ISSN 2220-3176 (online)

World Journal of *Clinical Infectious Diseases*

World J Clin Infect Dis 2016 May 25; 6(2): 6-27

Published by Baishideng Publishing Group Inc

WJCID

World Journal of Clinical Infectious Diseases

A peer-reviewed, online, open-access journal of clinical infectious diseases

Editorial Board

2016-2019

The *World Journal of Clinical Infectious Diseases* Editorial Board consists of 284 members, representing a team of worldwide experts in infectious diseases. They are from 55 countries, including Argentina (5), Australia (8), Austria (1), Belgium (2), Bosnia and Herzegovina (1), Brazil (6), Brunei Darussalam (1), Bulgaria (1), Cameroon (1), Canada (7), China (17), Colombia (1), Costa Rica (1), Cuba (1), Denmark (2), Egypt (2), Ethiopia (1), Finland (1), France (11), Germany (3), Greece (8), Hungary (5), India (14), Iran (6), Israel (10), Italy (20), Japan (3), Jordan (1), Kosovo (1), Kuwait (1), Lebanon (3), Lithuania (1), Malawi (1), Mexico (5), Morocco (2), Netherlands (3), Nigeria (1), Pakistan (1), Peru (1), Portugal (5), Russia (1), Saudi Arabia (2), Singapore (3), South Africa (2), South Korea (5), Spain (24), Switzerland (2), Tanzania (1), Thailand (4), Tunisia (1), Turkey (5), United Arab Emirates (1), United Kingdom (10), United States (57), and Venezuela (1).

Costa Rica

Adriano Gerardo Arguedas Mohs, *San José*

Cuba

Maria G Guzman, *Havana*

Denmark

Janne K Klitgaard, *Odense*
Henrik Torkil Westh, *Hvidovre*

Egypt

Tarek M Diab, *Giza*
Olfat G Shaker, *Cairo*

Ethiopia

Solomon A Yimer, *Oslo*

Finland

Jari TJ Nuutila, *Turku*

France

Hassane Adakal, *Bobo-Dioulasso*
Pascal Bigey, *Paris*
Philippe Brouqui, *Marseille*
Christophe Chevillard, *Marseille*
Raphaele Girard, *Pierre* Bénite
Vincent Jarlier, *Paris*
Sandrine Marquet, *Marseille*
Thierry Naas, *Bicetre*
Saad Nseir, *Lille*
Philippe Seguin, *Rennes*
Muriel Vayssier-Taussat, *Maisons-Alfort*

Germany

Stefan Borgmann, *Ingolstadt*
Georg Harter, *Ulm*
Matthias Imohl, *Aachen*

Greece

Apostolos Beloukas, *Athens*
Alex P Betrosian, *Athens*
George L Daikos, *Athens*
Helena C Maltezou, *Athens*
Argyris S Michalopoulos, *Athens*
Maria A Moschovi, *Goudi*
George Petrikkos, *Athens*
Athanassios Tragiannidis, *Thessaloniki*

Hungary

László Galgoczy, *Szeged*
Ferenc Orosz, *Budapest*
Ferenc Rozgonyi, *Budapest*
Jozsef Soki, *Szeged*
Dezso P Virok, *Szeged*

India

Ritesh Agarwal, *Chandigarh*
Syed I Alam, *Gwalior*
Atmaram H Bandivdekar, *Mumbai*
Runu Chakravarty, *Kolkata*
Dipshikha Chakravortty, *Bangalore*
Mamta Chawla-Sarkar, *Kolkata*
Sanjay Chhibber, *Chandigarh*
Belgode N Harish, *Pondicherry*
Triveni Krishnan, *Kolkata*
Rashmi Kumar, *Lucknow*
Mohammad Owais, *Aligarh*
Banwarilal Sarkar, *Kolkata*
Akashdeep Singh, *Ludhiana*

Iran

Parissa Farnia, *Tehran*
Seyed M Jazayeri, *Tehran*
Morteza Pourahmad, *Jahrom*
Mohammad R Pourshafie, *Tehran*
Mohammad H Salari, *Tehran*
Hasan Shojaei, *Isfahan*

Israel

Jacob Amir, *Petach Tikvah*
Shai Ashkenazi, *Petach Tikva*
Gadi Borkow, *Gibton*
Raul Colodner, *Afula*
Jacob M Gilad, *Tel-Aviv*
Noah Isakov, *Beer Sheva*
Michal Mandelboim, *Tel-Hashomer*
Shifra Shvarts, *Omer*
Oshri Wasserzug, *Tel-Aviv*
Pablo V Yagupsky, *Beer-Sheva*

Italy

Giuseppe Barbaro, *Rome*
Paolo Bonilauri, *Reggio Emilia*
Guido Calleri, *Torino*
Mario Cruciani, *Verona*
Antonella d'Arminio Monforte, *Milan*
Silvano Esposito, *Salerno*
Marco Falcone, *Rome*
Antonio Fasanella, *Foggia*
Daniele Focosi, *Pisa*
Delia Goletti, *Roma*
Guido Grandi, *Siena*
Fabio Grizzi, *Rozzano*
Giuseppe Ippolito, *Rome*
Roberto Manfredi, *Bologna*
Claudio M Mastroianni, *Rome*
Ivano Mezzaroma, *Rome*
Giuseppe Micali, *Catania*
Annamaria Passantino, *Messina*
Mariagrazia Perilli, *L'Aquila*
Patrizia Pontisso, *Padova*

Japan

Emoto Masashi, *Gunma*
Toshi Nagata, *Hamamatsu*
Ryohei Yamasaki, *Tottori*

Jordan

Asem A Shehabi, *Amman*

Kosovo

Lul Raka, *Prishtina*

Kuwait

Willias Masocha, *Safat*

Lebanon

Ziad Daoud, *Tripoli*
Ghassan M Matar, *Beirut*
Sami Ramia, *Beirut*

Lithuania

Gazim Bizanov, *Vilnius*

Malawi

Adamson S Muula, *Zomba*

Mexico

Agnes Fleury, *Mexico City*
Guadalupe García-Elorriaga, *Mexico City*
Alejandro Macias, *Mexico City*
Mussaret Zaidi, *Merida*
Roberto Zenteno-Cuevas, *Veracruz*

Morocco

Redouane Abouqal, *Rabat*
Sayeh Ezzikouri, *Casablanca*

Netherlands

John Hays, *Rotterdam*

Nisar A Khan, *Rotterdam*
Rogier Louwen, *Rotterdam*

Nigeria

Samuel S Taiwo, *Osogbo*

Pakistan

Muhammad Idrees, *Lahore*

Peru

Salim Mohanna, *Lima*

Portugal

Ricardo Araujo, *Porto*
Manuela Canica, *Lisboa*
Francisco Esteves, *Lisboa*
Fernando Rodrigues, *Braga*
Nuno Taveira, *Lisbon*

Russia

Alexander M Shestopalov, *Novosibirsk*

Saudi Arabia

Jaffar A Al-Tawfiq, *Dhahran*
Atef M Shibl, *Riyadh*

Singapore

Yee S Leo, *Singapore*
Laurent CS Renia, *Singapore*
Richard J Sugrue, *Singapore*

South Africa

Carolina H Pohl, *Bloemfontein*
Natasha Potgieter, *Louis Trichardt*

South Korea

Yong-Hyun Cho, *Seoul*
Sang-Ho Choi, *Seoul*
Ju-Young Chung, *Seoul*
Jung Mogg Kim, *Seoul*
Kyongmin Kim, *Suwon*

Spain

Alberto Arnedo-Pena, *Castellon*
Alfredo Berzal-Herranz, *Granada*
Vicente Boix, *Alicante*
Enrique Calderon, *Seville*
Rafael Canton, *Madrid*
Jose M Cuevas, *Valencia*
Laila Darwich, *Barcelona*
Pere Domingo, *Barcelona*
Tahia D Fernandez, *Malaga*
Lucia Gallego, *Leioa*
Adela González de la Campa, *Madrid*
Luis I Gonzalez-Granado, *Madrid*
Bruno Gonzalez-Zorn, *Madrid*
Eduardo Lopez-Collazo, *Madrid*
Miguel Marcos, *Salamanca*
Antonio T Martí, *Barcelona*
Andrés Moya, *València*
Rafael Najera, *Madrid*
Maria-Mercedes Nogueras-Mas, *Sabadell*
Jose A Oteo, *La Rioja*
Pilar Perez-Romero, *Sevilla*
Ruth G Prieto, *Alcorcon*
Eduardo Reyes, *Alcala de Henares*
Francisco Soriano, *Madrid*

Switzerland

Stephen Hawser, *Epalinges*
Andrew Hemphill, *Bern*

Tanzania

John PA Lusingu, *Tanga*

Thailand

Kosum Chansiri, *Bangkok*
Subsai Kongsaengdao, *Bangkok*
Niwat Maneekarn, *Chiang Mai*
Viroj Wiwanitkit, *Bangkok*

Tunisia

Aouni Mahjoub, *Monastir*

Turkey

Oguz Karabay, *Sakarya*
Uner Kayabas, *Malatya*
Gokhan Metan, *Kayseri*
Oral Oncul, *Uskudar*
Mesut Yilmaz, *Istanbul*

United Arab Emirates

Muhammad Mukhtar, *Ras Al Khaimah*

United Kingdom

Zainab Al-Doori, *Glasgow*
David Carmena, *London*
Ronald A Dixon, *Lincoln*
Vanya A Gant, *London*
Robin Goodwin, *Coventry*
Andrew C Hayward, *London*
Laura A Hughes, *Cheshire*
Michele E Murdoch, *Herts*
Harunor Rashid, *London*
Craig W Roberts, *Glasgow*

United States

Majdi N Al-Hasan, *Lexington*
Ibne KM Ali, *Charlottesville*
Hossam M Ashour, *Detroit*
Joseph U Becker, *Palo Alto*
M Eric Benbow, *Dayton*
Eliahu Bishburg, *Newark*
Luz P Blanco, *Ann Arbor*
Robert Bucki, *Philadelphia*
Steven D Burdette, *Dayton*
Archana Chatterjee, *Omaha*
Pai-Lien Chen, *Durham*
Pawel S Ciborowski, *Omaha*
Michael Cynamon, *Syracuse*
Siddhartha Das, *El Paso*
Ralph J DiClemente, *Atlanta*
Noton K Dutta, *Baltimore*
Garth D Ehrlich, *Pittsburgh*
Michael S Firstenberg, *Akron*
Walter A Hall, *Syracuse*
Yongqun He, *Ann Arbor*
Brenda L Helms, *Plano*
Joseph U Igietseme, *Atlanta*
Mohammad K Ijaz, *Montvale*
Suresh G Joshi, *Philadelphia*
Christian Joukhadar, *Boston*
Thomas F Kresina, *Rockville*
Alain B Labrique, *Baltimore*
Shenghan Lai, *Baltimore*
Benfang Lei, *Bozeman*
Jeff G Leid, *Flagstaff*
Vladimir Leontiev, *St.Louis*
James M McMahon, *Rochester*
Geraldine M McQuillan, *Hyattsville*
Lawrence F Muscarella, *Ivyland*
Daniel Musher, *Houston*
Stella Nowicki, *Nashville*
M Jacques Nsuami, *New Orleans*
Phillipe N Nyambi, *New York*
Raymund R Razonable, *Rochester*
Anand Reddi, *Denver*
William R Schwan, *La Crosse*
Richard A Slayden, *Fort Collins*
Theodore J Standiford, *Ann Arbor*
William M Switzer, *Atlanta*
Ashutosh Tamhane, *Birmingham*
Giorgio E Tarchini, *Weston*
Carmen Taype, *New York*
Barbara Van Der Pol, *Bloomington*
Jose A Vazquez, *Detroit*
Fernando Villalta, *Nashville*
Haider J Warraich, *Boston*
Xianfu Wu, *Atlanta*
X Frank Yang, *Indianapolis*
Genyan Yang, *Atlanta*
Hong Zhang, *Rockville*
Lyna Zhang, *Atlanta*

Venezuela

Alfonso J Rodriguez-Morales, *Caracas*

WJCID World Journal of Clinical Infectious Diseases

Contents

Quarterly Volume 6 Number 2 May 25, 2016

REVIEW

MINIREVIEWS

World Journal of Clinical Infectious Diseases
Volume 6 Number 2 May 25, 2016

Contents

NAME OF JOURNAL
World Journal of Clinical Infectious Diseases

ISSN
ISSN 2220-3176 (online)

LAUNCH DATE
December 30, 2011

FREQUENCY
Quarterly

EDITORIAL OFFICE
Jin-Lei Wang, Director
Xiu-Xia Song, Vice Director
World Journal of Clinical Infectious Diseases
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

PUBLICATION DATE
May 25, 2016

INSTRUCTIONS TO AUTHORS
Full instructions are available online at http://www.wjgnet.com/bpg/g_info_20160116143427.htm

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

World Journal of Clinical Infectious Diseases

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.5495/wjcid.v6.i2.6

World J Clin Infect Dis 2016 May 25; 6(2): 6-21
ISSN 2220-3176 (online)

REVIEW

Systems biology applications to study mechanisms of human immunodeficiency virus latency and reactivation

Cory H White, Bastiaan Moesker, Angela Ciuffi, Nadejda Beliakova-Bethell

Cory H White, Graduate Program in Bioinformatics and Systems Biology, University of California San Diego, La Jolla, CA 92093, United States

Cory H White, San Diego VA Medical Center and Veterans Medical Research Foundation, San Diego, CA 92161, United States

Bastiaan Moesker, Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton, Southampton, Hants SO16 6YD, United Kingdom

Angela Ciuffi, Institute of Microbiology, University Hospital of Lausanne (CHUV) and University of Lausanne, 1011 Lausanne, Switzerland

Nadejda Beliakova-Bethell, Department of Medicine, University of California San Diego, La Jolla, CA 92093, United States

Author contributions: All authors contributed equally to this paper with conception and design of the study, literature review and interpretation, manuscript preparation and approval of the final version.

Supported by The grant from the National Institutes of Health, Martin Delaney Collaboratory of AIDS Researchers for Eradication (CARE, U19 AI 096113); the Swiss National Science Foundation (grant 31003A_146579); and the University of California, San Diego Fellowships for Graduate Researchers, Frontiers of Innovation Scholars Program.

Conflict-of-interest statement: No potential conflict of interest.

Correspondence to: Nadejda Beliakova-Bethell, PhD, Department of Medicine, University of California San Diego, Stein Clinical Research Building, Rm. 303, 9500 Gilman Drive, #0679, La Jolla, CA 92093, United States. nbeliako@ucsd.edu
Telephone: +1-858-5528585
Fax: +1-858-5527445

Received: September 30, 2015
Peer-review started: October 7, 2015
First decision: November 30, 2015
Revised: January 15, 2016
Accepted: March 7, 2016
Article in press: March 9, 2016
Published online: May 25, 2016

Abstract

Eradication of human immunodeficiency virus (HIV) in infected individuals is currently not possible because of the presence of the persistent cellular reservoir of latent infection. The identification of HIV latency biomarkers and a better understanding of the molecular mechanisms contributing to regulation of HIV expression might provide essential tools to eliminate these latently infected cells. This review aims at summarizing gene expression profiling and systems biology applications to studies of HIV latency and eradication. Studies comparing gene expression in latently infected and uninfected cells identify candidate latency biomarkers and novel mechanisms of latency control. Studies that profiled gene expression changes induced by existing latency reversing agents (LRAs) highlight uniting themes driving HIV reactivation and novel mechanisms that contribute to regulation of HIV expression by different LRAs. Among the reviewed gene expression studies, the common approaches included identification of differentially expressed genes and gene functional category assessment. Integration of transcriptomic data with other biological data types is presently scarce, and the field would benefit from increased adoption of these methods in future studies. In addition, designing prospective studies that use the same methods of data acquisition and statistical analyses will facilitate a more reliable

identification of latency biomarkers using different model systems and the comparison of the effects of different LRAs on host factors with a role in HIV reactivation. The results from such studies would have the potential to significantly impact the process by which candidate drugs are selected and combined for future evaluations and advancement to clinical trials.

Key words: Gene expression; Microarrays; RNA-Seq; Systems biology; Human immunodeficiency virus; Viral latency; Disease eradication; Biomarkers; Molecular mechanisms; Latency reversing agents

Core tip: Gene expression profiling and systems biology methods are reviewed with respect to their possible application in the field of human immunodeficiency virus (HIV) research. Studies profiling gene expression in latently infected and uninfected cells are summarized to illustrate application of these methods to identification of latency biomarkers and the molecular mechanisms contributing to regulation of HIV expression. Studies that measure changes in host and HIV gene expression upon treatment with latency reversing agents (LRAs) highlight uniting themes driving HIV reactivation and identify novel mechanisms of action of LRAs. The field will further benefit from increased adoption of systems biology methods in future studies.

White CH, Moesker B, Ciuffi A, Beliakova-Bethell N. Systems biology applications to study mechanisms of human immunodeficiency virus latency and reactivation. *World J Clin Infect Dis* 2016; 6(2): 6-21 Available from: URL: http://www.wjgnet.com/2220-3176/full/v6/i2/6.htm DOI: http://dx.doi.org/10.5495/wjcid.v6.i2.6

INTRODUCTION

In the present era of combination anti-retroviral therapy (cART), the persistence of cellular human immunodeficiency virus (HIV) reservoir is considered to be the major barrier to a cure[1]. This cellular reservoir mainly consists of latently infected resting CD4+ T cells bearing HIV integrated provirus. It is highly stable[2-5] and inducible, necessitating life-long adherence to cART to prevent rebound of viremia. In a search for therapeutic strategies to eradicate this latent reservoir, mechanisms leading to latency have been extensively studied and include transcriptional and post-transcriptional blocks[1,6-14].

The main strategies directed toward a cure are reviewed elsewhere[6,7,9,12,15-17] and include the inactivation of replication-competent virus and the elimination of latently infected cells. An essential milestone to HIV reservoir eradication is the identification of biomarkers of latently infected cells[18,19], so that these cells can be specifically targeted by immunotoxins[20]. Currently, the foremost strategy for elimination of latently infected cells is controlled virus reactivation in the presence of continuing cART ("shock and kill")[21,22]. For this purpose, small molecule compound latency reversing agents (LRAs) are currently tested. The first LRAs used were histone deacetylase (HDAC) inhibitors (HDACi), which progressed to clinical trials[23-27] and demonstrated the ability to induce expression of HIV RNA. Unfortunately, none of the studies that followed the reservoir size post-treatment reported a significant reduction[23,25,27]. The multiplicity of molecular mechanisms involved in latency control suggests that a combination approach will likely be required to achieve the degree of reactivation necessary for the infected cell to be recognized by the immune system[28-30]. Indeed, some of the tested LRA combinations demonstrated synergy for HIV reactivation[31-35].

Gene expression profiling techniques and systems biology applications may be extremely useful in the identification of biomarkers of latency, further delineating mechanisms of regulation of HIV expression in a search for novel strategies of latency reversal, and for our understanding of the mechanisms of action of existing LRAs. Methods of analysis of gene expression data have been reviewed previously[36-40], including application of bioinformatics methods to HIV integration site analysis and the assessment of transcriptome and proteome changes induced in cells infected with HIV[41]. The present review provides a broader perspective on the use of gene expression profiling and systems biology applications in the field of HIV latency and eradication. Specifically, the objectives of the present review are: (1) to review the existing gene expression profiling and systems biology methods and their potential in the field of HIV research. We focus on the transcriptomic methods, and progress from simple approaches of differential gene expression to more complex types of analyses that integrate transcriptomic data with other biological data types, including proteomic analyses, integration site distribution, epigenetic modifications and transcription factor databases; and (2) to systematically demonstrate how methods of gene expression profiling and systems biology have been applied to answer specific questions in the fields of HIV latency and eradication. In this section we summarize specific findings that were obtained using gene expression profiling and systems biology methods, as described in existing literature.

GENE EXPRESSION PROFILING AND SYSTEMS BIOLOGY APPROACHES APPLIED IN THE FIELD OF HIV LATENCY AND ERADICATION

In this section, we describe the major methods of gene expression analysis and systems biology approaches and outline specific questions that can be addressed in

Table 1 Methods of gene expression profiling and systems biology and their applications in the field of human immunodeficiency virus latency and eradication

Method	Applications to discovery of latency biomarkers and mechanisms of regulation of HIV expression	Applications to studying the LRA mechanisms of action and evaluating combination therapies
Differential gene expression	Identification of latency biomarkers	Identification of genes responsive to LRA treatment
GO term/pathway enrichment	(1) Focusing study efforts upon gene groups of interest (*e.g.*, membrane proteins as biomarkers) (2) Identification of the mechanisms behind gene expression alterations (3) Delineating the molecular mechanisms contributing to latency control	(1) Elucidation of mechanisms of action of LRAs (2) Selection of gene targets for combination therapy based on gene function in enriched pathway
Network-based analysis	Identification of major regulators involved in HIV latency control, which may be only slightly dysregulated but significantly affect downstream molecules and pathways	(1) Elucidation of mechanisms of action of LRAs; (2) Prioritization of targets for combination therapies based upon type of connectivity (include if it regulates HIV-related processes; exclude if it regulates general intracellular processes)
Consolidating gene expression with other biological data (proteome, integration sites, chromatin features, *etc.*)	(1) Identification of latency biomarkers with transient RNA, but stable protein expression; (2) Identification of mechanisms of latency control by correlating chromatin features to gene expression	(1) Identification of post-transcriptional mechanisms of action of LRAs; (2) Assessment of chromatin features of genes and HIV integration sites responsive to LRA treatment
HIV expression and transcript type	Potential biomarker of latency	Assessment of the effectiveness of LRAs for HIV reactivation

LRA: Latency reversing agent; GO: Gene ontology; HIV: Human immunodeficiency virus.

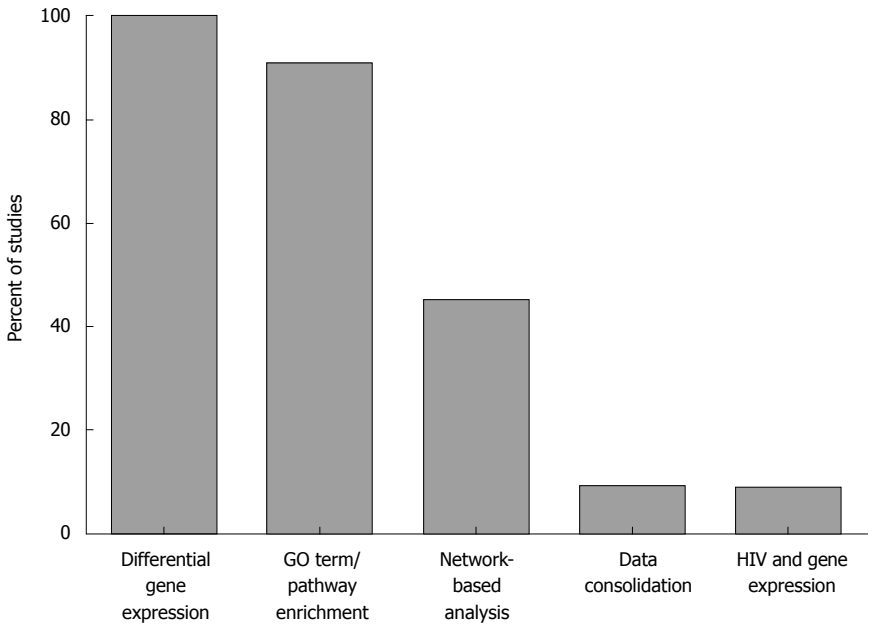

Figure 1 Summary of methods used across gene expression profiling studies in the field of human immunodeficiency virus latency and eradication. Identification of DEGs and functional analysis of GO terms and pathways enriched for DEGs are the methods that are most commonly used across studies. Network-based analyses are used in a subset of studies; while methods that consolidate host gene expression with other data types (*e.g.*, proteomics or HIV expression data) are scarce. DEGs: Differentially expressed genes; GO: Gene ontology; HIV: Human immunodeficiency virus.

the fields of HIV latency research and eradication using LRAs by each major type of application (Table 1). Where applicable, we highlight advantages and disadvantages of using individual methods over other methods for HIV latency related studies.

Differential gene expression

This basic analysis, common in all gene expression studies (Figure 1), aims at identifying genes that are expressed at different levels among the conditions tested. Gene expression can be compared in latently infected and uninfected cells to identify biomarkers of latency, and between cells treated with LRAs and untreated cells to identify genes that are responsive to LRA treatment.

To obtain gene expression data, two primary technologies are available: Microarrays and RNA-Seq. The majority of the published studies in the HIV latency field utilized microarrays, which is a well-developed technology with a fully established data analysis

pipeline. However, because microarrays use specific oligonucleotide probes, the detection is limited to only known genes. In addition, most of the microarray platforms are species-specific, which does not allow for simultaneous detection of host and pathogenic RNAs present in a sample. With advances in RNA-Seq technology and per sample cost reduction, gene expression profiling by RNA-Seq is more increasingly used. RNA-Seq allows measuring viral and cellular transcripts concomitantly in the same sample[42]. Other benefits of using RNA-Seq include increased sensitivity towards rare transcripts (as may be the case for HIV transcripts in latent state); detection of novel splice variants; and the wide dynamic range (reviewed in[43]). Numerous methods exist to analyze microarray (reviewed in[36,37,44]) and RNA-Seq datasets (reviewed in[38,39]), including methods of data processing, normalization and identification of differentially expressed genes (DEGs).

While methods of identification of DEGs are relatively straightforward, their application to mechanistic studies is limited. First, these methods usually generate far more DEGs that can be meaningfully discussed due to the lack of existing knowledge of their role in regulation of HIV expression. The second major issue in such studies is multiple comparisons. As more genes are included in either microarrays or RNA-Seq studies, the threshold for differential expression becomes much harder to reach due to the increased chance of type 1 error. Finally, a third issue arises with regards to the ranking of importance for genes which are differentially expressed. These can be ranked based upon fold change or a ranking system based upon prior knowledge of the gene. However, a gene product which is an important player of a pathway may not be well characterized, nor be heavily dysregulated, but may still cause large downstream changes.

Functional analyses to identify gene ontology terms and pathways enriched for DEGs

These frequently used methods (Figure 1) are designed to identify groups of genes sharing a common functional category or purpose that is significantly altered by gene dysregulation. Functional gene annotation may be useful for biomarker discovery to identify genes that encode membrane proteins. These proteins represent more feasible targets for antibody-bound immunotoxins as compared to intracellular proteins. Mainly, though, gene ontology (GO) term and pathway enrichment analysis is used to identify the mechanisms behind gene expression alterations in latency and during LRA treatment. Finally, specific pathways may be identified for targeting in combinatorial reactivation strategies, based on enrichment for DEGs.

There are numerous databases of annotated GO terms and pathways, and methods to analyze these functional categories, many of which are publicly available tools (reviewed in[40]). Gene set enrichment analysis (GSEA) approaches are the most commonly used method to identify GO terms and pathways that are enriched for DEGs[45-47]. Among these, ToppGene[47] has several advantages, including a user-friendly interface, allowing multiple input codes for genes, and performing both GO term and pathway enrichment analyses. Many similar functions are available in the DAVID Bioinformatics Resources tool[46]. GoSeq tool was developed specifically for RNA-Seq data and quantifies gene length bias present in the data[48]. In cases when an intervention significantly alters the expression of an extremely large number of genes, as may be the case for some LRAs, GSEA approaches may not work as most categories are enriched. An alternative method, Functional Analysis of Individual Microarray Expression[49] utilizes an exponentially decreasing weighted expression to generate a score for each GO category or pathway in both experimental and control conditions. A *t*-test, or other statistical test can be then performed to determine if the scores are significantly different. One drawback of this method is the importance placed upon highly expressed genes. However, lowly expressed genes may play other roles through post-translational modifications or hub roles which are not detected by this method or differential expression methods in general. To address these issues, network analysis techniques are extremely useful.

Network-based gene expression analyses

These tools, used in about half of the studies in the field of HIV latency (Figure 1), are designed to identify key functional regulators among DEGs, and to evaluate gene network differences among experimental conditions. In the network-based analyses, the function of a single gene may be elucidated through a "guilt by association" approach. High connectivity between a known and unknown gene may shed light upon their function. Additionally, a group of highly connected genes may indicate that a biologically relevant pathway is at work in the altered state. These pathways or networks of genes can be tested for differential expression without the high type 1 error rate, which is common when testing many thousands of individual genes. Heavily connected genes whose importance may have been missed in a standard differential expression test would show up in a network method as a hub (highly connected) gene. In this way, additional genes with a role in latency control or reactivation may be identified, which would be missed in other types of analysis. Finally, genes may be selected as therapeutic targets based on the network analysis, if they are connected to other factors with roles in HIV latency control. Conversely, if a gene is connected to genes that encode proteins with broad cellular functions, it may be selected against as side effects from a therapeutic intervention would be expected.

One well-developed network analysis tool is Weighted Gene Co-expression Network Analysis (WGCNA)[50]. In this method, the connectivity between genes is determined by correlating the expression of these genes

across samples, independent of known protein-protein and protein-DNA interactions. First, an adjacency matrix is constructed based on correlations between each gene pair, followed by creating a topological overlay map (TOM) that utilize information not only from the direct interaction between two genes, but also their neighboring nodes. Once this TOM is created, genes may be subdivided into highly connected groups or modules. The eigengene of this module represents the mathematically optimal summary of the expression profiles of all genes within the module as determined by their expression variation across samples. This eigengene may then be correlated to any trait of interest, such as the expression of specific HIV transcripts, or the degree of HIV reactivation upon treatment with LRAs. Genes with unknown function may be explored through both the behavior of the module as a whole and within the module itself (peripheral gene or a primary hub gene). Highly connected genes often represent key players in pathways and shed light upon the mechanistic differences between the two conditions being compared, such as uninfected CD4+ T-cells *vs* HIV-infected CD4+ T cells. Another network-based method, the "Active modules" algorithm[51], utilizes a different approach to network analysis by determining which portions of the network contain an unexpectedly high occurrence of genes with significant changes in expression. In contrast with WGCNA, the "active modules" algorithm utilizes protein interaction data from available databases, which allows incorporating information about the host and HIV interactions[51]. Available software packages for network analysis usually use literature curated protein-protein and protein-DNA interactions databases, but do not take into account enrichment of specific clusters for DEGs (*e.g.,* Metacore, Ingenuity, iRefWeb). A major advantage of utilizing known interactions is independence from differential expression (*i.e.,* all known protein-protein and protein-DNA interactions will be displayed for each DEG). A drawback of literature-based networks is the dependency on the accuracy of annotated sources and the robustness of the algorithms for network generation.

Integrating gene expression with other types of biological data

Methods of transcriptomics are well-developed and capture the majority of annotated genes. However, previous studies have shown that the transcriptome only partially correlates with the proteome[52-54]; therefore, assessment of gene expression at the functional (protein) level may be necessary to validate the role of specific genes in HIV latency control and reactivation. In addition, proteomics methods identify the effects that are not reflected or captured at the RNA level; for example, due to an increase of translation from existing messenger RNA[55], or because of the transient RNA expression. Thus, proteome profiling may be used to identify latency biomarkers that are stably expressed at the protein level. In addition, profiling of post-transcriptional effects of LRAs is beneficial to capture those effects that would be missed if only the transcriptome profiling were performed. Analysis of the proteome may thus shed light on the mechanisms by which LRAs regulate gene expression[56], including, possibly, transcriptional activation of HIV.

Other biological data types may be integrated with gene expression profiling data to further understand the mechanisms of HIV latency and reactivation. The activity of the HIV promoter may depend on the characteristics of the site of proviral integration[57]. Chromatin features surrounding an integration site may contribute to the levels of HIV transcription, including histone acetylation and methylation, and DNA methylation. For example, latent inducible proviruses have a tendency to be integrated into highly expressed genes, gene deserts, or alphoid repeats[58]. The transcription level of nearby genes as well as viral genome orientation may influence transcription of viral genes by RNA interference mechanisms[59-61]. However, to date, no clear feature of integration sites could be identified when comparing 5 different models of HIV latency[62]. Integration of HIV into specific genes, such as genes associated with cell cycle, may provide advantage to the maintenance of the latent reservoir through clonal expansion[63].

Depending on the type of data, different modeling methods may be used. The study described below was done with cancer cell lines; however, their method of integrating datasets would be applicable for many types of HIV latency related data. The aims of the study were to determine how DNA methylation in different genomic regions contribute to gene expression in cancer cell lines, and whether methylation of transcription factor binding sites impact transcription factor recruitment and therefore gene expression[64]. Gene expression was measured by Affymetrix microarrays, and DNA methylation by methyl-CpG binding domain-based capture (MbDCap)-Seq[65]. Pearson correlation analysis and decision tree learning were used to determine the effect of methylation in various genomic regions (promoters, first and second exons, and first introns) on the breast cancer subtype differential gene expression. To determine the role of methylation in transcription factor binding, cell line-specific consensus sequences were generated by assembling reads that mapped to the significantly hypermethylated regions and then matching these sequences to candidate transcription factors using the TRANSFAC package[66]. Similar approaches can be used to determine the role of chromatin features such as DNA methylation, as well as histone acetylation and methylation, in regulation of the expression levels of genes that control HIV latency, in the latent state and during reactivation using LRAs.

Evaluating the levels of HIV RNA using RNA-Seq datasets

HIV full length unspliced (US) genomic RNA can be spliced into different mRNA species, 47 identified in an early study[67], and 78 more recently[68]. The major classes of transcripts constitute multiply spliced (MS)

transcripts that encode regulatory and accessory proteins Tat, Rev, and Nef; and singly spliced (SS) transcripts that encode one-exon Tat, Vpr, Vif, Vpu, and Env. The US transcripts encode Gag and Gag-Pol polyproteins. In cell line models of latency (ACH-2 and U1), MS and SS transcripts were detected at early stages of replication cycle, when little or no genomic (US) RNA was produced[69]. Both MS and US transcripts were detected at low levels in resting CD4+ T cells from the HIV-infected individuals, while the majority of detected transcripts represented abortive HIV transcripts lacking polyA tail[70]. As was suggested previously[71], HIV RNA itself may represent a biomarker of latency. While multiple assays have been developed to detect HIV RNA using PCR-based methods[72,73], they require design of specific primers to detect various forms of HIV RNA, and may be plagued by inability to detect HIV RNA in a subset of patients due to virus mutations. RNA-Seq technology allows for concomitant detection and quantification of various HIV RNA species from the same samples as host transcripts, regardless of the viral sequence. Total HIV transcripts, including the abortive transcripts, can be measured by RNA-Seq using total RNA (ribo-depleted) libraries that capture non-polyadenylated RNAs.

RNA-Seq can also be used to evaluate induction of HIV expression using LRAs. In this case, libraries enriched for polyA (polyadenylated) RNAs would be a more appropriate choice, since induction of abortive transcripts or read-through transcripts from the neighboring genes is not relevant to the success of the "shock and kill" strategy, as no viral proteins will be produced. Specifically, induction of polyA US transcripts would need to be monitored, as it is indicative of productive infection (that will result in production of virions). Unfortunately, none of the existing RNA-Seq data analysis packages have reliable tools for precise splice variant measurement from standard RNA-Seq datasets (50-100 base pair reads), in particular, complex overlapping sequences as in the case of HIV[67]. Precise measurement of splice variants require longer read capacity (10 kb)[74]; otherwise, expression of the major splice variants, MS and SS, and the US genomic RNA can be only estimated. Mohammadi *et al*[42] developed a method that allows the approximation of the proportions of different HIV transcripts in the RNA-Seq data. The method is based on determining the number of reads that pass through the splice junctions D1 [directly after the long terminal repeat (LTR) region] and D4 (splice junction between Tat-Rev and Vpu) that define MS, SS, and US transcripts. If a read passes through the junction D1, then it belongs to the US transcript. Reads which align to the left of the D1 junction but are broken at D1 and align to another segment of the HIV genome correspond to reads from either SS or MS transcripts (SS + MS). Reads overlapping the D4 junction correspond to reads from either US transcripts or SS transcripts (US + SS). Finally, reads which are broken at the D4 junction correspond to reads from MS transcripts. The SS read percentage is then estimated by subtracting the US and MS percentages from 100.

USING TRANSCRIPTOME PROFILING TO IDENTIFY BIOMARKERS OF HIV LATENCY

A recent study[20] provided a proof of principle that immunotoxins can be used to target cells expressing a specific surface molecule; however, the choice of CCR5 co-receptor resulted in killing of both HIV-infected and uninfected CCR5-expressing cells. This choice of target would not be optimal for therapeutic applications, since CD4+ T cells are usually already compromised in HIV-infected individuals. Therefore, identification of a unique biomarker signature of latently infected cells is warranted to target these cells for eradication with high specificity. These biomarkers may have additional applications; for example, reliable quantification of latently infected cells *in vivo* to follow the size of the latent reservoir in patients post-treatment, and enrichment for latently infected cells for further studies.

The proof of principle that latently infected cells may have a distinct gene expression signature was provided in an early study comparing gene expression in resting CD4+ T cells from aviremic HIV-infected individuals and HIV seronegative donors as controls using microarrays[75]. Whilst less than 0.1% of cells from aviremic patients were latently HIV-infected (as determined by presence of HIV-1 proviral DNA), 165 genes showed differential expression between CD4+ T cells from aviremic patients as compared to HIV-seronegative donors. The limitations of this study were the low prevalence of latently infected cells and the confounding effect of antiretroviral therapy on gene expression. Later studies aimed at characterizing the gene expression profile of latently HIV-infected cells using chronically HIV-infected cell lines or *in vitro* infected primary resting CD4+ T cells and reporter viruses, allowing for strategies to enrich or select for latently HIV-infected cells.

Table 2 summarizes the four studies comparing gene expression in latently infected cells *vs* their uninfected counterparts. To estimate the proportions of latently infected cells present in each model, provirus expression is reactivated following establishment of latency, using strong agents that induce T cell activation, such as phorbol myristate acetate[18], anti-CD3/anti-CD28 + IL-2[42], or phytohemagglutinin and feeder peripheral blood mononuclear cells[76]. The percentage of uninfected cells may be estimated by subtracting the percentage of latently infected cells from the total (100%), assuming that all latent proviruses were induced. The percentage of cells expressing HIV Gag protein (p24+) or GFP reporter is also measured before the stimulation, to determine whether there is background expression of HIV in each latency model. These p24+ or GFP+ cells may represent productively infected cells present due

Table 2 Features of gene expression studies comparing latently infected *vs* uninfected cells

Study characteristics	Krishnan and Zeichner[18]	Iglesias-Ussel *et al*[19]	Mohammadi *et al*[42]	Evans *et al*[76]
Cells used	Cell lines ACH-2, A3.01, J1.1	Primary CD4+ T cells	Primary CD4+ T cells co-cultured with feeder H80 human brain tumor cell line	Primary resting CD4+ T cells co-cultured with dendritic cells
Virus used	CXCR4 tropic HIV-1 LAV strain	CXCR4 tropic GFP reporter virus (GFP inserted in place of Nef)	CXCR4 tropic GFP reporter virus with mutations in Gag, Vif, Vpr, Vpu, Env and Nef	CCR5 tropic GFP reporter virus (GFP inserted into the Nef open reading frame)
Proportion of uninfected cells	⩽ 1.1%	0%	8%-18%	99.7%
Proportion of GFP+ or p24+ cells	8.20%	8.15%	Approximately 16%	0% (removed by sorting)
Proportion of latently infected cells	98.9%	100%	Approximately 82%-92%	Approximately 0.3%
Time of culture	N/A (chronically infected)	20-22 d	13 wk	5 d
Experiment replicates	8	4	Not reported	4
Gene expression profiling platform	Microarrays (Hs. UniGem2)	Microarrays (Agilent-012391 Whole Human Genome Oligo Microarray G4112A)	RNA-Seq (polyA RNA library; Illumina HiSeq2000)	Microarrays (Illumina Human-Ref8)
Method to identify DEGs	Parametric one-sample random variance *t*-test (BRB-Array Tools, $P < 0.001$)	Linear modeling and using an empirical Bayes method with FDR correction (limma)	Generalized linear modeling (DESeq, FDR < 0.05)	Linear modeling and using an empirical Bayes method (limma, FDR < 0.05)
Databases used for functional analyses	NIH mAdb	GO consortium; MsigDb; KEGG pathways	Reactome pathways Ver.40; MsigDb	IPA
Total number of DEGs	32	875	227	Not reported

CXCR4: Chemokine (C-X-C motif) receptor 4; LAV: Lymphadenopathy-associated virus; CCR5: Chemokine (C-C motif) receptor 5 (gene/pseudogene); GFP: Green fluorescent protein; polyA: Polyadenylated; DEGs: Differentially expressed genes; BRB: Biometric Research Branch; FDR: False discovery rate; NIH: National Institutes of Health; mAdb: Mad Bee; GO: Gene ontology; MsigDb: Molecular Signature database; KEGG: Kyoto Encyclopedia of Genes and Genomes; IPA: Ingenuity Pathway Analysis; N/A: Not applicable.

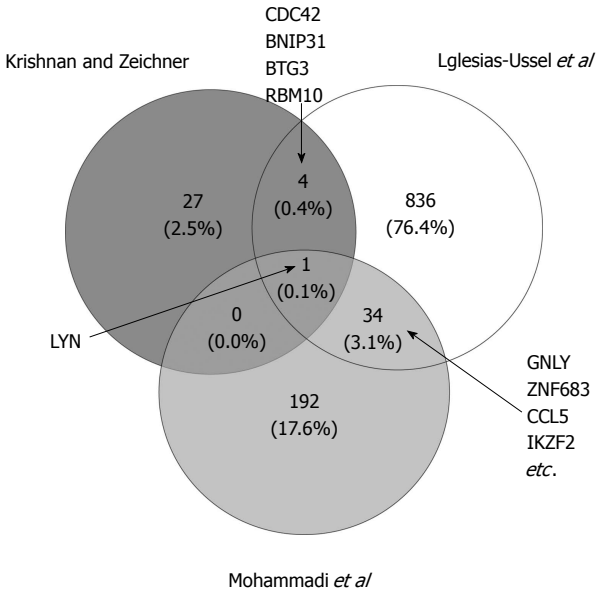

Figure 2 Venn diagram depicting differentially expressed genes across three latency models. The overlapping genes were identified using the online tool Venny (http://bioinfogp.cnb.csic.es/tools/venny/index.html). Shown are the total number of differentially expressed genes and percent of total identified across all models[18,19,42]. For each overlap, gene symbols are listed. For the overlap between Iglesias-Ussel *et al*[19] and Mohammadi *et al*[42] studies, the four genes with the highest average absolute fold change are listed.

the leakiness of a model, or be reflective of the viral entry in the absence of *de novo* viral production. Of note, Krishnan and Zeichner[18] provided these estimates only for one of the cell lines studied, ACH-2. The proportions of each cell type need to be taken into account when evaluating the results from differential expression analysis.

Table 2 presents additional characteristics that differed among the studies, including cells that were used (proliferating cell lines, resting CD4+ T cells or total CD4+ T cells), the duration of time in culture and viruses used to infect the cells. Finally, gene expression profiling platforms and statistical approaches to analyze the data were also different.

In order to assess whether biomarkers of latency can be reliably identified using gene expression profiling, we compared the DEG lists, where available (all studies except for Evans *et al*[76]). Krishnan and Zeichner[18] reported 32 genes that were consistently changed in latency in all three cell lines that were tested, and this list of DEGs was used. The number of DEGs from each study that participated in this analysis is indicated in Table 2 (bottom row). If consistent changes across model systems could be detected, these genes would represent strong latency biomarker candidates.

Figure 2 depicts the result of comparison of DEGs between latently infected and uninfected cells available from three published studies[18,19,42]. A total of 1094 DEGs were identified. Only one gene, LYN proto-oncogene, Src family tyrosine kinase (*LYN*), was dysregulated in latency in all three models. Not surprisingly, there were

Table 3 Limitations of the present studies that identify differentially expressed genes between latently infected and uninfected cells and possible solutions that may enable identification of solid candidate biomarkers of latency

Limitations	Solutions
Small percentage of latently infected cells	Isolate latently infected cells using reporter system OR perform gene expression profiling on a single-cell level
Effect from the exposure to the virus without infection	Use aldrithiol-2 inactivated virus[123] instead of mock-infection to compare to latently infected cell model
Identified differentially expressed genes are ubiquitously expressed on all CD4+ T cells	Identify a panel of biomarkers that best differentiates between latently infected and uninfected cells
Different models represent different aspects of latency establishment	Include additional models into analysis; use same statistical approaches to ensure differences in biomarkers are biological, not technical differences
Gene expression profiling can only identify candidate biomarkers	Perform experimental validation that latently infected cells can be detected using these biomarkers

fewer similarities between the cell lines and each of the primary cell models. In addition to *LYN*, only four genes were in common between Krishnan and Zeichner[18] and Iglesias-Ussel *et al*[19] studies. More similarities were found when comparing the two studies that performed gene expression profiling using primary CD4+ T cells (Iglesias-Ussel *et al*[19] and Mohammadi *et al*[42]): 34 genes were found in common, with the majority (29 of 34) consistently up- or down- regulated in latency in both models. The remaining genes were unique for any given study (27 of 32, or 84% for Krishnan and Zeichner[18], 836 of 875, or 96% for Iglesias-Ussel *et al*[19], and 192 of 227, or 85% for Mohammadi *et al*[42]).

This comparison indicated that despite the small proportion of overlapping genes between models, genes whose products may be able to differentiate between latently infected and uninfected cells can be identified using gene expression profiling, especially when comparing models established in primary cells. However, these studies have several limitations that presently preclude from achieving a consensus on what genes may represent suitable biomarkers of latency. These limitations and potential solutions that may advance this field are summarized in Table 3.

TRANSCRIPTOME PROFILING AND SYSTEMS BIOLOGY APPROACHES TO IDENTIFY MOLECULAR MECHANISMS OF REGULATION OF HIV EXPRESSION

Understanding the mechanisms of establishment and maintenance of HIV latency has greatly contributed to the development of strategies for eradication. It has become apparent that multiple cellular processes and pathways contribute to the control of HIV latency at both the transcriptional and post-transcriptional levels[1], suggesting that combination strategies will likely be needed to achieve eradication of the latent reservoir[28]. Block of viral transcription from the LTR is the most studied mechanism, which occurs through several proposed routes: Inhibition of transcription though histone and DNA modifications[77-79]; absence of necessary transcriptional activators and presence of transcriptional repressors in resting CD4+ T cells[80,81]; integration into inactive transcription sites[57]; or premature termination of viral transcripts in the absence of Tat and Tat-associated host factors[82]. Another mechanism suggests that latency may be maintained due to post-transcriptional blocks. HIV could be transcribed, but could fail to export MS HIV transcripts, contributing to non-productive infection in resting CD4+ T cells[83]. Finally, discoveries in the field of inhibitory micro RNAs (miRNAs) suggest a possibility of transcriptional inhibition of HIV by miRNAs encoded in HIV genome[84] and translational inhibition by host miRNAs[85].

Gene expression profiling data can be used to identify gene categories that describe cellular processes and pathways, as well as key regulatory factors with a role in HIV latency control, thus contributing to our understanding of the mechanisms that regulate HIV expression. The same studies described in Table 2 performed functional category analysis by identifying pathways and GO terms enriched for DEGs. Though these four studies utilized different cell types and viruses (Table 2), some uniting themes were observed in the mechanisms contributing to HIV latency control. We utilized the lists of GO terms and pathways that were reported in each of the four studies, to compare the gene categories dysregulated in different latency models. The reported terms were assigned to two major categories: Transcriptional regulation, including signaling pathways that regulate activity and localization of transcription factors, and functional categories related to RNA synthesis; and post-transcriptional regulation, both at the RNA and protein levels (Figure 3); terms that could not be assigned to these categories are not shown. Not surprisingly, the specific GO terms and pathways in each category were different between the studies, which was at least in part attributable to the usage of different annotated databases to obtain these terms (Table 2). However, terms associated with both transcriptional and post-transcriptional control of HIV latency were reported in more than one study. These GO terms and pathways comprise both well-established (*e.g.,* NFκB signaling and transcriptional regulation[86,87]) and novel mechanisms of regulation of HIV expression (*e.g.,* proteasome[18]).

Network-based approaches can also be utilized to identify genes that may have a role in regulation of HIV

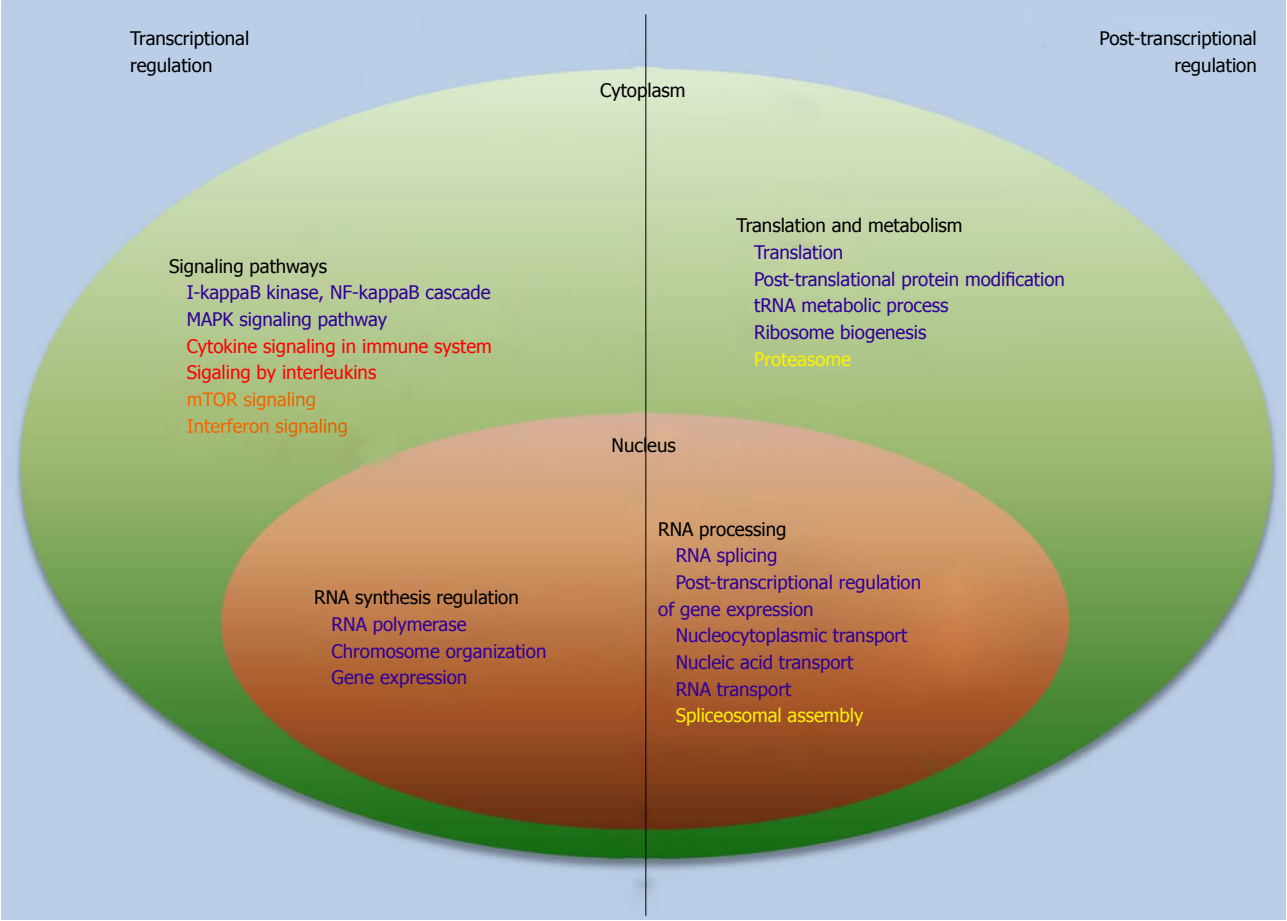

Figure 3 Transcriptional and post-transcriptional mechanisms of regulation of human immunodeficiency virus expression. Pathway and GO term categories related to transcriptional and post-transcriptional regulation of HIV expression, identified in gene expression studies that compared latently infected and uninfected cells, are shown. Dark blue, Iglesias-Ussel *et al*[19]; Red, Mohammadi *et al*[42]; Brown, Evans *et al*[76]; Yellow, Krishnan and Zeichner[18]. GO: Gene ontology; HIV: Human immunodeficiency virus; mTOR: Mammalian target of rapamycin.

expression, despite not being detected as differentially expressed in latency. For example, tubulin alpha 3 (*TUBA3)* was a well-connected gene in a network constructed by Bandyopadhyay *et al*[51] who utilized the Krishnan and Zeichner dataset[18]. *TUBA3* was connected to both Tat and Rev in the network, suggesting a possible yet unknown post-transcriptional role for this gene in regulation of HIV expression, one which would not have been detected in non-network-based approaches.

Taken together, functional studies using systems biology approaches to analyze host gene expression in the *in vitro* models of HIV latency suggest that maintenance of HIV quiescence in T cells involves basic cellular mechanisms beyond those traditionally implicated in transcriptional repression of the HIV-1 provirus.

TRANSCRIPTOME PROFILING AND SYSTEMS BIOLOGY APPROACHES TO IDENTIFY MOLECULAR MECHANISMS OF HIV REACTIVATION USING LRAS

HDACis have been the most studied LRAs, with a number of these compounds progressing to clinical trials[23-27]. The primary mechanism of action proposed for HIV reactivation using HDACis was histone acetylation and chromatin decondensation, which provide a transcriptionally favorable environment[88]. However, the results from gene expression profiling studies following the discovery of anti-cancer properties of HDACis (reviewed in[89]) strongly suggest the existence of secondary mechanisms of action of HDACis beyond chromatin remodeling. In particular, despite chromatin decondensation, as many genes were downregulated by HDACis as were upregulated. Over the years, studies using HDACis demonstrated that transformed cells responded to treatment differently as compared to primary cells[90-93]. Therefore, gene expression profiling of HDACis using primary CD4+ T cells is more relevant for delineating the mechanisms driving HIV reactivation. Most of the gene expression studies using HDACis in primary cells up-to-date have utilized the HDACi vorinostat/suberoylanilide hydroxamic acid (SAHA), which was the first of the FDA-approved HDACis for treatment of cutaneous T cell lymphoma[94]. These studies are summarized in Table 4. In addition to SAHA, the effects on gene expression were profiled for another HDACi, valproic acid (VPA) in primary CD4+ T cells

Table 4 Features of gene expression studies comparing suberoylanilide hydroxamic acid -treated and untreated primary cells

Study characteristics	Beliakova-Bethell *et al*[96]	Reardon *et al*[100]	White *et al*[99]	Mohammadi *et al*[42]	Elliott *et al*[25]
Cells used	Primary CD4+ T cells	Primary CD4+ T cells	Primary CD4+ T cells	*In vitro* primary CD4+ T cell latency model	Total blood from HIV-infected individuals on cART
Concentration or dose of SAHA	0.34 μmol/L	0.34, 1, 3, 10 μmol/L	1 μmol/L	0.5 μmol/L	400 mg orally once daily
Time of treatment	24 h	24 h	24 h	8 h and 24 h	14 d (samples analyzed at 2, 8 h; 1, 14 and 84 d)
Experiment replicates	9	6	6	Not reported	9
Gene expression profiling platform	Microarrays (Illumina HT12 Beadchips version 3)	Microarrays (Illumina HT12 Beadchips version 3)	Microarrays (Illumina HT12 Beadchips version 3)	RNA-Seq (polyA RNA library; Illumina HiSeq2000)	Microarrays (Illumina Human HT12 version 4)
Methods to identify DEGs	Multivariate permutation test (BRB-Array tools)	Dose-response analysis using likelihood ratio test (Isogene) with Bonferroni correction ($P < 0.05$)	Linear modeling (limma, FDR $P < 0.05$)	Generalized linear modeling (DESeq, FDR < 0.05)	Linear modeling (limma, $P < 0.05$)
Databases used for functional analyses	GO consortium, KEGG and Biocarta pathways (BRB-Array Tools), MetaCore networks	GO consortium, KEGG and Biocarta pathways (BRB-Array Tools), MetaCore networks	GO consortium, KEGG pathways (FAIME), MetaCore networks	Reactome pathways Ver.40; MsigDb	IPA, MsigDb
Total number of DEGs	1847	3477	2982	1289	Not reported

cART: Combination antiretroviral therapy; polyA: Polyadenylated; DEGs: Differentially expressed genes; BRB: Biometric Research Branch; FDR: False discovery rate; GO: Gene ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; MsigDb: Molecular Signature database; FAIME: Functional Analysis of Individual Microarray Expression; IPA: Ingenuity Pathway Analysis; HIV: Human immunodeficiency virus.

Table 5 Features of gene expression studies comparing cells treated with latency reversing agents of different functional classes and untreated cells

Study characteristics	Jiang *et al*[95]	Mohammadi *et al*[42]	Sung and Rice[97]	Banerjee *et al*[98]
Cells used	Primary cells from HIV-infected individuals on cART	*In vitro* primary CD4+ T cell latency model	Primary resting CD4+ T cells	J-Lat 10.6 T cell line
LRA (functional class)	Valproic acid (HDACi)	Disulfiram (alcohol dehydrogenase inhibitor)	Prostratin (PKC agonist)	JQ1 (bromodomain inhibitor)
Concentration	1 mmol/L (+20 U/mL IL-2)	0.5 μmol/L	250 ng/mL	0.1 μmol/L, 1 μmol/L
Time of treatment	6 h	8 and 24 h	48 h	24 h
Experiment replicates	4	Not reported	3	Not reported
Gene expression profiling platform	Microarrays (Agilent)	RNA-Seq (polyA RNA library; Illumina HiSeq2000)	Microarrays (Affymetrix Human Genome U133 Plus 2.0)	Microarrays (Affymetrix ST 1.0)
Methods to identify DEGs	Rosetta Resolver system ($P < 0.01$)	Generalized linear modeling (DESeq, FDR < 0.05)	*t*-test with FDR correction	ANOVA ($P < 1E-5$)
Databases used for functional analyses	Not used	Reactome pathways Ver.40; MsigDb	GO consortium, KEGG pathways	GO consortium
Total number of DEGs	199 (fold change > 3)	189	2514 (fold change > 1.5)	Not reported

cART: Combination antiretroviral therapy; LRA: Latency reversing agent; HDACi: Histone deacetylase inhibitor; PCK: Protein kinase C; polyA: Polyadenylated; DEGs: Differentially expressed genes; FDR: False discovery rate; ANOVA: Analysis of variance; MsigDb: Molecular Signature database; GO: Gene ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; LRAs: Latency reversing agents.

from HIV-infected individuals. Treatment with either SAHA or VPA resulted in downregulation of V-Myc avian myelocytomatosis viral oncogene homolog (*MYC*)[95,96]. Among other LRA classes, the effects of alcohol dehydrogenase inhibitor Disulfiram and protein kinase C (PKC) agonist Prostratin on host gene expression were assessed using primary CD4+ T cells[42,97], while the effects of a bromodomain inhibitor, JQ1, on gene expression were assessed in a cell line model of HIV latency (J-Lat 10.6 T cell line)[98] (see Table 5 for the summary of the studies).

For all classes of compounds tested, Disulfiram appeared to induce minimal changes to host gene expression[42], while SAHA and Prostratin modulated thousands of genes[42,96,97,99,100]. Gene expression studies were able to identify novel mechanisms contributing to HIV reactivation out of latency by LRAs, besides their primary mechanisms of action. For example, in addition to chromatin decondensation, SAHA upregulated specific HIV transcriptional activators [*e.g.,* immunity-related GTPase family, M (*IRGM*)[101], heat shock protein 70 (HSP70, gene symbol *HSPA2*)[102,103] and lysine (K)-

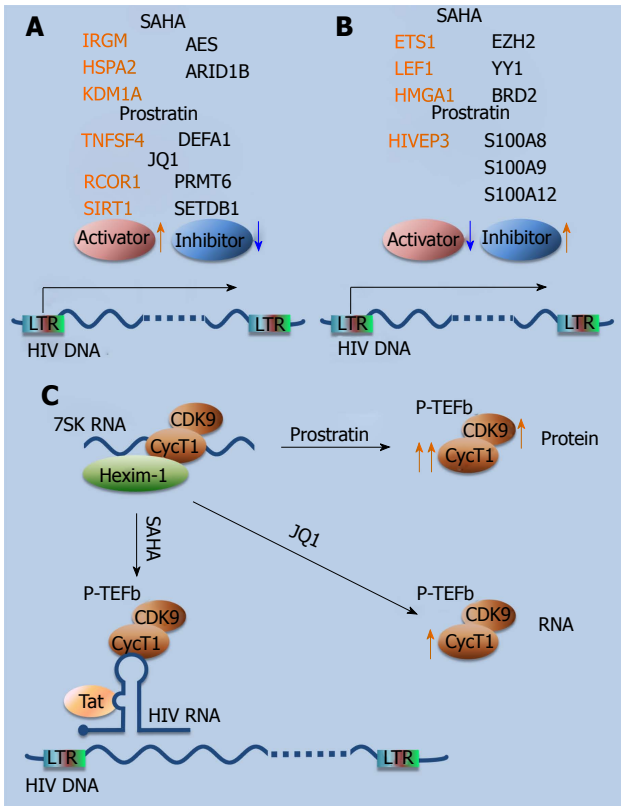

Figure 4 Main findings from gene expression studies using Latency reversing agents. A: Novel mechanisms of HIV reactivation besides primary mechanisms of action of LRAs. These include upregulation (red arrow) of HIV activators (red oval) and downregulation (blue arrow) of repressors (blue oval). Examples for LRAs from 3 functional classes (HDACi, SAHA; PKC agonist, Prostratin; and bromodomain inhibitor, JQ1) are listed; B: Effects of LRAs on host genes that are inhibitory for HIV reactivation. These include upregulation (red arrow) of HIV repressors (blue oval) and downregulation (blue arrow) of activators (red oval). Examples for LRAs from 2 functional classes (HDACi, SAHA; and PKC agonist, Prostratin) are shown; C: LRAs of different classes act on components of p-TEFb complex *via* different mechanisms, contributing to HIV reactivation. SAHA induced dissociation of p-TEFb from the inactive 7SK RNA complex and facilitated its recruitment to the HIV LTR. Prostrain and JQ1 upregulated components of p-TEFb complex at the protein and RNA level, respectively (red arrows indicate upregulation). LRA: Latency reversing agent; HDACi: Histone deacetylase inhibitor; PKC: Protein kinase C; SAHA: Suberoylanilide hydroxamic acid; IGRM: Immunity-related GTPase family, M; HSPA2: Heat shock 70 kDA protein 2; KDM1A: Lysine (K)-specific demethylase; TNFSF4: Tumor necrosis factor (ligand) superfamily, member 4; RCOR1: REST coreceptor 1; SIRT1: Sirtuin 1; AES: Amino-terminal enhancer of split; ARID1B: AT rich interactive domain 1B, SWI1-like; DEFA1: Defensin alpha 1; PRMT6: Protein arginine methyltransferase 6; SETDB1: SET domain, bifurcated 1; ETS1: V-Ets avian erythroblastosis virus E26 oncogene homolog 1; LEF1: Lymphoid enhancer-binding factor 1; HMGA1: High mobility group AT-hook 1; HIVEP3: HIV type I enhancer binding protein 3; EZH2: Enhancer of zeste 2 polycomb repressive complex 2 subunit; YY1: YY1 transcription factor; BRD2: Bromodomain protein containing 2; S100A8: S100 Calcium Binding Protein A8; S100A9: S100 Calcium Binding Protein A9; S100A12: S100 Calcium Binding Protein A12; CDK9: Cyclin-dependent kinase 9; P-TEFb: Positive transcription elongation factor; CycT1: Cyclin T1; Hexim-1: Hexamethylene Bis-Acetamide Inducible 1; LTR: Long terminal repeat; Tat: Transactivator of transcription.

specific demethylase (*KDM1A*)[104]], and downregulated repressors [amino-terminal enhancer of split[105] and AT rich interactive domain 1B, SWI1-like (*ARID1B*, or BAF250)[106]][25,99,100] (Figure 4A). Sung and Rice[97] found that Prostratin upregulated HIV activator, tumor necrosis factor (ligand) superfamily, member 4 (*TNFSF4*)[107], and downregulated defensin alpha 1, which interferes with PKC signaling[108]. Among genes with a role in regulation of HIV expression that were modulated by JQ1, Banerjee *et al*[98] noted upregulation of activators REST coreceptor 1 (*RCOR1*)[104] and the class Ⅲ deacetylase sirtuin 1 (*SIRT1*)[109], and downregulation of repressor methyltransferases, protein arginine methyltransferase 6 (*PRMT6*) and SET domain, bifurcated 1 (*SETDB1*)[110,111].

In addition to the effects of LRAs on gene expression that may promote HIV reactivation, possible inhibitory effects were also observed in gene expression studies that used SAHA and Prostratin-treated primary cells (Figure 4B). Genes encoding factors that activate HIV transcription, V-Ets avian erythroblastosis virus E26 oncogene homolog 1(*ETS1*), CCAAT/enhancer binding protein, Beta (*CEBPB*), and lymphoid enhancer-binding factor 1 (*LEF1*)[112-114], were downregulated by SAHA in primary CD4+ T cells[100]. Enhancer of zeste 2 polycomb repressive complex 2 subunit (*EZH2*), a methyltransferase implicated in HIV LTR silencing[115], was upregulated[100]. Genes encoding HIV transcriptional repressors YY1[116] and bromodomain protein containing 2 (*BRD2*)[117] were upregulated by SAHA in blood cells from HIV-infected individuals on cART[25]. Downregulation of *ETS1* and *LEF1* and upregulation of *BRD2* were confirmed at the protein level in primary CD4+ T cells[99]. In addition, a network-based approach integrating transcriptomics and proteomics datasets highlighted upregulation of high mobility group AT-hook 1[99], which represses HIV transcription by competing with Tat for TAR binding[118] and by recruiting inactive positive transcription elongation factor (p-TEFb) to the HIV LTR[119]. Possible inhibitory effects of Prostratin with respect to HIV reactivation identified by Sung and Rice[97] were upregulation of a repressor, HIV type Ⅰ enhancer binding protein 3[120], and downregulation of the three genes encoding S100 calcium-binding proteins (*S100A8, S100A9,* and *S100A12*), shown to enhance HIV-1 transcription in a NFκB-dependent manner[121].

Finally, gene expression profiling studies using LRAs of different functional classes highlighted uniting themes driving HIV reactivation, such as importance of the components of p-TEFb complex (Figure 4C). Cyclin T1 (CycT1) was upregulated at the RNA level by JQ1[98]; both CycT1 and cyclin-dependent kinase 9 were upregulated at the protein level by Prostratin[97], while SAHA induced dissociation of p-TEFb from the inactive 7SK RNA complex and facilitated its recruitment to the HIV LTR[122]. Though through different mechanisms, p-TEFb function appears to be enhanced *via* action of several classes of LRAs.

CONCLUSION AND PERSPECTIVES

This review discusses how methods of gene expression profiling and systems biology can be applied to address specific questions in the field of HIV latency and eradication. It presents a systematic analysis of the

application of these methods to discover biomarkers of latency, identify molecular mechanisms of latency control and reactivation using LRAs. Identification of DEGs and functional category assessment are the most common methods currently used in the field (Figure 1). Network-based approaches are utilized in a subset of more recent studies. Advances in RNA-Seq technologies allow for integration of HIV expression analysis with the changes in expression of host genes in a single experiment. Integration of transcriptomic data with other biological data types in the field of HIV latency is presently scarce; and the field would benefit from increased adoption of these methods in future studies.

Gene expression analysis of latently infected and uninfected cells has been used to identify candidate biomarkers of latency and to delineate the molecular mechanisms that contribute to regulation of HIV expression. Studies comparing gene expression in HIV latency models to uninfected cells have several limitations that presently preclude from achieving a consensus on what genes may represent suitable biomarkers (Table 3). Improved bioinformatics approaches (*e.g.,* using the same methods of data acquisition and statistical analyses across models) and experimental validation of candidate biomarkers would be extremely useful in future studies to more reliably identify biomarkers of latency. Studies profiling gene expression changes induced by LRAs identified novel mechanisms of action of the LRAs and their inhibitory effects with respect to HIV reactivation out of latency, as well as highlighted uniting themes driving HIV reactivation. Using similar statistical approaches in prospective studies using LRAs would facilitate prediction of whether the inhibitory effects of different LRAs on HIV reactivation could be cancelled out in a combination strategy. The results from such studies would have the potential to significantly impact the process by which candidate drugs are selected and combined for future evaluations and advancement to clinical trials.

ACKNOWLEDGMENTS

This material is based upon work supported in part by the Department of Veterans Affairs (VA), Veterans Health Administration, Office of Research and Development. The views expressed in this article are those of the authors and do not necessarily reflect the position or policy of the Department of VA or the United States government. We thank Dr. Andrew Rice for sharing the list of DEGs that were identified in the study of the effects of Prostratin in primary CD4+ T cells.

REFERENCES

1 **Richman DD**, Margolis DM, Delaney M, Greene WC, Hazuda D, Pomerantz RJ. The challenge of finding a cure for HIV infection. *Science* 2009; **323**: 1304-1307 [PMID: 19265012 DOI: 10.1126/science.1165706]

2 **Siliciano JD**, Kajdas J, Finzi D, Quinn TC, Chadwick K, Margolick JB, Kovacs C, Gange SJ, Siliciano RF. Long-term follow-up studies confirm the stability of the latent reservoir for HIV-1 in resting CD4+ T cells. *Nat Med* 2003; **9**: 727-728 [PMID: 12754504 DOI: 10.1038/nm880]

3 **Strain MC**, Günthard HF, Havlir DV, Ignacio CC, Smith DM, Leigh-Brown AJ, Macaranas TR, Lam RY, Daly OA, Fischer M, Opravil M, Levine H, Bacheler L, Spina CA, Richman DD, Wong JK. Heterogeneous clearance rates of long-lived lymphocytes infected with HIV: intrinsic stability predicts lifelong persistence. *Proc Natl Acad Sci USA* 2003; **100**: 4819-4824 [PMID: 12684537 DOI: 10.1073/pnas.0736332100]

4 **Crooks AM**, Bateson R, Cope AB, Dahl NP, Griggs MK, Kuruc JD, Gay CL, Eron JJ, Margolis DM, Bosch RJ, Archin NM. Precise quantification of the latent HIV-1 reservoir: implication for eradication strategies. *J Infect Dis* 2015; **212**: 1361-1365 [PMID: 25877550 DOI: 10.1093/infdis/jiv218]

5 **Finzi D**, Blankson J, Siliciano JD, Margolick JB, Chadwick K, Pierson T, Smith K, Lisziewicz J, Lori F, Flexner C, Quinn TC, Chaisson RE, Rosenberg E, Walker B, Gange S, Gallant J, Siliciano RF. Latent infection of CD4+ T cells provides a mechanism for lifelong persistence of HIV-1, even in patients on effective combination therapy. *Nat Med* 1999; **5**: 512-517 [PMID: 10229227 DOI: 10.1038/8394]

6 **Van Lint C**, Bouchat S, Marcello A. HIV-1 transcription and latency: an update. *Retrovirology* 2013; **10**: 67 [PMID: 23803414 DOI: 10.1186/1742-4690-10-67]

7 **Battistini A**, Sgarbanti M. HIV-1 latency: an update of molecular mechanisms and therapeutic strategies. *Viruses* 2014; **6**: 1715-1758 [PMID: 24736215 DOI: 10.3390/v6041715]

8 **Donahue DA**, Wainberg MA. Cellular and molecular mechanisms involved in the establishment of HIV-1 latency. *Retrovirology* 2013; **10**: 11 [PMID: 23375003 DOI: 10.1186/1742-4690-10-11]

9 **Tyagi M**, Bukrinsky M. Human immunodeficiency virus (HIV) latency: the major hurdle in HIV eradication. *Mol Med* 2012; **18**: 1096-1108 [PMID: 22692576 DOI: 10.2119/molmed.2012.00194]

10 **Mbonye U**, Karn J. Transcriptional control of HIV latency: cellular signaling pathways, epigenetics, happenstance and the hope for a cure. *Virology* 2014; **454-455**: 328-339 [PMID: 24565118 DOI: 10.1016/j.virol.2014.02.008]

11 **Colin L**, Van Lint C. Molecular control of HIV-1 postintegration latency: implications for the development of new therapeutic strategies. *Retrovirology* 2009; **6**: 111 [PMID: 19961595 DOI: 10.1186/1742-4690-6-111]

12 **Dahabieh MS**, Battivelli E, Verdin E. Understanding HIV latency: the road to an HIV cure. *Annu Rev Med* 2015; **66**: 407-421 [PMID: 25587657 DOI: 10.1146/annurev-med-092112-152941]

13 **Siliciano RF**, Greene WC. HIV latency. *Cold Spring Harb Perspect Med* 2011; **1**: a007096 [PMID: 22229121 DOI: 10.1101/cshperspect.a007096]

14 **Liu RD**, Wu J, Shao R, Xue YH. Mechanism and factors that control HIV-1 transcription and latency activation. *J Zhejiang Univ Sci B* 2014; **15**: 455-465 [PMID: 24793763 DOI: 10.1631/jzus.B1400059]

15 **Choudhary SK**, Margolis DM. Curing HIV: Pharmacologic approaches to target HIV-1 latency. *Annu Rev Pharmacol Toxicol* 2011; **51**: 397-418 [PMID: 21210747 DOI: 10.1146/annurev-pharmtox-010510-100237]

16 **Katlama C**, Deeks SG, Autran B, Martinez-Picado J, van Lunzen J, Rouzioux C, Miller M, Vella S, Schmitz JE, Ahlers J, Richman DD, Sekaly RP. Barriers to a cure for HIV: new ways to target and eradicate HIV-1 reservoirs. *Lancet* 2013; **381**: 2109-2117 [PMID: 23541541 DOI: 10.1016/S0140-6736(13)60104-X]

17 **Chan CN**, Dietrich I, Hosie MJ, Willett BJ. Recent developments in human immunodeficiency virus-1 latency research. *J Gen Virol* 2013; **94**: 917-932 [PMID: 23364195 DOI: 10.1099/vir.0.049296-0]

18 **Krishnan V**, Zeichner SL. Host cell gene expression during human immunodeficiency virus type 1 latency and reactivation and effects of targeting genes that are differentially expressed in viral latency. *J Virol* 2004; **78**: 9458-9473 [PMID: 15308739 DOI: 10.1128/jvi.78.17.9458-9473.2004]

19 **Iglesias-Ussel M**, Vandergeeten C, Marchionni L, Chomont N, Romerio F. High levels of CD2 expression identify HIV-1 latently infected resting memory CD4+ T cells in virally suppressed subjects.

J Virol 2013; **87**: 9148-9158 [PMID: 23760244 DOI: 10.1128/JVI.01297-13]
20 **Rawlings SA**, Alonzo F, Kozhaya L, Torres VJ, Unutmaz D. Elimination of HIV-1-infected primary T cell reservoirs in an in vitro model of latency. *PLoS One* 2015; **10**: e0126917 [PMID: 25993666 DOI: 10.1371/journal.pone.0126917]
21 **Ylisastigui L**, Archin NM, Lehrman G, Bosch RJ, Margolis DM. Coaxing HIV-1 from resting CD4 T cells: histone deacetylase inhibition allows latent viral expression. *AIDS* 2004; **18**: 1101-1108 [PMID: 15166525 DOI: 10.1097/00002030-200405210-00003]
22 **Demonté D**, Quivy V, Colette Y, Van Lint C. Administration of HDAC inhibitors to reactivate HIV-1 expression in latent cellular reservoirs: implications for the development of therapeutic strategies. *Biochem Pharmacol* 2004; **68**: 1231-1238 [PMID: 15313421 DOI: 10.1016/j.bcp.2004.05.040]
23 **Archin NM**, Liberty AL, Kashuba AD, Choudhary SK, Kuruc JD, Crooks AM, Parker DC, Anderson EM, Kearney MF, Strain MC, Richman DD, Hudgens MG, Bosch RJ, Coffin JM, Eron JJ, Hazuda DJ, Margolis DM. Administration of vorinostat disrupts HIV-1 latency in patients on antiretroviral therapy. *Nature* 2012; **487**: 482-485 [PMID: 22837004 DOI: 10.1038/nature11286]
24 **Archin NM**, Bateson R, Tripathy MK, Crooks AM, Yang KH, Dahl NP, Kearney MF, Anderson EM, Coffin JM, Strain MC, Richman DD, Robertson KR, Kashuba AD, Bosch RJ, Hazuda DJ, Kuruc JD, Eron JJ, Margolis DM. HIV-1 expression within resting CD4+ T cells after multiple doses of vorinostat. *J Infect Dis* 2014; **210**: 728-735 [PMID: 24620025 DOI: 10.1093/infdis/jiu155]
25 **Elliott JH**, Wightman F, Solomon A, Ghneim K, Ahlers J, Cameron MJ, Smith MZ, Spelman T, McMahon J, Velayudham P, Brown G, Roney J, Watson J, Prince MH, Hoy JF, Chomont N, Fromentin R, Procopio FA, Zeidan J, Palmer S, Odevall L, Johnstone RW, Martin BP, Sinclair E, Deeks SG, Hazuda DJ, Cameron PU, Sékaly RP, Lewin SR. Activation of HIV transcription with short-course vorinostat in HIV-infected patients on suppressive antiretroviral therapy. *PLoS Pathog* 2014; **10**: e1004473 [PMID: 25393648 DOI: 10.1371/journal.ppat.1004473]
26 **Wei DG**, Chiang V, Fyne E, Balakrishnan M, Barnes T, Graupe M, Hesselgesser J, Irrinki A, Murry JP, Stepan G, Stray KM, Tsai A, Yu H, Spindler J, Kearney M, Spina CA, McMahon D, Lalezari J, Sloan D, Mellors J, Geleziunas R, Cihlar T. Histone deacetylase inhibitor romidepsin induces HIV expression in CD4 T cells from patients on suppressive antiretroviral therapy at concentrations achieved by clinical dosing. *PLoS Pathog* 2014; **10**: e1004071 [PMID: 24722454 DOI: 10.1371/journal.ppat.1004071]
27 **Rasmussen TA**, Tolstrup M, Brinkmann CR, Olesen R, Erikstrup C, Solomon A, Winckelmann A, Palmer S, Dinarello C, Buzon M, Lichterfeld M, Lewin SR, Østergaard L, Søgaard OS. Panobinostat, a histone deacetylase inhibitor, for latent-virus reactivation in HIV-infected patients on suppressive antiretroviral therapy: a phase 1/2, single group, clinical trial. *Lancet HIV* 2014; **1**: e13-e21 [PMID: 26423811 DOI: 10.1016/S2352-3018(14)70014-1]
28 **Margolis DM**, Hazuda DJ. Combined approaches for HIV cure. *Curr Opin HIV AIDS* 2013; **8**: 230-235 [PMID: 23446138 DOI: 10.1097/COH.0b013e32835ef089]
29 **De Crignis E**, Mahmoudi T. HIV eradication: combinatorial approaches to activate latent viruses. *Viruses* 2014; **6**: 4581-4608 [PMID: 25421889 DOI: 10.3390/v6114581]
30 **Mohammadi P**, Ciuffi A, Beerenwinkel N. Dynamic models of viral replication and latency. *Curr Opin HIV AIDS* 2015; **10**: 90-95 [PMID: 25565177 DOI: 10.1097/COH.0000000000000136]
31 **Tripathy MK**, McManamy ME, Burch BD, Archin NM, Margolis DM. H3K27 demethylation at the proviral promoter sensitized latent HIV to the effects of vorinostat in ex vivo cultures of resting CD4+ T cells. *J Virol* 2015; **89**: 8392-8405 [PMID: 26041287 DOI: 10.1128/jvi.00572-15]
32 **Pérez M**, de Vinuesa AG, Sanchez-Duffhues G, Marquez N, Bellido ML, Muñoz-Fernandez MA, Moreno S, Castor TP, Calzado MA, Muñoz E. Bryostatin-1 synergizes with histone deacetylase inhibitors to reactivate HIV-1 from latency. *Curr HIV Res* 2010; **8**: 418-429 [PMID: 20636281 DOI: 10.2174/157016210793499312#sthash.6eLI6koQ.dpuf]
33 **Laird GM**, Bullen CK, Rosenbloom DI, Martin AR, Hill AL, Durand CM, Siliciano JD, Siliciano RF. Ex vivo analysis identifies effective HIV-1 latency-reversing drug combinations. *J Clin Invest* 2015; **125**: 1901-1912 [PMID: 25822022 DOI: 10.1172/JCI80142]
34 **Reuse S**, Calao M, Kabeya K, Guiguen A, Gatot JS, Quivy V, Vanhulle C, Lamine A, Vaira D, Demonte D, Martinelli V, Veithen E, Cherrier T, Avettand V, Poutrel S, Piette J, de Launoit Y, Moutschen M, Burny A, Rouzioux C, De Wit S, Herbein G, Rohr O, Collette Y, Lambotte O, Clumeck N, Van Lint C. Synergistic activation of HIV-1 expression by deacetylase inhibitors and prostratin: implications for treatment of latent infection. *PLoS One* 2009; **4**: e6093 [PMID: 19564922 DOI: 10.1371/journal.pone.0006093]
35 **Burnett JC**, Lim KI, Calafi A, Rossi JJ, Schaffer DV, Arkin AP. Combinatorial latency reactivation for HIV-1 subtypes and variants. *J Virol* 2010; **84**: 5958-5974 [PMID: 20357084 DOI: 10.1128/JVI.00161-10]
36 **Simon R**. Analysis of DNA microarray expression data. *Best Pract Res Clin Haematol* 2009; **22**: 271-282 [PMID: 19698933 DOI: 10.1016/j.beha.2009.07.001]
37 **Tarca AL**, Romero R, Draghici S. Analysis of microarray experiments of gene expression profiling. *Am J Obstet Gynecol* 2006; **195**: 373-388 [PMID: 16890548 DOI: 10.1016/j.ajog.2006.07.001]
38 **Oshlack A**, Robinson MD, Young MD. From RNA-seq reads to differential expression results. *Genome Biol* 2010; **11**: 220 [PMID: 21176179 DOI: 10.1186/gb-2010-11-12-220]
39 **Chen G**, Wang C, Shi T. Overview of available methods for diverse RNA-Seq data analyses. *Sci China Life Sci* 2011; **54**: 1121-1128 [PMID: 22227904 DOI: 10.1007/s11427-011-4255-x]
40 **Werner T**. Bioinformatics applications for pathway analysis of microarray data. *Curr Opin Biotechnol* 2008; **19**: 50-54 [PMID: 18207385 DOI: 10.1016/j.copbio.2007.11.005]
41 **Ciuffi A**, Mohammadi P, Golumbeanu M, di Iulio J, Telenti A. Bioinformatics and HIV latency. *Curr HIV/AIDS Rep* 2015; **12**: 97-106 [PMID: 25586146 DOI: 10.1007/s11904-014-0240-x]
42 **Mohammadi P**, di Iulio J, Muñoz M, Martinez R, Bartha I, Cavassini M, Thorball C, Fellay J, Beerenwinkel N, Ciuffi A, Telenti A. Dynamics of HIV latency and reactivation in a primary CD4+ T cell model. *PLoS Pathog* 2014; **10**: e1004156 [PMID: 24875931 DOI: 10.1371/journal.ppat.1004156]
43 **Malone JH**, Oliver B. Microarrays, deep sequencing and the true measure of the transcriptome. *BMC Biol* 2011; **9**: 34 [PMID: 21627854 DOI: 10.1186/1741-7007-9-34]
44 **Simon R**, Lam A, Li MC, Ngan M, Menenzes S, Zhao Y. Analysis of gene expression data using BRB-ArrayTools. *Cancer Inform* 2007; **3**: 11-17 [PMID: 19455231]
45 **Subramanian A**, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA* 2005; **102**: 15545-15550 [PMID: 16199517 DOI: 10.1073/pnas.0506580102]
46 **Dennis G**, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, Lempicki RA. DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol* 2003; **4**: P3 [PMID: 12734009 DOI: 10.1186/gb-2003-4-5-p3]
47 **Chen J**, Bardes EE, Aronow BJ, Jegga AG. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic Acids Res* 2009; **37**: W305-W311 [PMID: 19465376 DOI: 10.1093/nar/gkp427]
48 **Young MD**, Wakefield MJ, Smyth GK, Oshlack A. Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol* 2010; **11**: R14 [PMID: 20132535 DOI: 10.1186/gb-2010-11-2-r14]
49 **Yang X**, Regan K, Huang Y, Zhang Q, Li J, Seiwert TY, Cohen EE, Xing HR, Lussier YA. Single sample expression-anchored mechanisms predict survival in head and neck cancer. *PLoS Comput Biol* 2012; **8**: e1002350 [PMID: 22291585 DOI: 10.1371/journal.pcbi.1002350]
50 **Langfelder P**, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 2008; **9**: 559

[PMID: 19114008 DOI: 10.1186/1471-2105-9-559]
51 **Bandyopadhyay S**, Kelley R, Ideker T. Discovering regulated networks during HIV-1 latency and reactivation. *Pac Symp Biocomput* 2006; 354-366 [PMID: 17094252 DOI: 10.1142/9789812701626_0033]
52 **Fu X**, Fu N, Guo S, Yan Z, Xu Y, Hu H, Menzel C, Chen W, Li Y, Zeng R, Khaitovich P. Estimating accuracy of RNA-Seq and microarrays with proteomics. *BMC Genomics* 2009; **10**: 161 [PMID: 19371429 DOI: 10.1186/1471-2164-10-161]
53 **Ghazalpour A**, Bennett B, Petyuk VA, Orozco L, Hagopian R, Mungrue IN, Farber CR, Sinsheimer J, Kang HM, Furlotte N, Park CC, Wen PZ, Brewer H, Weitz K, Camp DG, Pan C, Yordanova R, Neuhaus I, Tilford C, Siemers N, Gargalovic P, Eskin E, Kirchgessner T, Smith DJ, Smith RD, Lusis AJ. Comparative analysis of proteome and transcriptome variation in mouse. *PLoS Genet* 2011; **7**: e1001393 [PMID: 21695224 DOI: 10.1371/journal.pgen.1001393]
54 **Low TY**, van Heesch S, van den Toorn H, Giansanti P, Cristobal A, Toonen P, Schafer S, Hübner N, van Breukelen B, Mohammed S, Cuppen E, Heck AJ, Guryev V. Quantitative and qualitative proteome characteristics extracted from in-depth integrated genomics and proteomics analysis. *Cell Rep* 2013; **5**: 1469-1478 [PMID: 24290761 DOI: 10.1016/j.celrep.2013.10.041]
55 **Schwanhäusser B**, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, Chen W, Selbach M. Global quantification of mammalian gene expression control. *Nature* 2011; **473**: 337-342 [PMID: 21593866 DOI: 10.1038/nature10098]
56 **Mackmull MT**, Iskar M, Parca L, Singer S, Bork P, Ori A, Beck M. Histone deacetylase inhibitors (HDACi) cause the selective depletion of bromodomain containing proteins (BCPs). *Mol Cell Proteomics* 2015; **14**: 1350-1360 [PMID: 25755299 DOI: 10.1074/mcp.M114.042499]
57 **Jordan A**, Defechereux P, Verdin E. The site of HIV-1 integration in the human genome determines basal transcriptional activity and response to Tat transactivation. *EMBO J* 2001; **20**: 1726-1738 [PMID: 11285236 DOI: 10.1093/emboj/20.7.1726]
58 **Lewinski MK**, Bisgrove D, Shinn P, Chen H, Hoffmann C, Hannenhalli S, Verdin E, Berry CC, Ecker JR, Bushman FD. Genome-wide analysis of chromosomal features repressing human immunodeficiency virus transcription. *J Virol* 2005; **79**: 6610-6619 [PMID: 15890899 DOI: 10.1128/JVI.79.11.6610-6619.2005]
59 **Lenasi T**, Contreras X, Peterlin BM. Transcriptional interference antagonizes proviral gene expression to promote HIV latency. *Cell Host Microbe* 2008; **4**: 123-133 [PMID: 18692772 DOI: 10.1016/j.chom.2008.05.016]
60 **Han Y**, Lin YB, An W, Xu J, Yang HC, O'Connell K, Dordai D, Boeke JD, Siliciano JD, Siliciano RF. Orientation-dependent regulation of integrated HIV-1 expression by host gene transcriptional readthrough. *Cell Host Microbe* 2008; **4**: 134-146 [PMID: 18692773 DOI: 10.1016/j.chom.2008.06.008]
61 **Shan L**, Yang HC, Rabi SA, Bravo HC, Shroff NS, Irizarry RA, Zhang H, Margolick JB, Siliciano JD, Siliciano RF. Influence of host gene transcription level and orientation on HIV-1 latency in a primary-cell model. *J Virol* 2011; **85**: 5384-5393 [PMID: 21430059 DOI: 10.1128/jvi.02536-10]
62 **Sherrill-Mix S**, Lewinski MK, Famiglietti M, Bosque A, Malani N, Ocwieja KE, Berry CC, Looney D, Shan L, Agosto LM, Pace MJ, Siliciano RF, O'Doherty U, Guatelli J, Planelles V, Bushman FD. HIV latency and integration site placement in five cell-based models. *Retrovirology* 2013; **10**: 90 [PMID: 23953889 DOI: 10.1186/1742-4690-10-90]
63 **Maldarelli F**, Wu X, Su L, Simonetti FR, Shao W, Hill S, Spindler J, Ferris AL, Mellors JW, Kearney MF, Coffin JM, Hughes SH. HIV latency. Specific HIV integration sites are linked to clonal expansion and persistence of infected cells. *Science* 2014; **345**: 179-183 [PMID: 24968937 DOI: 10.1126/science.1254194]
64 **Rhee JK**, Kim K, Chae H, Evans J, Yan P, Zhang BT, Gray J, Spellman P, Huang TH, Nephew KP, Kim S. Integrated analysis of genome-wide DNA methylation and gene expression profiles in molecular subtypes of breast cancer. *Nucleic Acids Res* 2013; **41**: 8464-8474 [PMID: 23887935 DOI: 10.1093/nar/gkt643]
65 **Brinkman AB**, Simmer F, Ma K, Kaan A, Zhu J, Stunnenberg HG. Whole-genome DNA methylation profiling using MethylCap-seq. *Methods* 2010; **52**: 232-236 [PMID: 20542119 DOI: 10.1016/j.ymeth.2010.06.012]
66 **Kel AE**, Gössling E, Reuter I, Cheremushkin E, Kel-Margoulis OV, Wingender E. MATCH: A tool for searching transcription factor binding sites in DNA sequences. *Nucleic Acids Res* 2003; **31**: 3576-3579 [PMID: 12824369 DOI: 10.1093/nar/gkg585]
67 **Purcell DF**, Martin MA. Alternative splicing of human immunodeficiency virus type 1 mRNA modulates viral protein expression, replication, and infectivity. *J Virol* 1993; **67**: 6365-6378 [PMID: 8411338]
68 **Sherrill-Mix S**, Ocwieja KE, Bushman FD. Gene activity in primary T cells infected with HIV89.6: intron retention and induction of genomic repeats. *Retrovirology* 2015; **12**: 79 [PMID: 26377088 DOI: 10.1186/s12977-015-0205-1]
69 **Pomerantz RJ**, Trono D, Feinberg MB, Baltimore D. Cells nonproductively infected with HIV-1 exhibit an aberrant pattern of viral RNA expression: a molecular model for latency. *Cell* 1990; **61**: 1271-1276 [PMID: 2364429 DOI: 10.1016/0092-8674(90)90691-7]
70 **Lassen KG**, Bailey JR, Siliciano RF. Analysis of human immunodeficiency virus type 1 transcriptional elongation in resting CD4+ T cells in vivo. *J Virol* 2004; **78**: 9105-9114 [PMID: 15308706 DOI: 10.1128/jvi.78.17.9105-9114.2004]
71 **Adams M**, Sharmeen L, Kimpton J, Romeo JM, Garcia JV, Peterlin BM, Groudine M, Emerman M. Cellular latency in human immunodeficiency virus-infected individuals with high CD4 levels can be detected by the presence of promoter-proximal transcripts. *Proc Natl Acad Sci USA* 1994; **91**: 3862-3866 [PMID: 8171003 DOI: 10.1073/pnas.91.9.3862]
72 **Hilldorfer BB**, Cillo AR, Besson GJ, Bedison MA, Mellors JW. New tools for quantifying HIV-1 reservoirs: plasma RNA single copy assays and beyond. *Curr HIV/AIDS Rep* 2012; **9**: 91-100 [PMID: 22215419 DOI: 10.1007/s11904-011-0104-6]
73 **Bruner KM**, Hosmane NN, Siliciano RF. Towards an HIV-1 cure: measuring the latent reservoir. *Trends Microbiol* 2015; **23**: 192-203 [PMID: 25747663 DOI: 10.1016/j.tim.2015.01.013]
74 **Ocwieja KE**, Sherrill-Mix S, Mukherjee R, Custers-Allen R, David P, Brown M, Wang S, Link DR, Olson J, Travers K, Schadt E, Bushman FD. Dynamic regulation of HIV-1 mRNA populations analyzed by single-molecule enrichment and long-read sequencing. *Nucleic Acids Res* 2012; **40**: 10345-10355 [PMID: 22923523 DOI: 10.1093/nar/gks753]
75 **Chun TW**, Justement JS, Lempicki RA, Yang J, Dennis G, Hallahan CW, Sanford C, Pandya P, Liu S, McLaughlin M, Ehler LA, Moir S, Fauci AS. Gene expression and viral prodution in latently infected, resting CD4+ T cells in viremic versus aviremic HIV-infected individuals. *Proc Natl Acad Sci USA* 2003; **100**: 1908-1913 [PMID: 12552096 DOI: 10.1073/pnas.0437640100]
76 **Evans VA**, Kumar N, Filali A, Procopio FA, Yegorov O, Goulet JP, Saleh S, Haddad EK, da Fonseca Pereira C, Ellenberg PC, Sekaly RP, Cameron PU, Lewin SR. Myeloid dendritic cells induce HIV-1 latency in non-proliferating CD4+ T cells. *PLoS Pathog* 2013; **9**: e1003799 [PMID: 24339779 DOI: 10.1371/journal.ppat.1003799]
77 **Du Chéné I**, Basyuk E, Lin YL, Triboulet R, Knezevich A, Chable-Bessia C, Mettling C, Baillat V, Reynes J, Corbeau P, Bertrand E, Marcello A, Emiliani S, Kiernan R, Benkirane M. Suv39H1 and HP1gamma are responsible for chromatin-mediated HIV-1 transcriptional silencing and post-integration latency. *EMBO J* 2007; **26**: 424-435 [PMID: 17245432 DOI: 10.1038/sj.emboj.7601517]
78 **Bednarik DP**, Mosca JD, Raj NB. Methylation as a modulator of expression of human immunodeficiency virus. *J Virol* 1987; **61**: 1253-1257 [PMID: 3469417]
79 **Van Lint C**, Emiliani S, Ott M, Verdin E. Transcriptional activation and chromatin remodeling of the HIV-1 promoter in response to histone acetylation. *EMBO J* 1996; **15**: 1112-1120 [PMID: 8605881]
80 **Duh EJ**, Maury WJ, Folks TM, Fauci AS, Rabson AB. Tumor necrosis factor alpha activates human immunodeficiency virus type 1 through induction of nuclear factor binding to the NF-kappa B

sites in the long terminal repeat. *Proc Natl Acad Sci USA* 1989; **86**: 5974-5978 [PMID: 2762307 DOI: 10.1073/pnas.86.15.5974]
81 **He G**, Margolis DM. Counterregulation of chromatin deacetylation and histone deacetylase occupancy at the integrated promoter of human immunodeficiency virus type 1 (HIV-1) by the HIV-1 repressor YY1 and HIV-1 activator Tat. *Mol Cell Biol* 2002; **22**: 2965-2973 [PMID: 11940654 DOI: 10.1128/mcb.22.9.2965-2973.2002]
82 **Herrmann CH**, Rice AP. Lentivirus Tat proteins specifically associate with a cellular protein kinase, TAK, that hyperphosphorylates the carboxyl-terminal domain of the large subunit of RNA polymerase II: candidate for a Tat cofactor. *J Virol* 1995; **69**: 1612-1620 [PMID: 7853496]
83 **Lassen KG**, Ramyar KX, Bailey JR, Zhou Y, Siliciano RF. Nuclear retention of multiply spliced HIV-1 RNA in resting CD4+ T cells. *PLoS Pathog* 2006; **2**: e68 [PMID: 16839202 DOI: 10.1371/journal.ppat.0020068]
84 **Omoto S**, Fujii YR. Regulation of human immunodeficiency virus 1 transcription by nef microRNA. *J Gen Virol* 2005; **86**: 751-755 [PMID: 15722536 DOI: 10.1099/vir.0.80449-0]
85 **Huang J**, Wang F, Argyris E, Chen K, Liang Z, Tian H, Huang W, Squires K, Verlinghieri G, Zhang H. Cellular microRNAs contribute to HIV-1 latency in resting primary CD4+ T lymphocytes. *Nat Med* 2007; **13**: 1241-1247 [PMID: 17906637 DOI: 10.1038/nm1639]
86 **Nabel G**, Baltimore D. An inducible transcription factor activates expression of human immunodeficiency virus in T cells. *Nature* 1987; **326**: 711-713 [PMID: 3031512 DOI: 10.1038/326711a0]
87 **Osborn L**, Kunkel S, Nabel GJ. Tumor necrosis factor alpha and interleukin 1 stimulate the human immunodeficiency virus enhancer by activation of the nuclear factor kappa B. *Proc Natl Acad Sci USA* 1989; **86**: 2336-2340 [PMID: 2494664 DOI: 10.1073/pnas.86.7.2336]
88 **Matalon S**, Rasmussen TA, Dinarello CA. Histone deacetylase inhibitors for purging HIV-1 from the latent reservoir. *Mol Med* 2011; **17**: 466-472 [PMID: 21424110 DOI: 10.2119/molmed.2011.00076]
89 **Chueh AC**, Tse JW, Tögel L, Mariadason JM. Mechanisms of histone deacetylase inhibitor-regulated gene expression in cancer cells. *Antioxid Redox Signal* 2015; **23**: 66-84 [PMID: 24512308 DOI: 10.1089/ars.2014.5863]
90 **Bolden JE**, Shi W, Jankowski K, Kan CY, Cluse L, Martin BP, MacKenzie KL, Smyth GK, Johnstone RW. HDAC inhibitors induce tumor-cell-selective pro-apoptotic transcriptional responses. *Cell Death Dis* 2013; **4**: e519 [PMID: 23449455 DOI: 10.1038/cddis.2013.9]
91 **Papeleu P**, Vanhaecke T, Elaut G, Vinken M, Henkens T, Snykers S, Rogiers V. Differential effects of histone deacetylase inhibitors in tumor and normal cells-what is the toxicological relevance? *Crit Rev Toxicol* 2005; **35**: 363-378 [PMID: 15989141 DOI: 10.1080/10408440590935639]
92 **Chang J**, Varghese DS, Gillam MC, Peyton M, Modi B, Schiltz RL, Girard L, Martinez ED. Differential response of cancer cells to HDAC inhibitors trichostatin A and depsipeptide. *Br J Cancer* 2012; **106**: 116-125 [PMID: 22158273 DOI: 10.1038/bjc.2011.532]
93 **Ungerstedt JS**, Sowa Y, Xu WS, Shao Y, Dokmanovic M, Perez G, Ngo L, Holmgren A, Jiang X, Marks PA. Role of thioredoxin in the response of normal and transformed cells to histone deacetylase inhibitors. *Proc Natl Acad Sci USA* 2005; **102**: 673-678 [PMID: 15637150 DOI: 10.1073/pnas.0408732102]
94 **Mann BS**, Johnson JR, Cohen MH, Justice R, Pazdur R. FDA approval summary: vorinostat for treatment of advanced primary cutaneous T-cell lymphoma. *Oncologist* 2007; **12**: 1247-1252 [PMID: 17962618 DOI: 10.1634/theoncologist.12-10-1247]
95 **Jiang G**, Espeseth A, Hazuda DJ, Margolis DM. c-Myc and Sp1 contribute to proviral latency by recruiting histone deacetylase 1 to the human immunodeficiency virus type 1 promoter. *J Virol* 2007; **81**: 10914-10923 [PMID: 17670825 DOI: 10.1128/jvi.01208-07]
96 **Beliakova-Bethell N**, Zhang JX, Singhania A, Lee V, Terry VH, Richman DD, Spina CA, Woelk CH. Suberoylanilide hydroxamic acid induces limited changes in the transcriptome of primary CD4(+) T cells. *AIDS* 2013; **27**: 29-37 [PMID: 23221426 DOI: 10.1097/QAD.0b013e32835b3e26]
97 **Sung TL**, Rice AP. Effects of prostratin on Cyclin T1/P-TEFb function and the gene expression profile in primary resting CD4+ T cells. *Retrovirology* 2006; **3**: 66 [PMID: 17014716 DOI: 10.1186/1742-4690-3-66]
98 **Banerjee C**, Archin N, Michaels D, Belkina AC, Denis GV, Bradner J, Sebastiani P, Margolis DM, Montano M. BET bromodomain inhibition as a novel strategy for reactivation of HIV-1. *J Leukoc Biol* 2012; **92**: 1147-1154 [PMID: 22802445 DOI: 10.1189/jlb.0312165]
99 **White CH**, Johnston HE, Moesker B, Manousopoulou A, Margolis DM, Richman DD, Spina CA, Garbis SD, Woelk CH, Beliakova-Bethell N. Mixed effects of suberoylanilide hydroxamic acid (SAHA) on the host transcriptome and proteome and their implications for HIV reactivation from latency. *Antiviral Res* 2015; **123**: 78-85 [PMID: 26343910 DOI: 10.1016/j.antiviral.2015.09.002]
100 **Reardon B**, Beliakova-Bethell N, Spina CA, Singhania A, Margolis DM, Richman DR, Woelk CH. Dose-responsive gene expression in suberoylanilide hydroxamic acid-treated resting CD4+ T cells. *AIDS* 2015; **29**: 2235-2244 [PMID: 26258524 DOI: 10.1097/QAD.0000000000000839]
101 **Grégoire IP**, Richetta C, Meyniel-Schicklin L, Borel S, Pradezynski F, Diaz O, Deloire A, Azocar O, Baguet J, Le Breton M, Mangeot PE, Navratil V, Joubert PE, Flacher M, Vidalain PO, André P, Lotteau V, Biard-Piechaczyk M, Rabourdin-Combe C, Faure M. IRGM is a common target of RNA viruses that subvert the autophagy network. *PLoS Pathog* 2011; **7**: e1002422 [PMID: 22174682 DOI: 10.1371/journal.ppat.1002422]
102 **Agostini I**, Popov S, Li J, Dubrovsky L, Hao T, Bukrinsky M. Heat-shock protein 70 can replace viral protein R of HIV-1 during nuclear import of the viral preintegration complex. *Exp Cell Res* 2000; **259**: 398-403 [PMID: 10964507 DOI: 10.1006/excr.2000.4992]
103 **O'Keeffe B**, Fong Y, Chen D, Zhou S, Zhou Q. Requirement for a kinase-specific chaperone pathway in the production of a Cdk9/cyclin T1 heterodimer responsible for P-TEFb-mediated tat stimulation of HIV-1 transcription. *J Biol Chem* 2000; **275**: 279-287 [PMID: 10617616 DOI: 10.1074/jbc.275.1.279]
104 **Sakane N**, Kwon HS, Pagans S, Kaehlcke K, Mizusawa Y, Kamada M, Lassen KG, Chan J, Greene WC, Schnoelzer M, Ott M. Activation of HIV transcription by the viral Tat protein requires a demethylation step mediated by lysine-specific demethylase 1 (LSD1/KDM1). *PLoS Pathog* 2011; **7**: e1002184 [PMID: 21876670 DOI: 10.1371/journal.ppat.1002184]
105 **Tetsuka T**, Uranishi H, Imai H, Ono T, Sonta S, Takahashi N, Asamitsu K, Okamoto T. Inhibition of nuclear factor-kappaB-mediated transcription by association with the amino-terminal enhancer of split, a Groucho-related protein lacking WD40 repeats. *J Biol Chem* 2000; **275**: 4383-4390 [PMID: 10660609 DOI: 10.1074/jbc.275.6.4383]
106 **Rafati H**, Parra M, Hakre S, Moshkin Y, Verdin E, Mahmoudi T. Repressive LTR nucleosome positioning by the BAF complex is required for HIV latency. *PLoS Biol* 2011; **9**: e1001206 [PMID: 22140357 DOI: 10.1371/journal.pbio.1001206]
107 **Takahashi Y**, Tanaka Y, Yamashita A, Koyanagi Y, Nakamura M, Yamamoto N. OX40 stimulation by gp34/OX40 ligand enhances productive human immunodeficiency virus type 1 infection. *J Virol* 2001; **75**: 6748-6757 [PMID: 11435553 DOI: 10.1128/jvi.75.15.6748-6757.2001]
108 **Chang TL**, Vargas J, DelPortillo A, Klotman ME. Dual role of alpha-defensin-1 in anti-HIV-1 innate immunity. *J Clin Invest* 2005; **115**: 765-773 [PMID: 15719067 DOI: 10.1172/JCI200521948]
109 **Pagans S**, Pedal A, North BJ, Kaehlcke K, Marshall BL, Dorr A, Hetzer-Egger C, Henklein P, Frye R, McBurney MW, Hruby H, Jung M, Verdin E, Ott M. SIRT1 regulates HIV transcription via Tat deacetylation. *PLoS Biol* 2005; **3**: e41 [PMID: 15719057 DOI: 10.1371/journal.pbio.0030041]
110 **Xie B**, Invernizzi CF, Richard S, Wainberg MA. Arginine methylation of the human immunodeficiency virus type 1 Tat protein by PRMT6 negatively affects Tat Interactions with both cyclin T1 and the Tat transactivation region. *J Virol* 2007; **81**: 4226-4234 [PMID: 17267505 DOI: 10.1128/JVI.01888-06]

111 **Van Duyne R**, Easley R, Wu W, Berro R, Pedati C, Klase Z, Kehn-Hall K, Flynn EK, Symer DE, Kashanchi F. Lysine methylation of HIV-1 Tat regulates transcriptional activity of the viral LTR. *Retrovirology* 2008; **5**: 40 [PMID: 18498648 DOI: 10.1186/1742-4690-5-40]

112 **Posada R**, Pettoello-Mantovani M, Sieweke M, Graf T, Goldstein H. Suppression of HIV type 1 replication by a dominant-negative Ets-1 mutant. *AIDS Res Hum Retroviruses* 2000; **16**: 1981-1989 [PMID: 11153081 DOI: 10.1089/088922200750054710]

113 **Sheridan PL**, Sheline CT, Cannon K, Voz ML, Pazin MJ, Kadonaga JT, Jones KA. Activation of the HIV-1 enhancer by the LEF-1 HMG protein on nucleosome-assembled DNA in vitro. *Genes Dev* 1995; **9**: 2090-2104 [PMID: 7657162 DOI: 10.1101/gad.9.17.2090]

114 **Mukerjee R**, Sawaya BE, Khalili K, Amini S. Association of p65 and C/EBPbeta with HIV-1 LTR modulates transcription of the viral promoter. *J Cell Biochem* 2007; **100**: 1210-1216 [PMID: 17031851 DOI: 10.1002/jcb.21109]

115 **Friedman J**, Cho WK, Chu CK, Keedy KS, Archin NM, Margolis DM, Karn J. Epigenetic silencing of HIV-1 by the histone H3 lysine 27 methyltransferase enhancer of Zeste 2. *J Virol* 2011; **85**: 9078-9089 [PMID: 21715480 DOI: 10.1128/JVI.00836-11]

116 **Margolis DM**, Somasundaran M, Green MR. Human transcription factor YY1 represses human immunodeficiency virus type 1 transcription and virion production. *J Virol* 1994; **68**: 905-910 [PMID: 8289393]

117 **Boehm D**, Calvanese V, Dar RD, Xing S, Schroeder S, Martins L, Aull K, Li PC, Planelles V, Bradner JE, Zhou MM, Siliciano RF, Weinberger L, Verdin E, Ott M. BET bromodomain-targeting compounds reactivate HIV from latency via a Tat-independent mechanism. *Cell Cycle* 2013; **12**: 452-462 [PMID: 23255218 DOI: 10.4161/cc.23309]

118 **Eilebrecht S**, Wilhelm E, Benecke BJ, Bell B, Benecke AG. HMGA1 directly interacts with TAR to modulate basal and Tat-dependent HIV transcription. *RNA Biol* 2013; **10**: 436-444 [PMID: 23392246 DOI: 10.4161/rna.23686]

119 **Eilebrecht S**, Le Douce V, Riclet R, Targat B, Hallay H, Van Driessche B, Schwartz C, Robette G, Van Lint C, Rohr O, Benecke AG. HMGA1 recruits CTIP2-repressed P-TEFb to the HIV-1 and cellular target promoters. *Nucleic Acids Res* 2014; **42**: 4962-4971 [PMID: 24623795 DOI: 10.1093/nar/gku168]

120 **Hong JW**, Wu LC. ZAS3 represses NFκB-dependent transcription by direct competition for DNA binding. *BMB Rep* 2010; **43**: 807-812 [PMID: 21189157 DOI: 10.5483/BMBRep.2010.43.12.807]

121 **Ryckman C**, Robichaud GA, Roy J, Cantin R, Tremblay MJ, Tessier PA. HIV-1 transcription and virus production are both accentuated by the proinflammatory myeloid-related proteins in human CD4+ T lymphocytes. *J Immunol* 2002; **169**: 3307-3313 [PMID: 12218151 DOI: 10.4049/jimmunol.169.6.3307]

122 **Contreras X**, Schweneker M, Chen C-S, McCune JM, Deeks SG, Martin J, Peterlin BM. Suberoylanilide hydroxamic acid reactivates HIV from latently infected cells. *J Biol Chem* 2009; **284**: 6782-6789 [PMID: 19136668 DOI: 10.1074/jbc.M807898200]

123 **Rossio JL**, Esser MT, Suryanarayana K, Schneider DK, Bess JW, Vasquez GM, Wiltrout TA, Chertova E, Grimes MK, Sattentau Q, Arthur LO, Henderson LE, Lifson JD. Inactivation of human immunodeficiency virus type 1 infectivity with preservation of conformational and functional integrity of virion surface proteins. *J Virol* 1998; **72**: 7992-8001 [PMID: 9733838]

P- Reviewer: Beloukas AI, García-Elorriaga G, Kelesidis T, Pandey VN
S- Editor: Qiu S **L- Editor**: A **E- Editor**: Jiao XK

World Journal of Clinical Infectious Diseases

Submit a Manuscript: http://www.wjgnet.com/esps/
Help Desk: http://www.wjgnet.com/esps/helpdesk.aspx
DOI: 10.5495/wjcid.v6.i2.22

World J Clin Infect Dis 2016 May 25; 6(2): 22-27
ISSN 2220-3176 (online)

MINIREVIEWS

Osmolyte transport in *Staphylococcus aureus* and the role in pathogenesis

William R Schwan, Keith J Wetzel

William R Schwan, Keith J Wetzel, Department of Microbiology, University of Wisconsin-La Crosse, La Crosse, WI 54601, United States

Author contributions: All the authors contribute to the manuscript.

Supported by NIH grant, No. 1R15AI47801-01A.

Conflict-of-interest statement: Authors declare no conflicts of interest.

Correspondence to: William R Schwan, MD, Department of Microbiology, University of Wisconsin-La Crosse, 1725 State St, La Crosse, WI 54601, United States. wschwan@uwlax.edu
Telephone: +1-608-7856980

Received: October 23, 2015
Peer-review started: October 27, 2015
First decision: February 2, 2016
Revised: February 18, 2016
Accepted: April 7, 2016
Article in press: April 11, 2016
Published online: May 25, 2016

Abstract

Osmolyte transport is a pivotal part of bacterial life, particularly in high salt environments. Several low and high affinity osmolyte transport systems have been identified in various bacterial species. A lot of research has centered on characterizing the osmolyte transport systems of Gram-negative bacteria, but less has been done to characterize the same transport systems in Gram-positive bacteria. This review will focus on the previous work that has been done to understand the osmolyte transport systems in the species *Staphylococcus aureus* and how these transporters may serve dual functions in allowing the bacteria to survive and grow in a variety of environments, including on the surface or within humans or other animals.

Key words: *PutP*; *OpuD*; *Staphylococcus aureus*; Proline transport; Osmolyte

Core tip: *Staphylococcus aureus* (*S. aureus*) is the number one cause of skin and soft tissue infections. In the United States, *S. aureus* is usually the number one hospital-acquired pathogen. The skin and urinary tract organs are high osmotic stress environments. Osmolyte transport is essential for *S. aureus* survival in different environmental niches, such as within human skin abscesses or the human urinary tract.

Schwan WR, Wetzel KJ. Osmolyte transport in *Staphylococcus aureus* and the role in pathogenesis. *World J Clin Infect Dis* 2016; 6(2): 22-27 Available from: URL: http://www.wjgnet.com/2220-3176/full/v6/i2/22.htm DOI: http://dx.doi.org/10.5495/wjcid.v6.i2.22

INTRODUCTION

A well conserved, evolutionary strategy used by many organisms to adapt to high osmotic conditions is the transport of organic compounds, called compatible solutes[1]. These compatible solutes serve as cytoplasmic solutes that balance water relations, without interfering with normal cytoplasmic activities, within cells grown in high salt environments. Examination of the transport

systems in *Staphylococcus aureus (S. aureus)* may provide insight into how proline and glycine betaine may be transported into Gram-positive bacteria.

GENERAL OSMOLYTE TRANSPORT FEATURES IN *S. AUREUS*

Although osmolyte transport is best described in *E. coli*[1-3], there are also compatible solute transport systems in *S. aureus* to adapt to high salt environments[4]. Studies have shown that *S. aureus* cells grown in very high salt environments had increased intracellular levels of proline and glycine betaine[5-11]. Other intracellular molecules that also increased in high NaCl environments were choline, proline betaine, taurine, and glutamic acid[6,7,12]. Of these accumulated solutes, proline and glycine betaine were the most effective osmoprotectants of *S. aureus*, since *S. aureus* growth was observed when these solutes were excluded from defined high osmotic media[6,8,12].

Identification of genes that encode transport proteins and their importance for the survival of *S. aureus* coincides with previous observations that *S. aureus* requires several amino acids as a source of carbon and nitrogen[4]. Of these essential amino acids, proline and other amino acids are not synthesized by *S. aureus*[4,13,14]. The accumulation of most of the proline in *S. aureus* occurs because of proline transport proteins.

Although prior research performed using other Gram-positive bacteria may not have specifically addressed proline transport, it does help in uncovering commonly conserved mechanisms of compatible solute transport in *S. aureus*. Several studies that have examined compatible solutes accumulation in *S. aureus* grown at high osmotic environments showed increased intracellular levels of proline, aminobutyric acid, glutamic acid, choline, taurine, and glycine betaine[5-7,15,16]. Of these compatible solutes, only glutamic acid is synthesized by *S. aureus*, whereas the other compatible solutes have to be imported from the external environment[5,7,8,17-19]. To substantiate the osmoprotective importance of these transported compatible solutes, the growth rates of *S. aureus* grown in defined high osmotic media was observed to increase when supplemented with either proline or glycine betaine[8]. Although *S. aureus* normally possess relatively large concentrations of glycine betaine and potassium ions, compatible solute transport is believed to aid in creating high intracellular pressure that enables *S. aureus* to survive in high osmotic environments[15].

SPECIFIC PROLINE TRANSPORT SYSTEMS IN *S. AUREUS*

Initial proline uptake research using whole cell assays on *S. aureus* has shown the presence of at least two proline transport systems[10,17,20]: Both a low- and high-affinity system. These systems may be similar to the OpuE and OpuD transport systems found in *B. subtilis*[21,22] and they share properties with the PutP and ProP systems of *E. coli*[1]. They are both sodium-dependent transporters, since gramicidin D and monensin, which collapse Na^+ gradients, inhibit proline transport in both systems[10]. Proline transport in either system showed low susceptibility to inhibition by glycolysis and ATP formation by a combination of NaF and sodium iodoacetate or sodium arsenate, respectively. Lastly, alterations of pH from 5.5 to 8.5 had little effect on the transport rates of proline[10].

In *S. aureus*, proline transport kinetics is hard to interpret because of strain differences and the calculation setups used to determine the K_m and V_{max} values reported, one based on per mg protein and the other per mg dry weight. Reports have shown that the high-affinity proline transport system in *S. aureus* had a K_m ranging from 1.7 to 7.0 mol/L, with a V_{max} ranging from 1.1 nmol/min per milligram dry weight to 10 nmol/min per milligram protein[10,17]. Though these numbers are not directly comparative, they do give us a relative range of activity for this system, which correlates to a previously observed K_m value of 3.5 mol/L for proline uptake with vesicles prepared from *S. aureus* grown in a low-osmolarity medium[23] and K_m values of the PutP system in *E. coli*[1,17,24-26]. Moreover, like the PutP system of *E. coli*[1], the high-affinity proline transport system in *S. aureus* is specific for the transport of proline and it's activity increases when proline deprivation is encountered, suggesting that this system may also be involved in scavenging low concentrations of proline from the environment[10]. Further proof of the relatedness of these systems can be seen from the complementation of a genetic defect in proline transport within *E. coli* by the high-affinity proline transport system of *S. aureus*[27]. At the structural level, the PutP homolog of *S. aureus* shows a sodium-binding motif, the same ten conserved amino acids found in all other members of the sodium/solute symporters[28], and the predicted PutP protein of *S. aureus*[29] shares considerable similarity with the PutP protein of *E. coli*[1]. Although many similarities exist between the high-affinity proline transport systems in *S. aureus* and *E. coli*, major differences between these systems include: The concentration of NaCl appears to have no effect on proline transport in *S. aureus*[8,17]; the *S. aureus putP* gene is activated by high concentrations of osmolytes in the environment[30], whereas the *E. coli putP* gene is not[1,25,29]; and the *S. aureus putP* gene is regulated by SigB[30], which is similar to the regulation shown for *opuE* in *B. subtilis*[21]. Although PutP has a sodium binding motif and has homology with sodium/solute symporters, the concentration of NaCl does not affect proline transport[7,17], It is possible that when *S. aureus* is grown in an environment with a low sodium concentration that PutP behaves like other bacterial high affinity proline transporters that are driven by a sodium motive force. On the other hand, *S. aureus* grown in a high sodium environment may cause the PutP protein to use a proton motive force instead of a sodium motive

Table 1 Distribution of proline and glycine betaine transport genes in some sequenced

	S. aureus strains			
Gene	N315	MW2	COL	Mu50
putP	SA1718	MW1843	SACOL1963	SAV1902
putP	SA0531	MW0528	SACOL0620	SAV0573
opuD	SA1183	MW1236	Yes (2)[2]	SAV1349[4]
opuD1	-[1]	-[1]	SACOL1384	ND[3]
opuD2	-[1]	-[1]	SACOL2176	ND[3]
opuCA	SA2237	MW2372	ND[3]	SAV2448
opuCB	SA2236	MW2371	ND[3]	SAV2447
opuCC	SA2235	MW2370	ND[3]	SAV2446
opuCD	SA2234	MW2369	ND[3]	SAV2445

[1]Does not possess; [2]Multiple *opuD* genes in this species; [3]Not determined; [4]The gene appears to be fragmented into two pieces.

force to bring proline into the cell.

The low-affinity proline transport system of *S. aureus* also has similarities to the low-affinity proline transport system (ProP) of *E. coli*. For proline transport, the K_m value of *S. aureus* ATCC 12600 (K_m of 420 mol/L and V_{max} of 110 nmol/min per milligram protein) is similar to the K_m value of ProP in *E. coli* (approximately 300 mol/L)[17]. For *S. aureus* (K_m of 132 mol/L and V_{max} of 22 nmol/min per milligram dry weight), a greater difference in the K_m values for the low-affinity proline transport system can be seen between strains as compared to the difference in K_m values for the high-affinity system. Again, the K_m and V_{max} values from the ProP system of *E. coli* fit within the overall range found for *S. aureus*[1,31-33], but strain variation along with calculation setup differences may again be the cause of these divergent numbers. Excluding the differences of the K_m and V_{max} values between strains, the low-affinity proline transport systems of different *S. aureus* strains possess identical characteristics[10,17]. Many of these characteristics are similar to the regulatory and functional properties of the ProP system of *E. coli*[34] (*i.e.,* both of these systems transport proline and are stimulated by increasing osmolarity produced by either ionic or nonionic solutes)[17].

DIFFERENCES IN THE *S. AUREUS* OSMOLYTE TRANSPORT SYSTEMS COMPARED TO OTHER BACTERIA

Though these systems are similar, there are some major differences between the Gram-negative and Gram-positive low-affinity proline transport systems. One major difference is that the low-affinity proline transport systems in *S. aureus* are optimally activated at NaCl concentrations ranging from 0.75 to 1.0 mol/L[17,35], whereas the low-affinity proline transport systems in *E. coli* are inhibited by NaCl concentrations greater than 0.2 to 0.3 mol/L[29,36]. Other major differences include glycine betaine transport activity by the low-affinity proline transport system has not been conclusively established and there conflicting opinions and data presented for the glycine betaine transport activity for the low-affinity system[9,17,18,20,37]. In part, the previous lack of any low-affinity system mutants in those studies complicated the examination of glycine betaine transport activities. Since glycine betaine accumulation has been linked to proline transporters in Gram-negative bacteria[1] and *S. aureus* has been shown to transport glycine betaine from the external environment[38], this suggests that an additional glycine betaine transporter that is osmotically stimulated may be present in *S. aureus*. Moreover, *S. aureus* cells shocked with 0.5 mol/L NaCl in the presence and absence of chloramphenicol (100 g/mL) showed identical levels of transported proline, suggesting that new protein synthesis is not necessary for rapid proline uptake and that osmotic shock activates a pre-existing proline transport system[10].

BIOINFORMATIC TOOLS TO IDENTIFY OSMOLYTE TRANSPORT SYSTEMS IN *S. AUREUS*

Sequencing of several *S. aureus* genomes has provided a wealth of information on the existence of several putative osmolyte transport systems in *S. aureus*[14,39,40]. All of the strains appear to have a conserved *putP* gene for high affinity transport of proline, although there appears to be homologs for both a *proP* gene[1] and *opuD* gene[21,35] (Table 1). Additional analyses have shown that the *opuD* gene (encoding a low affinity proline transporter) is activated under osmotic stress conditions and OpuD transports proline under low affinity growth conditions[35]. Furthermore, a mutation in the *S. aureus proP* gene also causes lower proline transport in media with high concentrations of proline (Schwan WR unpublished data).

This is the first instance of both the ProP and OpuD low affinity proline/glycine betaine transport homologs being identified in one species and suggests the importance that proline transport must have in the survival of *S. aureus* cells in a variety of environments. Furthermore, the *opuC* system, which putatively transports glycine betaine/carnitine/choline, has also been observed. Together, the bioinformatic comparisons have uncovered some very interesting genomic features in *S. aureus* centered on osmolyte transport. A summary of the four osmolyte transport systems in *S. aureus* tied to proline transport and other known solutes is noted in Figure 1.

OSMOLYTE TRANSPORT TIED TO *S. AUREUS* SURVIVAL IN HUMANS AND MICE

The rationale of investigating proline and glycine betaine transport in *S. aureus* is not purely academic. In planktonic *S. aureus*, the glycine betaine level is high,

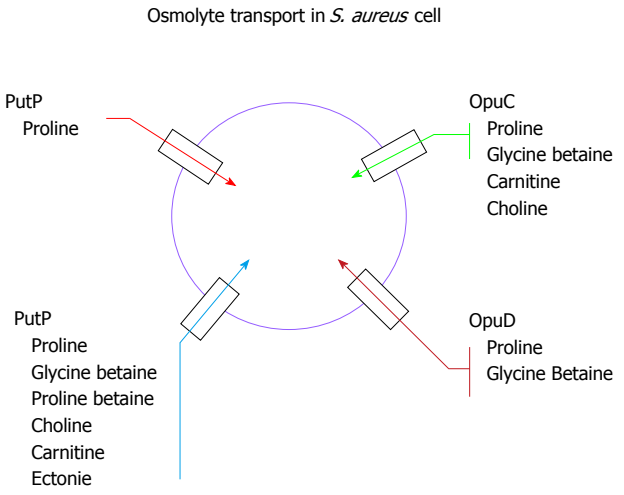

Figure 1 The four prominent osmolyte transport systems in *Staphylococcus aureus* tied to proline transport as well as other solutes.

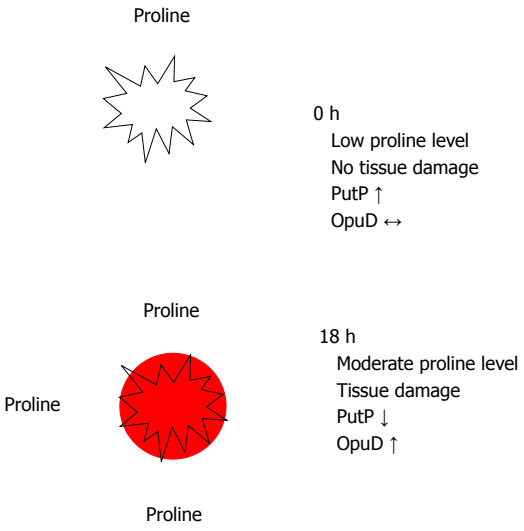

Figure 2 Model for the roles of proline transporters in *Staphylococcus aureus* pathogenesis within a murine abscess.

but lower in *S. aureus* found in biofilms[41]. Glycine betaine is the most effective osmoprotectant. To achieve the high glycine betaine level, an active glycine betaine transporter would need to be functioning in the planktonic *S. aureus* cells that are immersed in an environment of high osmotic stress, like the human skin.

Indirect effects on *S. aureus* survival have been tied to osmolyte transport systems. Defects in the cell wall caused by a *femAB* mutation caused an upregulation of *opuC* (glycine betaine/carnitine/choline transporter) and downregulation of *opuD* to compensate for the defect[42]. YhcSR encodes a two-component signal transduction system that is required for *S. aureus* survival. This two-component regulatory system regulates transcription of the *opuCABCD* operons affecting proline and glycine betaine levels in *S. aureus*[43]. One study examining daptomycin resistance revealed an accumulation of glycine betaine within *S. aureus* cells that was coupled with upregulation of the *cudT* (choline transporter) gene, a beta choline dehydrogenase gene, a *gbsA* gene (glycine betaine aldehyde dehydrogenase), an *opuD2* gene, and the *proP* gene[44]. Uptake of choline is needed to produce glycine betaine internally, the best osmoprotectant[19].

More directly, a transposon mutation in the gene for the high affinity (PutP) proline transport system of *S. aureus* rendered the bacteria less able to survive in several animal infection models[45-47]. Within cardiac vegetations, the viable *S. aureus* count was 1-3 logs lower than the wild-type parent strain[45]. Transcription of *putP* was shown to increase 105-fold shortly after *S. aureus* infection of murine kidneys[30]. In *S. aureus* infected murine bladders, spleen and livers, *putP* transcription was also elevated very quickly and then dropped markedly as the infection progressed. Proline levels in livers and spleens are very low[47] and the levels are likely low in the other organs (*e.g.*, bladder and kidney), but through tissue damage by staphylococcal toxins, the concentration of proline may increase substantially and in turn shut off transcription of the high affinity proline transport gene.

Conversely, transcription of the low affinity proline transport gene *opuD* was shown to be the highest after 4 h post-infection in murine bladders and 18 h post-infection in murine thigh abscesses[35]. Within murine bladders and kidneys, high osmotic conditions prevail. Initial observations demonstrated that at least one of the low-affinity proline transport systems of *S. aureus* was activated under moderate to high osmotic conditions[17], which has been subsequently confirmed[35].

Our model is that PutP is important in the early stages of an infection when proline concentrations are low, but OpuD expression is not as important (Figure 2). As the infection proceeds, tissue damage occurs, which releases free proline. By 18 h post-infection, the level of free proline is higher and OpuD becomes important at this stage of the infection.

These studies suggest that osmolyte transport systems may play essential roles in survival of *S. aureus* within humans or mice. Characterization of the proline and glycine betaine transport systems will provide us with experimental proof of the importance of these systems during growth in high osmotic conditions, how these systems are regulated, and will further our understanding of the significance of the proline/glycine betaine transport to the survival of *S. aureus in vivo*.

ACKNOWLEDGMENTS

I would like to thank the University of Wisconsin-La Crosse for grant support for my laboratory and also thank all of the undergraduate and graduate students whom I have mentored.

REFERENCES

1 **Wood JM**. Proline porters effect the utilization of proline as nutrient or osmoprotectant for bacteria. *J Membr Biol* 1988; **106**: 183-202

[PMID: 3072423 DOI: 10.1007/BF01872157]
2 **Csonka LN**. Physiological and genetic responses of bacteria to osmotic stress. *Microbiol Rev* 1989; **53**: 121-147 [PMID: 2651863]
3 **Fichman Y**, Gerdes SY, Kovács H, Szabados L, Zilberstein A, Csonka LN. Evolution of proline biosynthesis: enzymology, bioinformatics, genetics, and transcriptional regulation. *Biol Rev Camb Philos Soc* 2015; **90**: 1065-1099 [PMID: 25367752 DOI: 10.1111/brv.12146]
4 **Ruoff KL**. Algorithm for identification of aerobic Gram-positive cocci, p. 262-282. In Manual of clinical microbiology, 7th ed. Edited by Murray PR, Baron EJ, Pflaller MA, Tenover FC, Yolken RH, American Society for Microbiology, Washington, D.C, 1999
5 **Anderson CB**, Witter LD. Glutamine and proline accumulation by Staphylococcus aureus with reduction in water activity. *Appl Environ Microbiol* 1982; **43**: 1501-1503 [PMID: 7103493]
6 **Graham JE**, Wilkinson BJ. Staphylococcus aureus osmoregulation: roles for choline, glycine betaine, proline, and taurine. *J Bacteriol* 1992; **174**: 2711-2716 [PMID: 1556089]
7 **Koujima I**, Hayashi H, Tomochika K, Okabe A, Kanemasa Y. Adaptational change in proline and water content of Staphylococcus aureus after alteration of environmental salt concentration. *Appl Environ Microbiol* 1978; **35**: 467-470 [PMID: 637544]
8 **Miller KJ**, Zelt SC, Bae J-H. Glycine betaine and proline are the principle compatible solutes of Staphylococcus aureus. *Curr Micorobiol* 1991; **23**: 131-137 [DOI: 10.1007/BF02091971]
9 **Pourkomailian B**, Booth IR. Glycine betaine transport by Staphylococcus aureus: evidence for two transport systems and for their possible roles in osmoregulation. *J Gen Microbiol* 1992; **138**: 2515-2518 [PMID: 1487723 DOI: 10.1099/00221287-138-12-2515]
10 **Townsend DE**, Wilkinson BJ. Proline transport in Staphylococcus aureus: a high-affinity system and a low-affinity system involved in osmoregulation. *J Bacteriol* 1992; **174**: 2702-2710 [PMID: 1556088]
11 **Vijaranakul U**, Nadakavukaren MJ, Bayles DO, Wilkinson BJ, Jayaswal RK. Characterization of an NaCl-sensitive Staphylococcus aureus mutant and rescue of the NaCl-sensitive phenotype by glycine betaine but not by other compatible solutes. *Appl Environ Microbiol* 1997; **63**: 1889-1897 [PMID: 9143120]
12 **Amin US**, Lash TD, Wilkinson BJ. Proline betaine is a highly effective osmoprotectant for Staphylococcus aureus. *Arch Microbiol* 1995; **163**: 138-142 [PMID: 7710327 DOI: 10.1007/BF00381788]
13 **Iandolo JJ**, Worrell V, Groicher KH, Qian Y, Tian R, Kenton S, Dorman A, Ji H, Lin S, Loh P, Qi S, Zhu H, Roe BA. Comparative analysis of the genomes of the temperate bacteriophages phi 11, phi 12 and phi 13 of Staphylococcus aureus 8325. *Gene* 2002; **289**: 109-118 [PMID: 12036589 DOI: 10.1016/S0378-1119(02)00481-X]
14 **Kuroda M**, Ohta T, Uchiyama I, Baba T, Yuzawa H, Kobayashi I, Cui L, Oguchi A, Aoki K, Nagai Y, Lian J, Ito T, Kanamori M, Matsumaru H, Maruyama A, Murakami H, Hosoyama A, Mizutani-Ui Y, Takahashi NK, Sawano T, Inoue R, Kaito C, Sekimizu K, Hirakawa H, Kuhara S, Goto S, Yabuzaki J, Kanehisa M, Yamashita A, Oshima K, Furuya K, Yoshino C, Shiba T, Hattori M, Ogasawara N, Hayashi H, Hiramatsu K. Whole genome sequencing of meticillin-resistant Staphylococcus aureus. *Lancet* 2001; **357**: 1225-1240 [PMID: 11418146 DOI: 10.1016/S0140-6736(00)04403-2]
15 **Kunin CM**, Rudy J. Effect of NaCl-induced osmotic stress on intracellular concentrations of glycine betaine and potassium in Escherichia coli, Enterococcus faecalis, and staphylococci. *J Lab Clin Med* 1991; **118**: 217-224 [PMID: 1919294]
16 **Measures JC**. Role of amino acids in osmoregulation of non-halophilic bacteria. *Nature* 1975; **257**: 398-400 [PMID: 241020 DOI: 10.1038/257398a0]
17 **Bae JH**, Miller KJ. Identification of two proline transport systems in Staphylococcus aureus and their possible roles in osmoregulation. *Appl Environ Microbiol* 1992; **58**: 471-475 [PMID: 1610171]
18 **Bae JH**, Anderson SH, Miller KJ. Identification of a high-affinity glycine betaine transport system in Staphylococcus aureus. *Appl Environ Microbiol* 1993; **59**: 2734-2736 [PMID: 8368857]
19 **Kaenjak A**, Graham JE, Wilkinson BJ. Choline transport activity in Staphylococcus aureus induced by osmotic stress and low phosphate concentrations. *J Bacteriol* 1993; **175**: 2400-2406 [PMID: 8468298]
20 **Pourkomailian B**, Booth IR. Glycine betaine transport by Staphylococcus aureus: evidence for feedback regulation of the activity of the two transport systems. *Microbiology* 1994; **140** (Pt 11): 3131-3138 [PMID: 7812452 DOI: 10.1099/13500872-140-11-3131]
21 **Kappes RM**, Kempf B, Bremer E. Three transport systems for the osmoprotectant glycine betaine operate in Bacillus subtilis: characterization of OpuD. *J Bacteriol* 1996; **178**: 5071-5079 [PMID: 8752321]
22 **von Blohn C**, Kempf B, Kappes RM, Bremer E. Osmostress response in Bacillus subtilis: characterization of a proline uptake system (OpuE) regulated by high osmolarity and the alternative transcription factor sigma B. *Mol Microbiol* 1997; **25**: 175-187 [PMID: 11902719 DOI: 10.1046/j.1365-2958.1997.4441809.x]
23 **Short SA**, Kaback HR. Amino acid transport and staphylococcal membrane vesicles. *Ann N Y Acad Sci* 1974; **236**: 124-143 [PMID: 4371336 DOI: 10.1111/j.1749-6632.1974.tb41487.x]
24 **Bracher S**, Guérin K, Polyhach Y, Jeschke G, Dittmer S, Frey S, Böhm M, Jung H. Glu-311 in External Loop 4 of the Sodium/Proline Transporter PutP Is Crucial for External Gate Closure. *J Biol Chem* 2016; **291**: 4998-5008 [PMID: 26728461 DOI: 10.1074/jbc.M115.675306]
25 **Chen CC**, Wilson TH. Solubilization and functional reconstitution of the proline transport system of Escherichia coli. *J Biol Chem* 1986; **261**: 2599-2604 [PMID: 3512540]
26 **Myers RS**, Townsend D, Maloy S. Dissecting the molecular mechanism of ion-solute cotransport: substrate specificity mutations in the putP gene affect the kinetics of proline transport. *J Membr Biol* 1991; **121**: 201-214 [PMID: 1865487 DOI: 10.1007/BF01951554]
27 **Wengender PA**, Miller KJ. Identification of a PutP proline permease gene homolog from Staphylococcus aureus by expression cloning of the high-affinity proline transport system in Escherichia coli. *Appl Environ Microbiol* 1995; **61**: 252-259 [PMID: 7887605]
28 **Reizer J**, Reizer A, Saier MH. A functional superfamily of sodium/solute symporters. *Biochim Biophys Acta* 1994; **1197**: 133-166 [PMID: 8031825 DOI: 10.1016/0304-4157(94)90003-5]
29 **Grothe S**, Krogsrud RL, McClellan DJ, Milner JL, Wood JM. Proline transport and osmotic stress response in Escherichia coli K-12. *J Bacteriol* 1986; **166**: 253-259 [PMID: 3514577]
30 **Schwan WR**, Lehmann L, McCormick J. Transcriptional activation of the Staphylococcus aureus putP gene by low-proline-high osmotic conditions and during infection of murine and human tissues. *Infect Immun* 2006; **74**: 399-409 [PMID: 16368996 DOI: 10.1128/IAI.74.1.399-409.2006]
31 **Culham DE**, Lasby B, Marangoni AG, Milner JL, Steer BA, van Nues RW, Wood JM. Isolation and sequencing of Escherichia coli gene proP reveals unusual structural features of the osmoregulatory proline/betaine transporter, ProP. *J Mol Biol* 1993; **229**: 268-276 [PMID: 8421314 DOI: 10.1006/jmbi.1993.1030]
32 **MacMillan SV**, Alexander DA, Culham DE, Kunte HJ, Marshall EV, Rochon D, Wood JM. The ion coupling and organic substrate specificities of osmoregulatory transporter ProP in Escherichia coli. *Biochim Biophys Acta* 1999; **1420**: 30-44 [PMID: 10446288 DOI: 10.1016/S0005-2736(99)00085-1]
33 **Milner JL**, Grothe S, Wood JM. Proline porter II is activated by a hyperosmotic shift in both whole cells and membrane vesicles of Escherichia coli K12. *J Biol Chem* 1988; **263**: 14900-14905 [PMID: 3049595]
34 **Racher KI**, Culham DE, Wood JM. Requirements for osmosensing and osmotic activation of transporter ProP from Escherichia coli. *Biochemistry* 2001; **40**: 7324-7333 [PMID: 11401581 DOI: 10.1021/bi002331u]
35 **Wetzel KJ**, Bjorge D, Schwan WR. Mutational and transcriptional analyses of the Staphylococcus aureus low-affinity proline transporter OpuD during in vitro growth and infection of murine tissues. *FEMS Immunol Med Microbiol* 2011; **61**: 346-355 [PMID: 21231964 DOI: 10.1111/j.1574-695X.2011.00781.x]
36 **Faatz E**, Middendorf A, Bremer E. Cloned structural genes for the osmotically regulated binding-protein-dependent glycine betaine transport system (ProU) of Escherichia coli K-12. *Mol Microbiol*

1988; **2**: 265-279 [PMID: 2837616]
37 **Stimeling KW**, Graham JE, Kaenjak A, Wilkinson BJ. Evidence for feedback (trans) regulation of, and two systems for, glycine betaine transport by Staphylococcus aureus. *Microbiology* 1994; **140** (Pt 11): 3139-3144 [PMID: 7812453]
38 **Peddie BA**, Wong-She J, Randall K, Lever M, Chambers ST. Osmoprotective properties and accumulation of betaine analogues by Staphylococcus aureus. *FEMS Microbiol Lett* 1998; **160**: 25-30 [PMID: 9495008]
39 **Baba T**, Takeuchi F, Kuroda M, Yuzawa H, Aoki K, Oguchi A, Nagai Y, Iwama N, Asano K, Naimi T, Kuroda H, Cui L, Yamamoto K, Hiramatsu K. Genome and virulence determinants of high virulence community-acquired MRSA. *Lancet* 2002; **359**: 1819-1827 [PMID: 12044378 DOI: 10.1016/S0140-6736(02)08713-5]
40 **Gill SR**, Fouts DE, Archer GL, Mongodin EF, Deboy RT, Ravel J, Paulsen IT, Kolonay JF, Brinkac L, Beanan M, Dodson RJ, Daugherty SC, Madupu R, Angiuoli SV, Durkin AS, Haft DH, Vamathevan J, Khouri H, Utterback T, Lee C, Dimitrov G, Jiang L, Qin H, Weidman J, Tran K, Kang K, Hance IR, Nelson KE, Fraser CM. Insights on evolution of virulence and resistance from the complete genome analysis of an early methicillin-resistant Staphylococcus aureus strain and a biofilm-producing methicillin-resistant Staphylococcus epidermidis strain. *J Bacteriol* 2005; **187**: 2426-2438 [PMID: 15774886 DOI: 10.1128/JB.187.7.2426-2438.2005]
41 **Junka AF**, Deja S, Smutnicka D, Szymczyk P, Ziółkowski G, Bartoszewicz M, Młynarz P. Differences in metabolic profiles of planktonic and biofilm cells in Staphylococcus aureus - (1)H Nuclear Magnetic Resonance search for candidate biomarkers. *Acta Biochim Pol* 2013; **60**: 701-706 [PMID: 24432320]
42 **Hübscher J**, Jansen A, Kotte O, Schäfer J, Majcherczyk PA, Harris LG, Bierbaum G, Heinemann M, Berger-Bächi B. Living with an imperfect cell wall: compensation of femAB inactivation in Staphylococcus aureus. *BMC Genomics* 2007; **8**: 307 [PMID: 17784943 DOI: 10.1186/1471-2164-8-307]
43 **Yan M**, Hall JW, Yang J, Ji Y. The essential yhcSR two-component signal transduction system directly regulates the lac and opuCABCD operons of Staphylococcus aureus. *PLoS One* 2012; **7**: e50608 [PMID: 23226327 DOI: 10.1371/journal.pone.0050608]
44 **Song Y**, Rubio A, Jayaswal RK, Silverman JA, Wilkinson BJ. Additional routes to Staphylococcus aureus daptomycin resistance as revealed by comparative genome sequencing, transcriptional profiling, and phenotypic studies. *PLoS One* 2013; **8**: e58469 [PMID: 23554895 DOI: 10.1371/journal.pone.0058469]
45 **Bayer AS**, Coulter SN, Stover CK, Schwan WR. Impact of the high-affinity proline permease gene (putP) on the virulence of Staphylococcus aureus in experimental endocarditis. *Infect Immun* 1999; **67**: 740-744 [PMID: 9916085]
46 **Schwan WR**, Coulter SN, Ng EY, Langhorne MH, Ritchie HD, Brody LL, Westbrock-Wadman S, Bayer AS, Folger KR, Stover CK. Identification and characterization of the PutP proline permease that contributes to in vivo survival of Staphylococcus aureus in animal models. *Infect Immun* 1998; **66**: 567-572 [PMID: 9453610]
47 **Schwan WR**, Wetzel KJ, Gomez TS, Stiles MA, Beitlich BD, Grunwald S. Low-proline environments impair growth, proline transport and in vivo survival of Staphylococcus aureus strain-specific putP mutants. *Microbiology* 2004; **150**: 1055-1061 [PMID: 15073314 DOI: 10.1099/mic.0.26710-0]

P- Reviewer: García-Elorriaga G, Jung H, Krishnan T
S- Editor: Qiu S **L- Editor**: A **E- Editor**: Jiao XK

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