Lessons from the first host-genome analyses in HIV

Amalio Telenti

University Hospital Center &
University of Lausanne,
Institute of Microbiology,
CHUV 1011 Lausanne,
Switzerland
Tel.: +41 21 314 0550;
Fax: +41 21 314 4060;
amalio.telenti@
chuv.ch

'Among the several possible reasons to conduct this type of study ... two are of clear relevance to the field of HIV-1: first, understanding pathogenesis; and second, predicting risk and personalizing care.'

It is possible that 2007 will be remembered as the year when genome-wide association studies (GWAS) invaded the scientific journals – to the point of having a journal editor alert the readers that a rush of papers using GWA technologies would be published [1]. Also, in 2007, the first GWAS in infectious diseases was published on HIV-1 susceptibility [2]. This editorial will describe succinctly the general principles governing analysis of the whole genome of the host, the kind of results that are obtained and the applications, 'the fruits', of this type of research.

GWAS are the children of the genomics revolution that followed the publication of the first draft of the human genome in 2001 [3,4], the cataloguing of common human genetic variation (the so-called Haplotype Map or HapMap) in 2005 and 2007 [5,6], and the development of robust technologies capable of genotyping 500,000–1 million single nucleotide polymorphisms (SNPs) per individual. This wealth of genetic information can be related to a well-defined clinical phenotype to point ('map') to regions of the genome that may contain genes involved in variation of the trait of interest. GWAS generally identify anything from one to a handful of SNPs that explain from 1 to 10% of the variance of the study phenotype. Among the several possible reasons to conduct this type of study [7], two are of clear relevance to the field of HIV-1: first, understanding pathogenesis; and second, predicting risk and personalizing care [8].

The first GWAS in HIV-1 disease identified two genetic variants in the MHC locus that were associated with viral setpoint following seroconversion. One of these genetic variants, in the *HCP5* gene, tracks the well-known protective allele *HLA-B*5701*. The second variant is a SNP that modulates expression levels of *HLA-C*. A third group of

SNPs that are in the proximity of the *ZNRD1/RNF39* genes – an RNA polymerase I subunit – associates with differences in disease progression rates. These findings illustrate the usual outcome of GWAS; some results confirm existing knowledge (*HLA-B*5701*), some underscore the role of genes that no longer retain the attention of the scientific community (*HLA-C*), and some contribute unexpected information (*ZNRD1/RNF39*). In all cases, results can be measured and ranked in the context of all other possible genome influences. Findings from this first GWAS are by now confirmed and currently the object of functional analyses.

Other genomic approaches implemented in the field of HIV take advantage of cellular systems that are amenable to manipulation (infection) *in vitro*. Loeuillet *et al.* infected immortalized B-cell lines obtained from multi-generation, large pedigree families, using standardized conditions [9]. The family structure allows genetic mapping by linkage analysis. This work identified a significant heritability of the trait of cellular susceptibility to HIV. It also proposed a possible role for a family of glycosylphosphatidylinositol-anchored cell surface receptors or secreted cytotoxins – the LY6/uPAR family – in the observed interindividual variability of cellular permissiveness to HIV [9]. These cellular systems are also increasingly used to map the genome-wide determinants of gene expression [10].

GWAS are also advocated by the field of personalized medicine. One field that is already delivering clinical applications is that of pharmacogenetics – as illustrated by the strong association between *HLA-B*5701* and hypersensitivity reactions to abacavir [11]. However, in contrast to the highly predictive nature of the *HLA-B*5701* test, much of the genetic data that will be generated through GWA or large-scale pharmacogenomic studies [12] will not be of great value in isolation, but are meant to be interpreted within the context of a predictive score. This reflects the complexity of translating markers identified through population studies – generally associated with small size effects – into reliable predictors for the individual.

How would such GWAS or pharmacogenetics analyses take place in practice? A patient would be array- or chip-genotyped for a few dozen markers that would indicate the likelihood of disease progression, and a general recommendation on the best drugs to use would result from the analysis [13]. To some, the lack of 'strong effects' such as those of *HLA-B*5701* appears to be a critical limit to the future use of genetic markers in clinical care. To others, this is a matter of completeness in our understanding of the genes involved, and of the development of appropriate interpretative algorithms of genetic data in the context of other biological markers, such as CD4, viremia and laboratory tests of toxicity.

What does the future hold for the GWAS field? Undoubtedly, additional markers will emerge from new genome studies and from the meta-analysis of available studies, thus providing increasing statistical power. Luckily, the joint analysis of genome data is greatly facilitated by the

highly standardized nature of the genetic data that is being generated [14]. In addition, analysis will be enriched by new data-mining tools, by the genome-wide investigation of expression polymorphisms [15,16], copy number variation [17] and by the search for evolutionary signatures [18] that might result from the long-term exposure of our species to retroviruses and other retroelements [19]. Genomics presents a new set of research tools – and a great potential for innovation.

Financial & competing interests disclosure

The author has no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

No writing assistance was utilized in the production of this manuscript.

Bibliography

1. Drazen JM, Phimister EG: Publishing genomewide association studies. *N. Engl. J. Med.* 357, 496 (2007).
2. Fellay J, Shianna KV, Ge D *et al.*: A whole-genome association study of major determinants for host control of HIV-1. *Science* 317, 944–947 (2007).
3. Lander ES, Linton LM, Birren B *et al.*: Initial sequencing and analysis of the human genome. *Nature* 409, 860–921 (2001).
4. Venter JC, Adams MD, Myers EW *et al.*: The sequence of the human genome. *Science* 291, 1304–1351 (2001).
5. The International HapMap Consortium: a haplotype map of the human genome. *Nature* 437, 1299–1320 (2005).
6. Frazer KA, Ballinger DG, Cox DR *et al.*: A second generation human haplotype map of over 3.1 million SNPs. *Nature* 449, 851–861 (2007).
7. Gibson G, Goldstein DB: Human genetics: the hidden text of genome-wide associations. *Curr. Biol.* 17, R929–R932 (2007).
8. Telenti A, Goldstein DB: Genomics meets HIV-1. *Nat. Rev. Microbiol.* 4, 865–873 (2006).
9. Loeuillet C, Deutsch S, Ciuffi A *et al.*: *In vitro* whole genome analysis identifies a susceptibility locus for HIV-1. *PLoS Biol.* 6, e32 (2008).
10. Stranger BE, Nica AC, Forrest MS *et al.*: Population genomics of human gene expression. *Nat. Genet.* 39, 1217–1224 (2007).
11. Mallal S, Phillips E, Carosi G *et al.*: PREDICT-1: a novel randomised prospective study to determine the clinical utility of *HLA-B*5701* screening to reduce abacavir hypersensitivity in HIV-1 infected subjects (study CNA106030). Presented at: *4th International AIDS Conference on HIV Pathogenesis, Treatment and Prevention*. Sydney, Australia, 22–25 July 2007 (Abstract WESS101).
12. Lubomirov R, Csajka C, Telenti A: ADME pathway approach for pharmacogenetic studies of anti-HIV therapy. *Pharmacogenomics* 8, 623–633 (2007).
13. Telenti A, Zanger UM: Pharmacogenetics of anti-HIV drugs. *Annu. Rev. Pharmacol. Toxicol.* 48, 227–256 (2008).
14. Manolio TA, LL Rodriguez, Brooks L *et al.*: New models of collaboration in genome-wide association studies: the Genetic Association Information Network. *Nat. Genet.* 39, 1045–1051 (2007).
15. Göring HHH, Curran JE, Johnson MA *et al.*: Discovery of expression QTLs using large scale transcriptional profiling in human lymphocytes. *Nat. Genet.* 39, 1208–1216 (2007).
16. Dixon AL, Liang L, Moffatt MF *et al.*: A genome-wide association study of global gene expression. *Nat. Genet.* 39, 1202–1207 (2007).
17. Eichler EE, Nickerson DA, Altshuler D *et al.*: Completing the map of human genetic variation. *Nature* 447, 161–165 (2007).
18. Sabeti PC, Varilly P, Fry B *et al.*: Genome-wide detection and characterization of positive selection in human populations. *Nature* 449, 913–918 (2007).
19. Telenti A: Adaptation, co-evolution, and human susceptibility to HIV-1 infection. *Infect. Genet. Evol.* 5, 327–334 (2005).

Affiliation

- Amalio Telenti
University Hospital Center & University of Lausanne, Institute of Microbiology, CHUV 1011 Lausanne, Switzerland
Tél.: +41 21 314 0550;
Fax: +41 21 314 4060;
amalio.telenti@chuv.ch