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 http://www.wjgnet.com

Name of Journal: *World Journal of Clinical Infectious Diseases*

ESPS Manuscript NO: 22890

Manuscript Type: REVIEW

Response to reviewers' comments:

Reviewer 1.

At the foot of Table 2, there is a typo.

"Lamphadenopathy-associated virus" has been changed to "Lymphadenopathy-associated virus" (page 41).

Reviewer 2.

This manuscript reviews an important topic but needs to have better focus. More specifically the text should be reduced and additional summary figures need to be include. For example a figure summarizing the mechanisms of latency and how all the targets that are described in these tables can overlap to make meaningful conclusions. The authors should be better summarize the literature in a concise way and integrate conclusions and interpretation rather than just a list of targets that has been described. The reviewer feels that the authors can do a better job synthesizing the data.

The following changes have been implemented:

1. Biomarker section: Table 3, which listed genes that were differentially expressed between latently infected and uninfected cells using different models of latency, has been replaced with a Figure presenting an overlap of differentially expressed

genes across models (new Figure 2). Text was added: "**Figure 2** depicts the result of comparison of DEGs between latently infected and uninfected cells available from 3 published studies[18, 19, 42]. A total of 1094 DEGs were identified." (page 16). We believe that Figure 2 better delivers the point concerning variability of biomarkers identified depending on the model; however, shows a modest overlap, in particular when primary cell models were used. The limitations of present biomarker studies have been further consolidated in new Table 3 (page 43), and the text has been shortened (end of the Biomarker section, page 16): "However, these studies have several limitations that presently preclude from achieving a consensus on what genes may represent suitable biomarkers. These limitations and potential solutions that may advance this field are summarized in **Table 3.**" (See also response to Reviewer 4).

2. Mechanisms of latency section: Table 4, which listed all the pathways and GO terms enriched for differentially expressed genes in latency, has been replaced with a diagram presentation of the terms (new Figure 3) related to transcriptional and post-transcriptional regulation. Among these, the following categories have been highlighted: Signaling pathways; RNA synthesis regulation; RNA processing; and Translation and metabolism. The text was edited on page 17: "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." The text enumerating terms and pathways identified in each study has been cut out. The section describing results from the study by Bandyopadhyay and colleagues who used a network based approach has been shortened: "Network-based approaches can also be utilized to identify genes that may have a role in regulation of HIV 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 and colleagues[51] who utilized the Krishnan and Zeichner dataset[18]. *TUB3A* 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." (page 18).

3. LRA section: text has been shortened and re-structured to reflect the main conclusions of the section (novel mechanisms of HIV reactivation by LRAs; inhibitory effects of LRAs on HIV reactivation; common effects of LRAs of different classes via components of p-TEFb complex) (extensive edits on pages 19 through 21). A diagram presentation of major findings from the LRA studies have been added (new Figure 4). The authors felt that this particular section would benefit from providing specific examples of affected genes and explanations on how exactly HIV reactivation or inhibition may be achieved via the effects of LRAs; therefore, we have retained some of the detailed specific examples. The conclusions section has been edited to reflect the changes in the text: "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." (pages 21-22).
4. For clarity, we have also introduced Figure 1 in the methods review section rather than in the conclusions: "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" (page 7), "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" (page 8); "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." (page 9). Figure 1 has been edited to include percentages of studies in the field of HIV

latency and eradication that used each of the methods, rather than the total number of studies.

Reviewer 3.

The topic reviewed is of interest but the review ended to be huge and its text should be reduced. For example figures or summarizing tables could be used instead of extended text. The literature should be also summarized in concise way in order to present specific findings rather than all the targets that have been described.

See response to the critique from Reviewer 2.

The review does not contain information on the mechanisms and the different forms of latency (HIV-1 DNA) within the host cells (as for example reviewed in Current HIV Research, 2009, 7, 255-265). Those should be discussed/reviewed and how all the described targets act or correlate to them.

Our review has focused on the form of latency represented by integrated provirus (post-integration latency), which is long-lived and capable of producing replication-competent provirus. While pre-integration form of latency, LTR circles, was shown to be stable in some studies[1, 2], LTR circles do not serve as template for producing replication-competent provirus. Another form of pre-integration latency, unintegrated linear HIV DNA, may serve as a template for HIV expression[3, 4], but it is relatively short-lived[5] and not present after many years of viral suppression on combination antiretroviral therapy. In model systems, unintegrated linear DNA may be present and may contribute to HIV reactivation readout when using LRAs. However, model systems that have HIV protein (p24) or a reporter (e.g. GFP) readout will detect reactivation solely from integrated provirus, unless the reporter is cloned in place of Nef, as only RNA and early proteins Nef and Tat may be expressed from unintegrated linear DNA[4]. The reviewed gene expression studies that used GFP reporter did not discuss peculiarities of HIV expression from integrated vs unintegrated DNA. In addition, gene expression studies testing mode of action of LRAs reviewed in our

manuscript, often utilized uninfected cells. In these cases, we have discussed the effects relevant to the mechanisms of reactivation from integrated HIV genome, previously reported in literature. We have now clarified that the review discusses the long-lived inducible reservoir, which is represented by cells bearing integrated provirus (sentences 1-3 of Introduction): "In the present era of combination anti-retroviral therapy (cART), the persistence of cellular 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." (page 5).

In the segment of evaluating the levels of HIV RNA, it would be of interest to include not only the RNA-seq methodologies, but also the sensitive assays that can detect and quantify intracellular or plasma HIV-RNA down to 1 cop/ml (there is a long list of those assays, but also really nice reviews).

The focus of the present review is the systems biology approaches. Thus, we have not initially included any targeted approaches in our review. However, we agree that sensitive HIV RNA detection techniques can be used and the measurements can then be integrated with high throughput host gene expression profiling data. Therefore, we have now included 2 references to reviews of these methods in the segment dedicated to evaluating the levels of HIV RNA, per suggestion of this Reviewer: "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." (pages 12-13).

Reviewer 4.

The purpose of this review is to summarize gene expression profiling and systems biology applications to studies of HIV latency and eradication. They have listed many genes expressed in latently infected and uninfected cells but failed to make any consensus or identify any latency biomarkers. The entire review is well written but the objective of the review is not achieved. It would be appropriate if authors could have validated the review's objective by identify at last one biomarkers from the vast data available in the literature.

We agree with the reviewer that identification of latency biomarkers would be an important achievement for the field. While data from several laboratories show that differentially expressed genes between latently infected cells can be identified, there is a number of limitations currently impeding making a strong consensus regarding biomarkers of latency:

1. Percentage of latently infected cells, even when using *in vitro* models, is relatively small, suggesting that the gene expression differences in each individual cell have to be large. Genes with expression differences less than 40-fold will likely not be detected. Methods are needed to isolate latently infected cells or perform profiling at a single cell level.
2. Bystander effect in the models. Due to limitations of culturing primary cells, the existing models rely on short latency period preceding the experimental testing, thus it is possible that the exposure to the virus during latency establishment phase, even for those cells that did not get infected, may affect gene expression signature.
3. A biomarker has to be uniquely expressed on latently infected cells, so that not all CD4+ T cells are targeted in a potential cell killing approach.
4. Different models represent different aspects of latency establishment (e.g. activated cells coming to quiescence or direct infection of the resting cells). It is possible that biomarkers may be different based on the route of latency

establishment, and therefore more models representing different mechanisms need to be tested and compared before the consensus can be reached.

5. Gene expression studies provide information on what targets are potentially good candidates to be biomarkers; however, experimental validations are required both in the model systems and using cells from HIV-infected individuals *in vivo*, which has not been demonstrated yet in existing literature.

These limitations have now been concisely discussed in the form of a table (new Table 3), allowing to further consolidate the discussion and shorten the text of the manuscript per suggestions of the other reviewers. We hope that this clarifies our stand point on the biomarkers of latency. (Table numbers following Table 3 have been adjusted accordingly). The shortened text now reads: "However, these studies have several limitations that presently preclude from achieving a consensus on what genes may represent suitable biomarkers. These limitations and potential solutions that may advance this field are summarized in **Table 3.**" (page 16). The conclusions section has also been edited to emphasize limitations and possible solutions for biomarker discovery: "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." (page 21). (See also response to Reviewer 2).

REFERENCES

1 **Pierson T,** Kieffer TL, Ruff CT, Buck C, Gange SJ, Siliciano RF. Intrinsic stability of episomal circles formed during human immunodeficiency virus type 1 replication. *J Virol* 2002; **76**: 4138-4144 [PMID: 11907256 DOI: 10.1128/JVI.76.8.4138-4144.2002]

2 **Butler SL,** Johnson EP, Bushman FD. Human immunodeficiency virus cDNA metabolism: notable stability of two-long terminal repeat circles. *J Virol* 2002; **76**: 3739-3747 [PMID: 11907213 DOI: 10.1128/JVI.76.8.3739-3747.2002]

3 **Trinité B,** Ohlson EC, Voznesensky I, Rana SP, Chan CN, Mahajan S, Alster J, Burke SA, Wodarz D, Levy DN. An HIV-1 replication pathway utilizing reverse transcription products that fail to integrate. *J Virol* 2013; **87**: 12701-12720 [PMID: 24049167 DOI: 10.1128/JVI.01939-13]

4 **Poon B,** Chang MA, Chen ISY. Vpr Is required for efficient Nef expression from unintegrated human immunodeficiency virus type 1 DNA. *J Virol* 2007; **81**: 10515-10523 [PMID: 17652391 DOI: 10.1128/JVI.00947-07]

5 **Pierson TC,** Zhou Y, Kieffer TL, Ruff CT, Buck C, Siliciano RF. Molecular characterization of preintegration latency in human immunodeficiency virus type 1 infection. *J Virol* 2002; **76**: 8518-8531 [PMID: 12163571 DOI: 10.1128/JVI.76.17.8518-8531.2002]