, or RNA polymerase binding<sup>[85]</sup>. This technology could be applied to genome-wide screens for gene function.

Prior genome-editing technologies, ZFNs and TALENs, suggest new applications for CRISPR/Cas. For example, Konermann *et al.*<sup>[86]</sup> developed a light-inducible genome-editing system, using transcriptional effectors and the customizable TALE DNA-binding domain. They succeeded in transcriptional activation and epigenetic modification of endogenous genes using primary neurons as well as brain of living mice.

Bacterial DNA methyltransferases<sup>[87-91]</sup> and human DNA methyltransferase 3a and 3b subunits<sup>[92-94]</sup> have been fused to zinc-finger proteins and successfully demonstrated to perform targeted DNA methylation. Efficient targeting of DNA demethylation was also demonstrated using fusions of TALE repeat arrays and the TET1 hydroxylase catalytic domain (TALE-TET1)<sup>[95]</sup>. These targeted methylation and demethylation technologies will be applicable for gene therapy of cancer and other epigenetic diseases such as Beckwith-Wiedemann and Angelman syndromes mediated by abnormal DNA methylation, or of Huntington disease, which is caused by extra repetitive DNA sequences. In addition, Jiang *et al.*<sup>[96]</sup> inserted an inducible *XIST* transgene into chromosome 21 using ZFN in iPS cells derived from a Down's syndrome patient. In this system, chromosome 21 are coated with *XIST* non-coding RNA, followed by stable heterochromatin modifications, DNA methylation and chromosome-wide transcriptional silencing. This successful silencing of trisomy is the first step for chromosome therapy using genome engineering. These applications developed using ZFN and TALEN systems will be also applicable using the CRISPR/Cas technique.

In the future, CRISPR/Cas may be used to target the viral DNA that becomes integrated into the chromosomes of people with lifetime infections (*e.g.*, HIV). If this viral genetic material can be disrupted using the CRISPR/Cas system, this could negate the need for patients to continue taking antiviral drugs throughout their lives.

## CONCLUSION

CRISPR/Cas has already been applied to many species in which genome engineering has been difficult, because this technology has the advantages of a highly efficient

mutation rate and a simple system for design of target-specific sgRNA. Although improvements in the specificity of CRISPR/Cas will be necessary to eliminate off-target effects, the technique will be indispensable for researchers in both basic and applied science.

## REFERENCES

- 1 Kim YG, Cha J, Chandrasegaran S. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proc Natl Acad Sci USA* 1996; **93**: 1156-1160 [PMID: 8577732]
- 2 Smith J, Berg JM, Chandrasegaran S. A detailed study of the substrate specificity of a chimeric restriction enzyme. *Nucleic Acids Res* 1999; **27**: 674-681 [PMID: 9862996]
- 3 Bibikova M, Carroll D, Segal DJ, Trautman JK, Smith J, Kim YG, Chandrasegaran S. Stimulation of homologous recombination through targeted cleavage by chimeric nucleases. *Mol Cell Biol* 2001; **21**: 289-297 [PMID: 11113203 DOI: 10.1128/MCB.21.1.289-297.2001]
- 4 Bonas U, Stall RE, Staskawicz B. Genetic and structural characterization of the avirulence gene *avrBs3* from *Xanthomonas campestris* pv. *vesicatoria*. *Mol Gen Genet* 1989; **218**: 127-136 [PMID: 2550761]
- 5 Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, Lahaye T, Nickstadt A, Bonas U. Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* 2009; **326**: 1509-1512 [PMID: 19933107 DOI: 10.1126/science.1178811]
- 6 Moscou MJ, Bogdanove AJ. A simple cipher governs DNA recognition by TAL effectors. *Science* 2009; **326**: 1501 [PMID: 19933106 DOI: 10.1126/science.1178817]
- 7 Barnes DE. Non-homologous end joining as a mechanism of DNA repair. *Curr Biol* 2001; **11**: R455-R457 [PMID: 11448783]
- 8 Lieber MR. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annu Rev Biochem* 2010; **79**: 181-211 [PMID: 20192759 DOI: 10.1146/annurev.biochem.052308.093131]
- 9 van den Bosch M, Lohman PH, Pastink A. DNA double-strand break repair by homologous recombination. *Biol Chem* 2002; **383**: 873-892 [PMID: 12222678 DOI: 10.1515/bc.2002.095]
- 10 Wyman C, Kanaar R. DNA double-strand break repair: all's well that ends well. *Annu Rev Genet* 2006; **40**: 363-383 [PMID: 16895466 DOI: 10.1146/annurev.genet.40.110405.090451]
- 11 Friedland AE, Tzur YB, Esvelt KM, Colaiácovo MP, Church GM, Calarco JA. Heritable genome editing in *C. elegans* via a CRISPR-Cas9 system. *Nat Methods* 2013; **10**: 741-743 [PMID: 23817069 DOI: 10.1038/nmeth.2532]
- 12 Bassett AR, Tibbit C, Ponting CP, Liu JL. Highly efficient targeted mutagenesis of *Drosophila* with the CRISPR/Cas9 system. *Cell Rep* 2013; **4**: 220-228 [PMID: 23827738 DOI: 10.1016/j.celrep.2013.06.020]
- 13 Yu Z, Ren M, Wang Z, Zhang B, Rong YS, Jiao R, Gao G. Highly efficient genome modifications mediated by CRISPR/Cas9 in *Drosophila*. *Genetics* 2013; **195**: 289-291 [PMID: 23833182 DOI: 10.1534/genetics.113.153825]
- 14 Hwang WY, Fu Y, Reyon D, Maeder ML, Kaini P, Sander JD, Joung JK, Peterson RT, Yeh JR. Heritable and precise zebrafish genome editing using a CRISPR-Cas system. *PLoS One* 2013; **8**: e68708 [PMID: 23874735 DOI: 10.1371/journal.pone.0068708]
- 15 Jao LE, Wente SR, Chen W. Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. *Proc Natl Acad Sci USA* 2013; **110**: 13904-13909 [PMID: 23918387 DOI: 10.1073/pnas.1308335110]
- 16 Li D, Qiu Z, Shao Y, Chen Y, Guan Y, Liu M, Li Y, Gao N, Wang L, Lu X, Zhao Y, Liu M. Heritable gene targeting in the mouse and rat using a CRISPR-Cas system. *Nat Biotechnol* 2013; **31**: 681-683 [PMID: 23929336 DOI: 10.1038/nbt.2661]
- 17 Li W, Teng F, Li T, Zhou Q. Simultaneous generation and



- germline transmission of multiple gene mutations in rat using CRISPR-Cas systems. *Nat Biotechnol* 2013; **31**: 684-686 [PMID: 23929337 DOI: 10.1038/nbt.2652]
- 18 **Shan Q**, Wang Y, Li J, Zhang Y, Chen K, Liang Z, Zhang K, Liu J, Xi JJ, Qiu JL, Gao C. Targeted genome modification of crop plants using a CRISPR-Cas system. *Nat Biotechnol* 2013; **31**: 686-688 [PMID: 23929338 DOI: 10.1038/nbt.2650]
- 19 **Xie K**, Yang Y. RNA-guided genome editing in plants using a CRISPR-Cas system. *Mol Plant* 2013; **6**: 1975-1983 [PMID: 23956122 DOI: 10.1093/mp/sst119]
- 20 **Feng Z**, Zhang B, Ding W, Liu X, Yang DL, Wei P, Cao F, Zhu S, Zhang F, Mao Y, Zhu JK. Efficient genome editing in plants using a CRISPR/Cas system. *Cell Res* 2013; **23**: 1229-1232 [PMID: 23958582 DOI: 10.1038/cr.2013.114]
- 21 **Chiu H**, Schwartz HT, Antoshechkin I, Sternberg PW. Transgene-free genome editing in *Caenorhabditis elegans* using CRISPR-Cas. *Genetics* 2013; **195**: 1167-1171 [PMID: 23979577 DOI: 10.1534/genetics.113.155879]
- 22 **Katic I**, Großhans H. Targeted heritable mutation and gene conversion by Cas9-CRISPR in *Caenorhabditis elegans*. *Genetics* 2013; **195**: 1173-1176 [PMID: 23979578 DOI: 10.1534/genetics.113.155754]
- 23 **Tzur YB**, Friedland AE, Nadarajan S, Church GM, Calarco JA, Colaiácovo MP. Heritable custom genomic modifications in *Caenorhabditis elegans* via a CRISPR-Cas9 system. *Genetics* 2013; **195**: 1181-1185 [PMID: 23979579 DOI: 10.1534/genetics.113.156075]
- 24 **Waijers S**, Portegijs V, Kerver J, Lemmens BB, Tijsterman M, van den Heuvel S, Boxem M. CRISPR/Cas9-targeted mutagenesis in *Caenorhabditis elegans*. *Genetics* 2013; **195**: 1187-1191 [PMID: 23979586 DOI: 10.1534/genetics.113.156299]
- 25 **Dickinson DJ**, Ward JD, Reiner DJ, Goldstein B. Engineering the *Caenorhabditis elegans* genome using Cas9-triggered homologous recombination. *Nat Methods* 2013; **10**: 1028-1034 [PMID: 23995389 DOI: 10.1038/nmeth.2641]
- 26 **Jiang W**, Zhou H, Bi H, Fromm M, Yang B, Weeks DP. Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in *Arabidopsis*, tobacco, sorghum and rice. *Nucleic Acids Res* 2013; **41**: e188 [PMID: 23999092 DOI: 10.1093/nar/gkt780]
- 27 **Miao J**, Guo D, Zhang J, Huang Q, Qin G, Zhang X, Wan J, Gu H, Qu LJ. Targeted mutagenesis in rice using CRISPR-Cas system. *Cell Res* 2013; **23**: 1233-1236 [PMID: 23999856 DOI: 10.1038/cr.2013.123]
- 28 **Kondo S**, Ueda R. Highly improved gene targeting by germline-specific Cas9 expression in *Drosophila*. *Genetics* 2013; **195**: 715-721 [PMID: 24002648 DOI: 10.1534/genetics.113.156737]
- 29 **Seth A**, Stemple DL, Barroso I. The emerging use of zebrafish to model metabolic disease. *Dis Model Mech* 2013; **6**: 1080-1088 [PMID: 24046387 DOI: 10.1242/dmm.011346]
- 30 **Hisano Y**, Ota S, Kawahara A. Genome editing using artificial site-specific nucleases in zebrafish. *Dev Growth Differ* 2014; **56**: 26-33 [PMID: 24117409 DOI: 10.1111/dgd.12094]
- 31 **Upadhyay SK**, Kumar J, Alok A, Tuli R. RNA-guided genome editing for target gene mutations in wheat. *G3 (Bethesda)* 2013; **3**: 2233-2238 [PMID: 24122057 DOI: 10.1534/g3.113.008847]
- 32 **Blitz IL**, Biesinger J, Xie X, Cho KW. Biallelic genome modification in *F(0)* *Xenopus tropicalis* embryos using the CRISPR/Cas system. *Genesis* 2013; **51**: 827-834 [PMID: 24123579 DOI: 10.1002/dvg.22719]
- 33 **Nakayama T**, Fish MB, Fisher M, Oomen-Hajagos J, Thomson GH, Grainger RM. Simple and efficient CRISPR/Cas9-mediated targeted mutagenesis in *Xenopus tropicalis*. *Genesis* 2013; **51**: 835-843 [PMID: 24123613 DOI: 10.1002/dvg.22720]
- 34 **Sebo ZL**, Lee HB, Peng Y, Guo Y. A simplified and efficient germline-specific CRISPR/Cas9 system for *Drosophila* genomic engineering. *Fly (Austin)* 2013; **8**: 52-57 [PMID: 24141137]
- 35 **Wang Y**, Li Z, Xu J, Zeng B, Ling L, You L, Chen Y, Huang Y, Tan A. The CRISPR/Cas system mediates efficient genome engineering in *Bombyx mori*. *Cell Res* 2013; **23**: 1414-1416 [PMID: 24165890 DOI: 10.1038/cr.2013.146]
- 36 **Daimon T**, Kiuchi T, Takasu Y. Recent progress in genome engineering techniques in the silkworm, *Bombyx mori*. *Dev Growth Differ* 2014; **56**: 14-25 [PMID: 24175911 DOI: 10.1111/dgd.12096]
- 37 **Auer TO**, Durore K, De Cian A, Concordet JP, Del Bene F. Highly efficient CRISPR/Cas9-mediated knock-in in zebrafish by homology-independent DNA repair. *Genome Res* 2014; **24**: 142-153 [PMID: 24179142 DOI: 10.1101/gr.161638.113]
- 38 **Frøkjær-Jensen C**. Exciting prospects for precise engineering of *Caenorhabditis elegans* genomes with CRISPR/Cas9. *Genetics* 2013; **195**: 635-642 [PMID: 24190921 DOI: 10.1534/genetics.113.156521]
- 39 **Sung YH**, Kim JM, Kim HT, Lee J, Jeon J, Jin Y, Choi JH, Ban YH, Ha SJ, Kim CH, Lee HW, Kim JS. Highly efficient gene knockout in mice and zebrafish with RNA-guided endonucleases. *Genome Res* 2014; **24**: 125-131 [PMID: 24253447 DOI: 10.1101/gr.163394.113]
- 40 **Hruscha A**, Krawitz P, Rechenberg A, Heinrich V, Hecht J, Haass C, Schmid B. Efficient CRISPR/Cas9 genome editing with low off-target effects in zebrafish. *Development* 2013; **140**: 4982-4987 [PMID: 24257628 DOI: 10.1242/dev.099085]
- 41 **Gaj T**, Gersbach CA, Barbas CF. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol* 2013; **31**: 397-405 [PMID: 23664777 DOI: 10.1016/j.tibtech.2013.04.004]
- 42 **Menke DB**. Engineering subtle targeted mutations into the mouse genome. *Genesis* 2013; **51**: 605-618 [PMID: 23913666 DOI: 10.1002/dvg.22422]
- 43 **Wei C**, Liu J, Yu Z, Zhang B, Gao G, Jiao R. TALEN or Cas9-rapid, efficient and specific choices for genome modifications. *J Genet Genomics* 2013; **40**: 281-289 [PMID: 23790627 DOI: 10.1016/j.jgg.2013.03.013]
- 44 **Gatz SJ**, Cummings AM, Nguyen JN, Hamm DC, Donohue LK, Harrison MM, Wildonger J, O'Connor-Giles KM. Genome engineering of *Drosophila* with the CRISPR RNA-guided Cas9 nuclease. *Genetics* 2013; **194**: 1029-1035 [PMID: 23709638 DOI: 10.1534/genetics.113.152710]
- 45 **Wiedenheft B**, Sternberg SH, Doudna JA. RNA-guided genetic silencing systems in bacteria and archaea. *Nature* 2012; **482**: 331-338 [PMID: 22337052 DOI: 10.1038/nature10886]
- 46 **Horvath P**, Barrangou R. CRISPR/Cas, the immune system of bacteria and archaea. *Science* 2010; **327**: 167-170 [PMID: 20056882 DOI: 10.1126/science.1179555]
- 47 **Deltcheva E**, Chylinski K, Sharma CM, Gonzales K, Chao Y, Pirzada ZA, Eckert MR, Vogel J, Charpentier E. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* 2011; **471**: 602-607 [PMID: 21455174 DOI: 10.1038/nature09886]
- 48 **Jinek M**, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 2012; **337**: 816-821 [PMID: 22745249 DOI: 10.1126/science.1225829]
- 49 **Marraffini LA**, Sontheimer EJ. Self versus non-self discrimination during CRISPR RNA-directed immunity. *Nature* 2010; **463**: 568-571 [PMID: 20072129 DOI: 10.1038/nature08703]
- 50 **Mali P**, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM. RNA-guided human genome engineering via Cas9. *Science* 2013; **339**: 823-826 [PMID: 23287722 DOI: 10.1126/science.1232033]
- 51 **Cong L**, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F. Multiplex genome engineering using CRISPR/Cas systems. *Science* 2013; **339**: 819-823 [PMID: 23287718 DOI: 10.1126/science.1231143]
- 52 **Cho SW**, Kim S, Kim JM, Kim JS. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat Biotechnol* 2013; **31**: 230-232 [PMID: 23360966 DOI: 10.1038/nbt.2507]
- 53 **Jinek M**, East A, Cheng A, Lin S, Ma E, Doudna J. RNA-

- programmed genome editing in human cells. *Elife* 2013; **2**: e00471 [PMID: 23386978 DOI: 10.7554/eLife.00471]
- 54 **Wang H**, Yang H, Shivalila CS, Dawlaty MM, Cheng AW, Zhang F, Jaenisch R. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* 2013; **153**: 910-918 [PMID: 23643243 DOI: 10.1016/j.cell.2013.04.025]
- 55 **Yang H**, Wang H, Shivalila CS, Cheng AW, Shi L, Jaenisch R. One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell* 2013; **154**: 1370-1379 [PMID: 23992847 DOI: 10.1016/j.cell.2013.08.022]
- 56 **Leeb M**, Wutz A. Derivation of haploid embryonic stem cells from mouse embryos. *Nature* 2011; **479**: 131-134 [PMID: 21900896 DOI: 10.1038/nature10448]
- 57 **Yang H**, Shi L, Wang BA, Liang D, Zhong C, Liu W, Nie Y, Liu J, Zhao J, Gao X, Li D, Xu GL, Li J. Generation of genetically modified mice by oocyte injection of androgenetic haploid embryonic stem cells. *Cell* 2012; **149**: 605-617 [PMID: 22541431 DOI: 10.1016/j.cell.2012.04.002]
- 58 **Elling U**, Taubenschmid J, Wirsberger G, O'Malley R, Demers SP, Vanhaelen Q, Shukalyuk AI, Schmauss G, Schramek D, Schnuetgen F, von Melchner H, Ecker JR, Stanford WL, Zuber J, Stark A, Penninger JM. Forward and reverse genetics through derivation of haploid mouse embryonic stem cells. *Cell Stem Cell* 2011; **9**: 563-574 [PMID: 22136931 DOI: 10.1016/j.stem.2011.10.012]
- 59 **Horii T**, Morita S, Kimura M, Kobayashi R, Tamura D, Takahashi RU, Kimura H, Suetake I, Ohata H, Okamoto K, Tajima S, Ochiya T, Abe Y, Hatada I. Genome engineering of mammalian haploid embryonic stem cells using the Cas9/RNA system. *PeerJ* 2013; **1**: e230 [PMID: 24432195 DOI: 10.7717/peerj.230]
- 60 **Carbery ID**, Ji D, Harrington A, Brown V, Weinstein EJ, Liaw L, Cui X. Targeted genome modification in mice using zinc-finger nucleases. *Genetics* 2010; **186**: 451-459 [PMID: 20628038 DOI: 10.1534/genetics.110.117002]
- 61 **Geurts AM**, Cost GJ, Freyvert Y, Zeitler B, Miller JC, Choi VM, Jenkins SS, Wood A, Cui X, Meng X, Vincent A, Lam S, Michalkiewicz M, Schilling R, Foeckler J, Kalloway S, Weiler H, Ménoret S, Anegón I, Davis GD, Zhang L, Rebar EJ, Gregory PD, Urnov FD, Jacob HJ, Buelow R. Knockout rats via embryo microinjection of zinc-finger nucleases. *Science* 2009; **325**: 433 [PMID: 19628861 DOI: 10.1126/science.1172447]
- 62 **Sung YH**, Baek IJ, Kim DH, Jeon J, Lee J, Lee K, Jeong D, Kim JS, Lee HW. Knockout mice created by TALEN-mediated gene targeting. *Nat Biotechnol* 2013; **31**: 23-24 [PMID: 23302927 DOI: 10.1038/nbt.2477]
- 63 **Shen B**, Zhang J, Wu H, Wang J, Ma K, Li Z, Zhang X, Zhang P, Huang X. Generation of gene-modified mice via Cas9/RNA-mediated gene targeting. *Cell Res* 2013; **23**: 720-723 [PMID: 23545779 DOI: 10.1038/cr.2013.46]
- 64 **Fujii W**, Kawasaki K, Sugiura K, Naito K. Efficient generation of large-scale genome-modified mice using gRNA and CAS9 endonuclease. *Nucleic Acids Res* 2013; **41**: e187 [PMID: 23997119 DOI: 10.1093/nar/gkt772]
- 65 **Mashiko D**, Fujihara Y, Satouh Y, Miyata H, Isotani A, Ikawa M. Generation of mutant mice by pronuclear injection of circular plasmid expressing Cas9 and single guided RNA. *Sci Rep* 2013; **3**: 3355 [PMID: 24284873 DOI: 10.1038/srep03355]
- 66 **Szczepek M**, Brondani V, Büchel J, Serrano L, Segal DJ, Cathomen T. Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases. *Nat Biotechnol* 2007; **25**: 786-793 [PMID: 17603476 DOI: 10.1038/nbt1317]
- 67 **Grau J**, Boch J, Posch S. TALENoffer: genome-wide TALEN off-target prediction. *Bioinformatics* 2013; **29**: 2931-2932 [PMID: 23995255 DOI: 10.1093/bioinformatics/btt501]
- 68 **Fu Y**, Foden JA, Khayter C, Maeder ML, Reyon D, Joung JK, Sander JD. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol* 2013; **31**: 822-826 [PMID: 23792628 DOI: 10.1038/nbt.2623]
- 69 **Hsu PD**, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, Li Y, Fine EJ, Wu X, Shalem O, Cradick TJ, Marraffini LA, Bao G, Zhang F. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol* 2013; **31**: 827-832 [PMID: 23873081 DOI: 10.1038/nbt.2647]
- 70 **Ran FA**, Hsu PD, Lin CY, Gootenberg JS, Konermann S, Trevino AE, Scott DA, Inoue A, Matoba S, Zhang Y, Zhang F. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* 2013; **154**: 1380-1389 [PMID: 23992846 DOI: 10.1016/j.cell.2013.08.021]
- 71 **Gasiunas G**, Barrangou R, Horvath P, Siksnys V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc Natl Acad Sci USA* 2012; **109**: E2579-E2586 [PMID: 22949671 DOI: 10.1073/pnas.1208507109]
- 72 **Dianov GL**, Hübscher U. Mammalian base excision repair: the forgotten archangel. *Nucleic Acids Res* 2013; **41**: 3483-3490 [PMID: 23408852 DOI: 10.1093/nar/gkt076]
- 73 **Cho SW**, Kim S, Kim Y, Kweon J, Kim HS, Bae S, Kim JS. Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Res* 2014; **24**: 132-141 [PMID: 24253446 DOI: 10.1101/gr.162339.113]
- 74 **Jackson AL**, Bartz SR, Schelter J, Kobayashi SV, Burchard J, Mao M, Li B, Cavet G, Linsley PS. Expression profiling reveals off-target gene regulation by RNAi. *Nat Biotechnol* 2003; **21**: 635-637 [PMID: 12754523 DOI: 10.1038/nbt831]
- 75 **Saxena S**, Jónsson ZO, Dutta A. Small RNAs with imperfect match to endogenous mRNA repress translation. Implications for off-target activity of small inhibitory RNA in mammalian cells. *J Biol Chem* 2003; **278**: 44312-44319 [PMID: 12952966 DOI: 10.1074/jbc.M307089200]
- 76 **Snøve O**, Holen T. Many commonly used siRNAs risk off-target activity. *Biochem Biophys Res Commun* 2004; **319**: 256-263 [PMID: 15158470 DOI: 10.1016/j.bbrc.2004.04.175]
- 77 **Zwaka TP**, Thomson JA. Homologous recombination in human embryonic stem cells. *Nat Biotechnol* 2003; **21**: 319-321 [PMID: 12577066 DOI: 10.1038/nbt788]
- 78 **Hockemeyer D**, Wang H, Kiani S, Lai CS, Gao Q, Cassady JP, Cost GJ, Zhang L, Santiago Y, Miller JC, Zeitler B, Cherone JM, Meng X, Hinkley SJ, Rebar EJ, Gregory PD, Urnov FD, Jaenisch R. Genetic engineering of human pluripotent cells using TALE nucleases. *Nat Biotechnol* 2011; **29**: 731-734 [PMID: 21738127 DOI: 10.1038/nbt.1927]
- 79 **Choi SM**, Kim Y, Shim JS, Park JT, Wang RH, Leach SD, Liu JO, Deng C, Ye Z, Jang YY. Efficient drug screening and gene correction for treating liver disease using patient-specific stem cells. *Hepatology* 2013; **57**: 2458-2468 [PMID: 23325555 DOI: 10.1002/hep.26237]
- 80 **Hockemeyer D**, Soldner F, Beard C, Gao Q, Mitalipova M, DeKaveler RC, Katibah GE, Amora R, Boydston EA, Zeitler B, Meng X, Miller JC, Zhang L, Rebar EJ, Gregory PD, Urnov FD, Jaenisch R. Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. *Nat Biotechnol* 2009; **27**: 851-857 [PMID: 19680244 DOI: 10.1038/nbt.1562]
- 81 **Zou J**, Maeder ML, Mali P, Pruett-Miller SM, Thibodeau-Beganny S, Chou BK, Chen G, Ye Z, Park IH, Daley GQ, Porteus MH, Joung JK, Cheng L. Gene targeting of a disease-related gene in human induced pluripotent stem and embryonic stem cells. *Cell Stem Cell* 2009; **5**: 97-110 [PMID: 19540188 DOI: 10.1016/j.stem.2009.05.023]
- 82 **Soldner F**, Laganière J, Cheng AW, Hockemeyer D, Gao Q, Alagappan R, Khurana V, Golbe LI, Myers RH, Lindquist S, Zhang L, Guschin D, Fong LK, Vu BJ, Meng X, Urnov FD, Rebar EJ, Gregory PD, Zhang HS, Jaenisch R. Generation of isogenic pluripotent stem cells differing exclusively at two early onset Parkinson point mutations. *Cell* 2011; **146**: 318-331 [PMID: 21757228 DOI: 10.1016/j.cell.2011.06.019]
- 83 **Horii T**, Tamura D, Morita S, Kimura M, Hatada I. Genera-

- tion of an ICF syndrome model by efficient genome editing of human induced pluripotent stem cells using the CRISPR system. *Int J Mol Sci* 2013; **14**: 19774-19781 [PMID: 24084724 DOI: 10.3390/ijms141019774]
- 84 **Mali P**, Aach J, Stranges PB, Esvelt KM, Moosburner M, Kosuri S, Yang L, Church GM. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat Biotechnol* 2013; **31**: 833-838 [PMID: 23907171 DOI: 10.1038/nbt.2675]
- 85 **Qi LS**, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, Lim WA. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* 2013; **152**: 1173-1183 [PMID: 23452860 DOI: 10.1016/j.cell.2013.02.022]
- 86 **Konermann S**, Brigham MD, Trevino AE, Hsu PD, Heidenreich M, Cong L, Platt RJ, Scott DA, Church GM, Zhang F. Optical control of mammalian endogenous transcription and epigenetic states. *Nature* 2013; **500**: 472-476 [PMID: 23877069 DOI: 10.1038/nature12466]
- 87 **Xu GL**, Bestor TH. Cytosine methylation targetted to predetermined sequences. *Nat Genet* 1997; **17**: 376-378 [PMID: 9398832 DOI: 10.1038/ng1297-376]
- 88 **McNamara AR**, Hurd PJ, Smith AE, Ford KG. Characterisation of site-biased DNA methyltransferases: specificity, affinity and subsite relationships. *Nucleic Acids Res* 2002; **30**: 3818-3830 [PMID: 12202767]
- 89 **Carvin CD**, Parr RD, Kladde MP. Site-selective in vivo targeting of cytosine-5 DNA methylation by zinc-finger proteins. *Nucleic Acids Res* 2003; **31**: 6493-6501 [PMID: 14602907]
- 90 **Smith AE**, Ford KG. Specific targeting of cytosine methylation to DNA sequences in vivo. *Nucleic Acids Res* 2007; **35**: 740-754 [PMID: 17182629 DOI: 10.1093/nar/gkl1053]
- 91 **Nomura W**, Barbas CF. In vivo site-specific DNA methylation with a designed sequence-enabled DNA methylase. *J Am Chem Soc* 2007; **129**: 8676-8677 [PMID: 17583340 DOI: 10.1021/ja0705588]
- 92 **Li F**, Papworth M, Minczuk M, Rohde C, Zhang Y, Ragozin S, Jeltsch A. Chimeric DNA methyltransferases target DNA methylation to specific DNA sequences and repress expression of target genes. *Nucleic Acids Res* 2007; **35**: 100-112 [PMID: 17151075 DOI: 10.1093/nar/gkl1035]
- 93 **Rivenbark AG**, Stolzenburg S, Beltran AS, Yuan X, Rots MG, Strahl BD, Blancafort P. Epigenetic reprogramming of cancer cells via targeted DNA methylation. *Epigenetics* 2012; **7**: 350-360 [PMID: 22419067 DOI: 10.4161/epi.19507]
- 94 **Siddique AN**, Nunna S, Rajavelu A, Zhang Y, Jurkowska RZ, Reinhardt R, Rots MG, Ragozin S, Jurkowski TP, Jeltsch A. Targeted methylation and gene silencing of VEGF-A in human cells by using a designed Dnmt3a-Dnmt3L single-chain fusion protein with increased DNA methylation activity. *J Mol Biol* 2013; **425**: 479-491 [PMID: 23220192 DOI: 10.1016/j.jmb.2012.11.038]
- 95 **Maeder ML**, Angstman JF, Richardson ME, Linder SJ, Cascio VM, Tsai SQ, Ho QH, Sander JD, Reyon D, Bernstein BE, Costello JF, Wilkinson MF, Joung JK. Targeted DNA demethylation and activation of endogenous genes using programmable TALE-TET1 fusion proteins. *Nat Biotechnol* 2013; **31**: 1137-1142 [PMID: 24108092 DOI: 10.1038/nbt.2726]
- 96 **Jiang J**, Jing Y, Cost GJ, Chiang JC, Kolpa HJ, Cotton AM, Carone DM, Carone BR, Shivak DA, Guschin DY, Pearl JR, Rebar EJ, Byron M, Gregory PD, Brown CJ, Urnov FD, Hall LL, Lawrence JB. Translating dosage compensation to trisomy 21. *Nature* 2013; **500**: 296-300 [PMID: 23863942 DOI: 10.1038/nature12394]

**P-Reviewer:** Guo ZS, Vargas FR, Wei L, Xuei X  
**S-Editor:** Wen LL **L-Editor:** A **E-Editor:** Liu SQ





# World Journal of *Medical Genetics*

*World J Med Genet* 2014 November 27; 4(4): 77-109







**Contents**

Quarterly Volume 4 Number 4 November 27, 2014

**REVIEW**

- 77 Genome variation in the trophoblast cell lifespan: Diploidy, polyteny, depolytenization, genome segregation  
*Zybina TG, Zybina EV*
- 94 Role of SOX2 in foregut development in relation to congenital abnormalities  
*Schilders K, Ochieng JK, van de Ven CP, Gontan C, Tibboel D, Rottier RJ*

**MINIREVIEWS**

- 105 Preimplantation HLA typing: Practical tool for stem cell transplantation treatment of congenital disorders  
*Kuliev A, Rechitsky S*

## Contents

*World Journal of Medical Genetics*  
Volume 4 Number 4 November 27, 2014

**APPENDIX** I-V Instructions to authors

**ABOUT COVER** Editorial Board Member of *World Journal of Medical Genetics*, Alessio Squasina, PhD, Department of Neuroscience B.B. Brodie, University of Cagliari, Sp 8 Sestu-Monserrato, Km 0.700, 09042 Cagliari, Italy

**AIM AND SCOPE** *World Journal of Medical Genetics* (*World J Med Genet*, *WJMG*, online ISSN 2220-3184, DOI: 10.5496) is a peer-reviewed open access academic journal that aims to guide clinical practice and improve diagnostic and therapeutic skills of clinicians.

*WJMG* covers topics concerning genes and the pathology of human disease, molecular analysis of simple and complex genetic traits, cancer genetics, epigenetics, gene therapy, developmental genetics, regulation of gene expression, strategies and technologies for extracting function from genomic data, pharmacological genomics, genome evolution. The current columns of *WJMG* include editorial, frontier, diagnostic advances, therapeutics advances, field of vision, mini-reviews, review, topic highlight, medical ethics, original articles, case report, clinical case conference (Clinicopathological conference), and autobiography.

We encourage authors to submit their manuscripts to *WJMG*. We will give priority to manuscripts that are supported by major national and international foundations and those that are of great basic and clinical significance.

**INDEXING/ABSTRACTING** *World Journal of Medical Genetics* is now indexed in Digital Object Identifier.

**FLYLEAF** I-IV Editorial Board

## EDITORS FOR THIS ISSUE

Responsible Assistant Editor: *Xiang Li*  
Responsible Electronic Editor: *Su-Qing Liu*  
Proofing Editor-in-Chief: *Lian-Sheng Ma*

Responsible Science Editor: *Fang-Fang Ji*  
Proofing Editorial Office Director: *Xiu-Xia Song*

**NAME OF JOURNAL**  
*World Journal of Medical Genetics*

**ISSN**  
ISSN 2220-3184 (online)

**LAUNCH DATE**  
December 27, 2011

**FREQUENCY**  
Quarterly

**EDITOR-IN-CHIEF**  
**Hans van Bokhoven, Professor, PhD**, Department of Human Genetics and Cognitive Neurosciences, Radboud university Nijmegen Medical centre, PO Box 9101, 6500 HB Nijmegen, The Netherlands

**EDITORIAL OFFICE**  
Jin-Lei Wang, Director  
Xiu-Xia Song, Vice Director  
*World Journal of Medical Genetics*

Room 903, Building D, Ocean International Center, No. 62 Dongsihuan Zhonglu, Chaoyang District, Beijing 100025, China  
Telephone: +86-10-59080039  
Fax: +86-10-85381893  
E-mail: editorialoffice@wjgnet.com  
Help Desk: <http://www.wjgnet.com/esp/helpdesk.aspx>  
<http://www.wjgnet.com>

**PUBLISHER**  
Baishideng Publishing Group Inc  
8226 Regency Drive,  
Pleasanton, CA 94588, USA  
Telephone: +1-925-223-8242  
Fax: +1-925-223-8243  
E-mail: bpgoffice@wjgnet.com  
Help Desk: <http://www.wjgnet.com/esp/helpdesk.aspx>  
<http://www.wjgnet.com>

**PUBLICATION DATE**  
November 27, 2014

## COPYRIGHT

© 2014 Baishideng Publishing Group Inc. Articles published by this Open Access journal are distributed under the terms of the Creative Commons Attribution Non-commercial License, which permits use, distribution, and reproduction in any medium, provided the original work is properly cited, the use is non commercial and is otherwise in compliance with the license.

## SPECIAL STATEMENT

All articles published in journals owned by the Baishideng Publishing Group (BPG) represent the views and opinions of their authors, and not the views, opinions or policies of the BPG, except where otherwise explicitly indicated.

## INSTRUCTIONS TO AUTHORS

Full instructions are available online at [http://www.wjgnet.com/2220-3184/g\\_info\\_20100722180909.htm](http://www.wjgnet.com/2220-3184/g_info_20100722180909.htm)

## ONLINE SUBMISSION

<http://www.wjgnet.com/esp/>

## Genome variation in the trophoblast cell lifespan: Diploidy, polyteny, depolytenization, genome segregation

Tatiana G Zybyna, Eugenia V Zybyna

Tatiana G Zybyna, Eugenia V Zybyna, Laboratory of Cell Pathology, Institute of Cytology, Russian Academy of Sciences, 194064 St.-Petersburg, Russian Federation

Author contributions: Zybyna TG and Zybyna EV contributed equally in design and writing of the manuscript as well as to preparing illustrations for this paper.

Supported by The Program "Molecular and Cell Biology" of the Russian Academy of Sciences

Correspondence to: Tatiana G Zybyna, ScD, Laboratory of Cell Pathology, Institute of Cytology, Russian Academy of Sciences, Tikhoretsky ave., 4, 194064 St.-Petersburg, Russian Federation. [zybyna@mail.cytspb.rssi.ru](mailto:zybyna@mail.cytspb.rssi.ru)

Telephone: +7-812-2970341 Fax: +7-812-2970341

Received: January 16, 2014 Revised: June 24, 2014

Accepted: August 27, 2014

Published online: November 27, 2014

### Abstract

The lifespan of mammalian trophoblast cells includes polyploidization, its degree and peculiarities are, probably, accounted for the characteristics of placenta development. The main ways of genome multiplication—endoreduplication and reduced mitosis—that basically differ by the extent of repression of mitotic events, play, most probably, different roles in the functionally different trophoblast cells in a variety of mammalian species. In the rodent placenta, highly polyploid (512-2048c) trophoblast giant cells (TGC) undergoing endoreduplication serve a barrier with semiallogenic maternal tissues whereas series of reduced mitoses allow to accumulate a great number of low-ploid junctional zone and labyrinth trophoblast cells. Endoreduplication of TGC comes to the end with formation of numerous low-ploid subcellular compartments that show some signs of viable cells though mitotically inactive; it makes impossible their ectopic proliferation inside maternal tissues. In distinct from rodent trophoblast, deviation from (2<sup>n</sup>)c in human and silver fox trophoblast suggests a possibility of aneuploidy and other chromosome changes (aberrations, *etc.*). It suggests that in mammalian species with

lengthy period of pregnancy, polyploidy is accompanied by more diverse genome changes that may be useful to select a more specific response to stressful factors that may appear occasionally during months of intrauterine development.

© 2014 Baishideng Publishing Group Inc. All rights reserved.

**Key words:** Cell cycle; Endocycle; Polyploidy; Genome; Aneuploidy; Trophoblast; Placenta

**Core tip:** In rodent placenta, differentiation of secondary trophoblast giant cells give an example of the irreversible endoreduplication (up to 1024c and higher) that, however, results in formation of low-ploid subcellular compartments incapable of mitotic proliferation. In the mammalian species with lengthy period of pregnancy, more diverse genome changes may be useful to select a more specific response to stressful factors that may appear occasionally during months of intrauterine development.

Zybyna TG, Zybyna EV. Genome variation in the trophoblast cell lifespan: Diploidy, polyteny, depolytenization, genome segregation. *World J Med Genet* 2014; 4(4): 77-93 Available from: URL: <http://www.wjgnet.com/2220-3184/full/v4/i4/77.htm> DOI: <http://dx.doi.org/10.5496/wjmg.v4.i4.77>

### INTRODUCTION

Genome multiplication in the cells of placental trophoblast is a unique phenomenon among the mammalian and other vertebrate tissues. By now, ontogenesis of rodent trophoblast cells is fairly characterized and serves as a model for studying normal and pathological development of placenta. The main peculiarities of the trophoblast cell lifespan in the rodent placenta is a genome reproduction due to cell cycle reduction up to two phases-S and

G, the multifold repeat of which results in high level of cell polyploidization-up to 512-2048 and higher<sup>[1-4]</sup> that involves polytenization<sup>[5]</sup>. Beginning from the second half of pregnancy, a significant part of the secondary giant trophoblast cells undergo depolytenization and genome segregation with subsequent isolation of numerous small nuclear fragments detaching from the giant nucleus<sup>[5-8]</sup>.

A range of trophoblast cell populations do not leave mitotic cycle (a part of them probably represent the trophoblast stem cells); the cells and their derivatives form cell population of lower ploidy levels (2c-32c) at the account of uncompleted mitoses<sup>[5-7,9,10]</sup>. By now, a noticeable data are available in the polyploidy in the other mammalian groups, their modes of polyploidization being different from the rodent ones<sup>[11,12]</sup>. Simultaneously, the more and more data appear on the role of polyploidy in formation of different mammalian tissues<sup>[13-19]</sup>. In particular, in recent publications a great attention is drawn to the relationship of polyploidy and aneuploidy, the latter is considered as a factor of genetic variability that may be an adaptive under the stress conditions<sup>[17-21]</sup>. Therefore it seems to be interesting to compare regularities of genome multiplication in different mammalian species in accordance with possible role of trophoblast cells in placenta formation.

## THE WAYS OF SOMATIC CELL POLYPLOIDIZATION

At present a great number of data has been obtained that confirm the concept that modification (mostly, shortening) of the “archetypal” mitotic cell cycle results in genome multiplication<sup>[3,7,14,19,22-24]</sup>. Recently, there dominates a notion that such a modified cell cycle is characterized by alternating DNA synthesis (S) and Gap (G) phases in the absence of intervening mitoses, karyokinesis, and cytokinesis; a series of these shortened cycles allows cells to achieve high level of ploidy that may exceed 1000c<sup>[19,23,24]</sup>. Nevertheless, the extent of cell cycle shortening that results in polyploidization appears to differ significantly in different cell types and taxa. Therefore, it seems to be important to present short characteristics of different ways of polyploidization of somatic cells of different animals, plants, and human.

Switching off the last step of mitosis-cytokinesis-may be the first step to polyploidy. In this case, binucleate cell is formed ( $2c \times 2$ ). In the next cell cycle, both nuclei, as a rule, enter mitosis synchronously due to what the uniform metaphase plate is formed. If the mitosis comes to the end, it results in two mononucleated cells with tetraploid nuclei. Alternation of acytokinetic and following complete mitoses may result in formation of mono- and binucleate cells of the higher ploidy:

$$2c-(2c \times 2)-4c-(4c \times 2)-8c-(8c \times 2)$$

Noteworthy, in this case, mitosis is the key event that allows formation of polyploid nucleus. Such a way of polyploidization was demonstrated basing on the dynamics of transition of mono- and binucleate cells

with nuclei of different ploidy by using combination of cytophotometry and <sup>3</sup>H-thymidine DNA replication labeling<sup>[22]</sup>; this way was currently confirmed by using time-lapse video images<sup>[18]</sup>. Similar way of polyploidization was also demonstrated in some other mammalian cell types-cardiomyocytes<sup>[25,26]</sup> and aortic vascular smooth muscle cells<sup>[27]</sup>.

Block of mitoses at the meta- and anaphase also may result in polyploidization: the cell do not complete the mitotic division owing to what the nucleus with the doubled number of chromosomes is formed<sup>[3,7]</sup>. Such a way of polyploidization recently is often described under the name of endomitosis<sup>[14,19,23]</sup>. However, taking into account participation of mitosis (although uncompleted) in this cell cycle, we adhere to the used term “restitutional” or “uncompleted mitosis”<sup>[3,28]</sup>. And now, in our opinion, it makes sense to unify the cases of uncompleted and acytokinetic mitoses under the term “reduced mitoses”. Meantime, the term “endomitosis” includes prefix “endo” accounted for by the initial sense that implied chromosome/chromatid segregation inside the nucleus. Different forms of reduced mitoses may be observed simultaneously in the same cell type; in some cases they are accompanied by endocycles (see below); all these phenomena may result from disorder of many mitotic events. A good case in point is aortic vascular smooth cells studied by the time-lapse video<sup>[27]</sup>. In most cases there occurred normal anaphase chromosome segregation, but progress of cytokinesis was arrested; as a result, binucleate cells were formed. Sometimes such a binucleate cell renewed cytokinesis that came to the end with two daughter (probably diploid) cells formation. In other, rare, cases, mitoses resulted in mononuclear polyploid cells due to chromosome bridge(s) that did not allow forming two daughter nuclei. At last, there occurred some mitoses with shallow cleavage furrow and missegregation of sister chromatids; thereafter cleavage furrow disappeared and mononuclear polyploid cell was formed. All the phenomena were accompanied by downregulation of Survivin<sup>[27]</sup>; the disturbance of AuroraB/Survivin complex (a regulator of mitotic machinery) exerted such a pleiotropic effect on the progression of mitosis. The pleiotropic modifications of mitosis including endocycles (see below) were also observed in many cases of spontaneous and induced polyploidization, in particular, in endosperm and suspensor of higher plants<sup>[3,28]</sup>, in cancer cells<sup>[29]</sup>, in bovine trophoblast cells<sup>[17,30]</sup> and in some other cases.

The reduced mitoses in most cases can result in polyploidy of moderate level such as 4c, 8c, and some higher. As a rule, these modes of polyploidization allow cell to retain its mitotic potential<sup>[17,18]</sup>. In some cases they may result in “ploidy reversal” thereby generating diploid cells from the tetra- and octaploid ones as demonstrated by the time-lapse video<sup>[17,18,20]</sup>.

More profound reduction of mitotic cycle results in endocycles. Originally this term covered all cases of genome multiplication accomplished without the nuclear envelope disappearance<sup>[3,5-7,28,31-33]</sup>. Therefore, it may be



suggested that in these cases, everything or nearly all stages of mitosis are reduced: the chromosomes do not form the metaphase plate, and nucleus retains the traits of interphase or prophase. The giant cells with polytene chromosomes of salivary glands of Diptera that probably represent numerous copies of the tightly attached elongated sister chromatids with clear-cut chromomere structure most probably are formed by means of block of mitosis in prophase.

The cycle of polytene nucleus devoid of mitosis was described in the giant trophoblast cells of mouse and rat. Using  $^3\text{H}$ -thymidine labeling, two phases-S (endointerphase) and G (endoprophase) were identified in rat trophoblast cells. At the endointerphase (S-phase) the nucleus was filled with thin long paired Feulgen-positive threads, whereas at the endoprophase the bundles of non-classic polytene chromosomes were observed<sup>[5-7]</sup>. In mice, three phases of endoreduplication-G1, S, and G2 were discerned basing on the oscillative expression of the S-phase inhibitor p57<sup>kip2</sup><sup>[34]</sup>.

The so-called classic endomitosis costs independently from the point of view of classifications used here. It is also a cell cycle devoid of phase of mitosis, although it includes phases of endoprophase, endometaphase, and endoanaphase, when the chromosomes attached to the nuclear envelope undergo condensation and splitting into sister chromatids followed by decondensation<sup>[3,7,31,35]</sup>. Interestingly, the nuclear envelope in most cases is retained throughout the whole cycle of endomitosis. Using the  $^3\text{H}$ -thymidine delayed labeling, three phases of endointerphase-G1, S, and G2 were determined in the endomitotic cells of albumen gland of the snail *Succinea latta*<sup>[36]</sup>. All the data, in our opinion, suggest that classic endomitosis, most likely, belongs to endocycles, *i.e.*, modifications of cell cycle, in which the genome multiplication is accomplished due to reduction of most of mitotic stages.

The classic and some other types of endomitosis are widespread mostly in invertebrates and plant tissues<sup>[3,28]</sup>. Meanwhile, such a mode of genome multiplication is observed in the human placental trophoblast<sup>[7,32,37,38]</sup>, as well as in cancer cells<sup>[29,33,37]</sup>. The future investigation should elucidate the preferable meaning of the term “endomitosis”.

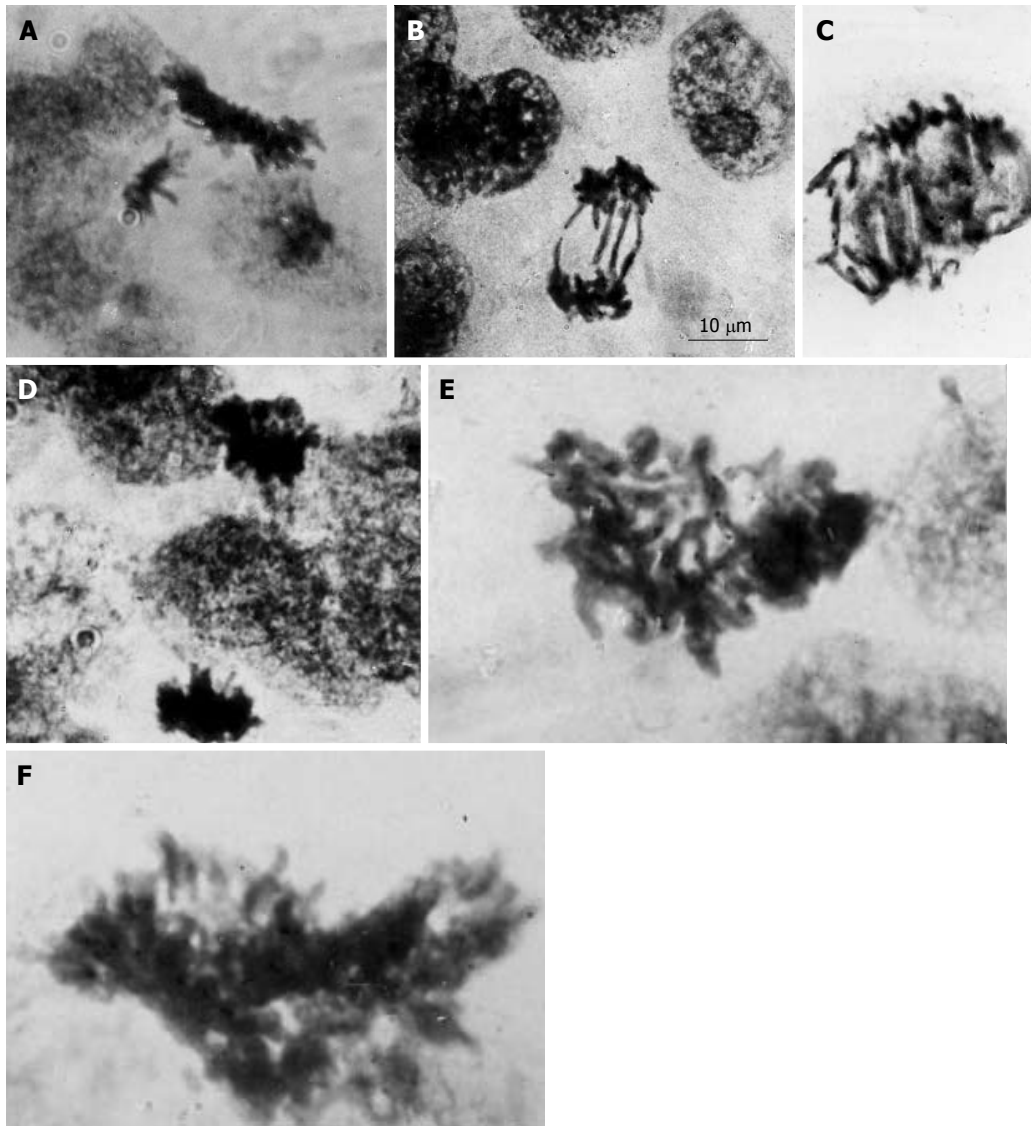
Numerous data suggest that endocycles allow cells to gain very high ploidy levels. Thus, a series of endoreduplication cycle polytene nuclei of *Drosophila* salivary glands may result in 1024c<sup>[3,39-41]</sup> that most frequently provides irreversibility of switch from mitotic cycle to endoreduplication<sup>[22,23]</sup>. That is why, probably, endocycles are characteristic of the majority of invertebrate and plant differentiated tissues<sup>[3,28,42]</sup>. By contrast, the uncompleted mitoses including the acytokinetic ones in most cases can result in polyploidy of moderate level such as 4c, 8c, and some higher. As a rule, these modes of polyploidization allow cells to retain their mitotic potential-for example, in hepatocytes<sup>[17,18]</sup> that enable tissue to undergo an effective regeneration and sometimes (mostly in cases of experiments), to reverse to the diploid state<sup>[17,18]</sup>.

## POLYPLOIDIZATION IN DIFFERENTIATION OF THE RAT AND MOUSE TROPHOBLAST CELLS

Beginning from the onset of differentiation, the primary and secondary trophoblast giant cells (TGC) in rat and mouse placenta undergo a series of endoreduplication cycles that result in a very high ploidy level<sup>[1,4,6,43]</sup> and polytenization of giant nuclei. In the non-classic polytene nuclei, chromatin underwent condensation in the G-phase showing bundles of thick and short chromonemes under the nuclear envelope or near the nucleolus and decondensation in the S-phase where nuclei were similar to the interphase ones<sup>[5,6]</sup>. Later on, *i.e.*, after the 12<sup>th</sup> gestational day in rat, the features of non-classic polytenic chromosomes become less expressed that the precede transformation of the giant nucleus into polygenomic and multinucleate<sup>[5]</sup>. Switch to endoreduplication in the murine TGC is connected to the switch of the cyclin D isoform expression from D3 to D1; the arrest of mitotic cycle and the onset of endoreduplication was most probably accounted for by the failure to assemble the cyclin B/p34<sup>cdk1</sup> complex during the first endocycle; in the subsequent endocycle the mitotic cyclin B was suppressed<sup>[43,44]</sup>.

As distinct from the primary and secondary TGC, the junctional zone (JZ) trophoblast cells in rat and mouse placenta represent a highly proliferative cell population; the mitotic activity of JZ trophoblast cells persists up to midgestation<sup>[6]</sup>. Simultaneously the JZ trophoblast cells undergo differentiation into a number of cell subtypes involved in glycogen store, hormone production, invasion, *etc.*<sup>[9,45-49]</sup>. Unlike the secondary TGC that form a barrier at the border between semiallogenic maternal and embryonic tissues, the JZ trophoblast cells undergo polyploidization *via* uncompleted polyploidizing mitoses up to 8c followed by some endoreduplication cycles allowing them to reach 16-32c<sup>[6]</sup>. The DNA content measurement in the mitotic figures showed 4c, 8c and 16c mitotic figures indicating the ability of 2c-8c polyploid cells to undergo the complete mitotic division<sup>[6]</sup>. However, the polyploid mitotic figures often were of abnormal shape; chromosome bridges were frequent in polyploid anaphases, the double and multiple bridges were commonly present (Figure 1). Therefore, in this case, polyploidization increases the possibility of mitotic arrest that may result in interphase renewal leading to further polyploidization.

Difference in endocycle and uncompleted mitosis is illustrated by the data on their regulation by the transcription factor family E2F<sup>[24,50]</sup>. The canonical and atypical transcriptional programs converge to control the endocycle through the regulation of cellular events important for mitosis, karyokinesis, and cytokinesis. Thus, in the murine trophoblast giant cells (TGC) the targeted gene inactivation of *E2f1*, *E2f2*, and *E2f3* transcription induced abnormally large nuclei, their ploidy exceeding their characteristic ploidy levels 8c-256c. In the *E2F7*<sup>-/-</sup>



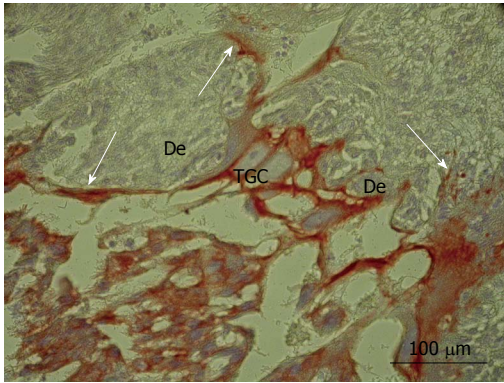
**Figure 1** Diploid and polyploid mitoses in the junctional zone trophoblast cells in rat placenta at the 14<sup>th</sup> day of gestation. A: Tetraploid metaphase and diploid anaphase; B: Diploid restitutional anaphase with multiple chromosome bridges; C: Polyploid restitutional anaphase with multiple chromosome bridges; D: Tetraploid normal anaphase; E: Octaploid metaphase; F: Hexadecaploid metaphase.

and E2F8<sup>-/-</sup> an unusual number of mitosis at different phases were observed, which suggest interruption of endocycle<sup>[24]</sup>. The ploidy did not exceed 64c that, according to our data is normally characteristic of the trophoblast cells of rat placenta junctional zone undergone polyploidization *via* polyploidizing mitoses with subsequent endocycles<sup>[6]</sup>. This phenotype of E2F7<sup>-/-</sup> and E2F8<sup>-/-</sup> mice included upregulation of mitotic cyclins A2 and B1 in TGC as well as mitotic marker P-H3. In addition, up to 40% of nuclei underwent separation of two approximately equal nuclei out of the single giant nucleus<sup>[24]</sup>. The data suggest that the E2F7/E2F8 ablation promotes switch on mitotic cycle that restricts endocycling in TGC.

In liver, the combined ablation of *E2f1*, *E2f2*, and *E2f3* resulted in an increase of ploidy level<sup>[24]</sup> that suggests a possibility of hepatocyte ectopic endocycling. By contrast, ablation of E2F7/E2F8 leads to hepatocyte diploidization and to a decrease of the binuclear cell number. Thus, the canonical and atypical E2Fs exerts

opposite effect: the canonical ones promote complete progression of mitosis, whereas the ancient, atypical one increases endocycling. Interestingly, study of the global gene expression in hepatocytes showed that majority of genes upregulated in E2F7/E2F8-deficient hepatocytes were downregulated in the E2F1/E2F2/E2F3-deficient ones, many of these genes had annotated to be the cell cycle function related, in particular, to bound to the G2/M transition or to mitosis<sup>[24]</sup>.

Loss of E2F1 suppressed some, but not all mitotic defects in TGC and hepatocytes caused by E2F7/E2F8-deficiency. Ablation of mitotic cyclins in the E2F7/E2F8-deficient mice resulted in the greater ploidy in TGC, whereas hepatocytes in mice of similar phenotype showed similar ploidy as wild type mice. Thus, inhibiting the transcriptional network that signals G2/M progression or interfering with mitotic machinery (by cyclins A1/A2 ablation) reestablished the mitotic block and reinitiated higher ploidy levels<sup>[24]</sup>, probably, *via* endocycle



**Figure 2 Secondary trophoblast giant cells of rat placenta at the 16<sup>th</sup> day of gestation.** Note massive long cyokeratin-positive sprouts that embrace wide zones of decidua (De). TGC: Trophoblast giant cells.

progression. The data demonstrate that an intricate E2F network involving balanced and antagonistic activities of canonical (E2F1-3) activators and atypical (E2F7/E2F8) repressors plays, most probably, one of the key roles in the mammalian endocycle control. Perhaps, most dramatic manifestation of altering the balance in the E2F network is ectopic mitoses in the *E2F7/E2F8* deficient TGC. Interestingly, such a phenotype was lethal, whereas inactivation of these genes did not produce any effect on the growth of liver. Thus, in the liver, mitosis is a prerequisite of polyploidization whereas in the highly endopolyploid murine TGC, in distinct from liver, such a key event is almost complete reduction of mitosis (*i.e.*, endoreduplication). It may be important for some functional peculiarities of these cell types: for example, retention of mitotic potential may be important for maintenance of regenerative ability of the liver.

Difference in the cell cycle machinery was also detected by using transcriptome analysis of the murine TGC and megakaryocytes<sup>[51]</sup>. The authors compared the expression level of orthologous genes in the DNA replication pathway and cell cycle gene controlling the G1-S transition and S-phase. The components of the DNA replication machinery including the origin recognition complex, minichromosome maintenance and proliferating cell nuclear antigen genes were strongly expressed in the TGC lineage. In contrast, polyploid megakaryocytes exhibited a reduction in the expression of these genes. The TGC similar to *Drosophila* polytene cells, on the contrary, showed a reduction in the expression of mitotic genes, as compared with diploid control embryonic cells<sup>[51]</sup>, thereby confirming endocycle origin of their high level of ploidy<sup>[5,34,42]</sup>. Meantime, megakaryocytes showed increased expression of M-phase factors as compared to TGC or *Drosophila*<sup>[51]</sup>. Such a comparison suggests that genes controlling G1-S transition might play a key role in genome multiplication *via* endoreplication (in murine trophoblast and *Drosophila* salivary glands). By contrast, upregulation of mitotic events rather than S-phase factors play a key role in genome multiplication of megakaryocytes *via* endomitosis (reduced mitosis).

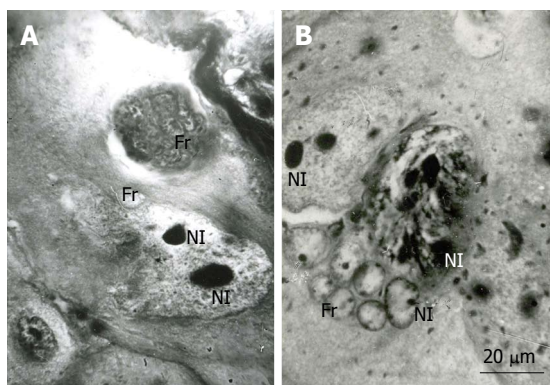
Significance of the two different ways of genome multiplication—endoreplication and reduced mitoses—can be shown on an example of functional organization of different trophoblast cell populations in the rodent placenta. According to our recent data, different trophoblast cell population show different patterns of intracellular cyokeratin localization. Long massive cyokeratin threads were found in the peripheral cytoplasm of the secondary TGC of rat and in their long sprouts by which they connected each other making continuous barrier at the border with decidua<sup>[7]</sup>. Similar cyokeratin immunostaining was observed in the trabecular spongiotrophoblast cells that line maternal blood sinuses. Clusters of low-ploidy proliferative and/or glycogen cells in the depth of JZ showed some weaker cyokeratin signal. Thus, the specific structure of giant trophoblast cells seems to provide a barrier between semiallogenic fetal and maternal tissues.

The east-european field vole *Microtus rossiaemeridionalis* provides another example of barrier function of TGC<sup>[52]</sup>. At the beginning of placentation TGC also form a continuous layer at the foeto-maternal interface. However, at midgestation, clusters of tightly attached low-ploid glycogen-rich junctional zone trophoblast cells progressively replace TGC thereby drawing them in the depth of the fetal part of placenta. Nevertheless, TGCs bound by their heavily cyokeratin-positive sprouts form a framework that holds other trophoblast cell populations and line lacunae with maternal blood.

Significance of specific TGC organization is confirmed by the data on the compound mutants on Cyokeratin 8 and 19<sup>[53]</sup>. In this case, there was an excessive number of TGC that however were not tightly attached to each other. Besides, *K8<sup>-/-</sup>*, *K8<sup>-/-</sup>/K19<sup>-/-</sup>* and *K18<sup>-/-</sup>/K19<sup>-/-</sup>* knockout conceptuses died by the moment of placenta formation and showed placental hemorrhaging<sup>[54-56]</sup>. This apparently caused flooding directly to fetal tissue where these trophoblast cells normally separate embryonic blood from maternal circulation.

The enormous sprouts of TGC probably play another role in rat placenta. Some secondary TGC protrude the decidua by means of large nipple-like highly cyokeratin-positive sprouts (Figure 2) that surround wide accumulations of decidua cells, the latter, probably, undergoing subsequent degradation<sup>[54]</sup>. This process was observed throughout the most of pregnancy and represents, most probably, a special kind of invasion that involves phagocytosis of decidua tissue by the trophoblast that may be called “group phagocytosis”. The trophoblast phagocytosis that provides histotrophic nutrition of embryo was described in detail in the endothelio- and hemochorial placenta of a range of mammalian species<sup>[6,57-59]</sup>. Thus, in mice, processes of trophoblast giant cells penetrate a layer of uterine epithelial cells and internalize the cells<sup>[58]</sup>; phagosomes with fragments of decidua cells are observed in TGC. Recently, transcriptome analysis showed high expression of scavenger receptor class B, member 1 (scarb 1) required for phagocytic activity of TGC<sup>[51]</sup>. Therefore, gigantism of TGC





**Figure 3** Fragmentation of the rat trophoblast giant cells. A and B: First the giant nuclei fall into two large nuclei, the latter then breaks down into numerous small fragments (Fr); B: The giant nucleolus falls down into several nucleoli (NI), then they form small nucleoli that seem to move into small nuclear fragments, some nucleoli in the nuclear fragments may be formed *de novo*. Heidenhain hematoxylin staining.

in this case probably allow them to perform invasion as well as partial degradation and phagocytosis of decidualized endometrium. It is interesting to note that one of the possible advantages of endoreduplication is that cell growth is accomplished without cell division that would imply a significant rearrangement of cytoskeleton<sup>[23]</sup>. Mitotic proliferation would prevent establishment of such a continuous system of the tightly attached phagocytosing TGC. Therefore, the endoreduplication allows TGC to combine growth at the restricted time period with formation of a barrier made of the tightly attached TGC that protect embryo from the immunological attack of the maternal organism and provide its histotrophic nutrition. It is also notable that the endoreduplication allows TGC growing without nuclear envelope disappearance. It also may be important, because isolation of genome inside the nuclear envelope may protect it from mutagenic effect of the degrading DNA of the phagocytosed maternal cells. In rare cases, in the field vole placenta we observed erythrocyte, *i.e.*, the anuclear cell, inside the phagocytic trophoblast cell undergone mitosis<sup>[6]</sup>.

An interesting example of the necessity of endopolyploidy for the “barrier” function is the recently obtained data on subperineural glia (SPG) of *Drosophila melanogaster* by Italic. These glial cells were highly polyploid, and ploidy correlated with brain mass. Inhibition of the SPG polyploidy caused rupture of the septate junctions necessary for the blood-brain barrier. Thus, the increased SPG cell size resulting from polyploidization is required to maintain the SPG envelope surrounding the growing brain<sup>[60]</sup>.

As to the trophoblast cell populations lying in the depth of the fetal part of placenta that first represent a proliferative pool of trophoblast, their primary steps of genome multiplication through reduced mitoses probably allow them to accumulate the great number of cells undergoing multidirectional differentiation into a number of subsets of cells within JZ and subsequently migrating into decidua (endovascularly and interstitially) as well as consisting trophoblast of labyrinth<sup>[6,7,61-63]</sup>.

## GENOME SEGREGATION: A TERMINAL STEP OF RODENT TROPHOBLAST GIANT CELL LIFESPAN

Beginning from the second half of pregnancy, the secondary giant trophoblast cells in rodent placenta undergo the so-called nuclear fragmentation when a part of giant nuclei break down into a number of the smaller nuclear fragments<sup>[5,7,64-66]</sup>. The fragmentation is preceded by depolytenization which is gained by disintegration of polytene chromosome bundles into the double chromosomes/endochromosomes; the process was described in detail earlier<sup>[5]</sup>. A similar process was recently investigated by using *in situ* hybridization the whole-chromosome labeling in the polytene chromosomes of *Calliphora erythrocephala* ovarian nurse cells<sup>[67]</sup>. In rat placenta, a portion of TGC undergoing nuclear fragmentation increases as the end of pregnancy approaches. Noteworthy, attenuation of DNA replication precedes nuclear fragmentation: this process begins in the nuclei that lost their capability for DNA synthesis<sup>[64]</sup>. Nevertheless, there were some reports about scarce murine <sup>3</sup>H-thymidine-labeled giant trophoblast cell nuclei undergoing fragmentation<sup>[66]</sup>.

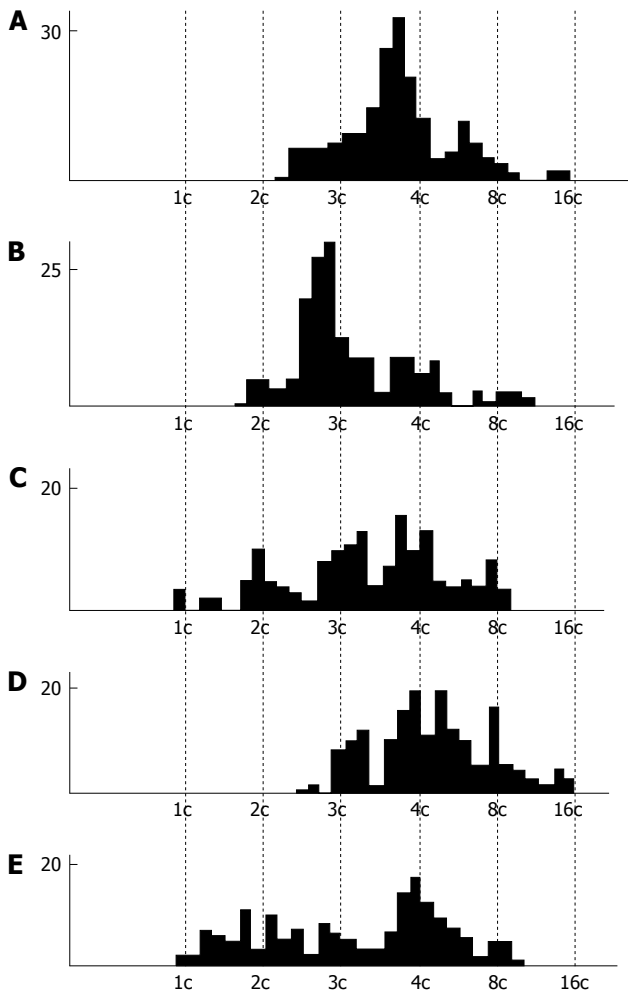
Interestingly, the process of nuclear fragmentation often begins with break down of the giant nuclei into two approximately equal parts (Figure 3). As a rule, one of these “parts” (“subnuclei”) undergoes more complete breakdown into numerous fragments. It is notable that ablation of mitosis progression regulators E2F1, E2F2, and E2F3 often resulted in break-down of mouse trophoblast giant nuclei into two approximately equal parts<sup>[24]</sup>. This suggests that although the process of giant nuclear segregation is not similar to mitosis, most probably, some elements of the mitotic machinery are involved in this mechanism.

Cytophotometric and cytofluorometric measurement of DNA content measurement showed a tendency to the whole-genome distribution of DNA into the nuclear fragments (Figure 4). In the rat secondary TGC, the histograms show quite clearly distinguishable peaks corresponding to different ploidy classes, *i.e.*, 1c, 2c, 4c, 8c, 16c, and 32c. However, there occurred incidence of some intermediate values of DNA content that could not be explained by the DNA synthesis, because in the nuclear fragments the S-phase is absent. In the mouse secondary TGC the tendency for the whole-genome distribution was also observed, but some peaks tend to be 3c and 6c values<sup>[5,64]</sup>. Similar data were obtained on the giant cells of the rabbit trophoblast<sup>[5]</sup>.

The DNA content measurement of nuclear fragments of the field vole secondary giant trophoblast cells showed more clear-cut correspondence to the distinct ploidy classes multiple to 2c<sup>[5,68]</sup>.

Behavior of natural chromosome markers, *i.e.*, sex chromatin body and nucleolus, were observed in the giant nuclei undergoing fragmentation. The inactivated X-chromosome that forms a condensed chromatin body in the interphase nucleus was undergone to successive





**Figure 4** DNA content in the nuclear fragments of secondary trophoblast giant cells of mouse (A and B) and rat (C-E) at the 17<sup>th</sup> (A), 19<sup>th</sup> (B), 14<sup>th</sup> (C), 16<sup>th</sup> (D), and 18<sup>th</sup> (E) day of gestation. Abscissa: The DNA content (arbitrary units, logarithmic scale), and ploidy, c; Ordinate: the number of nuclear fragments.

doubling in each cycle on endoreduplication of the rat secondary TGC of female embryos<sup>[1]</sup>. In the course of fragmentation, each nuclear fragment obtained a small condensed sex chromatin body attached to the nuclear envelope<sup>[5,65]</sup>. Similarly, each fragment contains a small nucleolus. At the onset of fragmentation, when the nucleus breaks down into several rather large fragments, the large nucleoli also break down into several ones that seem to be converted into separate fragments (Figure 3). Nevertheless, as to the final steps of fragmentation, it is not easy to decide whether the small nucleoli came from the initial giant nucleus or were synthesized *de novo*.

In the mouse secondary TGC, polytene chromosomes were characterized by the presence of large clear-cut heterochromatin blocks attached to the nucleolus (Figure 5). In the course of nuclear fragmentation, the heterochromatin blocks were separated into small blocks that were observed in the nuclear fragments near the nucleolus (Figure 5).

To determine whether the separate chromosomes are distributed into the nuclear fragments in correspondence

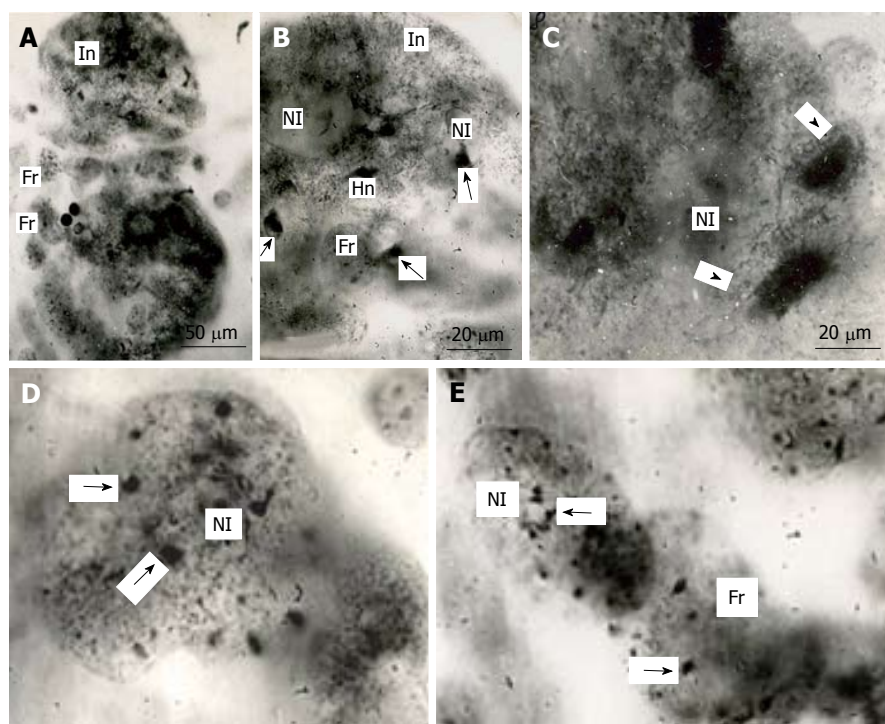
with ploidy level, we used the so-called gonosomal chromatin bodies (GCB) in the secondary TGC of the field vole *Microtus rossiaemeridionalis* by Italic that represent large heterochromatin blocks of X and Y chromosomes that form prominent condensed chromatin bodies in the interphase nucleus. According to the DNA content measurement in the nuclear fragments of the secondary TGC and their GCBs, each fragment contains GCB(s) in correspondence with its ploidy level<sup>[8]</sup>.

The above-mentioned data confirm that isolation of small “nuclei” inside the TGC is a process in which single chromosomes are regularly dispersed into nuclear fragments. Nevertheless, the data of cytophotometry leave open the possibility of some aneuploidy in the nuclear fragments. Thus, we can conclude that polyploid nuclei of TGC are transformed into the multinucleate giant cells, each of nuclei contains the euploid (1c, 2c, 3c, 4c, etc.) or the near-euploid set of endochromosomes.

The nuclear fragmentation seems to be a rare common process, and its mechanism is not elucidated in detail. Active participation of membranes of the nuclear envelope and its derivatives was observed in the rat secondary TGC<sup>[64]</sup>; bundles of intermediate filaments were also observed in invagination of nuclear envelope of nuclei into the process of fragmentation. However, such a process, most probably, makes possible separation of one or more genomes into separate fragments, and their number is not always divisible ( $2^n$ )c.

In the literature, very little data are known to draw analogies that could make an explanation of the above-mentioned chromosome distribution in nuclear fragments. Thus, in endomitotic nuclei of the seminal follicle wall of locust separation of some endochromosome groups corresponding to haploid, triploid, diploid, and hexadecaploid chromosome sets<sup>[69]</sup>, *i.e.*, segregation of a number of genomes not multiple to ( $2^n$ )c. The data also confirm the possibility of the spatial separation of genomes without their entering mitosis.

Another example of genome segregation is observed in polyploid nuclei of protists. In some Radiolarians, nuclei are polyploid, and nuclear division is achieved *via* genome segregation<sup>[70]</sup>. For example, in *Aulacantha* by Italic, there is a single polyploid, the so-called primary nucleus. In the course of nuclear division, a great number of large chromosomes are presumably gathered into chains. During endomitosis that precedes division, reproduction of the “gathered” chromosomes takes place. In the course of sporogenesis, fragmentation of the primary nucleus results in a number of secondary nuclei; the latter initially lie in the same cytoplasm, but later on the cytoplasm also undergoes subdivision into a number of secondary bolls. Then the secondary nuclei undergo a series of division and, as a result, the cytoplasm breaks down into mononuclear prespores that divide once again to give rise to zoospores. All this complicated process is considered as a breakdown of the primary nucleus into separate genomes (*via* their segregation) and depolyploidization is carried out in several steps.



**Figure 5** Distribution of heterochromatin blocks into nuclear fragments of mouse trophoblast giant cells. A: The initial nucleus (In) of trophoblast giant cells (TGC) in the process of fragmentation; B: The initial nucleus contains large clear-cut heterochromatin blocks (Hn, arrows) near nucleoli (NI); C: The nucleus contains non-classic polytene chromosomes with numerous distal loops (arrowheads); D and E: Nuclear fragments with small heterochromatin blocks (arrows) near nucleoli. Squash preparations, aceto-orcein staining. Fr: Nuclear fragments.

Since placenta is a provisory organ, it is felt that there should exist a mechanism of death of its cells, especially the peripheral ones) as the term approaches; it would enable placenta to separate at birth. In fact, the lifespan of the secondary rat TGC is strictly 22 d, which coincides with the length of pregnancy; it is not changed in culture and under conditions of transplantation under the kidney capsule<sup>[4]</sup>.

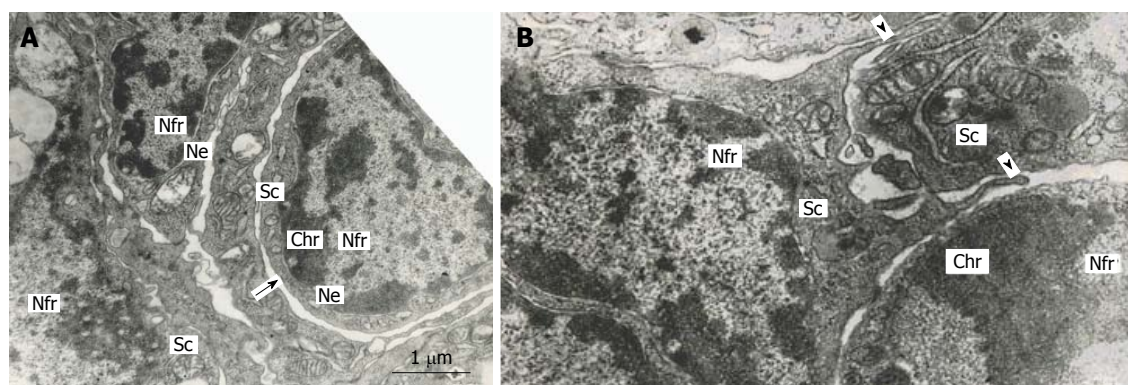
The multinucleate secondary TGC in rat placenta show some signs of apoptosis, attenuation of cell functions and degradation: condensation of chromatin located under the nuclear envelope of the nuclear fragments, inactivation of nucleolus, destruction of cytoplasmic organelles, *etc.*<sup>[5,7]</sup>. Nevertheless, the outer nuclear membrane of some nuclear fragments generates outgrowths continuous with the agranular endoplasmic reticulum, which, in turn, produces the double-membrane channels that delimit cytoplasmic mic territories around nuclear fragments. It is notable that the territories contain the whole set of cell organelles-mitochondria, Golgi complex, granular and agranular endoplasmic reticulum, numerous polysomes. In many cases, nuclear fragments with their cytoplasmic territories that can be called “subcellular compartments”, look quite viable; they are often isolated from the rest of cytoplasm of TGC that show signs of degradation. Moreover, sometimes this process results in isolation of subcellular compartments from other ones; however, in some cases the compartments connect each other by means of typical intercellular junctions (Figure 6). Occasionally it is possible to see as the pseudopodia

formed by one subcellular compartment as though try to surround another one (Figure 6). Therefore, the subcellular compartments of TGC may behave like full value cells<sup>[7]</sup>. Meantime, in other cases, these compartments look apoptotic.

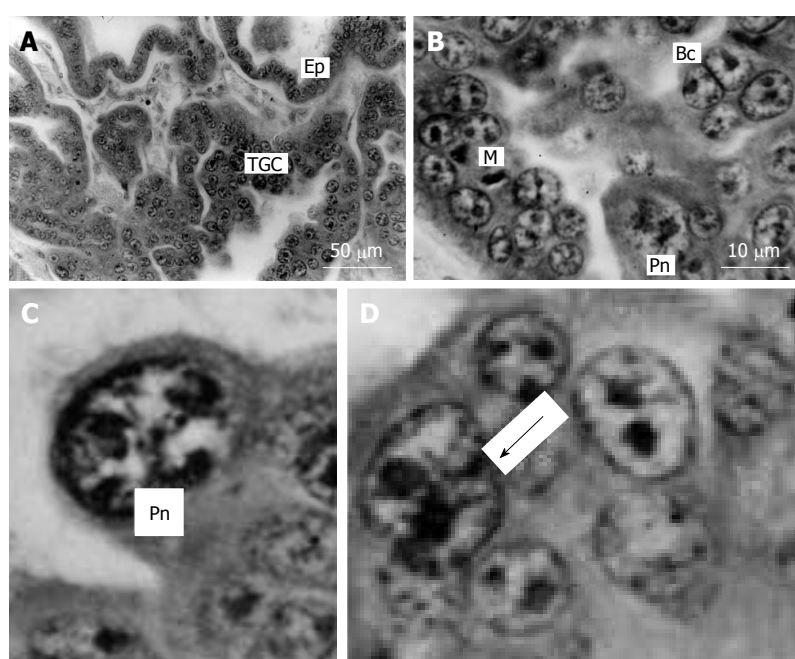
Thus, as the term approaches, TGC undergoing attenuation of reproductive, transcriptional, and other function simultaneously separates numerous subcellular compartments with near-euploid nuclei; the compartments may be both viable and apoptotic. It seems to be important that these fragments lose completely their capability for genome reproduction that would prevent renewal of proliferation of trophoblast cells. Nevertheless, the subcellular compartment formation may represent a reserve mechanism that preserve trophoblast genome for unknown functions. Further investigation may probably shed light on the significance of this phenomenon.

## POLYPLOIDIZATION OF SILVER FOX

In the silver fox the trophoblast invasiveness is manifested in other form as compared to rodents. In Carnivores, syncytiotrophoblast only partially destroys uterine epithelium and comes into contact with blood vessels, without destroying endothelium<sup>[71,72]</sup>. The main part of fox placental trophoblast in which active proliferation and polyploidization takes place is out of contact with the glandular zone of endometrium and forms the fetal part of placenta<sup>[73]</sup>. The giant trophoblast cells are scattered throughout the fetal part of placenta (Figure 7), forming



**Figure 6** Subnuclear compartments in the rat trophoblast giant cells are separated from each other by forming rather wide channels of endoplasmic reticulum (A, arrow); some of them show intercellular junctions (A and B), some compartments (B) produce pseudopodia-like outgrowths (arrowheads) moving to other compartments. Nfr: Nuclear fragments; Chr: Condensed chromatin; Sc: Subnuclear compartments; Ne: Nuclear envelope.



**Figure 7 Silver fox placenta.** A and B: Trabeculae of trophoblast and folds of uterine glandular epithelium (Ep) mutually contact each other, trophoblast giant cells (TGC) are scattered in the fetal part of placenta between accumulations of proliferative cells; B: Mitotic (M), binucleate (Bc) and cells with polytene nuclei (Pn); C: Nucleus with non-classic polytene chromosomes; D: A polyloid nucleus in the beginning of fragmentation (arrow). Meyer hematoxylin staining.

the largest accumulation near absorptional zones.

DNA content in the trophoblast cells corresponds predominantly to 4c-64c, the main peaks lying at 16 c and 32c; meantime, the highest ploidy corresponded to 64c<sup>[73]</sup>. Therefore, the ploidy level of the silver fox is lower as compared to the giant trophoblast cells in rodents. Another peculiarity is that there was a considerable deviation from (2<sup>n</sup>)c, with a tendency to  $2n \times 3c$  values and a great variety of the intermediate values suggesting a significant incidence of aneuploidy<sup>[73]</sup>.

Dynamics of polyploidization during placentation showed several steps of increase and decrease of ploidy levels. At the 20<sup>th</sup> day of pregnancy the trophoblast nuclei reach the highest ploidy levels (Figure 8). At the next, 21<sup>st</sup> day, percentage of nuclei of higher ploidy levels de-

creases. This trend is also seen at the 22<sup>nd</sup> day: the “divergence” of ploidy is observed, the percentage of diploid and tetraploid nuclei rises to 10% and 20%, respectively; simultaneously the percentage of 32c and 64c nuclei also increases as compared to that at the 21<sup>st</sup> day, and trophoblast cell population reaches the highest ploidy levels-128c and 256c. At the 23<sup>rd</sup> day, a new “wave” of polyploidization, similar to 20 d, takes place. The reason of such a fluctuation may be accounted for appearance of new zones of trophoblast cell proliferation. Therefore a part of trophoblast cells divide mitotically providing diploid and low-polyploid cells whereas other cell undergo endocycles to reach high ploidy levels. Simultaneously, pictures similar to nuclear fragmentation described in the rodent giant trophoblast cells (see above) were observed

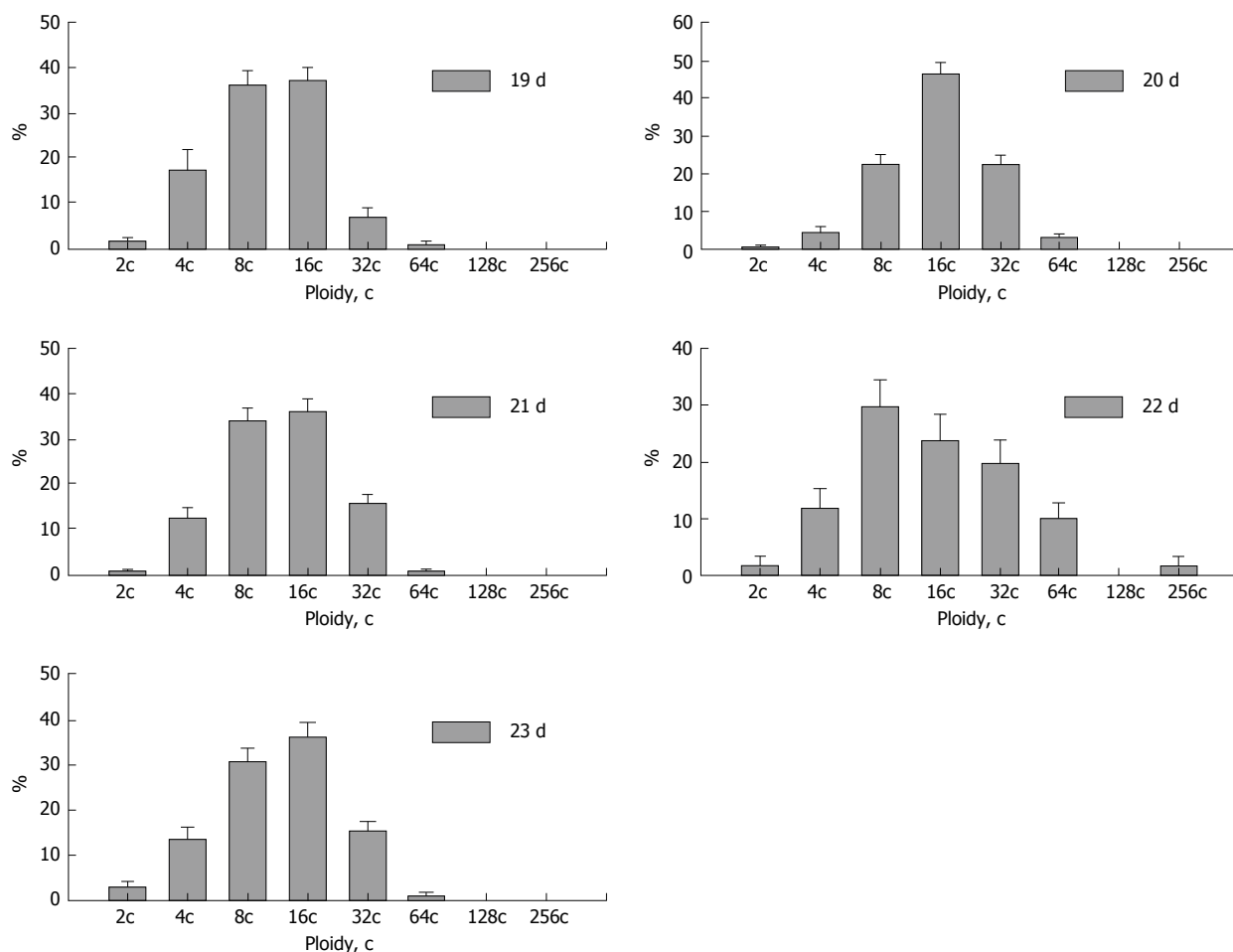


Figure 8 Dynamics of polyploidization of the trophoblast cells in silver fox placenta.

in the fox trophoblast<sup>[73]</sup>; such a genome segregation also may be a reason for a decrease of the ploidy level.

The DNA content measurement in mitotic figures showed 4c, 8c, and 16c mitotic figures that confirm the ability of 2c-8c to divide mitotically. The presence of binuclear cells, polyploid mitoses, and atypical metaphases and anaphases including multipolar mitoses indicates acytokinetic and uncompleted mitoses as a way by which the cells reach low (4c-8c) ploidy levels<sup>[6]</sup>. The higher ploidy levels may be attained by switch to endoreduplication cycle; it is confirmed by the presence of nuclei with characteristic non-classic polyteny (Figure 7). Interestingly, histograms of DNA content in the mitotic figures also show a tendency for “triploidy”. It should be noted that signs of aneu- or/and triploidy are intrinsic of fox trophoblast: the uterine epithelial cells that also polyploidize reaching 4c-8c show quite euploid DNA content histograms<sup>[73]</sup>.

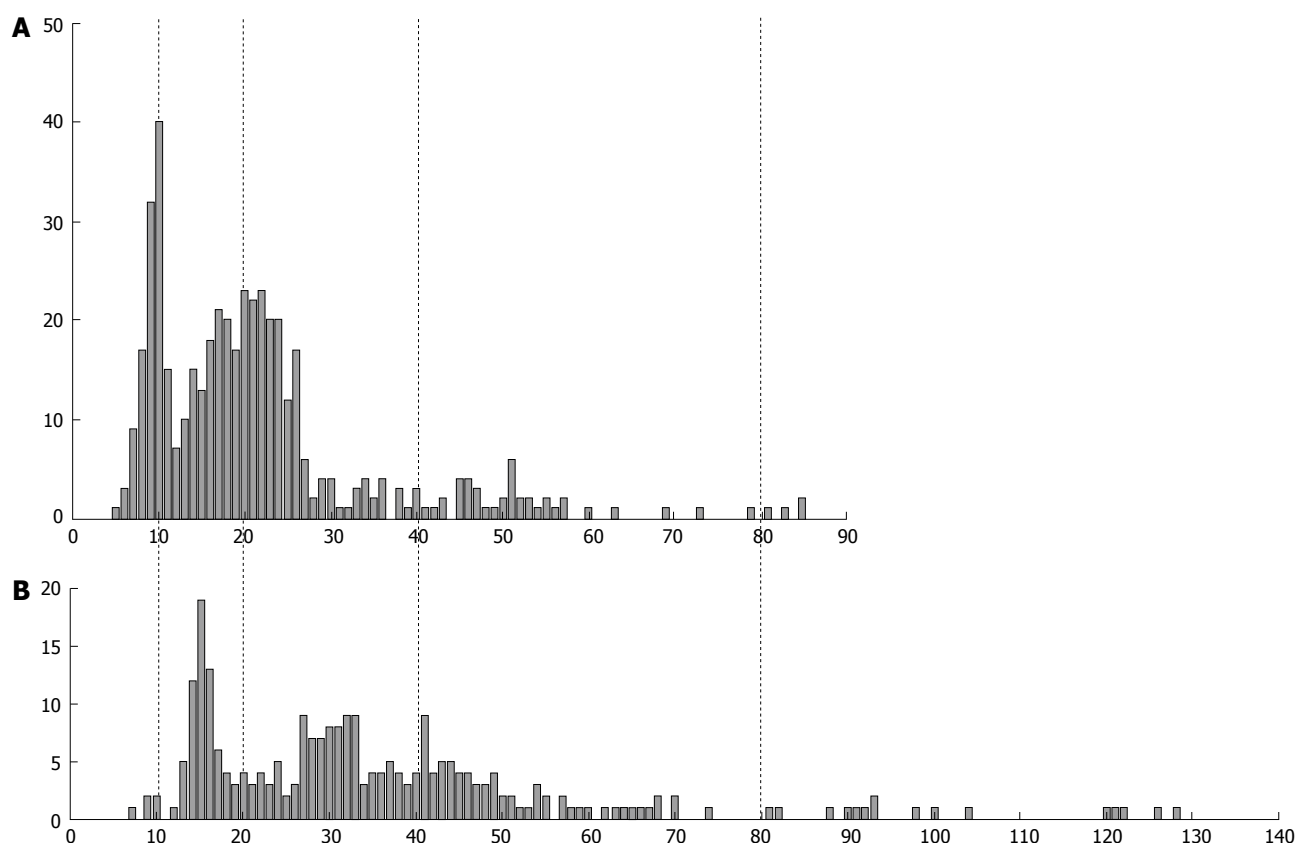
The reason for deviation from euploidy in the fox trophoblast cells was not clarified completely. However, it cannot be ruled out that it may result from depolyploidization as a result of the multipolar mitoses and the process similar to the nuclear fragmentation described in the previous chapter.

## PECULIARITIES OF HUMAN TROPHOBLAST GENOME MULTIPLICATION

Human trophoblast shows, at the first glance, a great difference in the trophoblast polyploidization as compared to rodents, which probably is accounted for by human placenta formation and growth characteristics. During the most of pregnancy, the trophoblast continuous layer at the border of decidua consists of a layer of syncytiotrophoblast, whereas several local zones of the intra-uterine invasion of extravillous trophoblast (EVT) are concentrated at tips of the anchoring villi<sup>[74]</sup>. Numerous data of caryologic analysis as well as DNA flow cytometry showed the prevalence of the diploid or, sometimes, near-diploid chromosome set in the human trophoblast cells<sup>[75-77]</sup>. Meantime, the human trophoblast invasion show many analogies in regularities of the genome reproduction to trophoblast of rodent placenta.

Thus, in human, like in rat placenta, the deep intrauterine interstitial and endovascular invasion is accomplished at the complete cessation of DNA replication<sup>[5,44,74,78]</sup>. Lack of genome replication of the invading





**Figure 9** The difference in DNA content in human extravillous trophoblast cells invaded endometrium in placentae of the first trimester between two individuals (A and B). Abscissa: The DNA content (arbitrary units); Ordinate: The number of cells.

trophoblast may prevent the ectopic proliferation of the trophoblast cells both within the uterus and in other parts of the maternal organism.

Polyploidization of the extravillous trophoblast cells also show similarities with rodent placenta. Thus, the proliferative EVT attached to the basal membrane of tip of villi are mostly diploid, but the ploidy increases progressively to 4-8c in process of approach to the border of decidua<sup>[7,79]</sup>. Thus, like in rat, mouse, and field vole, the highest ploidy is characteristic of the human trophoblast cell layer that borders the semiallogenic maternal tissues.

An interesting peculiarity of the human EVT genome reproduction is a ploidy divergence that takes place at the moment of intrauterine invasion (Figure 9)<sup>[7,79]</sup>. The fraction of 8c cells increased by up to 9.7%; besides, a few 16c nuclei appeared, whereas, on the other hand, the percentage of diploid cells also rise noticeably-on average, from 20% to 30%. Such a tendency was also found in the EVT invaded up to myometrium: the percentage of diploid nuclei exceeded 40%, whereas a number of highly polyploid 8c and 16c nuclei persisted in this zone.

One of peculiarities of the trophoblast cell polyploidization is a great variety of the genome multiplication between individual placentae<sup>[79]</sup>. Thus, Figure 9 presents histogram of the DNA content distribution in the extravillous trophoblast cells invading decidualized endometrium in placentae of two pregnant women. We can see that one of them shows more or less clear-cut peaks

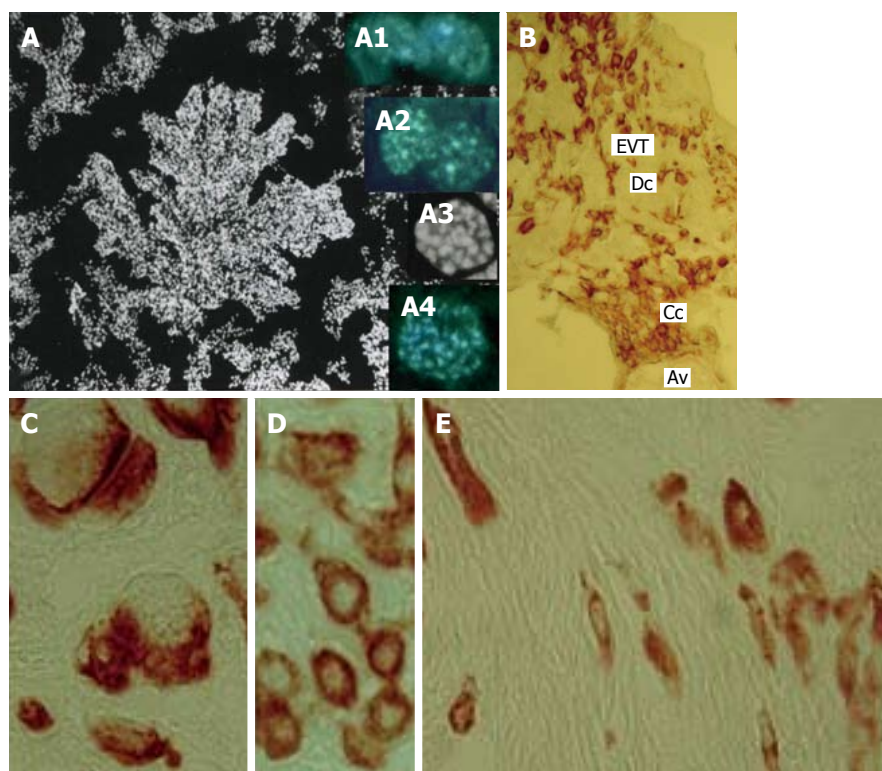
corresponding to 2c and 4c; the latter ploidy class prevails, some cells being able to reach as many as 16c. Another sample shows more extended histogram, in which the main peaks tend to 3c and 6c, the greatest number of cells being located between 4c and 8c. Besides, in the second sample, the noticeable number of cells exceeded 16c. A correlation between DNA content values not multiple to  $(2^n)c$  and the highest ploidy level is seen here; besides, a tendency for  $3c \times 2n$  values is characteristic of the samples of the higher ploidy level.

Endoreduplication of human cervical trophoblast with prevalence of tetraploid cells was also found by means of caryological analysis with fluorescent *in situ* hybridization (FISH) staining of the chromosomes X, Y, and 21<sup>[80]</sup>.

Endocycle progression in the human EVT is confirmed by downregulation of mitotic cyclins A and B1 alongside the invasion pathway and peculiar expression of Cyclins D and E as well as p57<sup>kip2</sup><sup>[7]</sup>.

According to our data, a great number of DNA content values not multiple to  $(2^n)c$  is observed in the beginning of the EVT invasion. When the cells penetrate endometrium and myometrium, more clear-cut peaks tending to  $(2^n)c$  are observed<sup>[7,79]</sup>.

Interestingly, some data suggest the necessity of aneuploidy for normal EVT trophoblast invasion. Thus, molecular cytogenetic data showing that approximately 20% to 60% of interphase EVT invasive cells in the normal



**Figure 10 Human trophoblast cells undergoing polyploidization.** A: A squash spread of chorionic villus stained with 4',6-diamidino-2-phenylindole; A1: Binucleate cell in interphase; A2: Binucleate cell in prophase; A3: Endometaphase; A4: Endoanaphase; B: A cell column (Cc) at the tip of anchoring villus (Av) generates a pool of extravillous trophoblast cells (EVT) capable for invasion of decidualized endometrium (Dc); C: Multinucleate EVT invaded decidua; D: Interstitial EVT of moderate ploidy; E: Small elongated low-ploidy EVT that reach myometrium.

pregnancies acquired aneusomies involving chromosomes X, Y, or 16<sup>[81]</sup>. The incidence of aneuploidy positively correlated with gestational age and differentiation to the invasive phenotype. Scoring 12 chromosomes in flow-sorted cytotrophoblasts showed that more than 95% of the cells were hyperdiploid. Thus, aneuploidy appears to be an important component of normal placentation, perhaps limiting the proliferative and invasive potential of cytotrophoblasts within the uterus<sup>[81]</sup>.

A series of recent investigations allow some authors to put forward a concept that transition from endoreduplication to polyploidy and then to aneuploidy represents a genetically diverse pool of cells<sup>[18-20]</sup>. The authors suggest that a set of these genetically changed cells may be useful and amplified under conditions of stress<sup>[18]</sup>. In the case of invading trophoblast we can assume that aneuploidy that accompanies the complicated processes of the trophoblast genome reproduction may rise several genotypes that may promote cells to survive under the stress conditions inside semiallogenic maternal tissues. It may be suggested that the most stressful condition for EVT cells is a moment of overriding the border with decidualized endometrium; at this moment the optimal genotypes are selected; later on, deviation of the euploidy decreases. It cannot also be ruled out that simultaneously the cells of higher ploidy are selected to invade the proximal part of endometrium and the small low-ploidy elongated cell to invade deep up to myometrium (Figure 9).

It is not easy to explain the possible significance of

tendency to triploidy in the invading human EVT (that may be analogous to the same tendency observed in silver fox). Triploidy of human trophoblast was reported in several papers<sup>[76,82]</sup>. Very often, it is observed under pathological conditions connected with ectopic or disturbed trophoblast invasion, for example, in the hydatidiform mole<sup>[76]</sup> or severe preeclampsia<sup>[82]</sup>. Meantime, basing on the above-mentioned data on the possibility of non-mitotic genome segregation (nuclear fragmentation) as a regular step of trophoblast cell lifespan, we can suggest that such a process may result in separation of the chromosome set non-multiple to (2)<sup>n</sup>c. This process, most probably, does not include the way of exact distribution of all chromosomes into the daughter cells characteristic of mitosis, that is why it may be suggested that such a way of depolyploidization of trophoblast cells may result in aneuploid genotypes.

The reason for such a variability may be accounted for by the ways of human EVT polyploidization. Endomitosis and non-classic polyteny were observed in the trophoblast cells in normal human pregnancies<sup>[32,37,83]</sup> as well as in hydatidiform moles<sup>[32]</sup>.

According to our data, the processes like endoreduplication or/and endomitosis, most probably, prevail in the EVT invasive pathway. The Figure 10 shows a chorionic villus tip stained with DAPI. A great number of cells with numerous chromocenters as well as nuclei at different stage of endomitosis are present there (Figure 10)<sup>[38]</sup>. Meantime, a number of nuclei with enlarged chromo-

centers, whose number was not increased with increasing ploidy, suggest a possibility of passages of several rounds of polytenization<sup>[38]</sup>. Thus, it cannot be ruled out that a relatively low ploidy level in the human placenta prone to aneuploidization may be linked with endoreduplication/endomitosis that, theoretically, may be involved in single/double/triple genome segregation processes.

## GENOME MULTIPLICATION IN THE TROPHOBLAST CELLS OF RUMINANTS

Ruminants represent a mammalian group with the so-called epitheliochorial placenta, in which trophoblast invasiveness is minimal: the trophoblast cells attach to the uterine epithelium mostly without its degradation<sup>[84]</sup>; meantime, the trophoblast cells express the same integrins as the highly invasive human trophoblast cells<sup>[85]</sup>. The bovine trophoblast cells that come into close contact to the uterine epithelium retain their mitotic activity and reach the ploidy level 4c-8c *via* reduced (restitutional) mitoses<sup>[11,30]</sup>. Tripolar mitoses suggesting a possibility to reverse to the lower ploidy level also are present there. It is quite probably that lack of deep invasion that does not imply phagocytosis of epithelium, blood or decidual cells by the trophoblast does not require switch to the endoreduplication cycle, although a low level of ploidy is still necessary for such a mode of feto-maternal interaction.

The most striking example of polyploid mitoses was described in alpaca trophoblast. Elevated nuclear DNA contents in the giant trophoblast cells of alpaca could be achieved by modified cell cycles with a complete lack of mitosis (endoreduplication) or with incomplete mitoses<sup>[12]</sup>. Electron microscope observation made on serial sections revealed that TGCs are truly multinucleate with several highly lobulated nuclei. Feulgen staining showed that TGC nuclei have the higher DNA content than nuclei of other trophoblast cells. The number of argyrophilic nucleolar organizer regions in nuclear profiles of TGC was between 15 and 100; numerous nucleoli suggest polyploidization, in which mitoses take part, as this was observed in the rat decidual cells<sup>[86]</sup>. In the latter case, numerous decidual cells are formed by active mitotic divisions; numerous binucleate cells were observed, there were many enlarged mitoses of irregular form, a part of mitoses were tri- and tetrapolar. Very often, bi-tri- and tetranuclear cells were observed that could result from multipolar mitoses, the nuclear ploidy reached 8c. Numerous nucleoli (up to 20) were also observed in the lobulated nuclei. As to alpaca trophoblast, even larger multipolar mitotic figures with maximal diameters of 80  $\mu\text{m}$  were observed in placentas on gestation days 264 and 347. No cytokinesis was seen in TGC<sup>[12]</sup>. The authors note that subsequent acytokinetic mitoses may lead to accumulation of chromosomes and centrioles in TGC. With increasing ploidy levels, the shape of these polyploidizing mitoses becomes more irregular. The restitution of nuclei after these complex multipolar mitoses is likely to result in the irregular nuclear shape in TGC.

Therefore, it is an exceptional example of polyploidization *via* restitutional (reduced) mitoses that probably may lead to the high ploidy level. It seems doubtful that in this case the multipolar mitoses may result in depolyploidization.

Thus, a conclusion can be made that shallow-invasive trophoblast cells of ruminants involve polyploidization *via* restitutional (reduced) mitoses.

## CONCLUSION

The data considered here demonstrate that trophoblast cells of different mammalian species are characterized by different modes of multiplication of their genome that, probably, is linked with their ploidy level, capability for further proliferation, necessity of irreversible or, on the contrary, of reversible polyploidization that, in turn, most probably, is accounted for by the trophoblast cell specific function.

One of the most important advantages of polyploidy for the trophoblast cells contacting semiallogenic maternal tissues may be the delay of proliferation to avoid segregation of the damaged chromosomes<sup>[19]</sup>. Besides, multifold genome doubling makes the endocycling cells more resistant to mutagens. Thus, the highly endopolyploid murine TGC are much more resistant to irradiation than the low-ploid trophoblast cells<sup>[42]</sup>. Interestingly, endoreduplication as a response to mutagens can be induced experimentally, and some regulatory pathways were recently revealed. Thus, following double strand breaks induction in the root tips of *Arabidopsis* by Italic the cells switch to endoreduplication<sup>[87]</sup>. This cell alteration requires the plant-specific transcription factor Suppressor gamma response 1 which transmits signals from the conserved Ataxia Telangiectasia mutated and Ataxia Telangiectasia-mutated and RAD3-related DNA damage sensor kinase<sup>[86]</sup>. This DNA break response produces transcriptional changes that are consistent with downregulation of mitotic factors and upregulation of cell cycle genes that promote endoreduplication.

Recently there were obtained some other confirmations of significance of non-mitotic polyploidization under condition of DNA damage. In *Drosophila* by Italic, endoreduplication cells acquire resistance to DNA damage through a mechanism involving the silencing of cell death genes<sup>[19,88]</sup>. Similarly, endoreduplication mouse trophoblast cells that undergo endoreduplication also downregulate the DNA damage response. During differentiation of trophoblast stem cells into polyploid TGC, the protein level of damage-responsive Chk1 is decreased providing for endoreduplication. This decrease in Chk1 enables polyploid trophoblast cells to evade apoptosis through suppression of the DNA damage pathway<sup>[89,90]</sup>.

It seems to be obvious that there are two main ways of genome multiplication: endoreduplication that involves downregulation of mitotic events and reduced mitosis ("endomitosis") in which entrance into mitosis is a prerequisite of genome multiplication. Endoreduplica-

tion allows to combine growth and specific functioning of cells that retain their peculiar organization. On the contrary, the cells that polyploidize *via* reduced mitoses retain their mitotic potential necessary, for example, for accumulation of a great number of cells, or for regeneration. In case of rodent trophoblast, the highly polyploid TGC undergoing endoreduplication serve a barrier between semiallogenic tissues whereas highly proliferative low-ploid trophoblast cell populations give rise to numerous JZ and labyrinth trophoblast cells.

Endoreduplication is a characteristic of highly invasive trophoblast cells. It is confirmed by comparison of placentation of mammalian species. Thus, highest invasiveness is characteristic of rodent and human placenta in which invasive trophoblast cells undergo endoreduplication whereas reduced mitoses are observed in the low invasive epitheliochorial placenta of ruminants.

Another advantage of polyploidization that is widely discussed now is a possibility to gain a variety of genome changes. Such a possibility was recently demonstrated in hepatocytes<sup>[17,18,21]</sup>. Apart from polyploidization by acytokinetic and subsequent normal (polyploidizing) mitoses, hepatocytes may undergo depolyploidization that result from multipolar mitoses<sup>[18]</sup>. Indeed, during multipolar mitosis, microtubules from different poles of spindle can be attached to a single kinetochore, and failure to repair such merotelic attachment can lead to incomplete chromosome segregation. In this case, chromosome bridges and laggings were observed. Karyotype analysis showed high frequency of aneuploidy in the normal murine liver: nearly 25% of hepatocytes from the 3-wk old mice were aneuploid, the aneuploidy increasing to 70% in the 4-15-mo old mice<sup>[18]</sup>. Interestingly, the entire chromosomes were gained and lost in this case, and structural rearrangement was rarely seen; besides, all chromosomes of genome were affected equally. One or more chromosomes were gained or lost by each aneuploid hepatocytes, and occasionally chromosome gains balanced losses mimicking the normal chromosome number. That is why, probably, aneuploidy did not lead to noticeable deviation from (2<sup>n</sup>)c seen in histograms of the DNA content distribution in hepatocytes measured by DNA cytometry<sup>[91]</sup>. The karyotype and FISH analysis also revealed a significant level (30%-90%) of aneuploidy in the human liver<sup>[17]</sup>. Strikingly, gain and loss of chromosomes in hepatocytes under stress conditions may result in selection of the specific karyotype that can result in adaptation to injury<sup>[20]</sup>.

The statement that polyploidy is often accompanied by aneuploidy may be accounted for by the fact that aneuploidy often results from depolyploidization that, in turn, requires polyploid cell formation. Meantime, it should be kept in mind that genome multiplication, in many cases, is irreversible. In this connection, a spectacular example is given by the ovarian follicle nurse cells<sup>[42]</sup>. When the cells reach 32c, centrioli leave their position near nucleus and move into oocyte through the cytoplasmic bridges. This is a way by which the cell "burns the bridges" to ploidy reversal and renewal of mitotic

divisions. In the majority of cells normally undergoing endocycles, the genome multiplication is irreversible<sup>[22]</sup>. Therefore, aneuploidy may arise in the modified cell cycles prone to depolyploidization. Actually, at present little is known as to which cell types may be capable for the spontaneous non-pathologic aneuploidization.

Endopolyploidization as an escape route from cell death has been investigated on cancer cells. It has been demonstrated that after irradiation the great majority of cells die due to a mitotic catastrophe<sup>[92]</sup>. However, some cells escape the mitotic catastrophe and polyploidize. A few endopolyploid cells undergo depolyploidization and create a set of para-diploid viable cells capable for mitoses that may give rise to subclones<sup>[92,93]</sup>. Recent data of this research group indicate that tumor cells can induce opposing processes of senescence and stem cell generation in response to these treatments whose biological significance and molecular regulation currently are poorly understood<sup>[94]</sup>. Although cellular senescence is typically considered a terminal cell fate, it was recently shown to be reversible in a small population of polyploid cancer cells induced after DNA damage. Overcoming genotoxic insults are associated with reversible polyploidy. The subsequent depolyploidization results in evoking the self-renewal potential in survived cells<sup>[94]</sup>.

In rodent placenta, the secondary TGC provide an example of irreversible differentiation that, however, results in formation of the low-ploidy subcellular compartments that may behave like viable cells except for renewal of their capability for proliferation. Theoretically, depolyploidization may take place in the JZ and labyrinth trophoblast cells of rat and field vole undergoing the polyploidizing mitoses, some amount of multipolar mitoses being observed here<sup>[6]</sup>. Deviation from (2<sup>n</sup>)c in human and silver fox trophoblast<sup>[7,73,79]</sup> suggests a possibility of aneuploidy and other chromosome changes (aberrations, *etc.*) that may be even more strongly pronounced in trophoblast cells of these species. It cannot be ruled out that in human trophoblast a high degree of genomic changes also may be important for selection of cells able to survive under conditions of deep intrauterine invasion. In silver fox, a series of endoreduplication and proliferation cycles involving depolyploidization *via* multipolar mitoses and a process similar to nuclear fragmentation would also form a multifunctional system resistant to injury from the maternal tissues.

It seems to be important that different mammalian species have different programs of genome multiplication. Thus, placenta of rodents with strict spatial location of high- and low-ploidy trophoblast cells, with low incidence of deviation from (2<sup>n</sup>)c, probably is accounted for by their strict developmental program allowing them to produce their progeny during a limited period of short pregnancy. The mammalian species with a long period of pregnancy probably stick to another strategy allowing them to generate a more specific response to stress factors that may appear occasionally during months of intrauterine development. In these cases, a



more diverse genome changes followed by selection of favorable genotypes may be useful to maintenance of placental functions.

# ACKNOWLEDGMENTS

The authors are grateful to Leonid Z Pevzner, MD, PhD, the chief interpreter of Institute of Cytology of the Russian Academy of Sciences for his help in translation and edition of the manuscript.

# REFERENCES

- 1 **Zybina EV**, Mosjan IA. Sex chromatin bodies during endomitotic polyploidization of trophoblast cells. *Tsitologiya* 1965; **9**: 265-272
- 2 **Zybina EV**, Kudryavtseva MV, Kudryavtsev BN. Polyploidization and endomitosis in giant cells of rabbit trophoblast. *Cell Tiss Res* 1975; **160**: 525-537 [DOI: 10.1007/BF00225769]
- 3 **Nagl W**. Endopolyploidy and polyteny in differentiation and evolution. Toward an understanding of quantitative and qualitative variation of nuclear DNA in ontogeny and phylogeny. Amsterdam-N.Y.- Oxford: North-Holland publishing Company, 1978: 283
- 4 **Barlow PW**, Sherman MJ. Cytological studies on the organization of DNA in giant trophoblast nuclei of the mouse and rat. *Chromosoma* 1972; **47**: 119-131 [DOI: 10.1007/BF00331800]
- 5 **Zybina EV**, Zybina TG. Polytenic chromosomes in mammalian cells. *Int Rev Cytol* 1996; **165**: 53-119 [PMID: 8900957 DOI: 10.1016/S0074-7696(08)62220-2]
- 6 **Zybina TG**, Zybina EV. Cell reproduction and genome multiplication in the proliferative and invasive trophoblast cell populations of mammalian placenta. *Cell Biol Int* 2005; **29**: 1071-1083 [PMID: 16316755 DOI: 10.1016/j.cellbi.2005.10.015]
- 7 **Zybina T**, Zybina E. Cell cycle modification in trophoblast cell population in the course of placenta formation. In: DNA replication and related cellular processes. Rijeka: InTech, 2011: 258. Available from: URL: <http://www.intechopen.com/articles/show/title/cell-cycle-modification-in-trophoblast-cell-populations-in-the-course-of-placenta-formation>.
- 8 **Zybina EV**, Bogdanova MS, Stein GI, Vlasova TD, Zybina TG. Endopolyploidization and the interstitial invasion of the supergiant trophoblast cells of the field vole *Microtus rossiaemeridionalis*. *Tissue Cell* 2009; **41**: 362-366 [PMID: 19329135]
- 9 **Zybina TG**, Severova EL, Zybina EV, Dyban AP. A study of silver-stained nucleolus organizer regions in metaphase chromosomes of cambial cells of mouse and rat placenta. *Chromosome Res* 1997; **5**: 142-144 [PMID: 9146918 DOI: 10.1023/A:1018474309639]
- 10 **Zybina TG**, Zybina EV, Bogdanova MS, Stein GI. Quantitative investigation of reproduction of gonosomal condensed chromatin during trophoblast cell polyploidization and endoreduplication in the East-European field vole *Microtus rossiaemeridionalis*. *Reprod Biol Endocrinol* 2003; **1**: 32 [PMID: 12725646 DOI: 10.1186/1477-7827-1-32]
- 11 **Klisch K**, Pfarrer C, Schuler G, Hoffmann B, Leiser R. Tripolar acytokinetic mitosis and formation of feto-maternal syncytia in the bovine placenta: different modes of the generation of multinuclear cells. *Anat Embryol (Berl)* 1999; **200**: 229-237 [PMID: 10424879 DOI: 10.1007/s004290050275]
- 12 **Klisch K**, Bevilacqua E, Olivera LV. Mitotic polyploidization in trophoblast giant cells of the alpaca. *Cells Tissues Organs* 2005; **181**: 103-108 [PMID: 16534204 DOI: 10.1159/000091099]
- 13 **Wang Z**, Zhang Y, Kamer D, Lees E, Ravid K. Cyclin D3 is essential for megakaryocytopoiesis. *Blood* 1995; **86**: 3783-3788
- 14 **Ravid K**, Lu J, Zimmet JM, Jones MR. Roads to polyploidy:

- the megakaryocyte example. *J Cell Physiol* 2002; **190**: 7-20 [PMID: 11807806 DOI: 10.1002/jcp.10035]
- 15 **Hixon ML**, Gualberto A. Vascular smooth muscle polyploidization--from mitotic checkpoints to hypertension. *Cell Cycle* 2003; **2**: 105-110 [PMID: 12695657 DOI: 10.4161/cc.2.2.341]
  - 16 **Gui Y**, He GH, Walsh MP, Zheng XL. Predisposition to tetraploidy in pulmonary vascular smooth muscle cells derived from the Eker rats. *Am J Physiol Lung Cell Mol Physiol* 2007; **293**: L702-L711 [PMID: 17575014 DOI: 10.1152/ajplung.00016.2007]
  - 17 **Duncan AW**. Aneuploidy, polyploidy and ploidy reversal in the liver. *Semin Cell Dev Biol* 2013; **24**: 347-356 [PMID: 23333793 DOI: 10.1002/hep.26233]
  - 18 **Duncan AW**, Taylor MH, Hickey RD, Hanlon Newell AE, Lenzi ML, Olson SB, Finegold MJ, Grompe M. The ploidy conveyor of mature hepatocytes as a source of genetic variation. *Nature* 2010; **467**: 707-710 [PMID: 20861837 DOI: 10.1038/nature09414]
  - 19 **Fox DT**, Duronio RJ. Endoreplication and polyploidy: insights into development and disease. *Development* 2013; **140**: 3-12 [PMID: 23222436 DOI: 10.1242/dev.080531]
  - 20 **Duncan AW**, Hanlon Newell AE, Bi W, Finegold MJ, Olson SB, Beaudet AL, Grompe M. Aneuploidy as a mechanism for stress-induced liver adaptation. *J Clin Invest* 2012; **122**: 3307-3315 [PMID: 22863619 DOI: 10.1172/JCI64026]
  - 21 **Pandit SK**, Westendorp B, de Bruin A. Physiological significance of polyploidization in mammalian cells. *Trends Cell Biol* 2013; **23**: 556-566 [PMID: 23849927]
  - 22 **Brodsky VYa**, Uryvaeva IV. Genome multiplication in growth and development. Cambridge: Univ, Press, 1985: 350
  - 23 **Edgar BA**, Orr-Weaver TL. Endoreplication cell cycles: more for less. *Cell* 2001; **105**: 297-306 [PMID: 11348589 DOI: 10.1016/S0092-8674(01)00334-8]
  - 24 **Chen HZ**, Ouseph MM, Li J, Pécot T, Chokshi V, Kent L, Bae S, Byrne M, Duran C, Comstock G, Trikha P, Mair M, Senapati S, Martin CK, Gandhi S, Wilson N, Liu B, Huang YW, Thompson JC, Raman S, Singh S, Leone N, Machiraju R, Huang K, Mo X, Fernandez S, Kalaszczynska I, Wolgemuth DJ, Sicinski P, Huang T, Jin V, Leone G. Canonical and atypical E2Fs regulate the mammalian endocycle. *Nat Cell Biol* 2012; **14**: 1192-1202 [PMID: 23064266 DOI: 10.1038/ncb2595]
  - 25 **Rumyantsev PP**. Growth and hyperplasia of cardiac muscle cells. Chur, Switzerland: Harwood Acad, Publ, 1991
  - 26 **Kudryavtsev B**, Anatskaya OV, Nilova VK, Komarov SA. Interrelation between myofibrillar and mitochondrial apparatus and cardiomyocyte polyploidy and hypertrophy in mammals. *Tsitologiya* 1997; **39**: 946-964
  - 27 **Nagata Y**, Jones MR, Nguyen HG, McCrann DJ, St Hilaire C, Schreiber BM, Hashimoto A, Inagaki M, Earnshaw WC, Tokokoro K, Ravid K. Vascular smooth muscle cell polyploidization involves changes in chromosome passenger proteins and an endomitotic cell cycle. *Exp Cell Res* 2005; **305**: 277-291 [PMID: 15817153 DOI: 10.1016/j.yexcr.2004.12.028]
  - 28 **Nagl W**. Cdc-2 kinases, cyclins and the switch from proliferation to polyploidization. *Protoplasma* 1995; **188**: 143-150 [DOI: 10.1007/BF01280365]
  - 29 **Therman E**, Kuhn E. Mitotic modifications and aberrations in cancer. *CRC Crit Rev Oncogenesis* 1989; **1**: 293-305
  - 30 **Klisch K**, Thomsen PD, Dantzer V, Leiser R. Genome multiplication is a generalised phenomenon in placental and interplacental trophoblast giant cells in cattle. *Reprod Fertil Dev* 2004; **16**: 301-306 [PMID: 15304202 DOI: 10.1071/RD03101]
  - 31 **Geitler L**. Endomitose und endomitotische Polyploidisierung. *Protoplasmatologia* 1953; **60**: 1-89
  - 32 **Therman E**, Sarto G, Stubblefield P. Endomitosis: a reappraisal. *Human Genetics* 1983; **63**: 13-18 [DOI: 10.1007/BF00285390]
  - 33 **Therman E**, Sarto GE, Kuhn EM. The course of endomitosis in human cells. *Cancer Genet Cytogenet* 1986; **19**: 301-310 [DOI: 10.1016/0165-4608(86)90059-2]
  - 34 **Hattori N**, Davies TC, Anson-Cartwright L, Cross JC. Pe-

- riodic expression of the cyclin-dependent kinase inhibitor p57(Kip2) in trophoblast giant cells defines a G2-like gap phase of the endocycle. *Mol Biol Cell* 2000; **11**: 1037-1045 [PMID: 10712518 DOI: 10.1091/mbc.11.3.1037]
- 35 **Anisimov AP**. Endopolyploidy as a morphogenic factor of development. *Cell Biol Int* 2005; **29**: 993-1004 [DOI: 10.1016/j.cellbi.2005.10.013]
- 36 **Anisimov AP**. A study of genome multiplication mechanisms in the development of albumen gland cells of *Succinea lauta* (Gastropoda, Pulmonata). VII. Transcriptive activity of the nuclei in the endomitotic cycle. *Tsitologiya* 1997; **39**: 237-243
- 37 **Kuhn EM**, Therman E. The behaviour of heterochromatin in mouse and human nuclei. *Cancer Genet. Cytogenet* 1988; **34**: 143-151 [DOI: 10.1016/0165-4608(88)90181-1]
- 38 **Zybina TG**, Kaufmann P, Frank H-G, Freed J, Biesterfeld S. Genome multiplication of extravillous trophoblast cells in human placenta in the course of differentiation and invasion into endometrium and myometrium. II. Mechanisms of polyploidization. *Tsitologiya* 2004; **46**: 640-648
- 39 **Calvi BR**, Lilly MA, Spradling AC. Cell cycle control of chorion gene amplification. *Genes Dev* 1998; **12**: 734-744 [DOI: 10.1101/gad.12.5.734]
- 40 **Lee HO**, Davidson JM, Duronio RJ. Endoreplication: polyploidy with purpose. *Genes Dev* 2009; **23**: 2461-2477 [PMID: 19884253 DOI: 10.1101/gad.1829209]
- 41 **Gunderina LI**, Sherudilo AI, Mitina RL. Cycle of DNA reduplication, In polytenization of the salivary gland cells of larvae of *Chironomus thummi*. I. Cytophotometric and autoradiographic analysis of DNA synthesis dynamics in the individual cells of salivary glands. *Tsitologiya* 1984; **26**: 784-801
- 42 **Machowald AP**, Caulton JH, Edwards MK, Floyd AD. Loss of centrioles and polyploidization of follicle cells of *Drosophila*. *Exp Cell Res* 1979; **118**: 404-410 [DOI: 10.1016/0014-4827(79)90167-8]
- 43 **MacAuley A**, Cross JC, Werb Z. Reprogramming the cell cycle for endoreduplication in rodent trophoblast cells. *Mol Biol Cell* 1998; **9**: 795-807 [PMID: 9529378 DOI: 10.1091/mbc.9.4.795]
- 44 **Hu D**, Cross JC. Development and function of trophoblast giant cells in the rodent placenta. *Int J Dev Biol* 2010; **54**: 341-354 [PMID: 19876834 DOI: 10.1387/ijdb.082768dh]
- 45 **Caluwaerts S**, Vercruyse L, Luyten C, Pijnenborg R. Endovascular trophoblast invasion and associated structural changes in uterine spiral arteries of the pregnant rat. *Placenta* 2005; **26**: 574-584 [PMID: 15993707 DOI: 10.1016/j.placenta.2004.09.007]
- 46 **Vercruyse L**, Caluwaerts S, Luyten C, Pijnenborg R. Interstitial trophoblast invasion in the decidua and mesometrial triangle during the last third of pregnancy in the rat. *Placenta* 2006; **27**: 22-33 [PMID: 16310034 DOI: 10.1016/j.placenta.2004.11.004]
- 47 **Simmons DG**, Rawn S, Davies A, Hughes M, Cross JC. Spatial and temporal expression of the 23 murine Prolactin/Placental Lactogen-related genes is not associated with their position in the locus. *BMC Genomics* 2008; **9**: 352 [PMID: 18662396 DOI: 10.1186/1471-2164-9-352]
- 48 **Soares MJ**, Chakraborty D, Katim Rumi MA, Konno T Renaud SJ. Rat placentation: an experimental model for investigating the hemochorial maternal-fetal interface. *Placenta* 2012; **33**: 233-243 [DOI: 10.1016/j.placenta.2011.11.026]
- 49 **Shankar K**, Zhong Y, Kang P, Blackburn ML, Soares MJ, Badger TM, Gomez-Acevedo H. RNA-seq analysis of the functional compartments within the rat placentation site. *Endocrinology* 2012; **153**: 1999-2011 [PMID: 22355068 DOI: 10.1210/en.2011-1833]
- 50 **Ouseph MM**, Li J, Chen H-Z, Pecot T, Wenzel P, Thompson JC, Comstock G, Chokshi V, Byrne M, Forde B., Chong J.L. Huang K, Machiraju R, De Bruin A, Leone G. Atypical E2F repressors and activators coordinate placental development. *Dev Cell* 2012; **22**: 849-862 [DOI: 10.1016/j.devcel.2012.01.013]
- 51 **Sher N**, Von Stetina JR, Bell GW, Matsuura S, Ravid K, Orr-Weaver TL. Fundamental differences in endoreplication in mammals and *Drosophila* revealed by analysis of endocycling and endomitotic cells. *Proc Natl Acad Sci USA* 2013; **110**: 9368-9373 [PMID: 23613587 DOI: 10.1073/pnas.1304889110]
- 52 **Zybina TG**, Stein GI, Pozharisski KM, Zybina EV. Invasion and genome reproduction of the trophoblast cells of placenta junctional zone in the field vole, *Microtus rossiaemeridionalis*. *Cell Biol Int* 2014; **38**: 136-143 [PMID: 24155276 DOI: 10.1002/cbin.10187]
- 53 **Tamai Y**, Ishikawa T, Bösl MR, Mori M, Nozaki M, Baribault H, Oshima RG, Taketo MM. Cytokeratins 8 and 19 in the mouse placental development. *J Cell Biol* 2000; **151**: 563-572 [DOI: 10.1083/jcb.151.3.563]
- 54 **Hesse M**, Franz T, Tamai Y, Taketo MM, Magin TM. Targeted deletion of keratins 18 and 19 leads to trophoblast fragility and early embryonic lethality. *EMBO J* 2000; **19**: 5060-5070 [PMID: 11013209 DOI: 10.1093/emboj/19.19.5060]
- 55 **Jaquemar D**, Kupriyanov S, Wankell M, Avis J, Benirschke K, Baribault H, Oshima RG. Keratin 8 protection of placental barrier function. *J Cell Biol* 2003; **161**: 749-756 [PMID: 12771125 DOI: 10.1083/jcb.200210004]
- 56 **Gauster M**, Blaschitz A, Siwetz M, Huppertz B. Keratins in the human trophoblast. *Histol Histopathol* 2013; **28**: 817-825 [PMID: 23450430]
- 57 **Zybina TG**, Stein GI, Zybina EV. Endopolyploid and proliferating trophoblast cells express different patterns of intracellular cytokeratin and glycogen localization in the rat placenta. *Cell Biol Int* 2011; **35**: 649-655 [PMID: 21299496 DOI: 10.1042/CBI20100278]
- 58 **Enders AC**, Carter AM. Comparative placentation: some interesting modifications for histotrophic nutrition. A review. *Placenta* 2006; **27** Suppl A: S11-S16
- 59 **Bevilacqua E**, Hoshida M-S, Amarante-Paffaro A, Albieri-Borges A, Gomes SZ. Trophoblast phagocytic program: roles in different placental system. *Int J Dev Biol* 2010; **54**: 495-505 [DOI: 10.1387/ijdb.082761eb]
- 60 **Unhavaithaya Y**, Orr-Weaver TL. Polyploidization of gliain neural development links tissue growth to blood-brain barrier integrity. *Genes Dev* 2012; **26**: 31-36 [DOI: 10.1101/gad.177436.111]
- 61 **Ain R**, Canham LN, Soares MJ. Gestation stage-dependent intrauterine trophoblast cell invasion in the rat and mouse: novel endocrine phenotype and regulation. *Devel Dynam* 2003; **260**: 176-190
- 62 **Coan PM**, Conroy N, Burton GJ, Ferguson-Smith AC. Origin and characteristics of glycogen cells in the developing murine placenta. *Devel Dynam* 2006; **235**: 3280-3294 [DOI: 10.1002/dvdy.20981]
- 63 **Konno T**, Rempel LA, Arroyo A, Soares M. Pregnancy in the brown Norway rat: a model for investigating the genetics of placentation. *Biol Reprod* 2007; **76**: 709-718 [DOI: 10.1095/biolreprod.106.056481]
- 64 **Zybina EV**, Zybina TG. Modifications of nuclear envelope during differentiation and depolyploidization of rat trophoblast cells. *Micron* 2008; **39**: 593-606 [PMID: 17627829 DOI: 10.1016/j.micron.2007.05.006]
- 65 **Zybina EV**, Kudryavtseva MV, Kudryavtsev BN. The distribution of chromosome material during giant nucleus division by fragmentation in the trophoblast of rodents. Morphological and cytophotometrical study. *Tsitologiya* 1979; **21**: 12-20
- 66 **Zavarzin AA**. Synthesis and kinetics of cell populations in the mammalian ontogenesis. Leningrad: Nauka, 1967 (in Russian)
- 67 **Wasserlauf IE**, Ananina TV, Unger MF, Karamysheva TV, Melnikova NN, Rubtsov NB, Stegnii VN. Chromosome organization and differential banding in endomitotic nuclei of nurse cells of *Calliphora erythrocephala*. *Genetika* 2003; **39**:

- 1193-1202
- 68 **Zybina TG**. DNA in the nuclear fragments that appear in the course of fragmentation of the secondary giant trophoblast cells in the field vole. *Tsitologiya* 1990; **32**: 806-810
- 69 **Kiknadze II**, Istomina AG. Endomitosis in grasshopper. I. Nuclear morphology and synthesis of DNA and RNA in the endopolyploid cells of the inner parietal layer of the testicular follicle. *Eur J Cell Biol* 1980; **21**: 122-133
- 70 **Raikov IB**. The protozoan nucleus. Morphology and evolution. Wien-New York: Springer-Verlag, 1985
- 71 **Mossman HW**. Fetal homeostasis. The New York Academy of Sciences (NY), 1967: 50-52
- 72 **Leiser R**, Koob B. Development and characteristics of placentation in a carnivore, the domestic cat. *J Exp Zool* 1993; **266**: 642-656 [PMID: 8371103 DOI: 10.1002/jez.1402660612]
- 73 **Zybina TG**, Zybina EV, Kiknadze II, Zhelezova AI. Polyploidization in the trophoblast and uterine glandular epithelium of the endotheliochorial placenta of silver fox (*Vulpes fulvus* Desm.), as revealed by the DNA content. *Placenta* 2001; **22**: 490-498 [PMID: 11373160 DOI: 10.1053/plac.2001.0675]
- 74 **Kaufmann P**, Castellucci M. Extravillous trophoblast in the human placenta. *Trophoblast Res* 1997; **10**: 21-65
- 75 **Wakuda K**, Yoshida Y. DNA ploidy and proliferative characteristics of human trophoblast. *Acta Obstet Gynecol Scand* 1992; **71**: 12-16 [DOI: 10.3109/00016349209007940]
- 76 **Redline RW**, Hassold T, Zaragoza M. Determinants of villous trophoblastic hyperplasia in spontaneous abortions. *Mod Pathol* 1998; **11**: 762-768 [PMID: 9720505]
- 77 **Pötgens AJ**, Gaus G, Frank HG, Kaufmann P. Characterization of trophoblast cell isolations by a modified flow cytometry assay. *Placenta* 2001; **22**: 251-255 [PMID: 11170831 DOI: 10.1053/plac.2000.0597]
- 78 **Bischof P**, Campana A. Trophoblast differentiation and invasion: its significance for human embryo implantation. *Earl Pregl Biol Med* 1997; **3**: 81-95
- 79 **Zybina TG**, Kaufmann P, Frank H-G, Freed J, Kadyrov M, Biesterfeld S. Genome multiplication of extravillous trophoblast cells in human placenta in the course of differentiation and invasion into endometrium and myometrium. I. Dynamics of polyploidization. *Tsitologiya* 2002; **44**: 1058-1067
- 80 **Biron-Shental T**, Fejgin MD, Sifakis S, Liberman M, Antsaklis A, Amiel A. Endoreduplication in cervical trophoblast cells from normal pregnancies. *J Matern Fetal Neonatal Med* 2012; **25**: 2625-2628 [PMID: 22877079 DOI: 10.3109/14767058.2012.717999]
- 81 **Weier JF**, Weier HU, Jung CJ, Gormley M, Zhou Y, Chu LW, Genbacev O, Wright AA, Fisher SJ. Human cytotrophoblasts acquire aneuploidies as they differentiate to an invasive phenotype. *Dev Biol* 2005; **279**: 420-432 [PMID: 15733669 DOI: 10.1016/j.ydbio.2004.12.035]
- 82 **Anev I**, Rajasri AG, Reddy K, Pillai M. Triploidy without molar change presenting as severe pre-eclampsia and left ventricular failure at 15 weeks. *J Obstet Gynecol* 2011; **3**: 659-660 [DOI: 10.3109/01443615.2011.598966]
- 83 **Sarto GE**, Stubblefield PA, Therman E. Endomitosis in human trophoblast. *Hum Genet* 1982; **62**: 228-232 [DOI: 10.1007/BF00333525]
- 84 **Hoffman LH**, Wooding FBP. Giant and binucleate trophoblast cells in mammals. *J Exper Zool* 1993; **266**: 559-577 [DOI: 10.1002/jez.1402660607]
- 85 **Pfarrer C**, Hirsch P, Guillomot M, Leiser R. Interaction of integrin receptors with extracellular matrix is involved in trophoblast giant cell migration in bovine placentomes. *Placenta* 2003; **24**: 588-597 [PMID: 12828918 DOI: 10.1016/S0143-4004(03)00059-6]
- 86 **Zybina T**, Zybina E. Quantitative changes of nucleolar organizer in the course of polyploidization and differentiation of rat decidua cells. In: Abstracts of the ... Meeting of the International Federation of Placenta Associations, Sorrento, Italy, 2001 Sept 19-23. *Placenta* 2001; **22**: A. 71
- 87 **Adachi S**, Minamisawa K, Okushima Y, Inagaki S, Yoshiyama K, Kondou Y, Kaminuma E, Kawashima M, Toyoda T, Matsui M, Kurihara D, Matsunaga S, Umeda M. Programmed induction of endoreduplication by DNA double-strand breaks in Arabidopsis. *Proc Natl Acad Sci USA* 2011; **108**: 10004-10009 [PMID: 21613568 DOI: 10.1073/pnas.1103584108]
- 88 **Mehrotra S**, Magbool SB, Kolpakas A., Murnen K., Calvi BR. Endocycling cells do not apoptose in response to DNA replication genotoxic stress. *Genes Dev* 2008; **22**: 3158-3171 [DOI: 10.1101/gad.1710208]
- 89 **Ullah Z**, Kohn MJ, Yagi R, Vassilev LT, DePamphilis ML. Differentiation of trophoblast stem cells into giant cells is triggered by p57/Kip2 inhibition of CDK1 activity. *Genes Dev* 2008; **22**: 3024-3036 [PMID: 18981479 DOI: 10.1101/gad.1718108]
- 90 **Ullah Z**, de Renty C, DePamphilis ML. Checkpoint kinase 1 prevents cell cycle exit linked to terminal cell differentiation. *Mol Cell Biol* 2011; **31**: 4129-4143 [PMID: 21791608]
- 91 **Kudryavtsev BN**, Kudryavtseva MV, Sakuta GA, Stein GI. Human hepatocyte polyploidization kinetics in the course of life cycle. *Virchows Arch B Cell Pathol Incl Mol Pathol* 1993; **64**: 387-393 [PMID: 8148960 DOI: 10.1007/BF02915139]
- 92 **Erenpreisa J**, Cragg MS. MOS, aneuploidy and the ploidy cycle of cancer cells. *Oncogene* 2010; **29**: 5447-5451 [DOI: 10.1038/ncr.2010.310.20676137]
- 93 **Salmina K**, Jankevics E, Huna A, Perminov D, Radovica I, Klymenko T, Ivanov A, Jascenko E, Scherthan H, Cragg M, Erenpreisa J. Up-regulation of the embryonic self-renewal network through reversible polyploidy in irradiated p53-mutant tumour cells. *Exp Cell Res* 2010; **316**: 2099-2112 [PMID: 20457152 DOI: 10.1016/j.yexcr.2010.04.030]
- 94 **Erenpreisa J**, Cragg MS. Three steps to the immortality of cancer cells: senescence, polyploidy and self-renewal. *Cancer Cell Int* 2013; **13**: 92 [PMID: 24025698 DOI: 10.1186/1475-2867-13-92]

**P- Reviewer:** Cascella R, Squassina A

**S- Editor:** Song XX **L- Editor:** A **E- Editor:** Liu SQ





## Role of SOX2 in foregut development in relation to congenital abnormalities

Kim Schilders, Joshua K Ochieng, Cornelis P van de Ven, Cristina Gontan, Dick Tibboel, Robbert J Rottier

Kim Schilders, Joshua K Ochieng, Cornelis P van de Ven, Cristina Gontan, Dick Tibboel, Robbert J Rottier, Departments of Pediatric Surgery of The Erasmus MC-Sophia Children's Hospital, 3000 CA Rotterdam, The Netherlands  
Cristina Gontan, The Department of Reproduction and Development of the Erasmus MC, 3015 CN Rotterdam, The Netherlands  
Robbert J Rottier, Cell Biology of the Erasmus MC, 3000 CA Rotterdam, The Netherlands

Author contributions: Schilders K, Ochieng JK, Gontan C and Rottier RJ wrote the first draft of the manuscript; van de Ven CP and Tibboel D added all patient information; Schilders K, Tibboel D and Rottier RJ provided critical input and comments; all authors edited the manuscript.

Correspondence to: Robbert J Rottier, PhD, Departments of Pediatric Surgery of The Erasmus MC-Sophia Children's Hospital, PO Box 2040, 3000 CA Rotterdam, The Netherlands. [r.rottier@erasmusmc.nl](mailto:r.rottier@erasmusmc.nl)

Telephone: +31-10-7044140 Fax: +31-10-7044468

Received: May 28, 2014 Revised: July 14, 2014

Accepted: September 4, 2014

Published online: November 27, 2014

© 2014 Baishideng Publishing Group Inc. All rights reserved.

**Key words:** SOX2; Congenital foregut abnormalities; Lung development

**Core tip:** Foregut abnormalities are complicated congenital diseases which still lack knowledge of the origin. This review highlights foregut development and associated abnormalities, specifically focussing on the transcription factor SOX2.

Schilders K, Ochieng JK, van de Ven CP, Gontan C, Tibboel D, Rottier RJ. Role of SOX2 in foregut development in relation to congenital abnormalities. *World J Med Genet* 2014; 4(4): 94-104  
Available from: URL: <http://www.wjgnet.com/2220-3184/full/v4/i4/94.htm> DOI: <http://dx.doi.org/10.5496/wjmg.v4.i4.94>

### Abstract

The uptake of the two essential ingredients for life, oxygen and nutrients, occurs primarily through the oral cavity, but these two lifelines need to be separated with high accuracy once inside the body. The two systems, the gas exchange pulmonary system and the gastro-intestinal feeding system, are derived from the same primitive embryonic structure during development, the foregut, which need to be separated before birth. In certain newborns, this separation occurs not or insufficiently, leading to life threatening conditions, sometimes incompatible with life. The development of the foregut, trachea and lungs is influenced and coordinated by a multitude of signaling cascades and transcription factors. In this review, we will highlight the development of the foregut and pulmonary system and focus on associated congenital abnormalities in light of known genetic alterations with specific attention to the transcription factor SOX2.

### FOREGUT DEVELOPMENT

Gastrulation is the process that adds complexity to the developing organism and results in a triploblastic animal by formation of the three germ layers, ectoderm, mesoderm and endoderm. At embryonic day 8 (E8.0) in mice (comparative to 3 wk in human), the sheet of endodermal cells starts to invaginate ventrally at the anterior and posterior intestinal portals, which subsequently migrate towards each other to form the primitive gut from the future mouth to anus<sup>[1]</sup>. At E9.0, the notochord delaminates from the dorsal endoderm and will eventually be situated between the primitive gut and the neural tube. The notochord serves in this phase of development as a strong signaling center, secreting morphogens like Sonic hedgehog (Shh) to pattern the endoderm as well as the neural tube<sup>[2]</sup>. Another signaling center associated with the early patterning and morphogenesis of the foregut is the heart mesoderm, which secretes Fibroblast growth factors (Fgf). High levels of Fgf signals activate lung specific genes while lower levels of Fgf activate liver specific



**Table 1** Gestational ages in human and mouse during the five stages of lung development<sup>[8]</sup>

Phases of lung development	Gestational age	
	Human	Mouse
Embryonic phase	Weeks 3-7	E9-11.5
Pseudoglandular phase	Weeks 5-17	E11.5-16.6
Canalicular phase	Weeks 16-25	E16.6-17.4
Saccular phase	Weeks 24-38	E17.4-PN5
Alveolar phase	Weeks 36 to maturity	PN5-30

E: Embryonic age; PN: Post natal age.

genes<sup>[3]</sup>. The prospective lung field, the area which will eventually lead to the emergence of the primitive lung bud, is subsequently patterned by retinoic acid (RA) signaling. The RA receptor (RAR $\alpha$ ) is required to maintain RA signaling and to assist the effects of RAR $\beta$ , which induces the expression of Fgf10. RA signaling integrates the Wnt and transforming growth factor beta (Tgf $\beta$ ) pathways by inhibiting the expression of the Wnt antagonist Dickkopf-1 and by preventing the expression of Tgf $\beta$ <sup>[4,5]</sup>. Overall, the foregut is regionalized as shown by the various dorsal-ventral gradients of morphogens and subsequent transcription factors. This pattern of expression is essential for proper development of the trachea, esophagus and lungs, and disturbances in these patterns result in various trachea-lung defects (see below).

### Lung development

The lung primordium arises from the ventral foregut as a primary bud, just anterior to the developing stomach around embryonic day 9.5 in mice or week 4 in humans<sup>[6,7]</sup>. The lung bud splits in two buds, the future left and right bronchus, elongates and the proximal part separates into oesophagus and trachea, while distally the bronchial tree is formed through a process called branching morphogenesis<sup>[6]</sup>.

Development of the lung can be divided into five distinct, but overlapping phases based on morphology (Table 1)<sup>[8]</sup>. During the earliest phase, the embryonic phase, the lung buds are formed from the primitive foregut, the mayor bronchi are formed and the tracheal-esophageal tube is dividing. Several signaling cascades direct the early embryonic morphogenetic events and cell fate decisions including Tgf $\beta$ , Bone Morphogenetic Proteins (BMPs), Shh, Wnt, and Fgf families, which will be discussed in more detail<sup>[9,10]</sup>. As development of the lung advances, the embryonic endoderm undergoes progressive fate decisions that generate epithelial progenitor cells with increasingly restricted developmental potential over time. The next phase, the pseudoglandular phase is characterized by the commencement of differentiation of epithelial cells. Also, the bronchial tree and all terminal bronchioles are formed. The pseudoglandular phase is followed by the canalicular phase and the saccular stage<sup>[6]</sup>. During the canalicular phase, the conducting airways are completed and the respiratory portions of the lung as well as the capillary bed are formed, while during the saccular phase

the terminal tubes narrow, giving rise to small saccules and the endoderm begins to differentiate into specialized alveolar type I and type II cells<sup>[6]</sup>. The last phase, the alveolar phase, is characterized by the establishment of secondary septa resulting into alveolar formation, which mainly takes places after birth<sup>[8,11]</sup>.

### Regulation of foregut and lung development

The morphogenesis of the foregut and lung is subsequently regulated by a myriad of transcription factors and signaling cascades. The molecular and cellular events contributing to lung development and the separation of the trachea and esophagus have been extensively described in recent reviews<sup>[6,10,12]</sup>. Regionalization of the different parts of the gut is controlled by the localized expression of Homeobox (*Hox*) genes<sup>[13]</sup>. *Hoxa3* and *Hoxb4* are expressed in the foregut endoderm, whereas *Hoxc5* and *Hoxa13* are expressed in the midgut and hindgut endoderm, respectively<sup>[13,14]</sup>. During tracheal-esophageal development, Shh is specifically and dynamically expressed during the patterning of the ventral foregut whereas its expression is transiently expressed in the tracheal endoderm<sup>[15]</sup>. During the early stages of branching morphogenesis, Shh is expressed in the epithelium, with the highest levels of expression in the tips. Later, there is downregulation of Shh in the proximal parts of the airways while distally the expression sustains<sup>[16]</sup>.

## SOX GENES AND FOREGUT DEVELOPMENT

Transcription factors that show specific expression profiles in the endoderm, include members of the SRY-related High-Mobility Group (HMG) transcription factors<sup>[17,18]</sup>. The Sex-determining region on the Y chromosome (*Sry*) gene, was the first identified member of the SOX family of transcription factors<sup>[19,20]</sup>. SOX family members are highly conserved across species and they were originally identified by homology, as they contain an HMG box closely related to that of the *Sry* gene<sup>[21]</sup>. Therefore, *Sry* gave the SOX gene family its name; *Sry*-related HMG box, hence "SOX", followed by a number corresponding to the order of discovery<sup>[19]</sup>. SOX proteins have properties of both classical transcription factors and architectural proteins<sup>[22]</sup>. They function as classical transcription factors, either activating or repressing specific target genes through interaction with different partner proteins.

All SOX factors bind DNA *via* their HMG domain and recognize the same consensus motif 5'-(A/T)(A/T)CAA(A/T)G-3'<sup>[23]</sup>. The transcriptional function of SOX proteins dependent on the cell type and the promoter context, and they often have functional redundancy among each other<sup>[22]</sup>. In contrast to other transcription factors which mainly target the major groove, SOX proteins interact with the minor groove of the DNA helix and, as a consequence, induce a sharp bend in the DNA<sup>[22]</sup>. The DNA bending capacity of SOX proteins can be functionally important for several reasons. It may

**Table 2** The role of *SOX* genes in diseases

<i>SOX</i> gene	Chromosome location	Disease
<i>SOX2</i>	3q26.3-q27	Microphthalmia, syndromic 3 optic nerve hypoplasia, abnormalities of the central nervous system, CHARGE-syndrome <sup>[65]</sup> , AEG-syndrome <sup>[57]</sup> , EA/TEF <sup>[58]</sup> , CPAM <sup>[39,40]</sup>
<i>SOX3</i>	Xq27.1	Mental retardation, X-linked with isolated growth hormone deficiency, infundibular hypoplasia, hypopituitarism <sup>[117]</sup>
<i>SOX9</i>	17q23	Campomelic dysplasia with autonomic XY sex reversal <sup>[117]</sup> , Pierre-Robin syndrome <sup>[118]</sup>
<i>SOX10</i>	22q13.1	Waardenburg-Shah syndrome, Yemenite deaf-blind hypopigmentation syndrome, peripheral demyelinating neuropathy, central dysmyelinating leukodystrophy, Waardenburg syndrome, Hirschprung's disease <sup>[117]</sup>
<i>SOX11</i>	2p25	Unknown
<i>SOX17</i>	8q11.23	Unknown
<i>SOX18</i>	20q13.33	Hypotrichosis-lymphedema-telangiectasia syndrome <sup>[117]</sup>

EA/TEF: Esophageal atresia/tracheoesophageal fistula; CPAM: Congenital pulmonary and airway malformations.

bring different regulatory regions of the target gene into close proximity. Thereby, it facilitates the formation of enhanceosomes, *i.e.*, functionally active complexes of transcription factors on different gene enhancer sequences<sup>[24]</sup>. It also allows the interaction of distant enhancer nucleoprotein complexes with the basal transcription machinery<sup>[25,26]</sup>. The local changes in chromatin structure induced by SOX proteins may facilitate the recruitment of higher-order architectural factors (like polycomb or trithorax protein groups)<sup>[27]</sup>. Bending of DNA by SOX proteins could also act in a negative way by preventing the binding of other factors to adjacent sites in the major groove<sup>[27]</sup>.

SOX proteins have been identified in all animal species (birds, reptiles, amphibians, fish, insects, and nematodes)<sup>[28]</sup> and mutations in several of the *SOX* genes have been implicated in the pathogenesis of human congenital anomalies and syndromes (Table 2).

*SOX* genes are expressed in diverse and dynamic patterns during embryogenesis. During development members of the SOX family are expressed in almost every tissue of the embryo, and also in a number of adult tissues<sup>[29,30]</sup>. The expression of a specific SOX transcription factor is not necessarily restricted to a particular cell type or lineage. Their expression pattern during development appears to correlate with early cell fate decisions. For example, Sry is expressed in the undifferentiated male gonad and is quickly down regulated once the decision is made to initiate male development<sup>[31-33]</sup>.

To date, four members from the *SOX* gene family are known to be involved in lung organogenesis, *SOX2*, *SOX9*, *SOX11* and *SOX17*<sup>[8,34-39]</sup>. *SOX9* was found throughout lung morphogenesis as a downstream gene of Shh and modulated by BMP4 and Noggin. Using epithelial specific gain and loss function mouse models, *SOX9* has been shown to play a crucial role in branching morphogenesis through controlling a balance between proliferation and differentiation<sup>[40]</sup>. In another study, knock out of *SOX9* in the mesenchyme demonstrated that it plays a crucial role in differentiation of the lung tracheal epithelium<sup>[41]</sup>. *SOX9* is required for formation and patterning of tracheal cartilage by a mechanism mediated by Fgf18<sup>[42,43]</sup>. *SOX9* promotes proper branching morphogenesis by controlling the balance between proliferation and differ-

entiation and regulating the extracellular matrix and can be used as a marker for the distal epithelium<sup>[40]</sup>. *SOX11* has been suggested to be involved in development and plays a key function in tissue remodeling, including the lung<sup>[35]</sup>. *SOX11* deficient mice die immediately after birth because of significant lung hypoplasia and other tissue defects<sup>[35]</sup>. *SOX17* was shown to be crucial early after gastrulation for the formation of definitive endoderm, which gives rise to the lung, liver, pancreas, stomach, and gastrointestinal tract<sup>[44]</sup>. In the lung, *SOX17* is expressed in the respiratory epithelial cells at embryonic day 18 in mice and becomes primarily restricted to ciliated cell in the postnatal and adult lung<sup>[36]</sup>. Ectopic expression of *SOX17* in lung epithelial cells inhibits peripheral epithelial cell differentiation and results in the activation of the cell cycle and the initiation of progenitor-like cell behavior in mature lung cells<sup>[36]</sup>. *SOX17* has been shown to impair the expression of Tgfβ1 responsive inhibitors, p15, p21 and p57, while inhibiting Tgfβ1 and Smad3 transcriptional activity<sup>[36]</sup>.

## SOX2

The mouse *SOX2* gene has been mapped to chromosome 3<sup>[45]</sup>. *SOX2* plays crucial roles during different stages of vertebrate embryonic development and its expression is temporally and spatially regulated<sup>[46]</sup>. *SOX2* expression starts at the morula-stage of embryo development. In blastocysts it is specifically detected within the cells of the inner cell mass (ICM). Expression continues in the epiblast, the tissue that will give rise to the embryo and germ cells<sup>[47]</sup>. *SOX2* is also expressed in embryonic stem cells, which are derived from the ICM. During early gastrulation, *SOX2* expression in the embryo is restricted to the anterior ectoderm, which gives rise to neuroectoderm and anterior surface ectoderm, while the extraembryonic expression becomes confined to the chorion<sup>[48,49]</sup>. At later stages of embryonic development, *SOX2* is expressed in the brain, neural tube, eyes, sensory placodes, branchial arches, gut endoderm, and the germ cells<sup>[47,50-55]</sup>. When the arches develop, *SOX2* continues to be expressed in the primitive foregut endoderm. Later, *SOX2* is present in the epithelium of foregut-derived organs, including the tongue, esophagus, trachea, proximal lung and stomach<sup>[39,47,50,56,57]</sup>.

The lack of SOX2 expression in mice results in early embryonic lethality<sup>[47]</sup>. SOX2 null mutant mouse embryos implant but fail to develop an egg cylinder or epiblast, and they die before gastrulation because SOX2 is required in the ICM of the blastocyst<sup>[47]</sup>. Other mutations that only affect SOX2 regulatory elements can cause deafness, defects in the inner ear, circling behavior, and a yellow coat color<sup>[45]</sup>. The use of two SOX2 hypomorphic mutants showed a dose-dependent role of SOX2 in the development of the retina and the differentiation of the foregut endoderm<sup>[54,58]</sup>. Heterozygous mutations in SOX2 have been associated in human with severe structural malformations of the eye, bilateral anophthalmia (absent eye) and microphthalmia (small eye), and anophthalmia-esophageal-genital (AEG) syndrome<sup>[57]</sup>. In AEG infants the esophagus and trachea fail to separate normally and the trachea is connected to the stomach by an abnormal distal esophagus<sup>[57,59,60]</sup>. These symptoms underwrite the developmental functions for SOX2, as found in SOX2 hypomorphic mice described above.

**SOX2 in foregut and lung development:** SOX2 is expressed throughout the early foregut epithelium, but becomes restricted to the dorsal epithelial cells at embryonic day 9.5, whereas Nkx2.1 is reciprocally expressed in the ventral epithelium<sup>[58,61]</sup>. SOX2 is expressed in the epithelial cells of the foregut at E9.5. From E11.5 until E14.5, SOX2 is exclusively expressed in the epithelial cells of the non-branching developing airways and it remains expressed in the epithelial cells of the conducting airways after birth. So SOX2 is exclusively expressed at the non-branching airways<sup>[39]</sup>.

Previously, it was shown that ectopic expression of SOX2 in epithelial cells of the lung result in abnormal alveolar formation, enlarged airspaces and a decrease in the number of airways, indicating that SOX2 modulates branching morphogenesis. Also, an increased number of neuroepithelial cells and (pre-) basal cells was observed. This indicates that SOX2 is important in cell fate choice and epithelial differentiation<sup>[39]</sup>. More recently it was shown that SOX2 regulates the emergence of lung basal cells by directly activating the transcription of the basal cell master gene Trp63, and the emergence of bronchioalveolar stem cells<sup>[62]</sup>.

The proper dorsal-ventral patterning of SOX2 and Nkx2.1 is critical for foregut morphogenesis. Down-regulation of SOX2 leads to the formation of esophageal atresia/tracheoesophageal fistula (EA/TEF) in SOX2 hypomorphic mutants<sup>[58]</sup>, whereas deletion of Nkx2.1 leads to defects in foregut separation and the formation of EA/TEF associated with high SOX2 expression in the epithelium<sup>[58,63]</sup>. Similarly, the epithelial cells in the fistula of SOX2 hypomorphic mutants express high levels of Nkx2.1 suggesting that low level of SOX2 is required for Nkx2.1 expression to expand dorsally and reprogram the dorsal epithelium to a respiratory fate<sup>[58]</sup>. These findings suggested that the dorsal-ventral arrangement of SOX2 and Nkx2.1 is essential for foregut separation and

the subsequent differentiation of epithelial progenitor cells into oesophageal and tracheal epithelium and lung buds<sup>[12]</sup>.

Using Chromatin Immuno Precipitation it was shown that SOX2 directly binds to the promoter region of the NKX2.1 gene in human embryonic stem cells and this binding resulted in the inhibition of NKX2.1 transcription<sup>[64]</sup>. Other interesting SOX2 target genes that are involved in early lung morphogenesis are members of the Notch (JAG1) pathway and Shh pathway (GLI2, GLI3)<sup>[65]</sup>. Since the activity of SOX2 depends on its interaction with other proteins it is of high importance to reveal its interacting partners. Recently some of these partners were identified in embryonic and neural stem cells<sup>[65,66]</sup>. One of the partners identified is Chromodomain-Helicase-DNA-Binding Protein7 (CHD7), which plays a major role in CHARGE syndrome. As mentioned before, SOX2 plays a role in AEG syndrome which shows many similarities with CHARGE syndrome. SOX2 and CHD7 also regulate common target genes, like MYCN, JAG1 and GLI2/3. These genes are involved in syndromes that are characterized by the same malformations as AEG and CHARGE syndromes. Gene networks like this SOX2-CHD7-regulated network can be used to better understand the molecular basis of various human diseases and therefore associating partners in the lung epithelium could help us to reveal the mechanisms underlying lung-related abnormalities<sup>[65]</sup>.

## FGF SIGNALING

Another study using *in vitro* organ cultures demonstrated that Fgf10 signaling inhibits SOX2 expression in the mouse foregut<sup>[58]</sup>. Mesenchymal expression of Fgf10 around the distal ends of the lung epithelium functions as a chemoattractant by binding to the epithelial expressed Fgf receptor 2b (Fgfr2b) leading to branching and outgrowth of the epithelium<sup>[67,68]</sup>. The functional interaction between Fgf10 and Fgfr2b was shown by the high similarity between the Fgf10-null and Fgfr2b-null mouse mutants<sup>[3,69]</sup>. Fgf10 knockout mice developed normal trachea, but completely lacked lung structures<sup>[69,70]</sup>, whereas targeted deletion of Fgfr2b prevented branching, causing the trachea to terminate as a blind-ended sac<sup>[71]</sup>. Conditional gene inactivation studies further demonstrate that both Fgf10 and Fgfr2b are required for a normal branching program and proper proximal-distal patterning of the lung<sup>[72]</sup>. Recently, it was shown that ubiquitous overexpression of Fgf10 throughout the lung could rescue lung agenesis in Fgf10 knockout mice, suggesting that precise localization of Fgf10 expression is not required for lung branching morphogenesis. Rather, Fgf10 signaling prevents cells from expressing SOX2 by the activation of  $\beta$ -catenin. As the lung bud grows, the cells become more distant from the Fgf10 source and start to adopt a more proximal cell fate expressing SOX2. When SOX2 is ectopically expressed in the distal epithelial cells of the developing airways, these cells are no longer responsive to

Fgf10 and differentiate into proximal cells, which results in reduced branching and formation of cyst-like structures<sup>[39,62]</sup>.

## WNT SIGNALLING

Receptor tyrosine kinases (RTKs), like the Fgfr, are able to activate Wnt/ $\beta$ -catenin through the Erk/MAPK mediated phosphorylation of the Wnt co-receptor Lrp6 on Ser1490 and Thr1572, leading to an increased cellular response to Wnt. Moreover, RTKs directly phosphorylate  $\beta$ -catenin on the Tyr142 residue, which causes its release from membrane bound cadherin complexes<sup>[73]</sup>. In turn, Fgfr2b expression is induced by activation of epithelial  $\beta$ -catenin activation, which results in an increase of Fgf10 signaling<sup>[74]</sup>. This regulation of distal epithelial progenitors by  $\beta$ -catenin suggests the progressive signaling cascade where Fgf10 regulates branching morphogenesis *via* Wnt signaling. Epithelial specific expression of Wntless, a cargo receptor protein important for directing Wnt ligands, has recently been shown to be important for lung differentiation and vasculature development probably by modulating the secretion of Wnt ligands<sup>[75]</sup>.

At embryonic day 9.5, Wnt signaling is active in the ventral side of the unseparated foregut tube, where the Wnt ligands Wnt2 and Wnt2b are highly expressed<sup>[61,76]</sup>. Wnt2 and Wnt2b are secreted by mesenchymal cells of the ventral foregut and signal through the canonical  $\beta$ -catenin pathway to specify lung progenitors in the foregut endoderm<sup>[61,76]</sup>. Conditional inactivation of  $\beta$ -catenin in the foregut endoderm results in the absence of both trachea and lung, whereas expression of a constitutively active  $\beta$ -catenin mutant results in the expansion of the earliest respiratory marker, Nkx2.1, and a loss of the SOX2 positive domain<sup>[61]</sup>. Later in development, Wnt/ $\beta$ -catenin signaling is required for proper proximal-distal patterning of the lung<sup>[74]</sup>. Wnt7b is expressed in the endoderm of the early foregut and its deletion does not disrupt foregut separation, but results in irregular lung branching morphogenesis and vasculature development<sup>[77]</sup>. Mesenchymal Wnt2 and epithelial Wnt7b cooperate with Pdgf signaling to promote mesenchymal differentiation<sup>[78]</sup>.

Respiratory endodermal specific expression of a constitutive active  $\beta$ -catenin isoform showed that canonical Wnt signaling is not required for the development of alveolar epithelium<sup>[79,80]</sup>. However, the formation of proximal epithelium was impaired, because ectopic Wnt signaling induced the expression of Tcf1 and Lef1 at the expense of SOX2 and Trp63<sup>[79]</sup>. On the other hand, conditional deletion of  $\beta$ -catenin in respiratory epithelium resulted in the loss of alveolar structures. Selective loss of bronchiolar lineages with continued proliferation may result in cystic lesions of the lung resembling an anomaly known in humans as congenital pulmonary and airway malformations (CPAM)<sup>[80]</sup>.

Ectopic expression of Wnt5a in the respiratory epithelium resulted in increased Fgf10 expression and a reduc-

tion in epithelial Shh expression<sup>[81]</sup>. The precise dose and timing of Fgf and Wnt signaling lead to the induction of Shh expression in the respiratory epithelium<sup>[3,76]</sup>. The paracrine effect of Shh on the surrounding mesenchyme results in the Foxf1 and Gli1/Gli3 mediated expression of BMP4.

## BMP SIGNALING

BMP signaling plays prominent roles in foregut separation and lung development, however the molecular mechanisms controlling temporal-spatial BMP signaling dynamics in foregut organogenesis are poorly understood<sup>[82]</sup>. In the unseparated foregut tube, BMP4 is expressed in the ventral mesenchyme, while BMP7 and the BMP antagonist Noggin are enriched in the dorsal endoderm<sup>[83]</sup>. Ablation of Noggin resulted in increased BMP signaling in the foregut and the formation of EA/TEF<sup>[84]</sup>. These embryos showed abnormal delamination of the notochord from the early definite endoderm epithelial sheet, resulting in epithelial cells of endodermal origin being present in the notochord<sup>[84]</sup>. Subsequent deletion of either BMP4 or BMP7 in these Noggin null mice rescued the separation defects<sup>[83,84]</sup>. Recently, it was shown that Noggin is required to attenuate BMP signaling in order to allow the notochord to delaminate from the dorsal foregut endoderm<sup>[85]</sup>. Tissue specific ablation of BMP4 in the early foregut endoderm resulted in tracheal agenesis accompanied by reduced cellular proliferation in the epithelial and mesenchymal compartments. However, the trachea does not separate from the foregut and Nkx2.1 expression is conserved in the ventral endodermal epithelium, suggesting that BMP4-mediated signaling is essential for separation but not for the initial specification of the tracheal epithelium<sup>[86]</sup>. Similarly, conditional inactivation of BMP4 and BMP7 in the foregut leads to tracheal agenesis, a decrease of Nkx2.1 expression and a ventral expansion of SOX2 and Trp63 expression. Subsequent activation of Wnt signaling did not promote respiratory differentiation. Deletion of SOX2 in the BMP4 deficient mouse rescued the foregut separation defect, showing that SOX2 is downstream of BMP signaling<sup>[87]</sup>. Ectopic expression of SOX2 in the distal lung buds showed that Fgf-Erk signaling was abrogated at the expense of BMP-Smad signaling<sup>[39]</sup>.

## RELATIONSHIP BETWEEN SOX2 AND CONGENITAL DEFECTS OF THE FOREGUT

Congenital malformations of the lung constitute a spectrum of lesions that originate during the embryonic period. Incidences of congenital defects of the foregut are in the range of 1:11.000-35.000 pregnancies (World Health Organization). Patients present a broad range of clinical manifestations ranging from intra uterine death to significant illnesses at birth with a variable severity of



respiratory symptoms and later on distress and repeated chest infections. However a number, although impressive at repeated prenatal ultrasound or magnetic resonance imaging may remain asymptomatic for long periods and significantly regress in the course of pregnancy. Congenital defects of the foregut occur either in isolated cases or as part of a complex syndrome<sup>[88]</sup>. The causes of most of these malformations as well as their molecular genetic background are still unknown. Different types of congenital lung malformations can be distinguished, which will be briefly discussed.

### CPAM

CPAM constitute a spectrum of lesions that originate during the embryonic period. The prevalence of congenital lung malformations has seemingly increased over the last decade probably due to better antenatal ultrasound screening and is estimated at 1 in 3000 pregnancies<sup>[89]</sup>. Although most newborns with antenatally diagnosed congenital lung malformations are asymptomatic at birth, approximately 10% show respiratory insufficiency. Secondary infections of these lesions occur in approximately 5% of unoperated children.

The different types of congenital pulmonary and airway malformations are classified in bronchopulmonary malformations, pulmonary hyperplasia, congenital lobar overinflation and other cystic lesions<sup>[90]</sup>.

**Bronchogenic cyst:** A bronchogenic cyst is often a solitary cyst in the mediastinum or in the lung parenchyma filled with fluid. Their structural lining resembles that of the bronchus, cartilage and bronchial-type glands included. Symptoms at birth are mostly due to compression of surrounding structures, especially bronchial structures resulting in hyperinflation of lung parenchyma distal to the obstruction. Symptoms at later age are mainly due to infection. Etiology is probably similar to other duplication cysts as aberrant bud formation from the foregut structures. The molecular biology has not been studied so far.

**Bronchial atresia:** Bronchial atresia, mostly asymptomatic often results in overinflation of a lobe, segment or even smaller part of the lung depending on the level of bronchus being atretic. Symptoms are rare. Etiology is unknown and may be similar to the reasons of bronchial blockage in congenital lobar overinflation.

**Cystic adenomatoid malformation stocker type 1 and type 2:** Although congenital cystic adenomatoid malformation (CCAM) pathogenesis is unknown, several authors have hypothesized that different types of CCAM originate at different stages of lung development. Abnormal airway development during branching morphogenesis probably results in specific areas of the lung where terminal bronchioles overgrow and alveolar formations are absent<sup>[88]</sup>. Another hypothesis postulates that CCAM originate as a result of imbalance between cell proliferation

and apoptosis during airway branching<sup>[62,91,92]</sup>. Type 1 CCAM consists of a few large cysts with bronchiolar configuration and a lining of respiratory epithelium overlying fibroelastic tissue and small amounts of smooth muscles. It may have a systemic arterial supply. Type 2 CCAM are multicystic lesions (cysts < 2 cm) often localized in one lobe, although multiple lobes can be affected.

Aberrant expression of genes involved in lung development has been shown to result in CCAM-like phenotypes. Transiently induced overexpression of SOX2<sup>[39,62]</sup>, Fgf10<sup>[93]</sup>, orthotopic overexpression of Fgf9 and heterotopic overexpression of Fgf7 show perturbations of lung morphogenesis some mimicking CCAM type 1 and 2 depending on the time of overexpression.

In human resection specimens increased levels of the transcription factors HOXB5, TTF1, Fgf9<sup>[94,95]</sup> as well as changed expression patterns of the adhesion molecules  $\alpha$ -2 integrins and E-cadherin, increased levels of Clara cell marker CC-10 and reduced expression of Fatty acid binding protein-7 have been described<sup>[96]</sup>. Moreover, microarray data revealed a 6 fold up-regulation of SOX2 in CCAM tissue compared with controls<sup>[96]</sup>. These findings correspond with recently published data, describing the generation of CCAM-like phenotype by overexpressing SOX2 in mouse. In a comparative study, expression of SOX2 in human CCAM tissue was identified<sup>[62]</sup>. Different etiology for type 1 and type 2 may be supported by recent findings that SOX2 is expressed in epithelial lining of cystic lesions in CCAM type 2, but not in CCAM type 1. Moreover, TRP63 was co-expression in SOX2 positive cells, suggesting that the epithelium had proximal characteristics (Ochieng *et al*<sup>[62]</sup>, 2014).

**Extralobar sequestration:** Extralobar sequestration (ELS) are characterized by normal, non-functioning lung tissue without connection with the bronchial tree and often receive blood supply from the systemic circulation. Mainly found in the left lower chest, these lesions can also be found in or below the diaphragm. In contrast to CCAM, associated anomalies like vertebral and chest wall deformities and congenital heart disease are described. In 5%-15% of patients with congenital diaphragmatic hernia an ELS is found at operation. Although often asymptomatic, antenatal diagnosis can be very helpful to detect congestive heart failure caused by the arteriovenous shunting through the anomalous systemic blood supply. Late symptoms are mainly infectious. One of the genes that is thought to be involved in the pathogenesis of this anomaly is the HOXB5 gene. It is previously shown that this gene is involved in airway branching<sup>[97,98]</sup>.

### Pulmonary hyperplasia and related lesions

**Laryngeal atresia:** Laryngeal atresia causes congenital high airway obstruction syndrome (CHAOS) at birth in the absence of a tracheoesophageal fistula. Prenatally, a polyhydramnios, large lung volume and inverted diaphragm are associated with fetal hydrops. Survivors are only described in those patients who have a tracheo-

esophageal fistula or a pinpoint laryngeal connection to relieve pressure from the lungs. Still, these patients may suffer from tracheobronchomalacia and diaphragmatic dysfunction due to increased lung extension during pregnancy. CHAOS can be associated with Fraser syndrome<sup>[99]</sup>. If a bigger tracheoesophageal connection exists, prenatal diagnosis is difficult and diagnosis is only made at birth due to severe dyspnea. As a cause of laryngeal atresia, failure of recanalization of the laryngeal membrane is described. No detailed molecular analysis of lungs of laryngeal atresia, either pre- or postnatally, has been reported.

**Solid or cystic adenomatoid malformation, stocker type 3:** The type 3 lesion, which accounts for 5%-10% of cases, occurs almost exclusively in males, and is associated with maternal polyhydramnios in nearly 80% of the cases. These are large, non-cystic bulky lesions, compressing the adjacent lung and mediastinum. Microscopically, randomly scattered bronchiolar/alveolar duct-like structures are lined by low cuboidal epithelium and surrounded by “alveoli” also lined by cuboidal epithelium. The virtual absence of any small, medium, or large pulmonary arteries in this type of lesion is remarkable.

### **Congenital lobar overinflation**

Congenital lobar overinflation, or Congenital lobar emphysema (CLE), is a rare lung malformation with an incidence ranging from 1:20000 births to 1:30000 births<sup>[100-102]</sup>. CLE is characterised by distended alveoli distal to the terminal bronchiole with destruction of the lining of the lobes, in contrast to a polyalveolar lobe where the number of alveoli is increased. CLE usually affects the left upper or right middle lobe<sup>[103]</sup>. Prenatal diagnosis can be made when a hyperechogenic lung is seen, but discrimination with CCAM and ELS is difficult at that point. Although respiratory failure can be present at birth due to compression of normal lung and displacement of the heart, many patients are asymptomatic. As a cause of CLE, a disruption of normal bronchopulmonary tree development is described with dysplastic cartilage, mucosal overgrowth, main stem bronchial atresia or external compression from abnormal cardiovascular structures. No specific genetic anomalies have been linked to this anomaly so far.

### **Bochdalek type of Congenital diaphragmatic hernia**

Bochdalek type of Congenital diaphragmatic hernia (CDH) is characterized by a posterolateral defect mostly in the left diaphragm, which results in herniation of the abdominal organs into the chest<sup>[104,105]</sup>. Subsequent pulmonary hypoplasia and pulmonary hypertension cause severe respiratory failure at birth. Lung hypoplasia is characterized by reduced alveolar air spaces lacking secondary septae, thickened alveolar walls and increased interstitial tissue. In the pulmonary vessels hyperplasia of the median and increased adventitial layer of the arterial wall is well described. The incidence of CDH is approxi-

mately one in 2500 births and the underlying cause of CDH is still unknown in a large number of patients.

Several links to gene loci have been found partly based on animal experiments<sup>[106]</sup> and several members of the vitamin A-RA pathway have been implicated in the occurrence of CDH, such as vitamin A deficiency, STRA6 and RALDH2<sup>[104,105,107]</sup>. Moreover, some genes that are downstream of this pathway, like COUP-TF II, FOG2, GATA4 and GATA6, have also been found to be associated with CDH review<sup>[107,108]</sup>. Recent exome sequencing identified a novel candidate gene, PIGN, aside from the known FOG2 involvement<sup>[109,110]</sup>. A direct link of CDH and SOX2 has not been described.

### **Esophageal atresia**

Esophageal atresia with or without TEF has an incidence ranging between one in 2500 to one in 4500 births<sup>[111]</sup>. In EA, the proximal esophagus is blunt-ended, while the distal part is connected to the trachea. This connection and the position of the atresia varies between patients and leads to the classification of five subtypes<sup>[112]</sup>. In 85% of cases the esophageal atresia (EA) is accompanied by a distal tracheoesophageal fistula. Approximately half of the patients suffer from associated anomalies, as recently been reviewed<sup>[113]</sup>.

The genetics of EA/TEF is complex, and several studies have indicated putative factors associated with either the multifactorial syndromes, or with EA/TEF. Based on murine models of this anomaly, several candidate genes and pathways have been identified, such as the receptors of the RAR $\alpha$ /RAR $\beta$ , members of the SHH-PTC-GLI pathway (Shh, Gli2/3, Foxf1), BMP signaling (Noggin) and some transcription factors (Hoxc4, Ttf-1, Pcsk5, Tbx4, SOX2)<sup>[113]</sup>.

Some of these candidate genes seem to be associated with human EA/TEF, such as FOXF1, PCSK5, SHH, NOG and SOX2. Moreover, other human genes have been associated with EA/TEF as part of several syndromes, such as Feingold (MYCN), Opitz G (MID1), Fanconi anemia (FANCA/C/D/G) and CHARGE (CHD7). Recently, CHD7 was shown to directly associate with SOX2, thereby linking CHARGE syndrome with AEG<sup>[65]</sup>. Moreover, it was shown that these two proteins activated the transcription of a number of genes that are implicated in related syndromes, like *Shh*, *Gli2/3*, *Myo*.

### **Alveolar capillary dysplasia**

Although alveolar capillary dysplasia with misalignment of pulmonary veins (ACD/MPV) is a very rare condition without known incidence it has a dismal prognosis. The diagnosis is most likely underreported because it can only be made by histological examination of lung tissue. Newborn patients present with respiratory failure, hypoxemia, metabolic acidosis, pulmonary hypertension and right ventricular failure. Chest X-ray can be interpreted as normal but might show diffuse haziness or ground-glass opacities. Associated congenital anomalies may be present, especially of the genitourinary, gastrointestinal

and cardiovascular system. Histologically, a decreased number of pulmonary capillaries is observed, distantly from the alveolar epithelium and thickened septae with a malposition of pulmonary veins close to pulmonary arteries. Often lymphangiectases are seen. Pulmonary arteries show medial hypertrophy with muscularization of distal arterioles. Although treatment response especially to therapy relieving pulmonary hypertension has been described, effects are transient. Late presenters and long-term survivors very rarely have been described and might be due to a lesser degree of histological changes.

Deletions in chromosomal region 16q24.1q24.2 have been described. The smallest region of overlap in these deletions contains the FOX transcription factor gene cluster, including *FOXF1*, *FOXC2* and *FOXL1*. Findings in the mouse model with heterozygous deficiency of *FOXF1* are similar to those in humans with ACD/MPV whereas mouse embryos with homozygous deficiency die at E9.5 due to pulmonary vascular abnormalities. A relationship with SOX2 never has been described<sup>[114-116]</sup>.

## CONCLUSION

SOX2 has only been found to be associated with a limited number of specific subsets of congenital anomalies. Modulating the expression levels of SOX2 during trachea and lung development in mice have led to abnormalities which relate to human conditions, such as CPAM and EA/TEF<sup>[39,58,62]</sup>. Interestingly, CHD7, which is linked to CHARGE syndrome, was recently shown to be a binding partner of SOX2<sup>[65]</sup>. SOX2 is linked to AEG syndrome, which is clinically related to CHARGE syndrome. In fact, in some the diagnosis of AEG or CHARGE is hard to distinguish from each other<sup>[57,58]</sup>. Thus, some of the congenital pulmonary abnormalities may be part of very complex syndromes, and it may be that the interactions between SOX2 and CHD7, or other proteins, may result in combinations of different clinical parameters. Therefore, mutations that change the interactions between SOX2 and other proteins, like CHD7, may result in the various clinical manifestations observed in syndromes.

## REFERENCES

- Wells JM, Melton DA. Vertebrate endoderm development. *Annu Rev Cell Dev Biol* 1999; **15**: 393-410 [PMID: 10611967 DOI: 10.1146/annurev.cellbio.15.1.393]
- Chamberlain CE, Jeong J, Guo C, Allen BL, McMahon AP. Notochord-derived Shh concentrates in close association with the apically positioned basal body in neural target cells and forms a dynamic gradient during neural patterning. *Development* 2008; **135**: 1097-1106 [PMID: 18272593 DOI: 10.1242/dev.013086]
- Serls AE, Doherty S, Parvatiyar P, Wells JM, Deutsch GH. Different thresholds of fibroblast growth factors pattern the ventral foregut into liver and lung. *Development* 2005; **132**: 35-47 [PMID: 15576401 DOI: 10.1242/dev.01570]
- Desai TJ, Chen F, Lü J, Qian J, Niederreither K, Dollé P, Chambon P, Cardoso WV. Distinct roles for retinoic acid receptors alpha and beta in early lung morphogenesis. *Dev Biol* 2006; **291**: 12-24 [PMID: 16427040 DOI: 10.1016/j.ydbio.2005.10.045]
- Chen F, Cao Y, Qian J, Shao F, Niederreither K, Cardoso WV. A retinoic acid-dependent network in the foregut controls formation of the mouse lung primordium. *J Clin Invest* 2010; **120**: 2040-2048 [PMID: 20484817]
- Morrissey EE, Hogan BL. Preparing for the first breath: genetic and cellular mechanisms in lung development. *Dev Cell* 2010; **18**: 8-23 [PMID: 20152174 DOI: 10.1016/j.devcel.2009.12.010]
- Herriges M, Morrissey EE. Lung development: orchestrating the generation and regeneration of a complex organ. *Development* 2014; **141**: 502-513 [PMID: 24449833 DOI: 10.1242/dev.098186]
- Warburton D, El-Hashash A, Carraro G, Tiozzo C, Sala F, Rogers O, De Langhe S, Kemp PJ, Riccardi D, Torday J, Belusci S, Shi W, Lubkin SR, Jesudason E. Lung organogenesis. *Curr Top Dev Biol* 2010; **90**: 73-158 [PMID: 20691848 DOI: 10.1016/S0070-2153(10)90003-3]
- Wells JM, Melton DA. Early mouse endoderm is patterned by soluble factors from adjacent germ layers. *Development* 2000; **127**: 1563-1572 [PMID: 10725233]
- Rackley CR, Stripp BR. Building and maintaining the epithelium of the lung. *J Clin Invest* 2012; **122**: 2724-2730 [PMID: 22850882 DOI: 10.1172/JCI60519]
- Burri PH. Structural aspects of postnatal lung development - alveolar formation and growth. *Biol Neonate* 2006; **89**: 313-322 [PMID: 16770071 DOI: 10.1159/000092868]
- Jacobs IJ, Ku WY, Que J. Genetic and cellular mechanisms regulating anterior foregut and esophageal development. *Dev Biol* 2012; **369**: 54-64 [PMID: 22750256 DOI: 10.1016/j.ydbio.2012.06.016]
- Roberts DJ, Johnson RL, Burke AC, Nelson CE, Morgan BA, Tabin C. Sonic hedgehog is an endodermal signal inducing Bmp-4 and Hox genes during induction and regionalization of the chick hindgut. *Development* 1995; **121**: 3163-3174 [PMID: 7588051]
- de Santa Barbara P, Roberts DJ. Tail gut endoderm and gut/genitourinary/tail development: a new tissue-specific role for Hoxa13. *Development* 2002; **129**: 551-561 [PMID: 11830557]
- Litingtung Y, Lei L, Westphal H, Chiang C. Sonic hedgehog is essential to foregut development. *Nat Genet* 1998; **20**: 58-61 [PMID: 9731532 DOI: 10.1038/1717]
- McMahon AP, Ingham PW, Tabin CJ. Developmental roles and clinical significance of hedgehog signaling. *Curr Top Dev Biol* 2003; **53**: 1-114 [PMID: 12509125 DOI: 10.1016/S0070-2153(03)53002-2]
- Moniot B, Biau S, Faure S, Nielsen CM, Berta P, Roberts DJ, de Santa Barbara P. SOX9 specifies the pyloric sphincter epithelium through mesenchymal-epithelial signals. *Development* 2004; **131**: 3795-3804 [PMID: 15240557 DOI: 10.1242/dev.01259]
- de Santa Barbara P, van den Brink GR, Roberts DJ. Development and differentiation of the intestinal epithelium. *Cell Mol Life Sci* 2003; **60**: 1322-1332 [PMID: 12943221 DOI: 10.1007/s00018-003-2289-3]
- Gubbay J, Collignon J, Koopman P, Capel B, Economou A, Münsterberg A, Vivian N, Goodfellow P, Lovell-Badge R. A gene mapping to the sex-determining region of the mouse Y chromosome is a member of a novel family of embryonically expressed genes. *Nature* 1990; **346**: 245-250 [PMID: 2374589]
- Sinclair AH, Berta P, Palmer MS, Hawkins JR, Griffiths BL, Smith MJ, Foster JW, Frischau AM, Lovell-Badge R, Goodfellow PN. A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. *Nature* 1990; **346**: 240-244 [PMID: 1695712]
- Pevny L, Placzek M. SOX genes and neural progenitor identity. *Curr Opin Neurobiol* 2005; **15**: 7-13 [PMID: 15721738]
- Wegner M. From head to toes: the multiple facets of Sox proteins. *Nucleic Acids Res* 1999; **27**: 1409-1420 [PMID: 10037800]
- Harley VR, Lovell-Badge R, Goodfellow PN. Definition of a consensus DNA binding site for SRY. *Nucleic Acids Res* 1994;



- 22: 1500-1501 [PMID: 8190643]
- 24 **Lefebvre V**, Dumitriu B, Penzo-Méndez A, Han Y, Pallavi B. Control of cell fate and differentiation by Sry-related high-mobility-group box (Sox) transcription factors. *Int J Biochem Cell Biol* 2007; **39**: 2195-2214 [PMID: 17625949]
- 25 **Werner MH**, Burley SK. Architectural transcription factors: proteins that remodel DNA. *Cell* 1997; **88**: 733-736 [PMID: 9118214]
- 26 **Wolffe AP**. Architectural transcription factors. *Science* 1994; **264**: 1100-1101 [PMID: 8178167]
- 27 **Pevny LH**, Lovell-Badge R. Sox genes find their feet. *Curr Opin Genet Dev* 1997; **7**: 338-344 [PMID: 9229109]
- 28 **Wilson M**, Koopman P. Matching SOX: partner proteins and co-factors of the SOX family of transcriptional regulators. *Curr Opin Genet Dev* 2002; **12**: 441-446 [PMID: 12100890]
- 29 **Sarkar A**, Hochedlinger K. The sox family of transcription factors: versatile regulators of stem and progenitor cell fate. *Cell Stem Cell* 2013; **12**: 15-30 [PMID: 23290134 DOI: 10.1016/j.stem.2012.12.007]
- 30 **Kamachi Y**, Kondoh H. Sox proteins: regulators of cell fate specification and differentiation. *Development* 2013; **140**: 4129-4144 [PMID: 24086078 DOI: 10.1242/dev.091793]
- 31 **Goodfellow PN**, Lovell-Badge R. SRY and sex determination in mammals. *Annu Rev Genet* 1993; **27**: 71-92 [PMID: 8122913]
- 32 **Hacker A**, Capel B, Goodfellow P, Lovell-Badge R. Expression of Sry, the mouse sex determining gene. *Development* 1995; **121**: 1603-1614 [PMID: 7600978]
- 33 **Koopman P**, Münsterberg A, Capel B, Vivian N, Lovell-Badge R. Expression of a candidate sex-determining gene during mouse testis differentiation. *Nature* 1990; **348**: 450-452 [PMID: 2247150]
- 34 **Kaplan F**. Molecular determinants of fetal lung organogenesis. *Mol Genet Metab* 2000; **71**: 321-341 [PMID: 11001825 DOI: 10.1006/mgme.2000.3040]
- 35 **Sock E**, Rettig SD, Enderich J, Bösl MR, Tamm ER, Wegner M. Gene targeting reveals a widespread role for the high-mobility-group transcription factor Sox11 in tissue remodeling. *Mol Cell Biol* 2004; **24**: 6635-6644 [PMID: 15254231 DOI: 10.1128/MCB.24.15.6635-6644.2004]
- 36 **Lange AW**, Keiser AR, Wells JM, Zorn AM, Whitsett JA. Sox17 promotes cell cycle progression and inhibits TGF-beta/Smad3 signaling to initiate progenitor cell behavior in the respiratory epithelium. *PLoS One* 2009; **4**: e5711 [PMID: 19479035 DOI: 10.1371/journal.pone.0005711]
- 37 **Tompkins DH**, Besnard V, Lange AW, Keiser AR, Wert SE, Bruno MD, Whitsett JA. Sox2 activates cell proliferation and differentiation in the respiratory epithelium. *Am J Respir Cell Mol Biol* 2011; **45**: 101-110 [PMID: 20855650]
- 38 **Park KS**, Wells JM, Zorn AM, Wert SE, Whitsett JA. Sox17 influences the differentiation of respiratory epithelial cells. *Dev Biol* 2006; **294**: 192-202 [PMID: 16574095 DOI: 10.1016/j.ydbio.2006.02.038]
- 39 **Gontan C**, de Munck A, Vermeij M, Grosveld F, Tibboel D, Rottier R. Sox2 is important for two crucial processes in lung development: branching morphogenesis and epithelial cell differentiation. *Dev Biol* 2008; **317**: 296-309 [PMID: 18374910 DOI: 10.1016/j.ydbio.2008.02.035]
- 40 **Rockich BE**, Hrycaj SM, Shih HP, Nagy MS, Ferguson MA, Kopp JL, Sander M, Wellik DM, Spence JR. Sox9 plays multiple roles in the lung epithelium during branching morphogenesis. *Proc Natl Acad Sci USA* 2013; **110**: E4456-E4464 [PMID: 24191021 DOI: 10.1073/pnas.1311847110]
- 41 **Turcatel G**, Rubin N, Menke DB, Martin G, Shi W, Warburton D. Lung mesenchymal expression of Sox9 plays a critical role in tracheal development. *BMC Biol* 2013; **11**: 117 [PMID: 24274029 DOI: 10.1186/1741-7007-11-117]
- 42 **Park J**, Zhang JJ, Moro A, Kushida M, Wegner M, Kim PC. Regulation of Sox9 by Sonic Hedgehog (Shh) is essential for patterning and formation of tracheal cartilage. *Dev Dyn* 2010; **239**: 514-526 [PMID: 20034104 DOI: 10.1002/dvdy.22192]
- 43 **Elluru RG**, Thompson F, Reece A. Fibroblast growth factor 18 gives growth and directional cues to airway cartilage. *Laryngoscope* 2009; **119**: 1153-1165 [PMID: 19358209 DOI: 10.1002/lary.20157]
- 44 **Kanai-Azuma M**, Kanai Y, Gad JM, Tajima Y, Taya C, Kurohmaru M, Sanai Y, Yonekawa H, Yazaki K, Tam PP, Hayashi Y. Depletion of definitive gut endoderm in Sox17-null mutant mice. *Development* 2002; **129**: 2367-2379 [PMID: 11973269]
- 45 **Dong S**, Leung KK, Pelling AL, Lee PY, Tang AS, Heng HH, Tsui LC, Tease C, Fisher G, Steel KP, Cheah KS. Circling, deafness, and yellow coat displayed by yellow submarine (ysb) and light coat and circling (lcc) mice with mutations on chromosome 3. *Genomics* 2002; **79**: 777-784 [PMID: 12036291]
- 46 **Liu K**, Lin B, Zhao M, Yang X, Chen M, Gao A, Liu F, Que J, Lan X. The multiple roles for Sox2 in stem cell maintenance and tumorigenesis. *Cell Signal* 2013; **25**: 1264-1271 [PMID: 23416461 DOI: 10.1016/j.cellsig.2013.02.013]
- 47 **Avilion AA**, Nicolis SK, Pevny LH, Perez L, Vivian N, Lovell-Badge R. Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes Dev* 2003; **17**: 126-140 [PMID: 12514105 DOI: 10.1101/gad.224503]
- 48 **Wood HB**, Episkopou V. Comparative expression of the mouse Sox1, Sox2 and Sox3 genes from pre-gastrulation to early somite stages. *Mech Dev* 1999; **86**: 197-201 [PMID: 10446282]
- 49 **Papanayotou C**, Mey A, Birot AM, Saka Y, Boast S, Smith JC, Samarut J, Stern CD. A mechanism regulating the onset of Sox2 expression in the embryonic neural plate. *PLoS Biol* 2008; **6**: e2 [PMID: 18184035]
- 50 **Ishii Y**, Rex M, Scotting PJ, Yasugi S. Region-specific expression of chicken Sox2 in the developing gut and lung epithelium: regulation by epithelial-mesenchymal interactions. *Dev Dyn* 1998; **213**: 464-475 [PMID: 9853967]
- 51 **Donner AL**, Episkopou V, Maas RL. Sox2 and Pou2f1 interact to control lens and olfactory placode development. *Dev Biol* 2007; **303**: 784-799 [PMID: 17140559]
- 52 **Schlosser G**, Ahrens K. Molecular anatomy of placode development in *Xenopus laevis*. *Dev Biol* 2004; **271**: 439-466 [PMID: 15223346]
- 53 **Uchikawa M**, Ishida Y, Takemoto T, Kamachi Y, Kondoh H. Functional analysis of chicken Sox2 enhancers highlights an array of diverse regulatory elements that are conserved in mammals. *Dev Cell* 2003; **4**: 509-519 [PMID: 12689590]
- 54 **Taranova OV**, Magness ST, Fagan BM, Wu Y, Surzenko N, Hutton SR, Pevny LH. SOX2 is a dose-dependent regulator of retinal neural progenitor competence. *Genes Dev* 2006; **20**: 1187-1202 [PMID: 16651659]
- 55 **Pan W**, Jin Y, Chen J, Rottier RJ, Steel KP, Kiernan AE. Ectopic expression of activated notch or SOX2 reveals similar and unique roles in the development of the sensory cell progenitors in the mammalian inner ear. *J Neurosci* 2013; **33**: 16146-16157 [PMID: 24107947 DOI: 10.1523/Jneurosci.3150-12.2013]
- 56 **Okubo T**, Pevny LH, Hogan BL. Sox2 is required for development of taste bud sensory cells. *Genes Dev* 2006; **20**: 2654-2659 [PMID: 17015430]
- 57 **Williamson KA**, Hever AM, Rainger J, Rogers RC, Magee A, Fiedler Z, Keng WT, Sharkey FH, McGill N, Hill CJ, Schneider A, Messina M, Turnpenny PD, Fantes JA, van Heyningen V, FitzPatrick DR. Mutations in SOX2 cause anophthalmia-esophageal-genital (AEG) syndrome. *Hum Mol Genet* 2006; **15**: 1413-1422 [PMID: 16543359]
- 58 **Que J**, Okubo T, Goldenring JR, Nam KT, Kurotani R, Morrisey EE, Taranova O, Pevny LH, Hogan BL. Multiple dose-dependent roles for Sox2 in the patterning and differentiation of anterior foregut endoderm. *Development* 2007; **134**: 2521-2531 [PMID: 17522155 DOI: 10.1242/dev.003855]
- 59 **Fantes J**, Ragge NK, Lynch SA, McGill NI, Collin JR, Howard-Peebles PN, Hayward C, Vivian AJ, Williamson K, van Heyningen V, FitzPatrick DR. Mutations in SOX2 cause an-



- ophthalmia. *Nat Genet* 2003; **33**: 461-463 [PMID: 12612584]
- 60 **Hagstrom SA**, Pauer GJ, Reid J, Simpson E, Crowe S, Maumenee IH, Traboulsi EI. SOX2 mutation causes anophthalmia, hearing loss, and brain anomalies. *Am J Med Genet A* 2005; **138A**: 95-98 [PMID: 16145681]
  - 61 **Harris-Johnson KS**, Domyan ET, Vezina CM, Sun X. beta-Catenin promotes respiratory progenitor identity in mouse foregut. *Proc Natl Acad Sci USA* 2009; **106**: 16287-16292 [PMID: 19805295 DOI: 10.1073/pnas.0902274106]
  - 62 **Ochieng JK**, Schilders K, Kool H, Boerema-De Munck A, Buscop-Van Kempen M, Gontan C, Smits R, Grosveld FG, Wijnen RM, Tibboel D, Rottier RJ. Sox2 regulates the emergence of lung basal cells by directly activating the transcription of Trp63. *Am J Respir Cell Mol Biol* 2014; **51**: 311-322 [PMID: 24669837 DOI: 10.1165/rcmb.2013-0419OC]
  - 63 **Minoo P**, Su G, Drum H, Bringas P, Kimura S. Defects in tracheoesophageal and lung morphogenesis in Nkx2.1(-/-) mouse embryos. *Dev Biol* 1999; **209**: 60-71 [PMID: 10208743 DOI: 10.1006/dbio.1999.9234]
  - 64 **Boyer LA**, Lee TI, Cole MF, Johnstone SE, Levine SS, Zucker JP, Guenther MG, Kumar RM, Murray HL, Jenner RG, Gifford DK, Melton DA, Jaenisch R, Young RA. Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* 2005; **122**: 947-956 [PMID: 16153702 DOI: 10.1016/j.cell.2005.08.020]
  - 65 **Engelen E**, Akinci U, Bryne JC, Hou J, Gontan C, Moen M, Szumska D, Kockx C, van Ijcken W, Dekkers DH, Demmers J, Rijkers EJ, Bhattacharya S, Philipsen S, Pevny LH, Grosveld FG, Rottier RJ, Lenhard B, Poot RA. Sox2 cooperates with Chd7 to regulate genes that are mutated in human syndromes. *Nat Genet* 2011; **43**: 607-611 [PMID: 21532573 DOI: 10.1038/Ng.825]
  - 66 **Gontan C**, Güttler T, Engelen E, Demmers J, Fornerod M, Grosveld FG, Tibboel D, Görlich D, Poot RA, Rottier RJ. Exportin 4 mediates a novel nuclear import pathway for Sox family transcription factors. *J Cell Biol* 2009; **185**: 27-34 [PMID: 19349578 DOI: 10.1016/j.ydbio.2008.05.306]
  - 67 **Bellusci S**, Grindley J, Emoto H, Itoh N, Hogan BL. Fibroblast growth factor 10 (FGF10) and branching morphogenesis in the embryonic mouse lung. *Development* 1997; **124**: 4867-4878 [PMID: 9428423]
  - 68 **Park WY**, Miranda B, Lebeche D, Hashimoto G, Cardoso WV. FGF-10 is a chemotactic factor for distal epithelial buds during lung development. *Dev Biol* 1998; **201**: 125-134 [PMID: 9740653]
  - 69 **Sekine K**, Ohuchi H, Fujiwara M, Yamasaki M, Yoshizawa T, Sato T, Yagishita N, Matsui D, Koga Y, Itoh N, Kato S. Fgf10 is essential for limb and lung formation. *Nat Genet* 1999; **21**: 138-141 [PMID: 9916808 DOI: 10.1038/5096]
  - 70 **Kato S**, Sekine K. [FGF10, a key factor for limb and lung formation]. *Seikagaku* 2000; **72**: 288-291 [PMID: 10853366]
  - 71 **Colvin JS**, White AC, Pratt SJ, Ornitz DM. Lung hypoplasia and neonatal death in Fgf9-null mice identify this gene as an essential regulator of lung mesenchyme. *Development* 2001; **128**: 2095-2106 [PMID: 11493531]
  - 72 **Abler LL**, Mansour SL, Sun X. Conditional gene inactivation reveals roles for Fgf10 and Fgfr2 in establishing a normal pattern of epithelial branching in the mouse lung. *Dev Dyn* 2009; **238**: 1999-2013 [PMID: 19618463 DOI: 10.1002/dvdy.22032]
  - 73 **Krejci P**, Aklia A, Kaucka M, Sevcikova E, Prochazkova J, Masek JK, Mikolka P, Pospisilova T, Spoustova T, Weis M, Paznekas WA, Wolf JH, Gutkind JS, Wilcox WR, Kozubik A, Jabs EW, Bryja V, Salazar L, Vesela I, Balek L. Receptor tyrosine kinases activate canonical WNT/ $\beta$ -catenin signaling via MAP kinase/LRP6 pathway and direct  $\beta$ -catenin phosphorylation. *PLoS One* 2012; **7**: e35826 [PMID: 22558232 DOI: 10.1371/journal.pone.0035826]
  - 74 **Shu W**, Guttentag S, Wang Z, Andl T, Ballard P, Lu MM, Piccolo S, Birchmeier W, Whitsett JA, Millar SE, Morrissey EE. Wnt/beta-catenin signaling acts upstream of N-myc, BMP4, and FGF signaling to regulate proximal-distal patterning in the lung. *Dev Biol* 2005; **283**: 226-239 [PMID: 15907834 DOI: 10.1016/j.ydbio.2005.04.014]
  - 75 **Cornett B**, Snowball J, Varisco BM, Lang R, Whitsett J, Sinner D. Wntless is required for peripheral lung differentiation and pulmonary vascular development. *Dev Biol* 2013; **379**: 38-52 [PMID: 23523683 DOI: 10.1016/j.ydbio.2013.03.010]
  - 76 **Goss AM**, Tian Y, Tsukiyama T, Cohen ED, Zhou D, Lu MM, Yamaguchi TP, Morrissey EE. Wnt2/2b and beta-catenin signaling are necessary and sufficient to specify lung progenitors in the foregut. *Dev Cell* 2009; **17**: 290-298 [PMID: 19686689 DOI: 10.1016/j.devcel.2009.06.005]
  - 77 **Shu W**, Jiang YQ, Lu MM, Morrissey EE. Wnt7b regulates mesenchymal proliferation and vascular development in the lung. *Development* 2002; **129**: 4831-4842 [PMID: 12361974]
  - 78 **Miller MF**, Cohen ED, Baggs JE, Lu MM, Hogenesch JB, Morrissey EE. Wnt ligands signal in a cooperative manner to promote foregut organogenesis. *Proc Natl Acad Sci USA* 2012; **109**: 15348-15353 [PMID: 22949635 DOI: 10.1073/pnas.1201583109]
  - 79 **Hashimoto S**, Chen H, Que J, Brockway BL, Drake JA, Snyder JC, Randell SH, Stripp BR.  $\beta$ -Catenin-SOX2 signaling regulates the fate of developing airway epithelium. *J Cell Sci* 2012; **125**: 932-942 [PMID: 22421361 DOI: 10.1242/jcs.092734]
  - 80 **Mucenski ML**, Nation JM, Thitoff AR, Besnard V, Xu Y, Wert SE, Harada N, Taketo MM, Stahlman MT, Whitsett JA. Beta-catenin regulates differentiation of respiratory epithelial cells in vivo. *Am J Physiol Lung Cell Mol Physiol* 2005; **289**: L971-L979 [PMID: 16040629 DOI: 10.1152/ajplung.00172.2005]
  - 81 **Li C**, Hu L, Xiao J, Chen H, Li JT, Bellusci S, Delanghe S, Minoo P. Wnt5a regulates Shh and Fgf10 signaling during lung development. *Dev Biol* 2005; **287**: 86-97 [PMID: 16169547]
  - 82 **Sountoulidis A**, Stavropoulos A, Giaglis S, Apostolou E, Monteiro R, Chuva de Sousa Lopes SM, Chen H, Stripp BR, Mummery C, Andreaskos E, Sideras P. Activation of the canonical bone morphogenetic protein (BMP) pathway during lung morphogenesis and adult lung tissue repair. *PLoS One* 2012; **7**: e41460 [PMID: 22916109 DOI: 10.1371/journal.pone.0041460]
  - 83 **Que J**, Choi M, Ziel JW, Klingensmith J, Hogan BL. Morphogenesis of the trachea and esophagus: current players and new roles for noggin and Bmps. *Differentiation* 2006; **74**: 422-437 [PMID: 16916379 DOI: 10.1111/j.1432-0436.2006.00096.x]
  - 84 **Li Y**, Litingtung Y, Ten Dijke P, Chiang C. Aberrant Bmp signaling and notochord delamination in the pathogenesis of esophageal atresia. *Dev Dyn* 2007; **236**: 746-754 [PMID: 17260385 DOI: 10.1002/dvdy.21075]
  - 85 **Fausett SR**, Brunet LJ, Klingensmith J. BMP antagonism by Noggin is required in presumptive notochord cells for mammalian foregut morphogenesis. *Dev Biol* 2014; **391**: 111-124 [PMID: 24631216 DOI: 10.1016/j.ydbio.2014.02.008]
  - 86 **Li Y**, Gordon J, Manley NR, Litingtung Y, Chiang C. Bmp4 is required for tracheal formation: a novel mouse model for tracheal agenesis. *Dev Biol* 2008; **322**: 145-155 [PMID: 18692041 DOI: 10.1016/j.ydbio.2008.07.021]
  - 87 **Domyan ET**, Ferretti E, Throckmorton K, Mishina Y, Nicolis SK, Sun X. Signaling through BMP receptors promotes respiratory identity in the foregut via repression of Sox2. *Development* 2011; **138**: 971-981 [PMID: 21303850 DOI: 10.1242/dev.053694]
  - 88 **Correia-Pinto J**, Gonzaga S, Huang Y, Rottier R. Congenital lung lesions--underlying molecular mechanisms. *Semin Pediatr Surg* 2010; **19**: 171-179 [PMID: 20610189 DOI: 10.1053/j.sempedsurg.2010.03.003]
  - 89 **Ng C**, Stanwell J, Burge DM, Stanton MP. Conservative management of antenatally diagnosed cystic lung malformations. *Arch Dis Child* 2014; **99**: 432-437 [PMID: 24406806 DOI: 10.1136/archdischild-2013-304048]
  - 90 **Langston C**. New concepts in the pathology of congenital

- lung malformations. *Semin Pediatr Surg* 2003; **12**: 17-37 [PMID: 12520470 DOI: 10.1053/spsu.2003.00001]
- 91 **Cass DL**, Quinn TM, Yang EY, Liechty KW, Crombleholme TM, Flake AW, Adzick NS. Increased cell proliferation and decreased apoptosis characterize congenital cystic adenomatoid malformation of the lung. *J Pediatr Surg* 1998; **33**: 1043-1046; discussion 1047 [PMID: 9694091]
- 92 **Fromont-Hankard G**, Philippe-Chomette P, Delezoide AL, Nessimann C, Aigrain Y, Peuchmaur M. Glial cell-derived neurotrophic factor expression in normal human lung and congenital cystic adenomatoid malformation. *Arch Pathol Lab Med* 2002; **126**: 432-436 [PMID: 11900567 DOI: 10.1043/0003-9985(2002)126<0432:GCDNFE>2.0.CO;2]
- 93 **Gonzaga S**, Henriques-Coelho T, Davey M, Zoltick PW, Leite-Moreira AF, Correia-Pinto J, Flake AW. Cystic adenomatoid malformations are induced by localized FGF10 overexpression in fetal rat lung. *Am J Respir Cell Mol Biol* 2008; **39**: 346-355 [PMID: 18421016 DOI: 10.1165/rcmb.2007-0290OC]
- 94 **Volpe MV**, Pham L, Lessin M, Ralston SJ, Bhan I, Cutz E, Nielsen HC. Expression of Hoxb-5 during human lung development and in congenital lung malformations. *Birth Defects Res A Clin Mol Teratol* 2003; **67**: 550-556 [PMID: 14632303 DOI: 10.1002/bdra.10086]
- 95 **Jancelewicz T**, Nobuhara K, Hawgood S. Laser microdissection allows detection of abnormal gene expression in cystic adenomatoid malformation of the lung. *J Pediatr Surg* 2008; **43**: 1044-1051 [PMID: 18558180 DOI: 10.1016/j.jpedsurg.2008.02.027]
- 96 **Wagner AJ**, Stumbaugh A, Tigue Z, Edmondson J, Paquet AC, Farmer DL, Hawgood S. Genetic analysis of congenital cystic adenomatoid malformation reveals a novel pulmonary gene: fatty acid binding protein-7 (brain type). *Pediatr Res* 2008; **64**: 11-16 [PMID: 18391847 DOI: 10.1203/PDR.0b013e318174eff8]
- 97 **Volpe MV**, Archavachotikul K, Bhan I, Lessin MS, Nielsen HC. Association of bronchopulmonary sequestration with expression of the homeobox protein Hoxb-5. *J Pediatr Surg* 2000; **35**: 1817-1819 [PMID: 11101743 DOI: 10.1053/jpsu.2000.19266]
- 98 **Volpe MV**, Vosatka RJ, Nielsen HC. Hoxb-5 control of early airway formation during branching morphogenesis in the developing mouse lung. *Biochim Biophys Acta* 2000; **1475**: 337-345 [PMID: 10913834]
- 99 **Berg C**, Geipel A, Germer U, Pertersen-Hansen A, Koch-Dörfler M, Gembruch U. Prenatal detection of Fraser syndrome without cryptophthalmos: case report and review of the literature. *Ultrasound Obstet Gynecol* 2001; **18**: 76-80 [PMID: 11489232 DOI: 10.1046/j.1469-0705.2001.00374.x]
- 100 **Epelman M**, Kreiger PA, Servaes S, Victoria T, Hellinger JC. Current imaging of prenatally diagnosed congenital lung lesions. *Semin Ultrasound CT MR* 2010; **31**: 141-157 [PMID: 20304322 DOI: 10.1053/j.sult.2010.01.002]
- 101 **Mei-Zahav M**, Konen O, Manson D, Langer JC. Is congenital lobar emphysema a surgical disease? *J Pediatr Surg* 2006; **41**: 1058-1061 [PMID: 16769334 DOI: 10.1016/j.jpedsurg.2006.02.011]
- 102 **Nayar PM**, Thakral CL, Sajwani MJ. Congenital lobar emphysema and sequestration--treatment by embolization. *Pediatr Surg Int* 2005; **21**: 727-729 [PMID: 15995872 DOI: 10.1007/s00383-005-1462-1]
- 103 **Kravitz RM**. Congenital malformations of the lung. *Pediatr Clin North Am* 1994; **41**: 453-472 [PMID: 8196988]
- 104 **Rottier R**, Tibboel D. Fetal lung and diaphragm development in congenital diaphragmatic hernia. *Semin Perinatol* 2005; **29**: 86-93 [PMID: 16050526]
- 105 **Klaassens M**, de Klein A, Tibboel D. The etiology of congenital diaphragmatic hernia: still largely unknown? *Eur J Med Genet* 2009; **52**: 281-286 [PMID: 19464395 DOI: 10.1016/j.jemg.2009.05.005]
- 106 **Beurskens N**, Klaassens M, Rottier R, de Klein A, Tibboel D. Linking animal models to human congenital diaphragmatic hernia. *Birth Defects Res A Clin Mol Teratol* 2007; **79**: 565-572 [PMID: 17469205 DOI: 10.1002/bdra.20370]
- 107 **Veenma DC**, de Klein A, Tibboel D. Developmental and genetic aspects of congenital diaphragmatic hernia. *Pediatr Pulmonol* 2012; **47**: 534-545 [PMID: 22467525 DOI: 10.1002/ppul.22553]
- 108 **Leeuwen L**, Fitzgerald DA. Congenital diaphragmatic hernia. *J Paediatr Child Health* 2014; **50**: 667-673 [PMID: 24528549 DOI: 10.1111/jpc.12508]
- 109 **Brady PD**, Moerman P, De Catte L, Deprest J, Devriendt K, Vermeesch JR. Exome sequencing identifies a recessive PIGN splice site mutation as a cause of syndromic congenital diaphragmatic hernia. *Eur J Med Genet* 2014; **57**: 487-493 [PMID: 24852103 DOI: 10.1016/j.jemg.2014.05.001]
- 110 **Brady PD**, Van Houdt J, Callewaert B, Deprest J, Devriendt K, Vermeesch JR. Exome sequencing identifies ZFPM2 as a cause of familial isolated congenital diaphragmatic hernia and possibly cardiovascular malformations. *Eur J Med Genet* 2014; **57**: 247-252 [PMID: 24769157 DOI: 10.1016/j.jemg.2014.04.006]
- 111 **Torfs CP**, Curry CJ, Bateson TF. Population-based study of tracheoesophageal fistula and esophageal atresia. *Teratology* 1995; **52**: 220-232 [PMID: 8838292 DOI: 10.1002/tera.1420520408]
- 112 **Spitz L**. Esophageal atresia. *Orphanet J Rare Dis* 2007; **2**: 24 [PMID: 17498283 DOI: 10.1186/1750-1172-2-24]
- 113 **de Jong EM**, Felix JF, de Klein A, Tibboel D. Etiology of esophageal atresia and tracheoesophageal fistula: "mind the gap". *Curr Gastroenterol Rep* 2010; **12**: 215-222 [PMID: 20425471 DOI: 10.1007/s11894-010-0108-1]
- 114 **Bishop NB**, Stankiewicz P, Steinhorn RH. Alveolar capillary dysplasia. *Am J Respir Crit Care Med* 2011; **184**: 172-179 [PMID: 21471096 DOI: 10.1164/rccm.201010-1697CI]
- 115 **Kalinichenko VV**, Lim L, Shin B, Costa RH. Differential expression of forkhead box transcription factors following butylated hydroxytoluene lung injury. *Am J Physiol Lung Cell Mol Physiol* 2001; **280**: L695-L704 [PMID: 11238010]
- 116 **Mahlpuu M**, Enerbäck S, Carlsson P. Haploinsufficiency of the forkhead gene Foxf1, a target for sonic hedgehog signaling, causes lung and foregut malformations. *Development* 2001; **128**: 2397-2406 [PMID: 11493558]
- 117 **Kiefer JC**. Back to basics: Sox genes. *Dev Dyn* 2007; **236**: 2356-2366 [PMID: 17584862 DOI: 10.1002/dvdy.21218]
- 118 **Rainger JK**, Bhatia S, Bengani H, Gautier P, Rainger J, Pearson M, Ansari M, Crow J, Mehendale F, Palinkasova B, Dixon MJ, Thompson PJ, Matarin M, Sisodiya SM, Kleinjan DA, Fitzpatrick DR. Disruption of SATB2 or its long-range cis-regulation by SOX9 causes a syndromic form of Pierre Robin sequence. *Hum Mol Genet* 2014; **23**: 2569-2579 [PMID: 24363063 DOI: 10.1093/hmg/ddt647]

P- Reviewer: Gunther T S- Editor: Ji FF

L- Editor: A E- Editor: Liu SQ



## Preimplantation HLA typing: Practical tool for stem cell transplantation treatment of congenital disorders

Anver Kuliev, Svetlana Rechitsky

Anver Kuliev, Svetlana Rechitsky, Reproductive Genetics Institute, Chicago, IL 60657, United States

Author contributions: Kuliev A and Rechitsky S contributed equally to this work, generated the design and wrote the manuscript.

Correspondence to: Anver Kuliev, MD, PhD, Reproductive Genetics Institute, 2825 N Halsted St, Chicago, IL 60657, United States. [anverkuliev@hotmail.com](mailto:anverkuliev@hotmail.com)

Telephone: +1-847-4001515 Fax: +1-847-4001516

Received: March 28, 2014 Revised: July 11, 2014

Accepted: August 27, 2014

Published online: November 27, 2014

### Abstract

It is well known that to achieve an acceptable engraftment and survival in stem cell therapy, an human leukocyte antigens (HLA) identical stem cell transplant is strongly required. However, the availability of the HLA matched donors even among family members is extremely limited, so preimplantation HLA typing provides an attractive practical tool of stem cell therapy for children requiring HLA matched stem cell transplantation. The present experience of preimplantation genetic diagnosis (PGD) for HLA typing of over one thousand cases shows that PGD provides the at-risk couples with the option to establish an unaffected pregnancy, which may benefit the affected member of the family with hemoglobinopathies, immunodeficiencies and other congenital or acquired bone marrow failures. Despite ethical issues involved in preimplantation HLA typing, the data presented below show an extremely high attractiveness of this option for the couples with affected children requiring HLA compatible stem cell transplantation.

© 2014 Baishideng Publishing Group Inc. All rights reserved.

**Key words:** Preimplantation HLA typing; Preimplantation genetic diagnosis; Stem cell transplantation; Hemoglobinopathies; Immunodeficiencies; Aneuploidy testing

**Core tip:** Human leukocyte antigens (HLA) identical stem cell transplant is the key in achieving an acceptable engraftment and survival in stem cell therapy. However, the availability of the HLA matched donors even among family members is extremely limited, so preimplantation HLA typing provides an attractive practical tool of stem cell therapy for children requiring HLA matched stem cell transplantation.

Kuliev A, Rechitsky S. Preimplantation HLA typing: Practical tool for stem cell transplantation treatment of congenital disorders. *World J Med Genet* 2014; 4(4): 105-109 Available from: URL: <http://www.wjgnet.com/2220-3184/full/v4/i4/105.htm> DOI: <http://dx.doi.org/10.5496/wjmg.v4.i4.105>

### INTRODUCTION

Preimplantation human leukocyte antigens (HLA) matching could not be indication for prenatal diagnosis because it is not acceptable to terminate a normal pregnancy only because the foetus is HLA unmatched. However, preimplantation genetic diagnosis (PGD) for this purpose is totally different, as maximum two embryos are transferred so these embryos may be pre-selected to be unaffected one and match to affected siblings to perform HLA matched stem cell transplantation. This was first introduced in combination with mutation analysis for Fanconi anemia (FA), with a total success<sup>[1,2]</sup>. In fact FA was also the first disorder for which cord blood stem cell transplantation has been first performed<sup>[3]</sup>. FA is genetically heterogeneous group of disorders, involving different complementation groups (FANCA, FANCB, FANCC, FANCD and FANCE)<sup>[4-6]</sup>, for which stem cell transplantation is the only treatment as the objective is to restore hematopoiesis which can be done only with HLA identical stem cells, to prevent severe graft vs host disease<sup>[7,8]</sup>.

As will be described in this paper, preimplantation

**Table 1** Experience in preimplantation genetic diagnosis with HLA typing

Disease	Patients	Cycles	No. of embryo transfers	No. embryos transferred	Pregnancy	Birth
Thalassemia/sickle cell disease	51	149	82	130	20	15
FANCA, FANCC, FANCD2, FANCE, FANCI, FANCF	17	53	34	52	7	4
WAS	2	2	2	4	1	1
X-ALD	2	5	1	1	0	0
Hyper IgM	5	8	6	9	3	2
HED-ID	2	9	6	8	2	3
DBA	3	5	3	6	2	2
Krabbe	1	1	1	2	1	2
MD	1	2	1	2	1	2
Chronic granulomatous disease	1	3	3	5	1	1
Total	85	238	139	219	38	32

WAS: Wiscott-aldrich syndrome; X-ALD: X-linked adrenoleukodystrophy; HED-ID: Hypohidrotic ectodermal dysplasia with immune deficiency; DBA: Diamond-Blackfan anemia; MD: Muscular dystrophy.

**Table 2** Chances for detection of disease free and HLA match embryo in preimplantation HLA typing

HLA MATCH only-1/4 (25%)
Autosomal recessive or X-linked free + HLA MATCH-3/4 $\times$ 1/4 = 3/16 (18.75%)
Autosomal dominant free + HLA MATCH-1/2 $\times$ 1/4 = 1/8 (12.5%)
Autosomal recessive or X-linked free + HLA MATCH + ANEUPLOIDY-free-3/4 $\times$ 1/4 $\times$ 1/2 = 3/32 (9.4%)
Autosomal dominant free + HLA MATCH + ANEUPLOIDY-free-1/2 $\times$ 1/4 $\times$ 1/2 = 1/16 (6.25%)

HLA: Human leukocyte antigens.

HLA testing is currently applied not only with PGD for single gene disorders, but also as a sole indication.

## PREIMPLANTATION HLA TYPING WITH AND WITHOUT PGD FOR SINGLE GENE DISORDERS

Our experience on PGD with HLA typing is presented in Table 1, showing that among conditions requiring HLA compatible stem cell transplantation, hemoglobinopathies are the major indication, representing the commonest autosomal recessive diseases in Mediterranean region, Middle East and South East Asia.

### Hemoglobinopathies

Hemoglobinopathies, including thalassemia and sickle cell disease, are autosomal recessive disorders with abnormal production of beta-globin chains that leads to a severe anaemia, requiring a life-long blood transfusion. Prevention of these disorders has been done using fetal diagnosis with reduction of new cases of thalassemia to up to 70% in many populations, including such large countries in the Eastern Mediterranean region, as Greece, Turkey and Iran<sup>[9-11]</sup>. There has been progress also in tin treatment by bone marrow transplantation<sup>[12]</sup>, but this is limited due to unavailability of HLA matched stem cells, that can be overcome by PGD. We introduced PGD for

thalassemia 18 years ago<sup>[13-15]</sup>, and HLA typing is actually a natural extension allowing couples to produce an unaffected child as a potential HLA matched donor for thalassemic sibling.

In our experience, of a total of 293 PGD cycles for 161 couples at risk for producing offspring with hemoglobinopathies, 149 cycles were performed for HLA typing. Polar body (PB) or embryo biopsy was used to identify hemoglobinopathy mutations, and embryo biopsy was also used for HLA testing, in order to identify the embryos containing the maternal and paternal chromosomes 6 identical to the affected siblings, as described in detail elsewhere<sup>[16-18]</sup>.

HLA typing was based on the methods described elsewhere<sup>[19-22]</sup>. The chances to identify unaffected embryos fully matched to thalassemic siblings is 18.75%, as for other autosomal recessive conditions, based on 25% chance of HLA match and 75% chance of having unaffected embryo (Table 2).

Of more than two dozens of different beta-globin gene mutations tested, the most frequent ones were IVSI-110 mutation -100 cases (33%), followed by IVSI-6-39 cases, IVSII-745-23 cases, Codon 8-20 cases, IVSI-1-18 cases, and codon 39 and IVSI-5-16 cases each. Among other mutations were IVSII-2, Codon 5, Codon 6, Codon 41/2, E121K, -29 (A-G)-87, R30T, Cap 1, deletion 69 kb and deletion 13.4 kb. Mutation testing resulted in detection and transfer of 476 unaffected embryos (approximately, 2 embryos per transfer) in 240 (81.9%) of 293 clinical cycles, yielding 67 (27.9%) unaffected pregnancies and birth of 70 thalassemia-free children<sup>[18]</sup>. PGD for thalassemias currently represents approximately 15% of our PGD series of 2028 cases, which is the world's largest series for monogenic conditions<sup>[23]</sup>.

A total of 149 of these PGD cycles were performed for HLA typing, which allowed detecting and transferring unaffected HLA matched embryos in 82 of them (Table 1). Of 824 embryos with conclusive results for testing of beta-globin gene mutations and HLA type, 602 (73.0%) were predicted to be unaffected carriers or normal, of which only 130 (15.8%) appeared to be HLA identical to



the affected siblings, which, as mentioned, is not significantly different from the expectation (Table 2)<sup>[18]</sup>. These embryos were replaced, yielding 20 healthy matched clinical pregnancies. Umbilical cord blood was collected at birth of these children, or bone marrow obtained at 1 year of age, and transplanted or pending, resulting in a successful hematopoietic reconstitution in all of them. Clearly the progress in radical treatment of hemoglobinopathies will depend on the availability of HLA identical donors<sup>[24]</sup>.

PGD for HLA typing has currently been applied as an efficient tool for couples at risk in many PGD centres to ensure having thalassemia-free children who are HLA identical to the affected siblings, to serve a potential donor for stem cells for transplantation treatment. This currently is a practical tool for a use in communities where hemoglobinopathies are endemic and will improve the access to HLA matched bone marrow transplantation of these prevalent conditions.

The other large series of PGD for HLA typing for thalassemia was reported from Turkey, where 236 PGD cycles were performed resulting in birth of 70 thalassemia-free children. Of 48 affected children transplanted (in addition to thalassemia, morbid children with 9 other different conditions was transplanted), successful outcome was observed in 44 of them with a total hematopoietic reconstitution, while the graft failure occurred only in 4 of them<sup>[25-27]</sup>.

### Immunodeficiencies

Preimplantation HLA typing appeared to be of great utility for severe congenital immunodeficiencies (SCID), which is the key in finding matched stem cell transplant to save live of SCID patients. Our accumulated series on PGD for SCID is presented in Table 1<sup>[28]</sup>. A total of 19 PGD cycles for 9 couples for producing affected progeny with the above conditions (this does not include PGD cycles for FA, which will be described below) were performed, including 8 cycles for Hyper IgM (HIGM), 2 for wiscott-aldrich syndrome (WAS), and 9 for hypohidrotic ectodermal dysplasia with immune deficiency (HED-ID). The Table 1 does not include three cases of PGD for AT and one for omen syndrome (OMS), which were performed without HLA typing, as the affected children did not survive by the time of performing PGD. PGD for OMS was the world's first case, which resulted in transfer of two unaffected embryos, yielding the birth of healthy twins. As mentioned, there was no need for HLA typing in this particular case, but the couples with previous OMS children will definitely be potential candidates for performing PGD with HLA typing to provide also an identical HLA donor progeny for stem cell transplantation. This is also highly relevant to ataxia-telangiectasia (AT), which is a progressive, neurodegenerative childhood disease that affects the brain and other body systems. A weakened immune system makes the patients susceptible to recurrent respiratory infections. Although the currently used symptomatic and supportive treatment, including high-dose vitamin regimens, physical

and occupational therapy and gamma-globulin injections to supplement a weakened immune system may be helpful, the prognosis is very poor, patients still dying in their teens.

A single case PGD for AT has been reported previously for a Saudi patient with 3 affected children<sup>[29]</sup>. The disease was caused by a large deletion of more than two thirds of the *AT* gene, which was detected by amplification of one of the deleted exons (exon 19). Of three embryos available for biopsy and testing, one was a deletion free and transferred, resulting in an unaffected pregnancy.

Of 17 couples at risk for producing a progeny with FA, in addition to two carriers of IVS 4+4 A-T mutation in *FANCC* gene, three were carriers of *FANCD2*, *FANCF*, *FANCI*, *FAMCCJ*, and *FANCA* gene mutations. Overall, 52 unaffected HLA matched embryos were transferred in 34 of 53 initiated cycles, resulting in seven unaffected pregnancies and 4 FA free and HLA matched children, as potential donors for their siblings.

Five cycles were performed for X-linked Adrenoleukodystrophy, which affects the nervous system and the adrenal cortex. Of special interest is preimplantation HLA typing for HIGM, which is a rare immunodeficiency characterized by normal or elevated serum IgM levels, with absence of IgG, IgA and IgE, which results in an increased susceptibility to infections.

Of 5 couples with HIGM for whom PGD was performed, one was with C218X mutation in exon 5 of CD40 ligand gene, 3 with maternal mutations C218X exon 4 c.437\_38 ins A, and one with exon 4 c.397 ins T. The maternal mutations were analyzed by PB1 and PB2, followed by HLA and aneuploidy testing in biopsied blastomeres. Of 8 PGD cycles for HLA performed, 9 unaffected HLA matched embryos were transferred in 6 cycles, resulting in 3 clinical pregnancies and birth of 2 healthy babies, as potential donors of HLA compatible stem cells for their siblings.

The first attempt of cord blood transplantation from one of the babies did not result in acceptable engraftment, so the second transplantation was performed one year later, using bone marrow mixed with the remaining portion of the cord blood sample, which provided better results in achieving successful engraftment and reconstitution of the sibling's bone marrow, and resulting in a total cure of the patient.

A total of 11 cycles were performed for WAS and X-linked HED-ID, in which 12 embryos were detected to be unaffected and HLA matched (8 for HED-ID and 4 for WAS), and transferred in 8 cycles, resulting in birth of 4 unaffected babies (3 free of HED-ID and 1 free of WAS), confirmed to be HLA matched to affected sibling.

### Preimplantation HLA typing as a sole indicator

As presented in Table 3, in addition to 238 PGD for HLA cycles, 98 cycles were performed for preimplantation HLA matching without testing for causative gene. These couples were wishing to have another child anyway, but requested that if these children could become a

**Table 3** Preimplantation HLA typing with and without preimplantation genetic diagnosis

Preimplantation testing	Patients	Cycles	No. of embryo transfers	No. embryos transferred	Pregnancy/birth
HLA TESTING ONLY	46	98	65	99	24/19
HLA + MUTATION	85	238	139	219	38/ 32
Total	131	336	204	318	62/51

HLA: Human leukocyte antigens.

source of stem cell transplant to save live of siblings with acquired bone marrow failures, such as sporadic Diamond-Blackfan anemia<sup>[30]</sup>.

There was no difference in performing preimplantation HLA testing without PGD, except limiting the analysis of the day 3 or day 5 embryos to only HLA typing, with the sibling requiring stem cell transplantation, using a multiplex hemi-nested PCR system.

In a total of 98 clinical cycles from 46 couples performed with a primary indication of HLA typing, 99 HLA matched embryos were pre-selected for transfer. Proportion of embryos predicted to be HLA matched to the affected siblings was 21.5%, not significantly different from the expected 25% (Table 2). The transfer of 99 HLA matched embryos in 65 clinical cycles, resulted in 24 pregnancies and 19 HLA identical deliveries, with already available results of complete cure<sup>[30]</sup>.

## LIMITATIONS AND FUTURE PROSPECT OF PGD FOR HLA TYPING

A relatively high frequency of recombination in the HLA region is one of the major limitations of PGD for HLA typing, which may affect not only the accuracy of preimplantation HLA typing, but also the outcome of stem cell transplantation. In our experience, of 1713 embryos tested for HLA, 1634 (95.5%) were non-recombinant, 52 (3%) with maternal, and 27 (1.5%) with paternal recombination. The major problem in performing PGD for HLA may be faced when the preparatory testing identified the sibling being with maternal recombination, so it could be unrealistic to identify the exact match, so the couples should be informed that only relatively close match may be possible, which may be discussed with paediatric haematologist in the pre-selection process of the embryos for transfer.

The other important limitation is that the majority of cases are in couples of advanced maternal age, so aneuploidy testing is usually an integral part of the procedure. Although the chances of pre-selecting unaffected HLA matched embryos that could be also euploid is quite low, our preliminary results of the reproductive outcome comparison between the groups of combined PGD/HLA with and without aneuploidy testing showed a significant difference. Despite transferring a lower number of embryos, the pregnancy rate was higher in the aneuploidy testing group, suggesting the potential utility of aneuploidy testing in preimplantation HLA typing, allowing the avoidance of transfer of those HLA identical embryos that are

chromosomally abnormal, which are destined to be lost anyway either before or after implantation.

Therefore, patients should be properly counseled to be aware of the limits of the procedure and even lower proportion of available embryos for transfer than may have been predicted, depending also on the maternal age.

## REFERENCES

- 1 **Verlinsky Y**, Rechitsky S, Schoolcraft W, Strom C, Kuliev A. Designer babies - are they a reality yet? Case report: simultaneous preimplantation genetic diagnosis for Fanconi anaemia and HLA typing for cord blood transplantation. *Reprod Biomed Online* 2000; **1**: 31 [PMID: 12804194 DOI: 10.1016/S1472-6483(10)61896-2]
- 2 **Verlinsky Y**, Rechitsky S, Schoolcraft W, Strom C, Kuliev A. Preimplantation diagnosis for Fanconi anemia combined with HLA matching. *JAMA* 2001; **285**: 3130-3133 [PMID: 11427142 DOI: 10.1001/jama.285.24.3130]
- 3 **Gluckman E**, Broxmeyer HA, Auerbach AD, Friedman HS, Douglas GW, Devergie A, Esperou H, Thierry D, Socie G, Lehn P. Hematopoietic reconstitution in a patient with Fanconi's anemia by means of umbilical-cord blood from an HLA-identical sibling. *N Engl J Med* 1989; **321**: 1174-1178 [PMID: 2571931 DOI: 10.1056/NEJM198910263211707]
- 4 **Strathdee CA**, Duncan AM, Buchwald M. Evidence for at least four Fanconi anaemia genes including FACC on chromosome 9. *Nat Genet* 1992; **1**: 196-198 [PMID: 1303234 DOI: 10.1038/ng0692-196]
- 5 **Strathdee CA**, Gavish H, Shannon WR, Buchwald M. Cloning of cDNAs for Fanconi's anaemia by functional complementation. *Nature* 1992; **356**: 763-767 [PMID: 1574115 DOI: 10.1038/356763a0]
- 6 **Whitney MA**, Saito H, Jakobs PM, Gibson RA, Moses RE, Grompe M. A common mutation in the FACC gene causes Fanconi anaemia in Ashkenazi Jews. *Nat Genet* 1993; **4**: 202-205 [PMID: 8348157 DOI: 10.1038/ng0693-202]
- 7 **Gluckman E**, Devergie A, Schaison G, Bussel A, Berger R, Sohier J, Bernard J. Bone marrow transplantation in Fanconi anaemia. *Br J Haematol* 1980; **45**: 557-564 [PMID: 7000153 DOI: 10.1111/j.1365-2141.1980.tb07178.x]
- 8 **Wagner J**, Davies SM, Auerbach AD. Hematopoietic stem cell transplantation in the treatment of Fanconi anemia. In: Forman SJ, Blum KG, Thomas ED, eds. *Hematopoietic Cell Transplantation*. 2nd ed. Malden, Mass: Blackwell Science Inc, 1999: 1204-1219
- 9 **Modell B**, Kuliev A. The history of community genetics: the contribution of the haemoglobin disorders. *Community Genet* 1998; **1**: 3-11 [PMID: 15178981 DOI: 10.1159/000016129]
- 10 **Canatan D**. Hemoglobinopathy prevention programme in Turkey. *Thalassemia Reports* 2011; **1**: 2
- 11 **Samavat A**. Thalassemia prevention programme: A Model from Iran. *Thalassemia Reports* 2011; **1**: 2
- 12 **Lucarelli G**, Andreani M, Angelucci E. The cure of thalassemia by bone marrow transplantation. *Blood Rev* 2002; **16**: 81-85 [PMID: 12127951 DOI: 10.1054/blre.2002.0192]
- 13 **Kuliev A**, Rechitsky S, Verlinsky O, Ivakhnenko V, Evsikov S, Wolf G, Angastiniotis M, Georgiou D, Kukharensko V,

- Strom C, Verlinsky Y. Preimplantation diagnosis of thalassemias. *J Assist Reprod Genet* 1998; **15**: 219-225 [PMID: 9604751]
- 14 **Kuliev A**, Rechitsky S, Verlinsky O, Ivakhnenko V, Cieslak J, Evsikov S, Wolf G, Angastiniotis M, Kalakoutis G, Strom C, Verlinsky Y. Birth of healthy children after preimplantation diagnosis of thalassemias. *J Assist Reprod Genet* 1999; **16**: 207-211 [PMID: 10224564 DOI: 10.1023/A:1020316924064]
  - 15 **Kanavakis E**, Vrettou C, Palmer G, Tzetis M, Mastrominas M, Traeger-Synodinos J. Preimplantation genetic diagnosis in 10 couples at risk for transmitting beta-thalassaemia major: clinical experience including the initiation of six singleton pregnancies. *Prenat Diagn* 1999; **19**: 1217-1222 [PMID: 10660958]
  - 16 **Rechitsky S**, Kuliev A, Tur-Kaspa I, Morris R, Verlinsky Y. Preimplantation genetic diagnosis with HLA matching. *Reprod Biomed Online* 2004; **9**: 210-221 [PMID: 15333254 DOI: 10.1016/S1472-6483(10)62132-3]
  - 17 **Kuliev A**, Rechitsky S, Verlinsky O, Tur-Kaspa I, Kalakoutis G, Angastiniotis M, Verlinsky Y. Preimplantation diagnosis and HLA typing for haemoglobin disorders. *Reprod Biomed Online* 2005; **11**: 362-370 [PMID: 16176679 DOI: 10.1016/S1472-6483(10)60845-0]
  - 18 **Kuliev A**, Pakhalchuk T, Verlinsky O, Rechitsky S. Preimplantation diagnosis: efficient tool for human leukocyte antigen matched bone marrow transplantation for thalassemia. *Thalassemia Reports* 2011; **1**: e1
  - 19 **Gibson RA**, Morgan NV, Goldstein LH, Pearson IC, Kesterton IP, Foot NJ, Jansen S, Havenga C, Pearson T, de Ravel TJ, Cohn RJ, Marques IM, Dokal I, Roberts I, Marsh J, Ball S, Milner RD, Llerena JC, Samochatova E, Mohan SP, Vasudevan P, Birjandi F, Hajianpour A, Murer-Orlando M, Mathew CG. Novel mutations and polymorphisms in the Fanconi anemia group C gene. *Hum Mutat* 1996; **8**: 140-148 [PMID: 8844212]
  - 20 **Mitsunaga S**, Tokunaga K, Kashiwase K, Akaza T, Tadokoro K, Juji T. A nested PCR-RFLP method for high-resolution typing of HLA-A alleles. *Eur J Immunogenet* 1998; **25**: 15-27 [PMID: 9587741 DOI: 10.1046/j.1365-2370.1998.00093.x]
  - 21 **Blasczyk R**, Hahn U, Wehling J, Huhn D, Salama A. Complete subtyping of the HLA-A locus by sequence-specific amplification followed by direct sequencing or single-strand conformation polymorphism analysis. *Tissue Antigens* 1995; **46**: 86-95 [PMID: 7482512 DOI: 10.1111/j.1399-0039.1995.tb02483.x]
  - 22 **Foissac A**, Salhi M, Cambon-Thomsen A. Microsatellites in the HLA region: 1999 update. *Tissue Antigens* 2000; **55**: 477-509 [PMID: 10902606 DOI: 10.1034/j.1399-0039.2000.550601.x]
  - 23 **Rechitsky S**, Kuliev A. 2010 Novel indications for preimplantation genetic diagnosis. 10th International Congress on Preimplantation Genetic Diagnosis, 5-8 May, Montpellier, France. *Reprod Biomed Online* 2010; **20** (Suppl): S1-2
  - 24 **Gaziev J**, Lucarelli G. Stem cell transplantation for thalassemia. *Reprod Biomed Online* 2005; **10**: 111-115 [PMID: 15705305]
  - 25 **Kahraman S**, Karlikaya G, Sertyel S, Karadayi H, Findikli N. Clinical aspects of preimplantation genetic diagnosis for single gene disorders combined with HLA typing. *Reprod Biomed Online* 2004; **9**: 529-532 [PMID: 15588472 DOI: 10.1016/S1472-6483(10)61637-9]
  - 26 **Kahraman S**, Beyazyurek C, Ekmekci CG. Seven years of experience of preimplantation HLA typing: a clinical overview of 327 cycles. *Reprod Biomed Online* 2011; **23**: 363-371 [PMID: 21782513 DOI: 10.1016/j.rbmo.2011.05.016]
  - 27 **Kahraman S**. PGD for HLA: Clinical Outcomes of HLA compatible transplantation following PGD. *Reprod Biomed Online* 2013; **26** (Suppl 1): S9-10 [DOI: 10.1016/S1472-6483(13)60030-9]
  - 28 **Verlinsky Y**, Rechitsky S, Sharapova T, Laziuk K, Barsky I, Verlinsky O, Tur-Kaspa I, Kuliev A. Preimplantation diagnosis for immunodeficiencies. *Reprod Biomed Online* 2007; **14**: 214-223 [PMID: 17298726 DOI: 10.1016/S1472-6483(10)60790-0]
  - 29 **Hellani A**, Laugé A, Ozand P, Jaroudi K, Coskun S. Pregnancy after preimplantation genetic diagnosis for Ataxia Telangiectasia. *Mol Hum Reprod* 2002; **8**: 785-788 [PMID: 12149412 DOI: 10.1093/molehr/8.8.785]
  - 30 **Verlinsky Y**, Rechitsky S, Sharapova T, Morris R, Taranissi M, Kuliev A. Preimplantation HLA testing. *JAMA* 2004; **291**: 2079-2085 [PMID: 15126435 DOI: 10.1001/jama.291.17.2079]

**P- Reviewer:** Eigens R, ML Chamayou S **S- Editor:** Song XX  
**L- Editor:** A **E- Editor:** Liu SQ





Published by **Baishideng Publishing Group Inc**

8226 Regency Drive, Pleasanton, CA 94588, USA

Telephone: +1-925-223-8242

Fax: +1-925-223-8243

E-mail: [bpgoffice@wjgnet.com](mailto:bpgoffice@wjgnet.com)

Help Desk: <http://www.wjgnet.com/esps/helpdesk.aspx>

<http://www.wjgnet.com>

