Paediatric HIV infection in the 'omics era: defining transcriptional signatures of viral control and vaccine responses

Nicola Cotugno[1,3], Lesley De Armas[2], Suresh Pallikkuth[2], Paolo Rossi[1,3], Paolo Palma[1], Savita Pahwa*[2]

[1]University Department of Pediatrics, DPUO, Unit of Immune and Infectious Diseases, Bambino Gesù Children's Hospital, Rome, Italy
[2]Miami Center for AIDS Research, Department of Microbiology and Immunology, Miller School of Medicine, University of Miami, Miami, Florida, USA
[3]Department of Systems Medicine, University of Rome, 'Tor Vergata', Italy

Abstract

Modern technologies and their increased accessibility have shifted 'benchtop' medical research to the larger dimension of 'omics. The huge amount of data derived from gene expression and sequencing experiments has propelled physicians, basic scientists and bioinformaticians towards a common goal to transform 'big data' into predictive constructs that are readily available and will offer clinical utility. Although most of the studies available in the literature have been performed on healthy subjects and in peripheral blood mononuclear cells (PBMC), which are a heterogenous and extremely variable pool of cells, scientists are now trying to address mechanistic questions in purified cell subsets in pathological conditions. In the field of HIV, few attempts have been made to comprehensively evaluate gene-expression profiles of infected patients with different disease status. With the view of discovering a path towards remission or viral eradication, perinatally HIV-infected children represent a unique model. In fact the well-defined time of infection and the resulting opportunity to start early treatment, thereby generating a smaller size of viral reservoir and a more intact immune system, allow for investigation of therapeutic strategies to defeat the virus. In this scenario, 'transcriptomic' or gene expression technologies and supporting bioinformatics applications need to be strategically integrated to provide novel information about immune correlates of virus control following treatment interruption. Here we review modern techniques for gene expression analysis and discuss the best transcriptomic strategies applicable to the field of functional cure in paediatric HIV infection.

Keywords: gene expression, paediatric HIV, systems biology, transcriptomics, cell subsets gene expression

Introduction

Transcriptional profiling has taken root over the last two decades in the fields of immunology and microbiology with the aim of defining molecular characteristics underlying the response of a complex immune system to perturbation by an antigen. Recently, 'systems biology' has been applied to vaccine research in an attempt to define and predict the efficacy of an immunisation in a specific host [1,2] thus moving modern medicine towards a personalised approach [3]. The field of vaccinology represents an interesting opportunity for using systems biology techniques because of easy accessibility to peripheral blood mononuclear cells (PBMCs) and certain known correlates of immune/memory responses. More importantly, it offers the possibility to study consequences of external perturbation in terms of *in vivo* changes in gene expression after immunisation [4,5]. An additional advantage of transcriptomic analysis is that a large amount of data (i.e. 'big data') or information can be obtained from very small blood volumes, increasing the significance in sample-limited settings such as paediatric studies.

In the last few years these technologies have been applied to infectious diseases, and specifically to HIV, in an attempt to define molecular profiles associated with different HIV clinical phenotypes in order to discover new markers of disease progression or potential therapeutic targets. Recently, molecular mechanisms underlying the ability of specific groups of HIV-infected patients to naturally control virus in the absence of antiretroviral therapy (e.g. elite controllers and long-term non-progressor patients) have been described [6]. These and other transcriptional analyses are needed in order to propel the field towards discovery of immune host characteristics that differentiate patients remaining asymptomatic after infection and those rapidly progressing to AIDS [7,8].

Although paediatric HIV infection and mother-to-child transmission represent a small portion (≈10%) of patients acquiring HIV globally [9], the study of perinatal HIV infection represents a unique model in the context of research for a functional cure. An immature immune system, predominantly naïve, restricts the ability of the virus to establish reservoirs in the infant. Furthermore, early treatment initiation following perinatal infection has provided 'proof of concept' for sustained virological control after therapy interruption and preserved function of the B cell compartment [10,11]. These findings highlight a substantial advantage for application of future 'therapeutic' immunisation strategies.

The well-known case of the 'Mississippi baby', whereby viral control was achieved for a long period in the absence of highly active antiretroviral therapy (HAART) [12], is unique, as a similar therapeutic approach (i.e. very early treatment in newborn period, within 48 hours after birth) in other HIV-infected infants was accompanied by a much earlier viral rebound following treatment interruption [13–16] (Figure 1). What defines the discrepancies in outcome between patients who receive the same treatment at the same time and yet differ in their ability to control the virus remains unknown, and is likely to be multifactorial (Figure 1). One hypothesis is that molecular characteristics in the host immune system may underlie the differences in terms of viral control. To test this possibility, transcriptomic analysis represents a viable research objective to: (i) define the molecular mechanisms underlying the ability of the host immune system to control the virus; (ii) predict the time to rebound in order to design effective strategies for structured therapy interruption; and (iii) design therapeutics to target particular gene pathways involved in virus control.

This review describes the currently available techniques for measuring gene expression and discusses strategies to integrate these innovative technologies into HIV cure research with a specific focus in perinatal HIV infection.
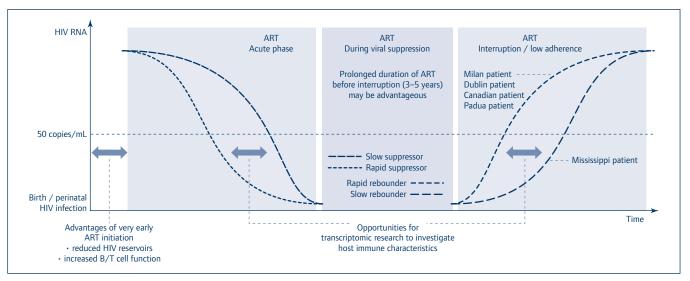
*Corresponding author: Savita Pahwa, Department of Microbiology and Immunology, Miller School of Medicine, University of Miami
1580 NW 10th Avenue, Miami, FL 33136, USA
Email: spahwa@med.miami.edu

Figure 1. Windows of opportunity for transcriptomic research to investigate host immune characteristics of HIV-infected children. Schematic shows hypothetical viral load in patients after very early ART and treatment interruption

Transcriptome analysis: RNA-Seq, and its application

Information contained within our deoxyribonucleic acid (DNA) in the form of genes becomes functional through transcription, which leads to the formation of active macromolecules, such as RNA and ultimately proteins. The real-time polymerase chain reaction (real-time PCR) and the reverse transcriptase-polymerase chain reaction (RT-PCR) are two techniques that provide quantitative transcriptional information. Although extremely accurate and sensitive, yielding transcript from nanolitre volumes, these techniques are historically limited to measure one or a few genes per reaction. On the other hand, technologies that quantify the whole transcriptome have been developed through hybridisation-based or sequence-based approaches such as tiling microarray. Hybridisation-based approaches have now been overshadowed by RNA-Seq where total cellular RNA is converted to a library of cDNA fragments attached at each end with an adaptor or 'barcode'. In a second step, the fragments are sequenced and the resulting 'reads' are aligned to a reference genome or exome, thereby providing information on both the transcriptome structure (known and novel splice variants) and the level of gene expression (number of reads). Importantly, the ability of this technique to match reads with a de novo reference genome reveals information about uncharacterised sequences [17]. Table 1 compares available transcriptional analysis techniques.

How can so much information regarding molecular networks orchestrating our systems help in clinical management? How can transcriptional profiles be used as predictive markers in healthy or pathological conditions? The comprehensive translational significance of the huge amount of data generated through analysis of the transcriptome can only be achieved by pooling resources between research sites and networks sharing clinical metadata, 'omics data and computational analysis [18,19].

In line with this, both the European Union [20] and the National Institute of Allergy and Infectious Diseases (NIAID-National Institute of Health) [21] have highlighted in their strategic plans the role of systems biology as a crucial player in discovery research regarding pathogen biology, interactions among pathogens, hosts and the environment. Well-curated and easy-to-use publicly available 'omics online databases have been created, such as the Immunological Genome Project and the ENCODE project, as well as online platforms and mathematical tools to define gene pathways or transcriptome modules [e.g. pathjam.org, genemania.org, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways enrichment, and Database for Annotation, Visualization and Integrated Discovery (DAVID)].

In the field of immunology, in the last decade much effort has been made to give translational value to 'big data' derived from transcriptomic analyses. Furthermore, in the context of vaccinology, immunological perturbation such as that induced by vaccination represents an attractive model to provide novel

Table 1. Description of the currently available transcriptional techniques

	cDNA sequencing	Microarray	RNA-seq	Fluidigm	Flow RNA
Principle	Sanger sequencing	Hybridisation	High-throughput sequencing	Hybridisation with amplification	Hybridisation with amplification
Throughput (number of samples)	Low	Medium	Medium	High (100+)	Medium
Throughput (number of genes)	Low	High (20–40K)	Very high (millions)	Medium (100+)	Low (3)
Dynamic range to quantify gene expression level	Not practical	Up to 300-fold	>8000-fold	Up to 10^5-fold	Up to 1000-fold
Ability to distinguish different isoforms	Yes	Limited	Yes	Limited	Not practical
Required amount of RNA	High	High	Low	Low	Low
Cost per sample	$10	$300–600	$600–900	$20–25	$45

discovery research data. In particular, influenza vaccination represents an extremely important and useful model for 'perturbation' analysis in both healthy and pathological conditions [2,22,23].

Transcriptomic analysis in HIV research: where do we stand?

In HIV research, transcriptomic studies have so far pursued the following approaches: (i) transcriptional profiling in rare subjects who are able to control the disease in the absence of treatment (elite controllers, EC; long-term non-progressors, LTNPs; or viraemic non-progressors, VNPs) as contrasted with rapid progressors (RP); (ii) evaluating therapeutic approaches through the analyses of changes in gene expression in HIV target cells after *in vitro* or *in vivo* infection/immunisation; and (iii) profiling for genes driving the reversal of latency in latent/replication-competent HIV-infected cells [24].

Gene expression in different HIV 'clinical phenotypes'

HIV-infected individuals demonstrate differences in virus replication and disease progression [25]. Gene expression profiles from EC have been recently characterised [6,26,27]. Indeed, in these patients, higher expression of gene pathways involved in cell survival and antiviral responses may contribute to spontaneous control of the virus in the absence of ART [28,29]. Another recent research study focused on identifying gene signatures associated with HIV reservoirs in memory CD4 T cells using *in vitro* latency models and cells isolated from EC [6]. Compared to RP, transcriptional profiles of resting T central memory (T_{CM}) cells from EC showed a pattern of differentially expressed genes (DEG) comprising three specific signatures, namely overexpression of T cell receptor (TCR) and co-stimulatory signalling pathways, overexpression of the PRDM-1/Blimp-1 transcriptional repressor, and downmodulation of type-I IFN-related genes. Among T cell subsets, the PRDM1/Blimp-1 upregulation was associated with lower levels of both cellular HIV DNA and HIV mRNA levels [6], implying a T cell-specific molecular signature that is characteristic of HIV controller status. These findings suggest that specific gene signatures have to be investigated on the scale of immune cell subsets rather than in PBMC. Another important study identified transcriptional signatures in CD4 T cells that characterised a pattern of progression to AIDS [29]. Thus, a list of genes involved in immune exhaustion/dysfunction (*CASP1*, *CD38*, *LAG3*) has been shown to be upregulated in HIV- and SIV-infected RP humans and macaques. On the other hand, genes that function to provide negative feedback signals to the loop of activation and inflammation (*SOCS1* and *EEF1D*) were upregulated in VNP, clearly showing a direct relationship between the transcriptional profile of these cells and their ability to control disease progression.

Based on the evidence that CD8 T cell proliferation is one of the strongest correlates of protection [30], transcriptional profiles of proliferating HIV-specific CD8 T cells have been investigated in both EC and RP. These studies have shown an inverse relationship between caspase 8 gene (*CASP8*) expression and the ability of CD8 T cells to proliferate after *in vitro* stimulation in EC. Accordingly, an increase in *CASP8* activation after *in vitro* stimulation was related to lower proliferation rate of HIV-specific CD8 T cells in chronic progressors [31]. Other studies have focused on immune exhaustion caused by HCV/HIV co-infection, comparing transcriptional profiles of CD8 T cells with HIV monoinfected patients [32]. These studies demonstrated a prominent decrease in genes regulating interleukin-2 (IL-2) and

interferon-gamma (IFN-γ) pathways in the HIV/HCV co-infected group compared to the HCV and HIV monoinfected patients. In support of these results, Sandler *et al*. described through the administration of a receptor agonist in SIV macaques that IFN-I responses are able to slow disease progression and IFN-induced genes such as *OAS2* and *MX1* are directly related to the virus control in the acute phase of the infection [33].

Although these data show a strong association between host transcript profiles of immune cells and disease progression and virus control, the molecular mechanisms associated with viral rebound following treatment interruption remain unknown. Such studies would be particularly informative in the setting of paediatric HIV where early treatment with HAART and the predominantly naïve T cell compartment lower the size of reservoirs [34].

Gene expression of HIV permissive cells and latency reversal

Transcriptional analysis of HIV-specific and HIV-permissive target cells have been performed, respectively in B and T cell compartments, as well as in monocytes, macrophages, dendritic cells and neurons (reviewed in [35]). It is known that HIV preferentially infects activated CD4 T cells; however, infection rates are very low even in this susceptible population [36]. In the last few years, systems biology has attempted to define the differences in transcriptional profiles of infected CD4 T cells compared to exposed, uninfected CD4 T cells in order to characterise the molecular mechanisms controlling infection and latency. These studies showed that profiles of HIV permissive cells overlap with those of highly activated effector T cells, and show a higher expression of genes involved in apoptosis, lymphocyte activation, p53 activation, and cytokine–cytokine receptor interaction [37].

Transcriptional profiles of latently infected cells have also been studied to find new strategies to 'kick' or 'shock' the virus towards active replication or conversely to define unique molecular characteristics governing latency in these cells in order for them to be therapeutically targeted. Genome-wide features of latently infected cell lines have demonstrated a role for genes in the protein acetylation pathway, such as histone deacetylase (*HDAC*), in establishing and maintaining viral latency [38]. Recent studies are targeting this pathway with novel therapeutic agents. Indeed, striking differences in gene expression were found 2 hours after vorinostat (HDAC inhibitor) administration in HIV-infected patients [39].

In an effort to discover a 'phenotypic' marker to define latently infected cells, Romerio *et al*. screened for DEGs encoding cell-surface markers. Out of 33 DEGs, *CD2* exhibited the highest expression in latently infected compared to uninfected cells [27]. Additionally, resting memory CD4+CD2^{hi} T cells from virally suppressed patients harboured higher HIV-1 DNA copy numbers than all other CD4 T cell subsets. These transcriptomic studies indicate the potential for identifying unique molecules in latently infected resting memory CD4 T cells that could be targeted therapeutically.

Gene expression of virus- and vaccine-induced cells

Transcriptional profiles of other non-HIV antigen-specific T cells (i.e. CMV-specific) have been analysed and several genes belonging to innate antiviral gene pathways (such as *IFIT1*, *IFIT3*, *MX1*, *TRIM5*) were shown to contribute to decreased susceptibility to HIV infection [40,41]. In order to test the efficacy of an HIV therapeutic vaccine, several groups are now shifting their focus towards molecular signatures of vaccine-induced cells, both in the cytotoxic and memory B cell

compartments. Changes in gene expression in distinct CD8 T cell subsets have been described in gene pathways such as T cell migration (*SELL, CCR7, XCL1, CCR5, CXCR3*), cytokine responsiveness (*IL2RA, IL2RB, IL7R*) and effector functions (*IFNG, TNF, PRF1, FASL, GZMM, GZMK*) following HIV DNA immunisation [42]. In the context of B cell memory directed to HIV immunisation, Env-specific switched memory B cells, sorted by flow cytometry using the same protein used for immunisation, have been isolated and sequenced in macaques [43]. Although these results need further confirmation in human studies, the approach to sort out and investigate the molecular characteristics of vaccine-induced HIV-specific memory B cells, along with their 'wild type' counterparts, may add significant value in studies aiming to define the efficacy of therapeutic vaccinations.

Transcriptional research approaches in paediatric HIV infection

With accumulating data and increasing accessibility of 'omics technologies, it is paramount to rationally apply these informative assays in the field of paediatric HIV infection in order to provide insight into pathways leading to a permanent remission of HIV or a cure. To date, only a few studies have been conducted in perinatally HIV-infected children and new strategies are needed in order to customise these assays for 'sample limited' settings, such as is the case in paediatric research. A study design aiming to perform transcriptomic analysis should follow a rationale which consists of: (i) strategies for selection of genes of interest in order to use inexpensive techniques requiring small amounts of cDNA; (ii) selection of cell subsets rather than bulk PBMC in order to reduce transcriptional 'noise' and highlight biological diversity between subsets; and (iii) *in vitro* stimulation to challenge the cells and obtain 'functional' transcriptional information about the cells of interest. These concepts are briefly elaborated below, and illustrated in Figure 2.

Gene selection and validation

Although RNA-Seq is becoming increasingly cost-effective, most research groups are still far from being able to conduct
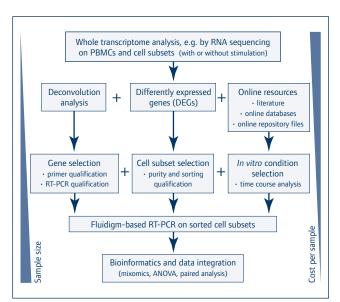

Figure 2. The flowchart defines an application strategy to investigate transcriptional profile in paediatric HIV infection research. In the first step, experimental variables should be selected according to results derived from RNA Seq analysis (high cost per sample). In the second step, selection of genes, conditions and cell subsets performed according to RNA-Seq results, literature and deconvolution analysis, should implement experimental design for multiplexed RT-PCR (low cost per sample) studies which, in turn, will select gene signatures of immune functions

population studies with this technique. As previously mentioned, RNA-Seq has the unique advantage to sequence known and, more importantly, unknown transcripts. In line with this, studies introducing 'omics into fields such as perinatal HIV infection should consider first, the exploration of transcriptional profiles in PBMC and cell subsets in a small number of individuals. As a second step, gene selection should be performed merging information derived from RNA-Seq analysis, deconvolution analysis and publicly available resources. Once genes as well as gene pathways of interest are defined, a validation phase should follow in which targeted transcriptional profiles are evaluated with multiplexed reverse transcriptase PCR (RT-PCR) that allows for analysis of cDNA from even a single cell (Fluidigm BioMark, San Francisco, CA, USA). Primer qualification to test for efficiency and sensitivity should be performed through multiplexed RT-PCR with two-fold dilutions of control cDNA [44].

Cell subsets, single cells or PBMC?

Most transcriptomic studies performed so far have been conducted on healthy subjects and in the heterogeneous pool of PBMC. This approach, while simpler, limits interpretation since in PBMC there is dilution of the gene expression in individual cell subsets or in single cells.

In the field of vaccinology, investigators have addressed this issue in experiments using PBMC, through complex statistical analysis that uses gene pathways and blood transcriptional modules in order to perform deconvolution calculations that normalise gene expression according to phenotype variability of cell subsets [1].

However, the problems with data interpretation from heterogeneous cell populations (e.g. PBMC) are exacerbated when analysing patients with perturbation of the immune phenotype [45]. In Figure 3 we show preliminary Fluidigm BioMark analysis demonstrating the transcriptional variation in different cell subsets within the B cell compartment of healthy and HIV-infected children using principal component analysis. Segregation amongst B cell subsets is altered in a cohort of paediatric HIV-infected children compared to uninfected children, even if virally suppressed (Figure 3b). This must be carefully considered in case–control human studies investigating diseases in which immune phenotypes are subverted and gene expression may need to be adjusted to specific subset frequencies within PBMC. An approach incorporating a validation phase on transcripts derived from sorted cells can help the deconvolution of PBMC transcriptomic data in patients affected by primary or acquired immune deficiencies.

Until newer technologies are accepted [46], flow-based cell sorting and magnetic-based cell enrichment represent the benchmarks of cell isolation. Multiway sorting through multicolour flow cytometry represents the most convenient way to sort cells. Although improvement in devices and equipment is still needed (e.g. multiway sorting in plates is not available; four-way sorting is the maximum if using tubes), this approach allows for efficient and simultaneous isolation of cells. Sorting efficiency before proceeding with transcriptome analysis should be tested twice: through a purity test post-sorting, and through PCR of reference genes.

In vitro stimulation: pathways towards efficacy prediction

RNA-Seq as well as RT-PCR are able to create a qualitative and quantitative snapshot of the RNA being transcribed by genes in a specific cell-type at a given moment. The possibility of studying the changes in gene expression in response to a specific stimuli gives one the opportunity to determine gene expression 'fitness' following *in vitro* stimulation with specific compounds. Several
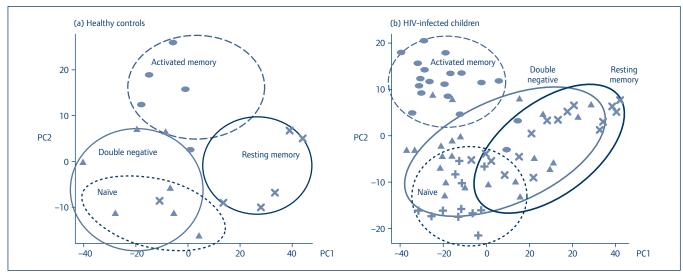
Figure 3. Principal component (PC) analysis of B cell subsets in (**a**) healthy controls and (**b**) HIV-infected children. Analysis performed with SingulaR: Gene Expression Analysis Software designed for Fluidigm (BioMark). Gene expression derived from analysis of 500 cells per subset and 96 B cell expressed genes. Cell subsets where sorted using Aria II Cell Sorter into PCR buffer-containing tubes. Activated Memory: alive, CD19, CD10-, IgD-, CD27+, CD21-; Resting memory: alive, CD19, CD10-, CD27+, CD21+; Naïve: alive, CD19, CD10-, CD27-, IgD+; Double negative: alive, CD19, CD10-, CD27-, IgD-

studies, as previously mentioned, have developed *in vitro* methods to elicit specific cellular responses such as viral replication from latently infected cells [27] or to study the *in vitro* effect of stimulation by HIV proteins [47]. Importantly, these results may be compared to *in vivo* gene expression analysed after *in vivo* immunisation with the same compounds, giving the opportunity to develop *in vitro* assays that are able to predict and model *in vivo* efficacy of therapeutic compounds. However, selection of the conditions to be used for *in vitro* experiments can be extremely complicated since different genes are differentially coordinated, in terms of magnitude and time, in response to a specific stimuli [48]. Gene selection should be supported by kinetic analyses designed according to specific biological queries. Experimental variables in this phase include not just length of stimulation, concentration and number of cells, but importantly, the choice of cellular pools to be used for cultures (e.g. PBMC/sorted cell subsets, co-cultures, proliferating/activated cells) (Figure 2).

Data integration

Interpretation of data derived from the aforementioned approaches needs to be rationally applied to the 'omic scale. The statistical programming language R has been developed and improved in the last 20 years to provide data analysis for gene expression or sequencing analysis (www.r-project.org). Through the application of packages this language can be applied to specific statistic needs, and software such as Bioconductor (Seattle, WA, USA) for RNA-Seq data or singulaR package (Biomark, San Francisco, CA, USA) can be used for basic or advanced statistical analysis [49]. Furthermore, other packages such as 'Mixomics' (QFAB Bioinformatics, St Lucia, Queensland, Australia) can perform correlation analysis between two datasets, such as gene expression data and clinical or functional immunology data, thereby increasing the clinical relevance of the gene expression analysis [47]. To improve robustness and power of transcriptomic data, gene set enrichment analyses (GSEA) have been developed in order to analyse genes within their functional group or as being part of the same signalling pathway. In line with this approach, increasing numbers of functional annotation tools available online free of charge can identify enriched biological themes – Gene Ontology (geneontology.org), DAVID (david.abcc.ncifcrf.gov), Pathjam.org, genemania.org – and functionally related gene groups [50]. Overall, other than

cost, the integration and the analysis of 'big data' seems to be the main limiting factor for clinical utility; therefore physicians, basic scientists and bioinformatics must collaborate in order to study gene expression at a wider scale.

Conclusions and future perspectives

Advances in 'omics have shifted discovery science to a different dimension. For paediatric research in general, and more specifically in the field of infectious diseases, we now have the opportunity to obtain 'big data' from very small samples, which overcomes the biggest limiting factor in studies involving children. However, even if these technologies are becoming increasingly accessible, sophisticated and precise, the best approach requires careful planning. Emerging literature and publicly available resources, coupled with experience in the field of HIV infection, provide the opportunity to rationally apply such technologies in the context of perinatal HIV infection in order to provide a proficient 'omics approach to the cause of finding a functional cure for HIV. The specific immune context of perinatally HIV-infected children treated early after infection is unique from the viewpoint of a therapeutic approach aimed at permanent HIV remission. An immunisation strategy may have a better chance to be effective in early treated children since their immune systems are relatively naïve to the virus infection, have an intact immune B cell compartment and lower size of viral reservoir. So far, no transcriptional studies or other 'omics technologies have been systematically applied to such patients who may represent an ideal study model for investigating potential strategies for a functional cure.

Disclosure of potential conflict of interest

References

1. Li S, Rouphael N, Duraisingham S *et al*. Molecular signatures of antibody responses derived from a systems biology study of five human vaccines. *Nat Immunol* 2014; **15**: 195–204.

2. Tsang JS, Schwartzberg PL, Kotliarov Y *et al*. Global analyses of human immune variation reveal baseline predictors of postvaccination responses. *Cell* 2014; **157**: 499–513.

3. Boyd SD, Jackson KJ. Predicting Vaccine Responsiveness. *Cell Host Microbe* 2015; **17**: 301–307.

4. Trautmann L, Sekaly RP. Solving vaccine mysteries: a systems biology perspective. *Nat Immunol* 2011; **12**: 729–731.

5. Rappuoli R, Aderem A. A 2020 vision for vaccines against HIV, tuberculosis and malaria. *Nature* 2011; **473**: 463–469.

6. de Masson A, Kirilovsky A, Zoorob R *et al.* Blimp-1 overexpression is associated with low HIV-1 reservoir and transcription levels in central memory CD4+ T cells from elite controllers. *AIDS* 2014; **28**: 1567–1577.

7. Blankson JN, Bailey JR, Thayil S *et al.* Isolation and characterization of replication-competent human immunodeficiency virus type 1 from a subset of elite suppressors. *J Virol* 2007; **81**: 2508–2518.

8. Blankson JN. Control of HIV-1 replication in elite suppressors. *Discov Med* 2010; **9**: 261–266.

9. WHO. Global Update on the Health Sector Response to HIV, 2014. Executive summary. 2014. Available at: http://apps.who.int/iris/bitstream/10665/128196/1/WHO_HIV_2014.15_eng.pdf (accessed May 2015).

10. Pensieroso S, Cagigi A, Palma P *et al.* Timing of HAART defines the integrity of memory B cells and the longevity of humoral responses in HIV-1-infected children. *Proc Natl Acad Sci U S A* 2009; **106**: 7939–7944.

11. Moir S, Fauci AS. Insights into B cells and HIV-specific B-cell responses in HIV-infected individuals. *Immunol Rev* 2013; **254**: 207–224.

12. Persaud D, Gay H, Ziemniak C *et al.* Absence of detectable HIV-1 viremia after treatment cessation in an infant. *N Engl J Med* 2013; **369**: 1828–1835.

13. Giacomet V, Trabattoni D, Zanchetta N *et al.* No cure of HIV infection in a child despite early treatment and apparent viral clearance. *Lancet* 2014; **384**: 1320.

14. Butler KM, Gavin P, Coughlan S *et al.* Rapid viral rebound after 4 years of suppressive therapy in a seronegative HIV-1 infected infant treated from birth. *Pediatr Infect Dis J* 2014; **34**: e48–51.

15. Zanchetta M, Anselmi A, Vendrame D *et al.* Early therapy in HIV-1-infected children: effect on HIV-1 dynamics and HIV-1-specific immune response. *Antivir Ther* 2008; **13**: 47–55.

16. Brophy J, Chun T-W, Samson L *et al.* Impact of early initiation of combination antiretroviral therapy on measures of virus in peripheral blood of vertically HIV-1-infected children. *International AIDS Conference*. July 2014. Melbourne, Australia. Abstract TUAB0206LB.

17. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* 2009; **10**: 57–63.

18. Germain RN, Schwartzberg PL. The human condition: an immunological perspective. *Nat Immunol* 2011; **12**: 369–372.

19. Germain RN, Meier-Schellersheim M, Nita-Lazar A, Fraser ID. Systems biology in immunology: a computational modeling perspective. Annu Rev Immunol 2011; 29: 527–585.

20. Strategic Report for Translational Systems Biology and Bioinformatics in the European Union. 2012. Available at: http://ec.europa.eu/digital-agenda/en/news/strategic-report-translational-systems-biology-and-bioinformatics-european-union (accessed May 2015).

21. NIAID. NIAID Strategic Plan 2013. 2013. Available at: www.niaid.nih.gov/about/whoWeAre/planningPriorities/Documents/NIAIDStrategicPlan2013.pdf (accessed May 2015).

22. Nakaya HI, Wrammert J, Lee EK *et al.* Systems biology of vaccination for seasonal influenza in humans. *Nat Immunol* 2011; **12**: 786–795.

23. Hoek KL, Samir P, Howard LM *et al.* A cell-based systems biology assessment of human blood to monitor immune responses after influenza vaccination. *PLoS One* 2015; **10**: e0118528.

24. Mehla R, Ayyavoo V. Gene array studies in HIV-1 infection. *Curr HIV/AIDS Rep* 2012; **9**: 34–43.

25. Okulicz JF, Marconi VC, Landrum ML *et al.* Clinical outcomes of elite controllers, viremic controllers, and long-term nonprogressors in the US Department of Defense HIV natural history study. *J Infect Dis* 2009; **200**: 1714–1723.

26. Fellay J, Shianna KV, Ge D *et al.* A whole-genome association study of major determinants for host control of HIV-1. *Science* 2007; **317**: 944–947.

27. Iglesias-Ussel M, Vandergeeten C, Marchionni L *et al.* 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.

28. Wu JQ, Dwyer DE, Dyer WB *et al.* Genome-wide analysis of primary CD4+ and CD8+ T cell transcriptomes shows evidence for a network of enriched pathways associated with HIV disease. *Retrovirology* 2011; **8**: 18.

29. Rotger M, Dalmau J, Rauch A *et al.* Comparative transcriptomics of extreme phenotypes of human HIV-1 infection and SIV infection in sooty mangabey and rhesus macaque. *J Clin Invest* 2011; **121**: 2391–2400.

30. Cortes FH, Passaes CP, Bello G *et al.* HIV controllers with different viral load cutoff levels have distinct virologic and immunologic profiles. *J Acquir Immune Defic Syndr* 2015; **68**: 377–385.

31. Gaiha GD, McKim KJ, Woods M *et al.* Dysfunctional HIV-specific CD8+ T cell proliferation is associated with increased caspase-8 activity and mediated by necroptosis. *Immunity* 2014; **41**: 1001–1012.

32. Zhao J, Yi L, Lu J *et al.* Transcriptomic assay of CD8+ T cells in treatment-naive HIV, HCV-mono-infected and HIV/HCV-co-infected Chinese. *PLoS One* 2012; **7**: e45200.

33. Sandler NG, Bosinger SE, Estes JD *et al.* Type I interferon responses in rhesus macaques prevent SIV infection and slow disease progression. *Nature* 2014; **511**: 601–605.

34. Ananworanich J, Puthanakit T, Suntarattiwong P *et al.* Reduced markers of HIV persistence and restricted HIV-specific immune responses after early antiretroviral therapy in children. *AIDS* 2014; **28**: 1015–1020.

35. Iglesias-Ussel MD, Romerio F. HIV reservoirs: the new frontier. *AIDS Rev* 2011; **13**: 13–29.

36. Chang ST, Sova P, Peng X *et al.* Next-generation sequencing reveals HIV-1-mediated suppression of T cell activation and RNA processing and regulation of noncoding RNA expression in a CD4+ T cell line. *MBio* 2011; **2**: e00134–00111.

37. Imbeault M, Giguere K, Ouellet M, Tremblay MJ. Exon level transcriptomic profiling of HIV-1-infected CD4(+) T cells reveals virus-induced genes and host environment favorable for viral replication. *PLoS Pathog* 2012; **8**: e1002861.

38. Park J, Lim CH, Ham S *et al.* Genome-wide analysis of histone modifications in latently HIV-1 infected T cells. *AIDS* 2014; **28**: 1719–1728.

39. Elliott JH, Wightman F, Solomon A *et al.* Activation of HIV transcription with short-course vorinostat in HIV-infected patients on suppressive antiretroviral therapy. *PLoS Pathog* 2014; **10**: e1004473.

40. Hu H, Nau M, Ehrenberg P *et al.* Distinct gene-expression profiles associated with the susceptibility of pathogen-specific CD4 T cells to HIV-1 infection. *Blood* 2013; **121**: 1136–1144.

41. Sakuma R, Noser JA, Ohmine S, Ikeda Y. Rhesus monkey TRIM5alpha restricts HIV-1 production through rapid degradation of viral Gag polyproteins. *Nat Med* 2007; **13**: 631–635.

42. Roychoudhuri R, Lefebvre F, Honda M *et al.* Transcriptional profiles reveal a stepwise developmental program of memory CD8(+) T cell differentiation. *Vaccine* 2015; **33**: 914–923.

43. Sundling C, Zhang Z, Phad GE *et al.* Single-cell and deep sequencing of IgG-switched macaque B cells reveal a diverse Ig repertoire following immunization. *J Immunol* 2014; **192**: 3637–3644.

44. Dominguez MH, Chattopadhyay PK, Ma S *et al.* Highly multiplexed quantitation of gene expression on single cells. *J Immunol Methods* 2013; **391**: 133–145.

45. Kardava L, Moir S, Shah N *et al.* Abnormal B cell memory subsets dominate HIV-specific responses in infected individuals. *J Clin Invest* 2014; **124**: 3252–3262.

46. Wyatt Shields IVth C, Reyes CD, Lopez GP. Microfluidic cell sorting: a review of the advances in the separation of cells from debulking to rare cell isolation. *Lab Chip* 2015; **15**: 1230–1249.

47. Liquet B, Le Cao KA, Hocini H, Thiebaut R. A novel approach for biomarker selection and the integration of repeated measures experiments from two assays. *BMC Bioinformatics* 2012; **13**: 325.

48. Chechik G, Koller D. Timing of gene expression responses to environmental changes. *J Comput Biol* 2009; **16**: 279–290.

49. Huber W, Carey VJ, Gentleman R *et al.* Orchestrating high-throughput genomic analysis with Bioconductor. *Nat Methods* 2015; **12**: 115–121.

50. Irizarry RA, Wang C, Zhou Y, Speed TP. Gene set enrichment analysis made simple. *Stat Methods Med Res* 2009; **18**: 565–575.