ISSN 2220-3184 (online)

World Journal of *Medical Genetics*

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

Published by Baishideng Publishing Group Inc

WJMG World Journal of Medical Genetics

A peer-reviewed, online, open-access journal of medical genetics

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).

WJMG World Journal of Medical Genetics

Contents

Quarterly Volume 4 Number 2 May 27, 2014

Contents

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

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.

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/esps/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/esps/helpdesk.aspx
http://www.wjgnet.com

PUBLICATION DATE
May 27, 2014

INSTRUCTIONS TO AUTHORS
Full instructions are available online at http://www.wjgnet.com/2220-3184/g_info_20100722180909.htm

ONLINE SUBMISSION
http://www.wjgnet.com/esps/

World Journal of Medical Genetics

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.5496/wjmg.v4.i2.6

World J Med Genet 2014 May 27; 4(2): 6-18
ISSN 2220-3184 (online)

TOPIC HIGHLIGHT

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
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.

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.wjgnet.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[1]. 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[2-4]. 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[5]. 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[5]. 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[6,7].

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 α 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[8]. These compounds can be administered in conjunction with pegylated-interferon α and ribavirin for a short period of time to achieve an improved sustained virological response[8] 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[9,10]. 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 quasispecies. This term defines the closely related sequences of a heterogeneous viral population infecting a single individual[11]. Quasispecies structure has been associated with the failure of infected people to clear the virus and the subsequent development of a chronic infection[12]. 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[13-15] 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[16,17] 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[18,19] (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[20-25]. 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[20-23,26-28], and also by the establish-

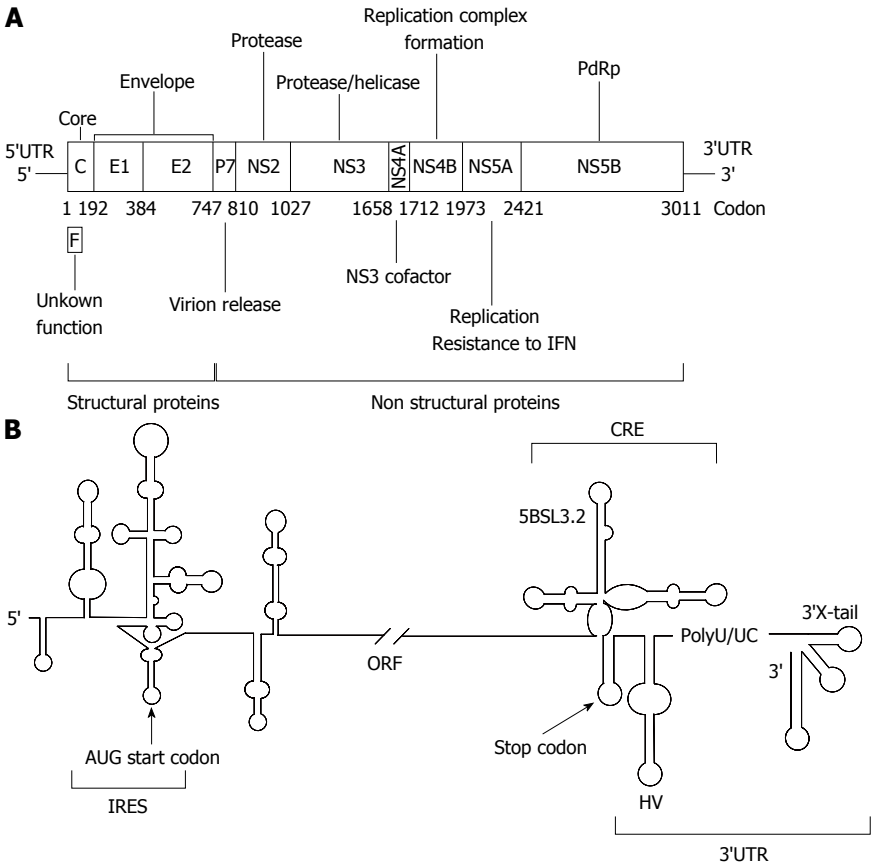

Figure 1 Genetic organization of the hepatitis C virus genomic RNA. A: A schematic view of the hepatitis C virus (HCV) genome, showing the 5' and 3' untranslatable regions (UTRs) and the genes encoding for the different viral proteins. Numbers allude to codon positions; B: A detailed diagram of the secondary 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[29-31]. 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[32,33]. 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[34-37]. 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[38,39]. The 5' core coding sequence helps in the preservation of structures important for IRES activity and replication (Figure 1B)[19,40-43]. 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)[44,45] and as a regulatory partner of the IRES function[25].

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[46-48]. 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[49]. The further binding of eIF3 aids the incorporation of the ternary complex eIF2-GTP-tRNA$_i^{Met}$ to yield the 48S particle[48,50]. The formation of the active translation complex 80S is assessed by the GTP hydrolysis for the concurrent release of eIF2 and eIF3[51] 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*[52] 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)[53], plus a short stem-loop containing the start codon (domain Ⅳ)[54]. 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)[49,55]. The 3D architecture and several single RNA structural elements are highly conserved among other closely related viruses from the *Flaviviridae* family[46,56,57].

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[56,58] and capped by an apical loop (Figure 2). Domain Ⅱ adopts an overall distorted L-shape conformation[59] because of the twist forced by the internal E-loop. This folding is conserved in HCV and related viruses[51].

While domain Ⅱ is not essential for 40S recruitment[47,60,61], 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[46,48,60,62,63]. Analysis by cryo-EM have demonstrated that the bend in domain Ⅱ is a requisite for changing the conformation of the 40S ribosomal subunit[49,64], in a reminiscent manner to that shown by eIF1 in the canonical cap-dependent translation initiation mechanism[65]. The apical loop placed in subdomain Ⅱb would also contribute to this structural rearrangement in the ribosome[66]. 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[67,68]. 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[48,51,63,64].

The large, highly branched domain Ⅲ consists of six hairpins (designated subdomains Ⅲa to Ⅲf) organized around three- and four-way junctions (Figure 2)[52], which can be identified as recruiting centers for the translational machinery. The apical Ⅲabc junction is the platform for the binding of eIF3[50,69]. 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[70]. 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 Ⅲd (Figure 2). This element is the core 40S binding center[60,71,72]. Its structure is that of a dynamic stem-loop with an internal E loop motif and an apical loop with typical U-turn geometry[73,74]. 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 Ⅲd seems to be a determinant partner in the acquisition of the functional folding of the surrounding domains[53].

The basal fragment of domain Ⅲ (subdomains Ⅲe and Ⅲf) includes the highly conserved, complex double-pseudoknot motif (PK1 and PK2; Figure 2)[55], 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[75]. 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[76]. 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[54].

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[33,77,78] and also may act as enhancer of the IRES function[20-24]. It is about 240 nts long sequence placed at the 3' end of the viral genome[79], 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[33,78]; (2) A polyU/UC tract, whose length and composition is a critical determinant of efficient HCV replication in cell culture[80]. It has been proved that a minimum of 26 U nts homopolymer is enough for efficient amplification of the viral RNA[33,78]. Further, it can act outside of its usual molecular context, thus suggesting that this is not only a linker region[80]. The polyU/UC stretch also interacts with host factors related to cellular protein synthesis, such as polypyrimidine tract-binding protein[81,82], the La autoantigen[83], heterogeneous nuclear ribonucleoprotein C (Gon-

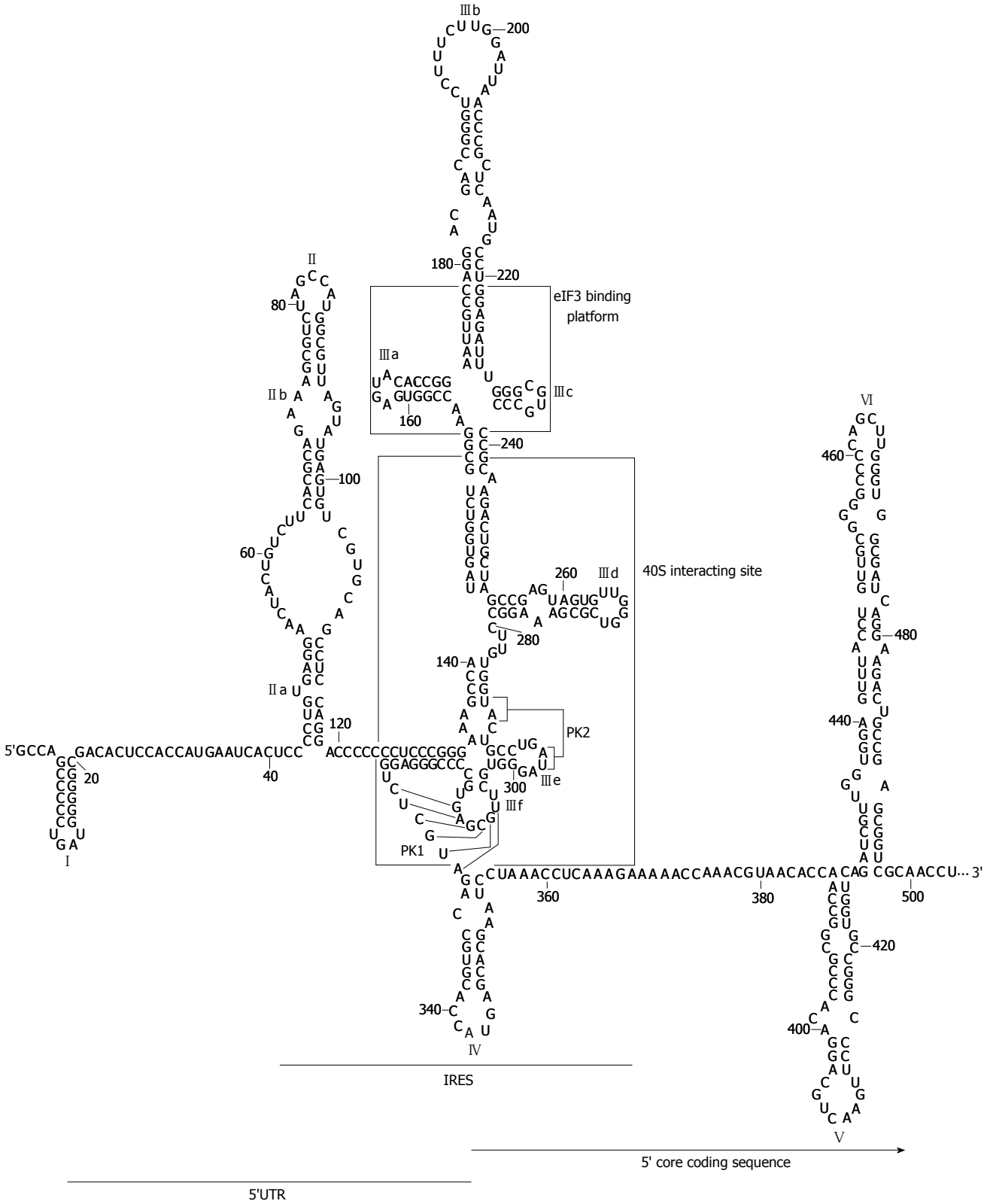

Figure 2 Secondary structure proposed for the hepatitis C virus internal ribosome entry site. The 5' untranslatable region (UTR) plus domains Ⅴ and Ⅵ located at the core coding sequence are included. Minimum region for internal ribosome entry site (IRES) activity is depicted. Domains involved in the 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[84], among others[85,86]. It seems likely that the recruitment of these factors could contribute to regulate viral translation mediated by the IRES region[20,21]; 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[35] (Figure 3). Both predicted structures pre-

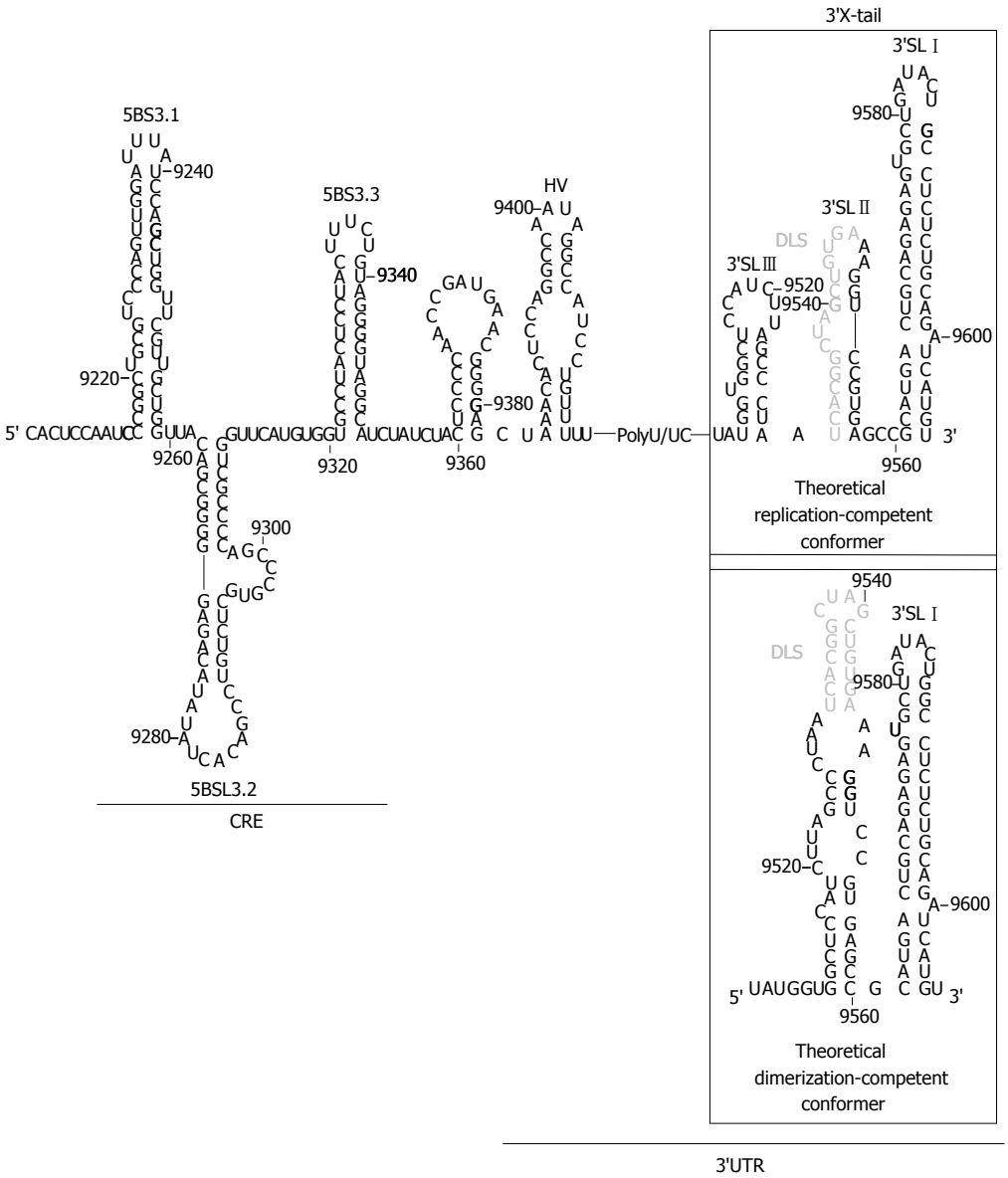

Figure 3 Theoretical folding of the 3' end of the hepatitis C virus genomic RNA. The figure shows the entire 3' end containing the cis-acting replicating element (CRE) region plus the 3' untranslatable region (UTR). The 3'X-tail folds into two different conformers with distinct functional roles. Dimer 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[87-89]. The 55 nts segment placed upstream of 3'SLI folds either as two stem-loops, named 3'SLⅡ and 3'SLⅢ, 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[81,90-97], 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[30,31,33,98,99].

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[100]. 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

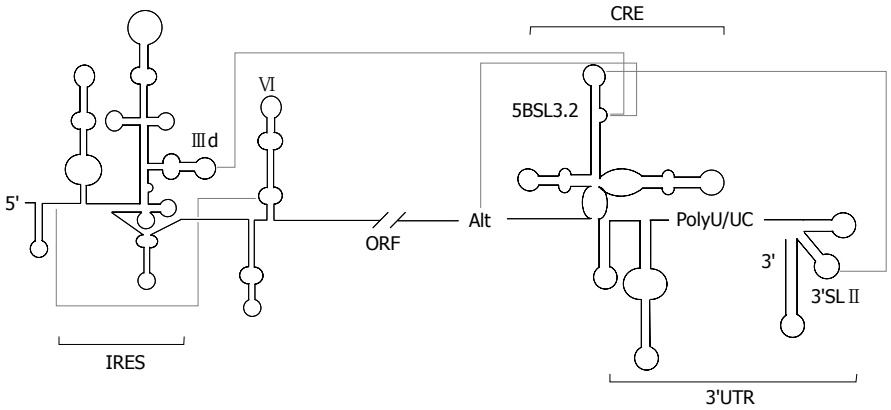

Figure 4 Long-range RNA-RNA interactions in the hepatitis C virus genome. Detailed diagram of the interacting network in the genomic hepatitis C virus RNA. The minimum region for internal ribosome entry site (IRES) activity is marked. The 3' end of the viral genome, containing both the cis-acting replicating 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[5].

constrain that was related to the presence of an alternative ORF coding for the so-called protein F[101,102] (Figure 1A); and to the existence of structural RNA domains with functional roles in the HCV cycle[38] (Figure 2). While the production and biological role of protein F is still a controversial issue[103], 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[19,40,42,43].

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)[40,41]. 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[104]. 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[105,106]. Alternatively, Roberts *et al*[107] 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[38,39,108,109]. One of these structural elements, the so-called domain 5BSL3.2 or SL9266, is embedded into a cruciform structure delimited 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[44,45,98,109], the role of the two additional domains 5BSL3.1 and 3.3 is still unclear[80].

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)[45,98]. Disruptions in either the sequence or its folding lead to replication-incompetent HCV genomes[45,98]. 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)[110] and, more likely, distal RNA functional elements[29,98,99,109]. 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[98]. Domain 5BSL3.2 has been also shown to act as an inhibitory element of the viral IRES function[25], 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[21-23,27,28,111], 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

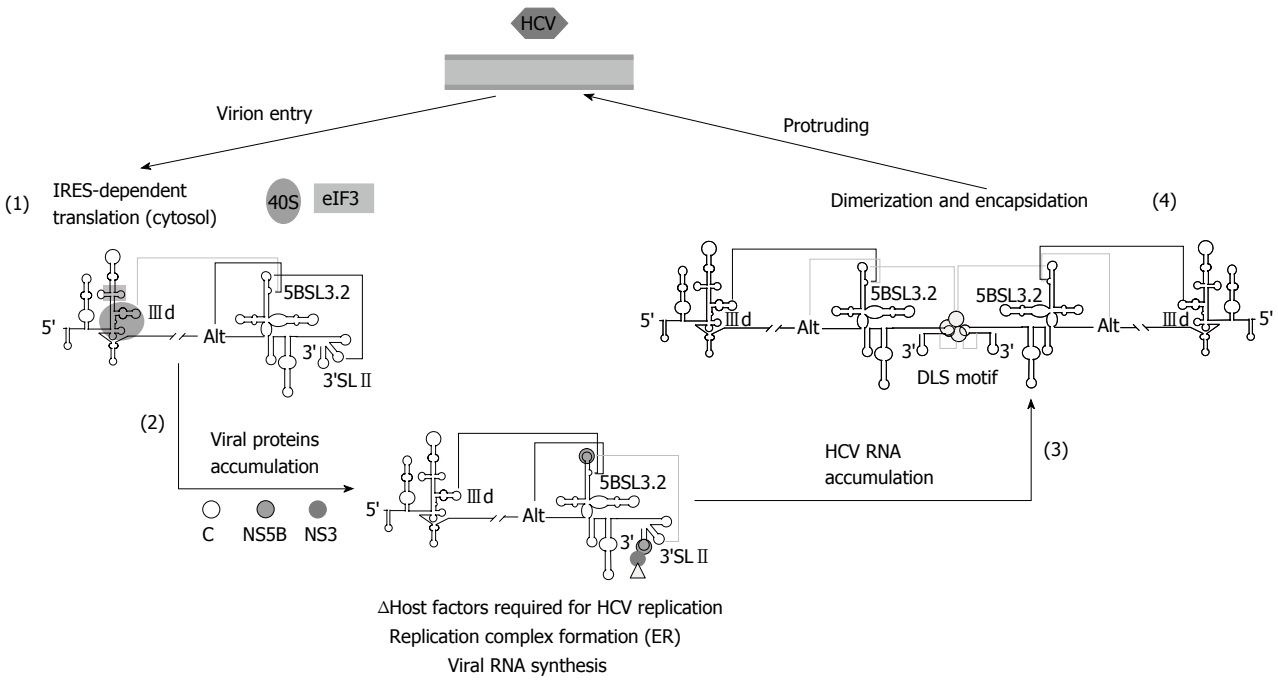

Figure 5 Proposed model for the participation of long-distant RNA-RNA interactions in the consecution of the hepatitis C virus infective cycle. Virions penetrate inside the cell and the genomic RNA is released in the cytoplasm to initiate viral translation (1). The 40S ribosomal subunit binds to the Ⅲd 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'SLⅡ. 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'SLⅡ domain. This would displace the interactions balance toward the Ⅲd-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 Ⅲd-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[31].

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'SLⅡ within the 3'X-tail[98,99,112]. The resulting kissing loop contact contributes to the structural organization of the 3'X-tail and is essential for HCV replication[98]. The 8-nts bulge may establish two different interactions: (1) one with the apical loop of the subdomain Ⅲd of the IRES region[29,112], which is related to the aforementioned translational inhibitory effect[25]; (2) the second with the Alt sequence, centred around position 9110, upstream of the CRE element[99,109,112]. This interaction is again critical for the synthesis of the viral genomic RNA[109]. 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[30,31,99]. 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[30], which could be associated to the regulation of viral translation[25]. Conversely, the interaction Ⅲd-5BSL3.2 induce structural rearrangements in the 3'X-tail that finally lead to the conformational transition of the essential domains 3'SLⅡ and 3'SLⅢ, which switch to a single stem-loop folding that exposes the DLS motif in an apical loop[31] (Figure 3). Importantly, it has been recently reported that all these interactions are equally probable[112]. 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[31], 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'SLⅡ, 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)[92,109,111,113-117]. In this context, both the 5BSL3.2-Ⅲd and 5BSL3.2-Alt interactions could be equally feasible. Swapping between them could contribute to the creation of a translational repressed state[25] and an enhanced replicative process[109]. The subsequent amplification of viral RNA molecules would displace the structural equilibrium between the 5BSL3.2-Ⅲd 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[34].

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[6,39]-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[7], 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[7]. 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[118]. 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, Dusheiko 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**, Bochkaeva 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, Lavergne 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, Lavergne JP, Ficheux 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, Lavergne 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 IIId 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 IIId 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.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

World Journal of Medical Genetics

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.5496/wjmg.v4.i2.19

World J Med Genet 2014 May 27; 4(2): 19-26
ISSN 2220-3184 (online)

REVIEW

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
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.

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[1]. 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)[2]. 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[3]. 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[4].

HBV has been classified into at least 9 genotypes (A through H and J) and shown to have a distinct geographical distribution[5,6]. 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[7,8], 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[9].

EPIDEMIOLOGY OF HBV INFECTION

The prevalence of chronic HBV infection varies greatly in different parts of the world and can be categorized as high (≥ 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[10]; has intermediate to high endemicity in Indonesia[11-13], Cambodia[10,14], Thailand[15,16], the Philippines[17-19], Vietnam[20,21] and Laos[10,24]; low to high endemicity in Malaysia[22,23] and China[25-27]; and intermediate endemicity in Singapore[28,29], Brunei[30,31] and South Korea[32,33]. Japan is the only country with low endemicity of HBV infection in Asia[22].

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

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

HBsAg: Hepatitis B surface antigen.

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[3].

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)[6,34,35], 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[36].

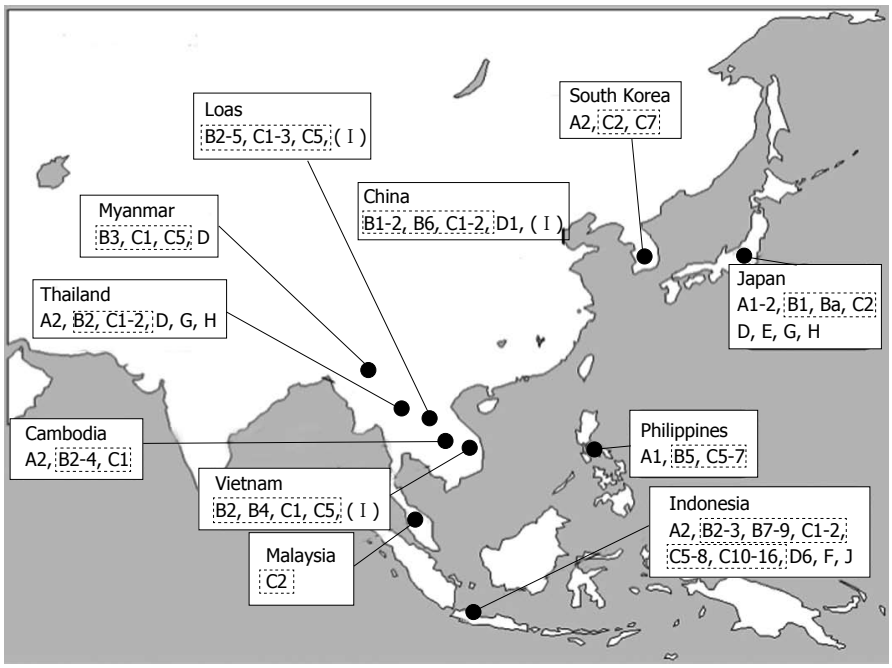

Figure 1 Genotype/subgenotype distribution in East and Southeast Asia. Subgenotypes of genotypes B and C commonly found in Asia are encircled with dotted lines.

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[37] (Figure 1) and co-infection has led to a frequent occurrence of recombination between these two genotypes[38,39]. 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[40,41], 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[42] while B4 is in Vietnam[5]. B5 was initially reported in 2006 from the Philippines[43]. B6 was identified in 2007 from the Arctic[39]. B7 to B9 were isolated in eastern Indonesia during the years 2007 to 2011[44-46]. 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[47] and C4 in the Aborigines from Australia[48]. C5 was initially reported in 2006 from the Philippines with B5[44]. C6 was identified from a Papuan population in Indonesia[13,49] and the Philippines[50] in 2008. Surprisingly, ten novel subgenotypes (C7 to C16) were isolated in Indonesia during 2009 to 2012[45,51-54]. Subgenotypes D1 to D4 of genotype D are widely distributed globally[5], D5 in India[55] and D6 in Papua, Indonesia[13]. 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[56], 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[35]. 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[57]. 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[58]. 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[59-62]. HBV genotypes are also related to the clinical characteristics[63]. 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[59,64-67]. On the other hand, genotypes D and A are prevalent in the southwest Asian countries, such as India and Pakistan[68]. 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	V173l, 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[69,70]. 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[71]. A study from north India also showed that the BCP A1762T/G1764A mutation was associated with progressive liver diseases among genotype D[72]. 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[73,74]. 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[42,75]. In addition, several studies from Taiwan and Japan showed that the pre-S mutation also contributed to the progressive liver disease and HCC[76,77]. The progression from acute hepatitis to chronic infection occurs more frequently in genotype A (23%) compared with genotypes B (11%) and C (7%)[78]. 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[79].

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[80,81].

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[82]. It was shown that patients infected with HBV genotypes A and B showed better response than those with genotypes C and D[83-87]. 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[88].

Currently, lamivudine, adefovir, entecavir, telbivudine and tenofovir have been approved for the treatment of chronic hepatitis B (Table 2). Lamivudine (Zeffix®) was first introduced in 1999 and the clinical efficacy was shown by a long-term follow-up study[93,94]. 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[95,96]. Entecavir (Baraclude®) 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[97,98]. 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[88]. 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**. C hanging 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, Magnius 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, Mishiro 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, Chongsrisawat 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 **Louisirirotchanakul 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, Hasnian 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/WPSAR.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, Samountry 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**, Jawaid 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, Brahmbhatt 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 E, Izumi N, Okuse 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, Condreay 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 intrauterine 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

World Journal of Medical Genetics

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.5496/wjmg.v4.i2.27

World J Med Genet 2014 May 27; 4(2): 27-33
ISSN 2220-3184 (online)

REVIEW

Acquisition and dissemination mechanisms of CTXΦ 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. rupakbhadra@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 (CTXΦ), 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 CTXΦ, genetic and ecological factors which support CTXΦ integration as well as production of virion from chromosomally integrated phage genome and interactions of CTXΦ with other genetic elements present in the genomes of *V. cholerae* is important for learning more about the biology of cholera pathogen.

Key words: *Vibrio cholerae*; Cholera toxin phage; VGJΦ; 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 (CTXΦ) 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 CTXΦ and related IMGEs of *V. cholerae*.

Das B, Nair GB, Bhadra RK. Acquisition and dissemination mechanisms of CTXΦ 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:

http://dx.doi.org/10.5496/wjmg.v4.i2.27

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[1]. 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 encodes specific dedicated recombinase for their movement[2]. Among several IMGEs, phages are the best-characterized genetic elements, which play major role in real-time evolution of toxigenic bacterial pathogens[3]. Most pathogenic bacterial cells acquired their virulence traits either from phages[3] or other IMGEs like, plasmids[4], transposons[5] or GIs[6]. 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 (CTXΦ) and evolution from non-pathogenic strains to toxigenic cholera pathogen[3].

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[7]. 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 CTXΦ either in large or in small or in both chromosomes. CTXΦ 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*[8].

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 CTXΦ in the *V. cholerae* chromosomes. The main emphasis is on the mechanistic part of CTXΦ integration, how (+)ssDNA genome of CTXΦ is recognized by the XerC and XerD enzymes and proceed for unusual irreversible integration. We also discussed about other IMGEs that follow CTXΦ like integration. Finally, we discussed how cooperative interactions between CTXΦ and other IMGEs could lead to dissemination of CTXΦ among closely or distantly related bacterial cells.

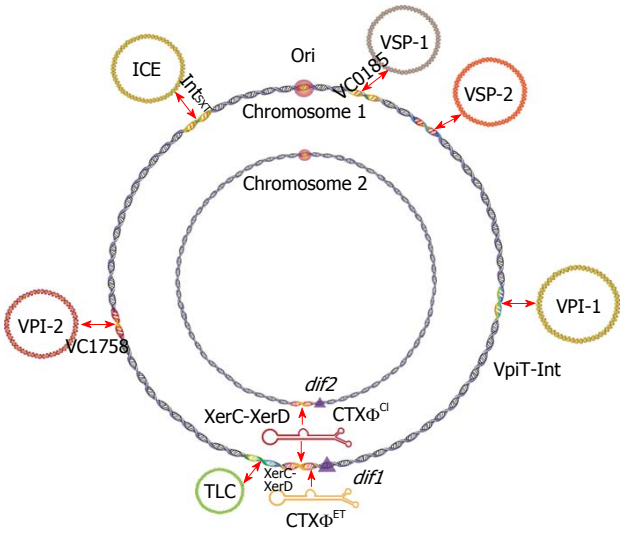

Figure 1 Schematic representation of the integrative mobile genetic elements present in the chromosome 1 (large) or 2 (small) of *Vibrio cholerae*. Except CTXΦ, all other IMGEs have unique *attB* site either in large or small chromosome. IMGEs present in the *dif* region exploit host-encoded recombinases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

CTXΦ AND OTHER IMGEs IN *V. CHOLERAE* 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[9]. 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[3].

As mentioned, CTXΦ and several other phages always integrate in the *dif* site present one in each of the two circular chromosomes of *V. cholerae*[10,11]. 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[12]. Strand exchange occurs immediately after binding sites of XerC and XerD. CTXΦ, RS1 and TLC element integrates as a single copy or in multiple tandemly arrayed copies in *dif1*[10,13]. Other IMGEs were found integrated at the *dif* sites as a single copy[9]. 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 7th 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[14] 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[15]. 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*[16,17]. VPI-1 is a 41-kb DNA segment physically linked to a tmRNA gene (*ssrA*), flanked by two nearly identical repeat sequences[18]. It is found preferentially in the toxigenic strains. VPI-1 carries two putative tyrosine recombinases, called Int_{vpi} and VpiT[15]. Sequence analyses of these two putative recombinases indicate that they are quite different from each other. While Int_{vpi} 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[15].

Presence or absence of VSPs could serve as distinct genetic signatures to differentiate the previous (6th) and the current (7th) 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*[18] 2000). A XerCD like putative tyrosine recombinase (Int_{VSP-1}) 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[19].

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[20]. 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*[21]. Wozniak *et al*[22] 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[23,24]. 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[8]. 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[25]. RstR acts as a transcriptional modulator and repressed transcription of *rstA* and *rstB* from P_{rstA}, the only phage promoter required for CTXΦ replication and integration[26]. Core region comprises of seven genes responsible for phage morphogenesis and toxin production. Five genes, namely, *psh*, *cep*, $gIII^{CTX}$, *ace* and *zot* encode proteins essential for phage morphogenesis and phage assembly. The Psh, Cep, OrfU ($pIII^{CTX}$) and Ace proteins are phage structural proteins, which encapsulate single stranded phage genome whereas Zot protein play central role in phage assembly[3]. 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[27]. 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*[28]. 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

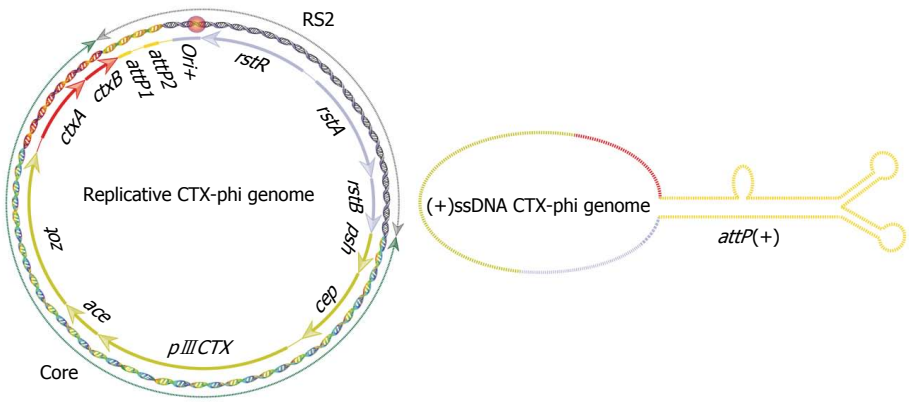

Figure 2 Replicative and integrative genomes of cholera toxin phage. Replicative genome of CTXΦ arranged in structurally and functionally distinct two modular structures RS2 and core. RS2 encoded proteins are essential for phage replication, integration and transcriptional regulation of phage genes. 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. CTXΦ: 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[29]. Strand stabilization between exchanged strands determines the tropism of phage integration. Although both Xer proteins are essential for integration reaction, CTXΦ integration needs only the XerC catalytic activity for one pair of strand exchange and final integration[28]. As a consequence, *dif* compatibility of CTXΦ 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 CTXΦ 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 CTXΦ excision is not detected under standard laboratory condition. Recently, CTXΦ excision from toxigenic *V. cholerae* isolates has been reported[30]. 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 CTXΦ and provides functional XerC-XerD binding sites, hence clearance of CTXΦ.

CTXΦ VARIANTS: CONVENTIONAL *VS* CURRENT GROUPING

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

On the other side, CTXΦ can be classified into three broad classes based on their *dif* compatibility (Table 1). CTXΦ isolated from 6th 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*[29]. In contrast, most well characterized CTXΦ from 7th pandemic El Tor cholera isolates could form such interaction only with the overlap region of *dif1* but not with *dif2*. Thus, the integration CTXΦ is specific for *dif1*. Recent *V. cholerae* isolates from Africa and India harbour CTXΦ in either chromosomes indicating their *attP* is similar to *attP* of CTXΦ isolated from 6th pandemic *V. cholerae*. In contrast, some environmental isolates carry *dif* sequenc-

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

CTXΦ variants	Isolated from	Chromosomal integration site	Ref.
CTXΦ^{d1}	O1 El Tor	*difA*, *dif1*	Das *et al*[29] Das *et al*[10]
CTXΦ$^{d1/2}$	O1 classical	*difA*, *dif1* and *dif2*	Das *et al*[29] Das *et al*[10]
CTXΦ^{dG}	Environmental	*difG*	Das *et al*[29] Das *et al*[10]

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

es[33], which support integration of only specific type of phage, CTXΦ^{dG}. 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 7th pandemic El Tor cholera isolates while the lowest integration efficiency is detected in between *difG* and CTXΦ^{dG} isolated from environmental *V. cholerae*[29]. 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. CHOLERAE 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*[10]. *difA* and *dif1* are identical except two bases in the XerC binding site[10,13]. 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[10]. 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[9]. Several *V. cholerae* cells have different arrangements of IMGE at *dif1* site, *e.g.*, single or multiple IMGE(s) in multiple combinations[9]. 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[8,28,29,34]. 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[13]. Further work in this

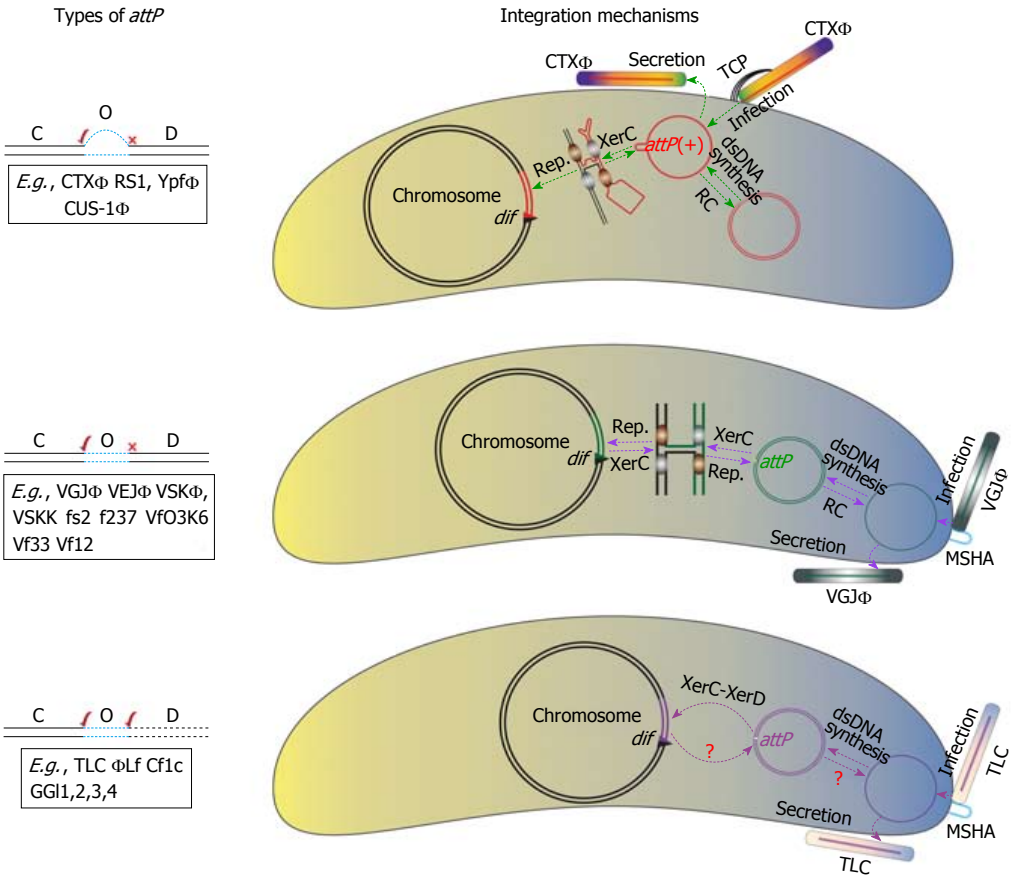

Figure 3 Schematic representation of the integration mechanisms of three classes of integrative mobile genetic elements present exclusively in the dimer resolution sites of bacterial genome. C-O-D denotes XerC-Overlap region-XerD. The XerC-XerD recombinases, attachment site of phage and bacterial genome 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. CTXΦ integration is irreversible. By contrast, VGJΦ and TLC integration is reversible. Key steps in the life cycle of CTXΦ, VGJΦ and TLC are also indicated. Host-encoded TCP and MSHA served as phage receptors for CTXΦ, VGJΦ 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; CTXΦ: Cholera toxin phage; TCP: Toxin co-regulated pilus; MSHA: Mannose sensitive hemagglutinin A.

direction is warranted to understand the intricate biology of TLC and CTXΦ.

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 TLCφ 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, Fleishmann RD, Nierman 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

World Journal of Medical Genetics

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.5496/wjmg.v4.i2.34

World J Med Genet 2014 May 27; 4(2): 34-38
ISSN 2220-3184 (online)

MINIREVIEWS

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
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

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

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

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[1]. All mutations either deleted the whole gene or resulted in the premature termination of translation[1].

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[1-3]. The WTX protein has been demonstrated to regulate the stability of β-catenin[2], a key effector of the WNT/β-catenin signaling pathway (reviewed in Clevers and Nusse 2012[4]). The critical importance of this pathway during embryogenesis is clearly demonstrated by the pleiotropic clinical presentation of OSCS 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[1,5]. 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[1,6-8].

Females affected with OSCS present a great variability of the phenotype and, while previous studies[9-11] suggested, at least in some cases, a nonrandom X-inactivation[9-11] that could explain this phenomenon[6], Jenkins *et al*[1] demonstrated that in 19 *WTX* mutation-bearing heterozygous females X-inactivation ratios were not skewed.

Among the features that constitute the OSCS 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[6,7,9,12].

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%)[12]. Coronal craniostenosis has been described in one patient only[9].

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[13-22]. Dental anomalies have been reported in 30% of patients[20]. 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[6,7,12,20,23]. 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[23]. 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[20,24]. 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[25] but it is also hypothesized that disruption of the nerve supporting vessels may lead to secondary cranial nerve deficiencies[20]. 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[26].

As already mentioned, OSCS 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[12,20]. 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[9,27,28]. Anal stenosis has been described in two girls only[29,30].

OSCS in males is more severe than in females because it follows an X-linked dominant pattern of inheritance that determines hemizygosity 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

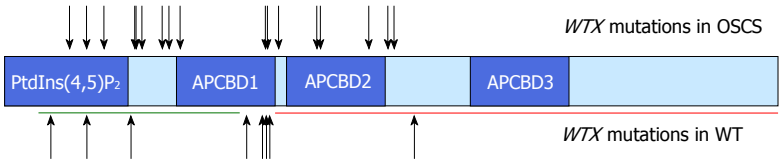

Figure 1 *WTX* mutations in osteopathia striata with cranial sclerosis and Wilms' tumor. The full length WTX protein possesses two phosphatidylinositol(4,5)-bisphosphate [PtdIns(4,5)P2] binding domains, three adenomatous polyposis coli binding domains (APCBD) and a β-catenin binding region (red line)[2,3]. The smaller WTX isoform lacks amino acids 50-326 (green line)[1,5]. 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[1,6-8,12,24,38,39,44,46,47].

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%)[8,31]. 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[9,10,21,27,32].

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[8], and by the absence of some anomalies such as fibular aplasia, duplicated phalanges, syndactyly[1,8], and gastrointestinal[9,10,21,27,32], cardiac[9,21,27,31,33,34] and genitourinary malformations[1,8,21,35,36].

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[8]. Milder bony sclerosis has been detected in males with mosaic mutation of *WTX*[37]. Striations of the long bones have also been observed in molecularly confirmed or suspected mosaics for *WTX* mutations[8,24,26,34,37,38].

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

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[39]. 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[40]. Approximately 40% of WTs occur in association with nephrogenic rests (NRs), embryonal remnants in the kidney which are known precursor lesions for WT[41]. 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[42] 2011, Royer-Pokora[43] 2013). Further genes involved in WT development include, in addition to *WTX*, *CTNNB1* and *TP53*[42,43]. *WTX* anomalies have been described in approximately 20% of WTs[39,44-47]. While WTX deletions and truncating mutations are somatically acquired, missense mutations of unknown functional relevance can be present in the germline (reviewed in Huff[42] 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[42] 2011). However, the spectrum of *WTX* truncating mutations in OSCS patients and in WTs is very similar[1,42] (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*[48] 2006). In contrast, individuals with OSCS do not seem to have any predisposition to develop either WT or other malignancies[1,22,35], although in a few of these patients the presence of bilateral multifocal NRs has been reported[8,35]. 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[40]. 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[49]. 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[46].

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

1 **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]

2 **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]

3 **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]

4 **Clevers H**, Nusse R. Wnt/β-catenin signaling and disease. *Cell* 2012; **149**: 1192-1205 [PMID: 22682243 DOI: 10.1016/j.cell.2012.05.012]

5 **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]

6 **Perdu B**, de Freitas F, Frints SG, Schouten M, Schrander-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]

7 **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]

8 **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]

9 **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]

10 **Rott HD**, Krieg P, Rütschle H, Kraus C. Multiple malformations in a male and maternal osteopathia strata with cranial sclerosis (OSCS). *Genet Couns* 2003; **14**: 281-288 [PMID: 14577672]

11 **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]

12 **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]

13 **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]

14 **Franklyn PP**, Wilkinson D. Two cases of osteopathia striata, deafness and cranial osteopetrosis. *Ann Radiol* (Paris) 1978; **21**: 91-93 [PMID: 677756]

15 **Cortina H**, Vallcanera A, Vidal J. Familial osteopathia striata with cranial condensation. *Pediatr Radiol* 1981; **11**: 87-90 [PMID: 7301452 DOI: 10.1007/BF00971786]

16 **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]

17 **Robinow M**, Unger F. Syndrome of osteopathia striata, macrocephaly, and cranial sclerosis. *Am J Dis Child* 1984; **138**: 821-823 [PMID: 6475869]

18 **Piechowiak H**, Goebel FD, Hirche U, Tyrell R. Cranial sclerosis with striated bone disease (osteopathia striata). *Klin Padiatr* 1986; **198**: 418-424 [PMID: 3784439]

19 **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]

20 **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]

21 **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]

22 **Behninger C**, Rott HD. Osteopathia striata with cranial sclerosis: literature reappraisal argues for X-linked inheritance. *Genet Couns* 2000; **11**: 157-167 [PMID: 10893667]

23 **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]

24 **Ciceri S**, Cattaneo E, Fossati C, Radice P, Selicorni A, Perotti D. First evidence of vertical paternal transmission of osteopatia striata with cranial sclerosis. *Am J Med Genet A* 2013; **161A**: 1173-1176 [PMID: 23494899 DOI: 10.1002/ajmg.a.35813]

25 **De Keyser J**, Bruyland M, De Greve J, Leemans J, Potvliege 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: 10.1016/0303-8467(83)90022-7]

26 **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]

27 **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]

28 **Koudstaal MJ**, Wolvius EB, Ongkosuwito 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 **Currarino 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

World Journal of Medical Genetics

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.5496/wjmg.v4.i2.39

World J Med Genet 2014 May 27; 4(2): 39-45
ISSN 2220-3184 (online)

RESEARCH REPORT

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, 4th Hospital of Hebei Medical University, Shijiazhuang 050011, Hebei Province, China
Deng-Gui Wen, Bao-En Shan, Shi-Jie Wang, Cancer Center, 4th 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, 4th Hospital of Hebei Medical University, Jiankanglu 12, Shijiazhuang 050011, Hebei Province, China. 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 4th 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 ± 8.39 *vs* 53.49 ± 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 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. 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.

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 ± 8.39 *vs* 53.49 ± 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[1]. 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[2,3]. 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[4-7].

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[8-16]. 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 4th Hospital of Hebei Medical University.

Subject selection

The high risk region and the Upper Gastrointestinal Cancer Center were previously described[4,5,7,9]. 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 4th Affiliated Hospital of Hebei Medical University (also the Hebei Province Cancer Center) from January 1 1975 to December 31 1989. Because the 4th 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[17-19] 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[20,21]. 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[17-19], 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 χ^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	*P t*-test
	n (%)	Mean age (yr) ± SD (1)	*n* (%)	Mean age (yr) ± SD (2)	(1)-(2)	(1)-(2)
Total	766	51.83 ± 8.39	1776	53.49 ± 8.23	-1.66	0.000[b]
Sex						
Female	229 (33.80)	52.46 ± 7.91	600 (29.90)	53.48 ± 7.81	-1.02	0.100
Male	537 (66.20)	51.56 ± 8.58	1176 (70.10)	53.49 ± 8.44	-1.93	0.000[b]
Tobacco						
Nonsmoker	282 (36.80)	52.66 ± 8.03	739 (41.60)	53.63 ± 8.27	-0.98	0.090
Smoker	484 (63.20)	51.35 ± 8.56	1037 (58.40)	53.38 ± 8.13	-2.03	0.000[b]
Alcohol						
Nondrinker	480 (62.70)	52.10 ± 8.37	1200 (67.60)	53.38 ± 8.21	-1.29	0.004[a]
Drinker	286 (37.30)	51.38 ± 8.42	578 (32.40)	53.70 ± 8.28	-2.32	0.000[b]
Surgery year						
1975-1979	224 (29.20)	50.90 ± 8.21	376 (21.20)	52.36 ± 8.00	-1.46	0.034[a]
1980-1984	216 (28.20)	51.69 ± 8.58	545 (30.70)	52.95 ± 8.18	-1.26	0.050[a]
1985-1989	326 (42.60)	52.56 ± 8.34	855 (48.10)	54.32 ± 8.28	-1.76	0.001[b]
TNM						
$T_{is,1}N_0M_0$	34 (4.40)	50.94 ± 9.78	66 (3.70)	55.52 ± 7.24	-4.57	0.010[a]
$T_{2,3}N_0M_0$	422 (55.10)	51.64 ± 8.30	983 (55.30)	53.30 ± 8.37	-1.67	0.001[b]
$T_{2,3,4}N_1M_0$	310 (40.50)	52.19 ± 8.37	727 (40.90)	53.55 ± 8.10	-1.36	0.010[a]
Resection						
Exploratory	16 (2.10)	49.88 ± 8.16	28 (1.60)	52.57 ± 9.18	-2.70	0.340
R1 or R2	70 (9.10)	49.33 ± 10.00	152 (8.60)	54.73 ± 8.49	-5.40	0.000[b]
R0	680 (88.80)	52.13 ± 8.17	1596 (89.90)	53.38 ± 8.18	-1.25	0.001[b]

Positive family history *vs* negative family history, [a]$P < 0.05$, [b]$P < 0.01$.

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[22].

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 subgroup analyses was significant for males, smokers, drinkers, non-drinkers, surgery during 1975-1979, 1980-1984, 1985-1989, tumor stage of $T_{is,1}N_0M_0$, $T_{2,3}N_0M_0$, $T_{2,3,4}N_1M_0$[23], complete resection (R_0) and partial resection (R_1/R_2) subgroups; 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 (χ^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, $T_{2,3}N_0M_0$, and $T_{2,3,4}N_1M_0$ groups; but not significant for the older onset age (over 55 years old), smoking, drinking and $T_{is,1}N_0M_0$ 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	χ^2	^aP
Overall		72 (2.8)	2542		
	Negative family history	30 (1.7)	1776		
	Positive family history	42 (5.5)	766	27.80	0.000[b]
Sex					
Male	Negative family history	20 (1.7)	1176		
	Positive family history	27 (5.0)	537	15.30	0.000[b]
Female	Negative family history	10 (1.7)	600		
	Positive family history	15 (6.6)	229	13.52	0.000[b]
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[b]
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[b]
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[b]
TNM					
$T_{is,1}N_0M_0$	Negative family history	2 (3.0)	66		
	Positive family history	2 (5.9)	34	0.48	0.480
$T_{2,3}N_0M_0$	Negative family history	8 (0.8)	983		
	Positive family history	20 (4.7)	422	23.29	0.000[b]
$T_{2,3,4}N_1M_0$	Negative family history	20 (2.8)	727		
	Positive family history	20 (6.5)	310	8.03	0.010[b]

Positive family *vs* negative family, $^aP < 0.05$, $^bP < 0.01$.

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 significantly 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[10]. 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[b]	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[b]	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[b]	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[b]	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[b]	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[b]	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[b]	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[b]	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[b]	0.974	0.963	0.985
Tobacco	0.240	0.096	6.245	0.012[a]	1.271	1.053	1.535
Multiple cancer	1.451	0.265	29.862	0.000[b]	4.265	2.535	7.176
Constant	0.108	0.337	0.104	0.747	1.114		

[a]$P < 0.05$, [b]$P < 0.01$ *vs* standard error.

tiple 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[8]. Instead, GCA and ESCC have identical epidemiological distributions[17-21]. Molecularly, gene polymorphisms associated with elevated risks for both cancers were found in Hebei Province[13,14], and identical DNA alterations in the two cancers were reported in Henan Province[15,16]. 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[17,18].

During surgery, the two cancers are often found to coincide in one patient: among the 2524 ESCC cases undergoing curative-intent surgery in the 4th 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[20,21], 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[23-25].

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[26], 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 ⩾ 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 clinicopathological 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, Tamaro S, Chia KS, Pompe-Kirn V, Kliewer 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 **Bignold 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

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
www.wjgnet.com

World J Med Genet 2014 May 27; 4(2): I-V
ISSN 2220-3184 (online)

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, mixedtreatment 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

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: editorialoffice@wjgnet.com
Help Desk: http://www.wjgnet.com/esps/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/esps/helpdesk.aspx
http://www.wjgnet.com

Instructions to authors
Full instructions are available online at http://www.wjgnet.com/2220-3184/g_info_20100722180909.htm.

Indexed and Abstracted in
Digital Object Identifier.

SPECIAL STATEMENT

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

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, Ridit, 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.

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 copyedit 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) 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, 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, University of California, Box 0538, San Francisco, CA 94143, United States. 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 ± 3.86 *vs* 3.61 ± 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. $^{a}P < 0.05$, $^{b}P < 0.01$ should be noted ($P > 0.05$ should not be noted). If there are other series of *P* values, $^{c}P < 0.05$ and $^{d}P < 0.01$ are used. A third series of *P* values can be expressed as $^{e}P < 0.05$ and $^{f}P < 0.01$. Other notes in tables or under illustrations should be expressed as ^{1}F, ^{2}F, ^{3}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[1,2]". If references are cited directly in the text, they should be put together within the text, for example, "From references[19,22-24], 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

Pleased 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-diarrhoea. *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 PMCID: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. *HRSA Careaction* 2002; 1-6 [PMID: 12154804]

Books

Personal author(s)

10 **Sherlock S**, Dooley J. Diseases of the liver and billiary 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**, Schorfheide AM. Adolescent pregnancy. 2nd ed. Wieczorek 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 ± SD or mean ± SE.

Statistical expression

Express *t* test as *t* (in italics), *F* test as *F* (in italics), chi square test as χ^2 (in Greek), related coefficient as *r* (in italics), degree of freedom as υ (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 pres-

sure, p (B) = 16.2/12.3 kPa; incubation time, t (incubation) = 96 h, blood glucose concentration, c (glucose) 6.4 ± 2.1 mmol/L; blood CEA mass concentration, p (CEA) = 8.6 24.5 μg/L; CO_2 volume fraction, 50 mL/L CO_2, not 5% CO_2; 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.

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: *Eco*RI, *Hin*dI, *Bam*HI, *Kbo* 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.

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.

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.

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: 698 USD per article. All invited articles are published free of charge.

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
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
http://www.wjgnet.com