Published in final edited form as:

Trends Genet. 2010 March ; 26(3): 119–131. doi:10.1016/j.tig.2010.01.002.

Host genes associated with HIV/AIDS: advances in gene discovery

Ping An and Cheryl A. Winkler*

¹Laboratory of Genomic Diversity, SAIC-Frederick, Inc., National Cancer Institute- Frederick, Frederick, MD

Abstract

Twenty five years after the discovery of HIV as the cause of AIDS, there is still no effective vaccine and no cure for this disease. HIV susceptibility shows a substantial degree of individual heterogeneity, much of which can be conferred by host genetic variation. In an effort to discover host factors required for HIV replication, identify critical pathogenic pathways, and reveal the full armament of host defenses, there has been a shift from candidate gene studies to unbiased genome wide genetic and functional studies. However, the number of securely identified host factors involved in HIV disease remains small, explaining only ~15–20% of the observed heterogeneity – most of which is attributable to HLA. Multidisciplinary approaches integrating genetic epidemiology to systems biology will be required to fully understand viral-host interactions to effectively combat HIV/AIDS.

Combating HIV/AIDS through host genetics

Identification of genes that affect susceptibility and resistance to HIV-1 is key to unravelling HIV-host interactions for drug and vaccine development. In 1996, a 32-base-pair deletion (*CCR5* 32) was discovered that provided near complete protection against HIV infection in homozygotes [1, 2]. This was the first convincing evidence that transmitted strains of HIV preferentially use the CCR5 co-receptor for cell entry and led to the development of new class of anti-HIV drugs inhibiting entry of HIV into the cell [1]. Recently, an HIV-1-positive patient was cured after the transplantation of *CCR5* 32/ 32 stem cells [3]. This isolated case illustrates the potential of translational genetic research to combat HIV/AIDS. Unlike HIV which can develop resistance through escape mutations, host targets are genetically stable. The search for host genetic restriction factors has intensified and expanded as a practical cure for HIV infection or an effective vaccine has failed to materialize 25 years after the discovery of HIV.

Here we review the insights gained into genetic variation in the host genome from a decade of candidate gene analysis and also from recent advances based on high throughput genome-wide scans utilizing single nucleotide polymorphism (SNP) arrays and small interfering RNA (siRNA) approaches.

© 2010 Elsevier Ltd. All rights reserved.

*To whom reprint requests should be addressed. winklerc@mail.nih.gov.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Epidemiologic evidence of genetic restriction of HIV/AIDS

There is considerable heterogeneity in viral control and progression rates (Figure 1), which not fully explained by environmental or viral factors. A subset of exceptional individuals deviate from the expected response to HIV exposure: (i) Exposed uninfected (EU) individuals show resistance to HIV acquisition even after multiple, high risk exposures (reviewed in [1, 4]); (ii) Long- term non-progressors (LTNP) maintain stable CD4 levels and low viral load (VL) for ten or more years; (iii) Fast progressors who cannot control viremia and develop AIDS within three years of infection [1]; and (iv) Elite controllers (EC), representing just 1% of HIV-infected persons, control HIV replication to <50 copies/mL [5]. Extensive studies have found that many EC are infected with replication-competent virus [6], pointing to a differences in host genetic background in HIV control. People with extreme phenotypes are highly informative in genetic association studies because they can be enriched for risk or protective alleles (Figure 2 and Box 1).

Study groups for HIV host genetic investigation

HIV-1 natural history cohorts established during 1980s enrolled subjects from the major HIV risk groups at the time: intravenous drug users, persons with hemophilia, and men who have sex with men (MSM) (Figure 2, Table S1 in the online supplementary material). Owing to the early demographics of HIV infection in the USA and Europe, these studies are over-represented by men of European descent infected with subtype B virus—not at all representative of the current global epidemic where more than half the HIV infections occur in sub-Saharan Africa with HIV subtype C. However, these studies remain of value because they are prospective and represent the natural history of HIV infection before the advent of effective HIV therapy and include HIV seroconverters (with dates of last negative and first positive HIV test). Table S1 lists the major HIV studies that have been used in genetic association studies.

Host genes associated with HIV/AIDS

Using a combination of candidate gene analysis (CGA) and unbiased genomewide association scans (GWAS), genetic variation has been identified that explain in part the observed variation in HIV-1/AIDS. These genes mainly fall into three classes: innate and adaptive immunity; HIV dependency factors (HDF) required by HIV for replication; and intrinsic anti-viral restriction factors. Genes identified by CGA are listed in Table 1.

The palace guards: HLA and KIR

HLA class I alleles were the first host genetic factors identified to affect AIDS. *HLA* class I genes encode cell surface molecules that differentially present viral antigenic epitopes to CD8+ T cytotoxic lymphocytes (CTL). Specific *HLA* Class I alleles, by differential HIV epitope binding, influence the effectiveness of immune response thus affecting HIV progression (reviewed in Refs [7, 8]). *HLA* homozygosity for 1, 2, or 3 class I loci were positively correlated to disease progression with individuals homozygous at *HLA-A*, *-B*, and *-C* showing the shortest AIDS-free survival [9]. The *HLA-B**35 Px allele is dominantly associated with more rapid progression to AIDS, presumably because of a limited repertoire of HIV epitope recognition [10]. However, the *HLA-B**57 and *B**27 alleles have been consistently associated with delayed disease progression [1, 8]. Studies of the effects of *HLA-B* alleles show that carriage of a protective *HLA-B* allele by either the child or the mother was associated with delayed progression to AIDS [11]; however, sharing of *HLA* alleles between the mother and child or maternal *HLA* class I homozygosity increased the risk of mother-to-child HIV transmission [11, 12].

Killing of infected cells by CD8⁺ CTL, following CTL recognition of HIV epitopes presented by HLA Class I receptors, is essential for immune control of HIV; mutation of the presented epitopes allows HIV to escape this control. Plausibly, *HLA* Class I alleles are susceptible or protective to the extent that they make this immune escape easier or more difficult. *HLA* Class I homozygosity decreases the number of *HLA* alleles, decreasing the number of epitopes presented and the extent of mutation needed to escape antigen presentation. However, *HLA-B*27* is notably nonselective in antigen presentation, making it harder for HIV epitopes to lose binding affinity by mutation. *HLA-B*57* is protective because the viral mutation required for the primary epitope to escape significantly decreases viral fitness. The specific mutational processes involved in HIV escape from *HLA-B*27* and *HLA-B*57* antigen presentation have been observed *in vivo* [13]. *HLA-B*5701* is also associated with severe hypersensitivity to abacavir, a protease inhibitor used in antiretroviral therapy. Genetic testing for *HLA-B*57* is required before prescribing Abacavir [14].

The intersection between innate and acquired immunity is manifested by interactive effects of HLA and the killer immunoglobulinlike receptors (KIR) on HIV disease; this complex subject has been reviewed by Carrington et al.[7]. An adaptive strategy shown by HIV and other viruses is to downregulate HLA Class I expression, thus blocking presentation of viral epitopes and CTL killing of the infected cells. One function of natural killer (NK) cells is to counter this strategy by eliminating cells which fail to display correct levels of Class I receptors, and one function of the NK cells' KIR is to determine whether the potential target cells carry the proper set of HLA receptors. In a study of ~1000 HIV seroconverters the activating KIR3DS1 receptor increased susceptibility to AIDS progression in the absence of its presumed ligand, a subset of the *HLA-Bw4* group of alleles, but was strongly protective in the presence of one or more of these *HLA-B* alleles [7, 8].

Modulation of the immune response to HIV-1 infection is critical in controlling HIV replication and in HIV-1 pathogenesis. A gene expression study in HIV-uninfected and -infected patients showed similar profiles in expression levels from CD8⁺ cells from both uninfected persons and HIV-infected individuals with control of viremia, but a distinctly different profile in early infection that persisted in chronic progressors, characterized by up-regulation of interferon-inducible genes [15]. Polymorphism in regulatory and coding regions of immune regulatory genes have been associated HIV acquisition and HIV progression (e.g. interferon regulatory factor 1 (*IRF-1*) [16], *IL10* [1], interferon gamma (*IFNG*) [1], IL8 receptor (*CXCR1*) [17]. Linking transcriptional profiles to underlying genetic variation might identify key players in determining normal from aberrant interferon responses following HIV infection.

Keepers of the gate: the chemokine-receptor nexus and HIV-cell entry

The chemokine receptors and their ligands have been extensively investigated since its discovery in 1996 as having a key role in HIV cell entry. HIV-1 enters cells by binding to CD4 and one of two major cell surface coreceptors: CCR5 used by transmitted forms of HIV-1 strains (R5) or CXCR4 used by X4 strains that arise in about 50% of individuals late during infection (Figure 3). The chemokines CCL5 (RANTES), CCL3 (Mip1⁻) and CCL4 (Mip1⁻) are ligands of CCR5 and competitively inhibit R5 HIV cell entry. A 32 base pair deletion (*CCR5*⁻³²) results in a truncated CCR5 protein that is not expressed on the cell surface [2, 18]. Only 1% of northern Europeans carry the *CCR5*⁻³² homozygous genotype—these individuals enjoy near complete protection against HIV infection [18]. Although *CCR5*⁻³² heterozygosity does not prevent infection, it does slow progression in both untreated [2] and treated—this protective effect is particularly strong in persons on antiretroviral therapy (ART) [19]. Genetic variation in the promoter region and infrequent nonsynonymous mutations in *CCR5* also modify progression rates [20-22]. The *CCR5*⁻³² to date is the only genetic mutation that completely blocks HIV infection in humans.

A compelling candidate gene is the Duffy Antigen Receptor for Chemokines (DARC) that binds HIV to the surfaces of erythrocytes and also binds pro-inflammatory cytokines and chemokines. A null mutation (*FY*-46T C) that is nearly fixed in subSaharan Africa but is absent elsewhere provides resistance to *Plasmodium vivax* and *P. knowlesi* malaria. A provocative study reported that the *DARC* null genotype was associated with a 40% increase in HIV acquisition and might account for up to 11% of the excess burden of HIV in Africa [23]. Subsequent studies have failed to replicate these results, probably due to inadequate control of population substructure in the original report [24-27].

Chemokines that bind HIV-1 coreceptors potentially inhibit HIV-1 replication in vitro and have been shown in multiple genetic association studies to modify HIV infection and pathogenesis. Gene expression of *CCL5*, a CCR5 ligand, is modulated by interacting regulatory SNPs [28, 29]. Down-regulation of *CCL5* levels by these regulatory SNPs are associated with HIV acquisition, progression to AIDS, and higher viral load in both Asian and European populations [28][30]. [31] [32] [33] [34]. These detrimental effects continue even after antiviral treatment [19]. An up-regulating allele (-28G) is associated with favorable outcomes in Asian populations [29, 32, 35]. Supporting the benefit of higher *CCL5* levels in resistance to HIV acquisition, *CCL5* levels in the genital tracts of EU commercial sex women were found to be 10-fold greater than that of un-exposed uninfected controls [36], attesting to the protective effects of abundant CCR5 ligand levels that compete with HIV for binding sites.

Copy number variation (CNV) is predicted to have a major role in human diseases. *CCRL3L1*, a ligand for CCR5, is encoded by 1 to 7 duplicated genes. It was reported that lower copy number of *CCL3L1* is risk factor for HIV acquisition and AIDS [37] due to lower levels of *CCL3L1* available to bind *CCR5*. Recently, three well-powered, independent studies failed to replicate these results using improved methods to quantify gene copy number [38-40].

The sentry guards: intrinsic HIV-1 restriction factors

The mammalian APOBEC3 protein family has emerged as a key mediator of intrinsic restriction to retroviruses including HIV and HBV (Reviewed in Chiu and **Greene** [41]). APOBEC3 proteins, a class of cytidine deaminase enzymes, are incorporated into virions. APOBEC3G edit newly synthesized viral DNA by deaminating dC to dU resulting in lethal G-to-A hypermutations. The anti-viral activity of APOBEC3 can also be deaminase-independent, through interference with viral transcription or integration. HIV has clearly out-manuevered its host—HIV-1 encoded viral infectivity factor (Vif) targets APOBEC3G and 3F for proteasomal degradation through an ubiquitination pathway using the host cellular proteins Cullin5, elongins B and C, and Rbx1 [42] (Figure 3). Seven *APOBEC3* family genes (A, B, C, DE, F, G and H) are located in tandem array on chromosome 22q12-q13.2 that spans 150 kb. Their anti-HIV activity and sensitivity to HIV-vif mediated degradation is variable (Table 2).

The role of *APOBEC3G* genetic variants on HIV-1 infection and progression has been extensively investigated [43-45]. The H186 R allele is associated with CD4 T cell decline and AIDS in homozygotes in Africa Americans [44], with high HIV-1 viral load in Africans with acute infection [46]. Multiple APOBEC3G haplotypes comprising promoter and intronic SNPs influence HIV progression [44]. An intronic SNP C40693T was found to increase risk to HIV-1 infection [43]. Because APOBEC3G is restricted by HIV-vif, it is unlikely that *APOBEC3G* variation will have a strong influence unless it increases APOBEC3G resistance to HIV-vif mediated degradation.

The impact of the *APOBEC3G* expression levels on HIV-1 infection and severity of HIV disease has been investigated. Most studies reported a positive correlation between expression levels and favorable outcomes [45, 47-49]. Higher *APOBEC3G* expression in PBMC and cervical tissues was observed in HIV-exposed seronegative women; their PBMC were also resistant to infection with R5 HIV strains [47]. *APOBEC3G* mRNA levels range from high to low in non-progressors > exposed uninfected > Progressors, respectively, and were inversely correlated with viral load and positively correlated with CD4⁺ T cell levels [50]. Taken together, the evidence indicate that *APOBEC3* genetic variation and expression levels modify HIV acquisition and disease progression *in vivo* suggesting that delivery of HIV-vif resistant APOBEC3 protein to HIV-1 susceptible cells could potentially restrict HIV replication.

A complete deletion of the *APOBEC3B* gene is common with allele frequencies ranging from greater than 27% (Asians) to 4% or less in Africans [51]. APOBEC3B is resistant to HIV-vif mediated degradation and shows moderate to strong anti-HIV activity. A recent study of the effect of the gene deletion on HIV-1/AIDS suggested that the null genotype was associated with greater risk of HIV acquisition, higher viral load, and shorter AIDS-free survival; no effect on HIV/AIDS was noted for heterozygotes [52]. This study warrants further investigation in other populations to confirm the association.

The interaction of HIV-1 Vif and the host factor Cullin5 (*CUL5*) is critical in disarming APOBEC3 anti-HIV activity. *CUL5* SNPs are strongly associated with CD4 T cell loss, one which differentially binds nuclear proteins suggesting altered regulation of *CUL5* transcription [53]. A significant interaction was also observed between *CUL5* SNPs and *APOBEC3G*-186R [53]. Cullin5 and APOBEC3G collectively contribute to HIV disease *in vivo* and provide support for the notion that increasing cellular levels of APOBEC3 protein by interfering with HIV-1 vif-cul5 mediated degradation by drugs targeting this pathway would have therapeutic benefit [53].

TRIM5

TRIM5 play a critical role in the primate anti-viral defense system (Figure 3). Trim5 in the Rhesus monkey completely blocks HIV-1 infection whereas human TRIM5 mediates a low level inhibition to HIV-1 replication [54]. The mechanism of HIV-1 restriction is not fully understood. Proposed models include binding of a multimer of TRIM5 to the incoming viral capsid that leads to premature uncoating of the capsid or by interaction between TRIM5 and the capsid with other undefined cofactors (e.g., CypA) [55]. Using an *in vitro* assay system, cells were transduced with different *TRIM5* missense polymorphisms; R136Q enhanced HIV restriction while H43Y, V112F, and G110E relaxed restriction on HIV replication [56-59] or had no effect (G249D, H419Y) [56-58] (Table 3).

Because *TRIM5* is a compelling candidate gene numerous associations have been done to determine if naturally occurring polymorphisms affects HIV acquisition or progression [56, 57, 59-61]. The results of these studies have been inconsistent, due to the low population frequencies of the missense mutations or small numbers of subjects. One study [61] found a rare haplotype (1.0%) carrying a 5' untranslated region allele (-2G) and 136Q was significantly higher in HIV-infected persons compared to EU, but this was not replicated in a second study [56]. Notably, the ancestral haplotype carrying more wildtype alleles for 12 SNPs was associated with HIV resistance [56]. SNP 43Y was associated with reduced susceptibility to HIV-1 infection in Asians, a result compatible with the *in vitro* studies (Table 2) [57]

Even a moderate over expression of wild-type human *TRIM5* provides substantial restriction to infection for HIV-1 [62]. It is conceivable that the alleles regulating TRIM5 protein

levels can have a major effect on HIV acquisition or HIV replication. Although *TRIM5* missense polymorphisms explain little of the variation in European descent populations tested to date, *TRIM5* expression might be important in HIV acquisition and should be investigated in diverse human populations. A South African study showed an inverse correlation between Human *TRIM5* expression levels in PBMC and susceptibility to HIV infection; HIV-infected persons had lower levels of *TRIM5* expression in PBMC compared from EU [63].

Aiding the enemy: cyclophilin facilitates infection

Unlike *TRIM5*, Cyclophilin A (CypA) encoded by *PPIA* is a cellular protein that increases viral infectivity by facilitating proper uncoating of the viral capsid (CA) (Figure 3) [55]. The genetic effects of *PPIA* stem from regulatory SNPs as no coding SNPs have been found. The promoter SNP1604G affects binding to transcription factor [64], increases HIV-1 infectivity in CD4 T cells [65], and was associated with the faster progression in Swiss and USA cohorts [65] and higher viral RNA levels in drug users in the Amsterdam Cohort [66]. Targeting CA–CypA interaction comprises another anti-HIV approach that is under development [67].

Genome-wide association scans (GWAS)

Although candidate gene studies have been used to identify several AIDS restriction genes (Table 1), GWAS have identified novel SNPs not previously implicated in HIV/AIDS. Four GWAS have been published to date on European or European descent individuals (Table 4) [68–71]. The Limou and Le Clerc studies used selected individuals at the tails of the distribution for AIDS-free survival—very rapid progressors and long term survivors [68, 69]. The Fellay study included participants with longitudinal data and used viral set-point or progression to CD4<350 as outcomes [71] and the Dalmasso GWAS assessed SNP association with plasma HIV-RNA and PBMC HIV-DNA levels during the primary infection [70]. Although the study designs were quite different, the *HLA* region emerged as the major influence on HIV/AIDS.

Few of SNP associations in GWAS studies reached a genome-wide significance of $P < 5 \times 10^{-8}$, but three of the four studies found strong association with rs2395029 (which is near HLA complex P5 (HCP5), a gene localized within the MHC class I region, but is not structurally related to MHC class I genes) [68, 70, 71], rs9264942 (associated with expression levels of *HLA-C*) and two found associations with *ZNDR1* (Zinc ribbon domain-containing protein 1, a DNA-dependent RNA polymerase) [68, 71]; these results clearly implicate chromosome 6p21.33 *HLA* region as having a major role in control of HIV. *HCP5* is near absolute linkage disequilibrium (LD) with *HLA-B*57* but the *HLA-C* SNP does not seem to be tracking through LD other *HLA* alleles. Although *HLA-B*57* is a well-known HIV restriction gene, *HLA-C* variation was not previously known to have a role in HIV/AIDS. *HCP5/HLA-B*57* and *HLA-C* explain 14% of the variance in HIV setpoint—a major predictor of AIDS. Forty-four of the top 100 SNPs strongest association in the Fellay [71] study were in the 6p21 *HLA* region attesting to the importance of HLA.

Other SNP associations with progression rates or viral load were observed in the various studies but were short of genome-wide significance and not replicated among the four GWAS (Table 4). Among the top 100 SNPs in the Fellay study [71] were several biologically plausible genes, including *DICER1* (generating miRNA and siRNA, interacting with HIV tat), *TRIM39* (same family as anti-HIV factor *TRIM5*), *IF* (which regulates the complement cascade), and *PROX1* (a negative regulator of IFN-gamma expression in T cells). The role of these SNPs is not yet determined—if these associations are replicated, they will provide new insights into HIV pathogenesis as none of these genes fall in the *HLA*

region or have been previously implicated in HIV/AIDS. Indeed, a multistage GWAS performed in > 700 US seroconverters has recently replicated the strong protective role of *PROX1* in AIDS progression [72].

HIV/AIDS GWAS using different study designs and phenotypes have been remarkably consistent in identifying variant alleles in the HLA region as major determinants of HIV viral load and progression. However, they have been less rewarding in revealing new or unexpected alleles or replicating previously published candidate gene associations. This is likely due to limitations inherent to the method. Because of the large number of comparisons and because there is no a priori hypothesis, stringent corrections for multiple tests are required to avoid false positive (type 1) errors—but this comes at the cost of increasing false negative (type 2) errors. The current standard for genome wide significance in GWAS is $<5 \times 10^{-8}$ for $p < 0.05$ or 1×10^{-8} for $p < 0.01$. This stringency provides assurance that implicated SNPs are probable true positives but small to moderate effect alleles, low frequency alleles, and/or recessive alleles are unlikely to be detected simply because of a lack of power due to sample size or to a loss of power due to genotyping or phenotyping errors. With sample sizes less than 1000 in published HIV/AIDS GWAS, only common alleles with moderate to large effect sizes would reach genome wide significance. The published GWAS scans suggest that larger effect alleles other than HLA controlling viral replication are less likely. Future GWAS will most likely identify multiple small effect polymorphisms that explain only a portion of the observed variance but might provide needed insight into biological pathways and provide useful insight for drug or vaccine development. It is also likely that host genetics will explain only a portion of the observed variance in HIV replication and pathogenesis with the remainder being explained by HIV and environmental factors.

The problem of identifying true positive associations among the sea of false negatives is daunting but not hopeless. Meta-analysis using combined phenotype-genotype data from multiple studies should identify new associations that fail to reach genome wide significance in any single study. Collaborative initiatives among investigators with DNA and clinical data are also critical in achieving adequate power and to validate associations through replication. As shown in Table S1, there are numerous case-control and longitudinal HIV/AIDS study groups with a combined enrollment of nearly 20 000 participants. With careful integration of phenotype data plus stringent assessment and correction of population substructure, it should be possible to identify and validate genetic predictors of HIV transmission and acquisition (not yet explored by GWAS); control of viral replication, kinetics of progression, response and adverse events to antiretroviral therapies, and specific AIDS-defining outcomes. In this era of highly effective antiretroviral treatment (HAART) leading to longer AIDS-free survival, there are increasing reports of non-AIDS-defining events (NADE) that include renal disease, diabetes, cardiovascular diseases, and metabolic syndromes leading to morbidity and mortality in HAART-treated cohorts. Predictors of NADEs are also amenable to discovery by GWAS, but, with the exception of HIV-associated nephropathy (HIVAN), have not yet been explored by genome-wide genetic studies.

Admixture mapping identifies *MYH9* as a major susceptibility gene for HIV-associated nephropathy

African descent individuals have long been recognized to having a strikingly increased risk for HIV-associated nephropathy; we have previously estimated at 20-fold compared to European descent individuals [73]. Studies of adults and children in the pre-ART era suggest that approximately 10% African Americans developed HIVAN, a leading cause of end stage kidney disease in African Americans. The predilection of African descent individuals to develop HIVAN as well as other major forms of chronic and end stage renal disease

suggested that a renal susceptibility locus would be more frequent on an African-origin chromosome and less frequent in non-African-origin chromosomes. Kopp et al. used a genome-wide ancestry admixture mapping strategy to identify a region of excess African ancestry (92%) on chromosome 22 compared to either the genome-wide average of 81% African ancestry or compared to the controls at the same chromosome 22 region [74]. Fine mapping identified *MYH9*, encoding non-muscle myosin IIA, to be strongly associated with HIV-associated nephropathy (HIVAN) in African Americans (odds ratio 5-7 range, $p < 10^{-8}$, attributable risk = 100%). It is notable that in the Kopp et al. study, every individual with HIVAN carried at least one risk haplotype and the number of homozygotes for the risk haplotype was 76% in the HIVAN cases compared to 36% in the controls. The risk alleles are extremely common in African descent individuals ($\approx 60\%$) and infrequent in non-African descent individuals ($< 4\%$) thus explaining a major USA and global health disparity [74]. The prevalence of *MYH9*-related kidney disease in sub-Saharan Africa is unknown but might be substantial in a setting of untreated HIV-1 infection.

Genome-wide siRNA functional scan

siRNA and small hairpin (sh)RNA knock-down of gene transcription providing a powerful tool to identify genes required for HIV-1 replication [75]. Three studies have used siRNA to knock down one by one $\sim 20\,000$ genes and then measured the ability of the cells to support transient HIV infection (Table 5). The fourth study [76] used shRNA to chronically silence each of 54 509 mRNAs in Jurket T cell clones (Table 5). Following infection, HIV-1 infected clones that survived represented silenced mRNA required for HIV-replication since active HIV replication leads to Jurket cell death. Notably, although each identified > 250 genes as HIV dependency factors, only three genes (*MED6*, mediator complex subunit 6 that involves in RNA polymerase II transcription, *MED7*, a cofactor required for Sp1 transcriptional activation and *RELA*, a component of NF-kappa B complex) overlapped among the four studies and 40 overlapped between 2 out of 4 studies; however, similar pathways were identified by each (Table 6) [75-78].

Differences in gene expression among the different cell lines, HIV strain, timing of infection to assay end-point, as well as off-target gene silencing might account the low level of replication among the RNAi studies. It is also possible that HIV is promiscuous in co-opting host proteins for its life cycle. The results of these studies are hypothesis generating and require further validation; however, they have the potential to identify novel host targets for anti-HIV drug design that, unlike viral genes, cannot develop escape mutations and might not be required for cell survival.

Other high-throughput approaches on functional genomic cDNA screening [79], mRNA profiling [80], proteomic [81], and microRNA (miRNA) [82] have also been applied to unravel the host factors in HIV pathogenesis (Table 5). Notably type III RNases Dicer and Drosha, PCAF (a histone acetylase associated with p300/CBP nuclear proteins) in the miRNA -silencing machinery was implicated in HIV-1/AIDS [82].

Concluding remarks

After more than two decades of genetic association studies using both candidate and more recently genome-wide association analysis, only a fraction of the variation ($\approx 15\text{-}20\%$) in HIV progression [1] and control of viral set-point [71] has been explained by genetic polymorphisms. The most important predictors are common variation in the chemokine-chemokine nexus involved in cell entry, *HLA* class I alleles involved in host defenses, and *MYH9* for HIVAN.

Figure 4 shows the overlap between different approaches for gene identification. At first glance the lack of replication among the methods is unsettling; however, these methods are each complementary (Table 6, Box 2). The GWAS studies enable comparisons the strength of association across the entire genome and to access the role of both intracellular and extracellular factors *in vivo*—and to provide quantification of attributable and relative risks important to clinical and public health policies. Functional and positional candidate gene studies have been remarkable successful at identifying rare and common genetic variation associated with HIV/AIDS through direct gene interrogation. Goldstein [83] has argued for a ‘shift from genome scans of ever larger samples to studies of rarer variants of larger effect’. With the advent of next generation sequencing and the careful selection of highly informative individuals from the extremes of the phenotype distribution (e.g. HIV set-point or rate of CD4+ T cell decline), it should be possible to identify rare variants in genes and pathways involved HIV pathogenesis or HIV replication. It should be pointed out that all the GWAS to date have had small samples sizes and were underpowered to detect small effects or infrequent alleles. International consortium pooling resources for large GWAS should identify additional genetic variation associated with HIV/AIDS.

The over-arching goals to prevent new HIV infections and cure established infections require an interdisciplinary, systems-based approach (Figure 5). Integrating knowledge gained from large scale genomic, transcriptomic, proteomic studies using computational modeling should lead to a comprehensive view of host-virus interactions and to the identification of vulnerable pathways that can be exploited for preventive or therapeutic interventions.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are in debt to Dr. George Nelson for many discussions and critical reading of the manuscript. This project has been funded in whole or in part with federal funds from the National Cancer Institute, National Institutes of Health, under contract HHSN261200800001E. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government. This Research was supported [in part] by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research.

References

1. O'Brien SJ, Nelson GW. Human genes that limit AIDS. *Nature genetics*. 2004; 36:565–574. [PubMed: 15167933]
2. Dean M, et al. Genetic restriction of HIV-1 infection and progression to AIDS by a deletion allele of the CCR5 structural gene. Hemophilia Growth and Development Study, Multicenter AIDS Cohort Study, Multicenter Hemophilia Cohort Study, San Francisco City Cohort, ALIVE Study. *Science*. 1996; 273:1856–1862. [PubMed: 8791590]
3. Hutter G, et al. Long-term control of HIV by CCR5 Delta32/Delta32 stem-cell transplantation. *N Engl J Med*. 2009; 360:692–698. [PubMed: 19213682]
4. Shacklett BL. Understanding the “lucky few”: the conundrum of HIV-exposed, seronegative individuals. *Current HIV/AIDS reports*. 2006; 3:26–31. [PubMed: 16522256]
5. Walker BD. Elite control of HIV Infection: implications for vaccines and treatment. *Top HIV Med*. 2007; 15:134–136. [PubMed: 17720999]
6. Miura T, et al. Genetic characterization of human immunodeficiency virus type 1 in elite controllers: lack of gross genetic defects or common amino acid changes. *J Virol*. 2008; 82:8422–8430. [PubMed: 18562530]

7. Carrington M, et al. KIR-*HLA* intercourse in HIV disease. *Trends in microbiology*. 2008; 16:620–627. [PubMed: 18976921]
8. Carrington M, O'Brien SJ. The influence of HLA genotype on AIDS. *Annual review of medicine*. 2003; 54:535–551.
9. Carrington M, et al. HLA and HIV-1: heterozygote advantage and B*35-Cw*04 disadvantage. *Science*. 1999; 283:1748–1752. [PubMed: 10073943]
10. Gao X, et al. Effect of a single amino acid change in MHC class I molecules on the rate of progression to AIDS. *N Engl J Med*. 2001; 344:1668–1675. [PubMed: 11386265]
11. Thobakgale CF, et al. Impact of HLA in mother and child on disease progression of pediatric human immunodeficiency virus type 1 infection. *J Virol*. 2009; 83:10234–10244. [PubMed: 19605475]
12. Mackelprang RD, et al. Maternal HLA homozygosity and mother-child HLA concordance increase the risk of vertical transmission of HIV-1. *J Infect Dis*. 2008; 197:1156–1161. [PubMed: 18462163]
13. Boutwell CL, et al. Reduced viral replication capacity of human immunodeficiency virus type 1 subtype C caused by cytotoxic-T-lymphocyte escape mutations in HLA-B57 epitopes of capsid protein. *J Virol*. 2009; 83:2460–2468. [PubMed: 19109381]
14. Colombo S, et al. The HCP5 single-nucleotide polymorphism: a simple screening tool for prediction of hypersensitivity reaction to abacavir. *J Infect Dis*. 2008; 198:864–867. [PubMed: 18684101]
15. Hycza MD, et al. Distinct transcriptional profiles in ex vivo CD4+ and CD8+ T cells are established early in human immunodeficiency virus type 1 infection and are characterized by a chronic interferon response as well as extensive transcriptional changes in CD8+ T cells. *J Virol*. 2007; 81:3477–3486. [PubMed: 17251300]
16. Ball TB, et al. Polymorphisms in IRF-1 associated with resistance to HIV-1 infection in highly exposed uninfected Kenyan sex workers. *AIDS*. 2007; 21:1091–1101. [PubMed: 17502719]
17. Vasilescu A, et al. A haplotype of the human CXCR1 gene protective against rapid disease progression in HIV-1+ patients. *Proc Natl Acad Sci U S A*. 2007; 104:3354–3359. [PubMed: 17360650]
18. Ioannidis JP, et al. Effects of CCR5-Delta32, CCR2-64I, and SDF-1 3 A alleles on HIV-1 disease progression: An international meta-analysis of individual-patient data. *Annals of internal medicine*. 2001; 135:782–795. [PubMed: 11694103]
19. Hendrickson SL, et al. Host genetic influences on highly active antiretroviral therapy efficacy and AIDS-free survival. *J Acquir Immune Defic Syndr*. 2008; 48:263–271. [PubMed: 18391751]
20. McDermott DH, et al. CCR5 promoter polymorphism and HIV-1 disease progression. Multicenter AIDS Cohort Study (MACS). *Lancet*. 1998; 352:866–870. [PubMed: 9742978]
21. Martin MP, et al. Genetic acceleration of AIDS progression by a promoter variant of CCR5. *Science*. 1998; 282:1907–1911. [PubMed: 9836644]
22. An P, et al. Influence of CCR5 promoter haplotypes on AIDS progression in African-Americans. *AIDS*. 2000; 14:2117–2122. [PubMed: 11061652]
23. He W, et al. Duffy antigen receptor for chemokines mediates trans-infection of HIV-1 from red blood cells to target cells and affects HIV-AIDS susceptibility. *Cell host & microbe*. 2008; 4:52–62. [PubMed: 18621010]
24. Horne KC, et al. Duffy antigen polymorphisms do not alter progression of HIV in African Americans in the MAC S cohort. *Cell host & microbe*. 2009; 5:415–417. author reply 418–419. [PubMed: 19454342]
25. Julg B, et al. Lack of Duffy antigen receptor for chemokines: no influence on HIV disease progression in an African treatment-naïve population. *Cell host & microbe*. 2009; 5:413–415. author reply 418–419. [PubMed: 19454341]
26. Walley NM, et al. The Duffy antigen receptor for chemokines null promoter variant does not influence HIV-1 acquisition or disease progression. *Cell host & microbe*. 2009; 5:408–410. author reply 418–409. [PubMed: 19454339]

27. Winkler CA, et al. Expression of Duffy antigen receptor for chemokines (DARC) has no effect on HIV-1 acquisition or progression to AIDS in African Americans. *Cell host & microbe*. 2009; 5:411–413. author reply 418–419. [PubMed: 19454340]
28. An P, et al. Modulating influence on HIV/AIDS by interacting RANTES gene variants. *Proc Natl Acad Sci U S A*. 2002; 99:10002–10007. [PubMed: 12114533]
29. Liu H, et al. Polymorphism in RANTES chemokine promoter affects HIV-1 disease progression. *Proc Natl Acad Sci U S A*. 1999; 96:4581–4585. [PubMed: 10200305]
30. Duggal P, et al. The effect of RANTES chemokine genetic variants on early HIV-1 plasma RNA among African American injection drug users. *J Acquir Immune Defic Syndr*. 2005; 38:584–589. [PubMed: 15793370]
31. Ahlenstiel G, et al. Distribution and effects of polymorphic RANTES gene alleles in HIV/HCV coinfection -- a prospective cross-sectional study. *World J Gastroenterol*. 2005; 11:7631–7638. [PubMed: 16437690]
32. Wichukchinda N, et al. Protective effects of IL4-589T and RANTES-28G on HIV-1 disease progression in infected Thai females. *AIDS*. 2006; 20:189–196. [PubMed: 16511411]
33. Liu XL, et al. [Preliminary study on the association of chemokine RANTES gene polymorphisms with HIV-1 infection in Chinese Han population]. *Zhonghua liu xing bing xue za zhi = Zhonghua liuxingbingxue zazhi*. 2003; 24:971–975. [PubMed: 14687494]
34. Rathore A, et al. Association of RANTES -403 G/A, -28 C/G and In1.1 T/C polymorphism with HIV-1 transmission and progression among North Indians. *Journal of medical virology*. 2008; 80:1133–1141. [PubMed: 18461620]
35. Koizumi Y, et al. RANTES -28G delays and DC-SIGN - 139C enhances AIDS progression in HIV type 1-infected Japanese hemophiliacs. *AIDS research and human retroviruses*. 2007; 23:713–719. [PubMed: 17530998]
36. Iqbal SM, et al. Elevated T cell counts and RANTES expression in the genital mucosa of HIV-1-resistant Kenyan commercial sex workers. *J Infect Dis*. 2005; 192:728–738. [PubMed: 16088822]
37. Gonzalez E, et al. The influence of CCL3L1 gene-containing segmental duplications on HIV-1/AIDS susceptibility. *Science*. 2005; 307:1434–1440. [PubMed: 15637236]
38. Urban TJ, et al. CCL3L1 and HIV/AIDS susceptibility. *Nat Med*. 2009; 15:1110–1112. [PubMed: 19812560]
39. Field SF, et al. Experimental aspects of copy number variant assays at CCL3L1. *Nat Med*. 2009; 15:1115–1117. [PubMed: 19812562]
40. Bhattacharya T, et al. Reply to: “CCL3L1 and HIV/AIDS susceptibility”. *Nat Med*. 2009; 15:1112–1115. [PubMed: 19812561]
41. Chiu YL, Greene WC. The APOBEC3 cytidine deaminases: an innate defensive network opposing exogenous retroviruses and endogenous retroelements. *Annual review of immunology*. 2008; 26:317–353.
42. Yu X, et al. Induction of APOBEC3G ubiquitination and degradation by an HIV-1 Vif-Cul5-SCF complex. *Science*. 2003; 302:1056–1060. [PubMed: 14564014]
43. Valcke HS, et al. APOBEC3G genetic variants and their association with risk of HIV infection in highly exposed Caucasians. *AIDS*. 2006; 20:1984–1986. [PubMed: 16988524]
44. An P, et al. APOBEC3G genetic variants and their influence on the progression to AIDS. *J Virol*. 2004; 78:11070–11076. [PubMed: 15452227]
45. Do H, et al. Exhaustive genotyping of the CEM15 (APOBEC3G) gene and absence of association with AIDS progression in a French cohort. *J Infect Dis*. 2005; 191:159–163. [PubMed: 15609224]
46. Reddy K, et al. APOBEC3G expression is dysregulated in primary HIV-1 infection and polymorphic variants influence CD4+ T-cell counts and plasma viral load. *AIDS*. 2010; 24:195–204. [PubMed: 19996938]
47. Biasin M, et al. Apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like 3G: a possible role in the resistance to HIV of HIV-exposed seronegative individuals. *J Infect Dis*. 2007; 195:960–964. [PubMed: 17330785]
48. Jin X, et al. APOBEC3G levels predict rates of progression to AIDS. *Retrovirology*. 2007; 4:20. [PubMed: 17374143]

49. Pace C, et al. Population level analysis of human immunodeficiency virus type 1 hypermutation and its relationship with APOBEC3G and vif genetic variation. *J Virol.* 2006; 80:9259–9269. [PubMed: 16940537]
50. Jin X, et al. APOBEC3G/CEM15 (hA3G) mRNA levels associate inversely with human immunodeficiency virus viremia. *J Virol.* 2005; 79:11513–11516. [PubMed: 16103203]
51. Kidd JM, et al. Population stratification of a common APOBEC gene deletion polymorphism. *PLoS genetics.* 2007; 3:e63. [PubMed: 17447845]
52. An P, et al. APOBEC3B Deletion and Risk of HIV-1 Acquisition. *J Infect Dis.* 2009; 200:1054–1058. [PubMed: 19698078]
53. An P, et al. Polymorphisms of CUL5 are associated with CD4+ T cell loss in HIV-1 infected individuals. *PLoS genetics.* 2007; 3:e19. [PubMed: 17257057]
54. Stremlau M, et al. The cytoplasmic body component TRIM5alpha restricts HIV-1 infection in Old World monkeys. *Nature.* 2004; 427:848–853. [PubMed: 14985764]
55. Sokolskaja E, Luban J. Cyclophilin, TRIM5, and innate immunity to HIV-1. *Current opinion in microbiology.* 2006; 9:404–408. [PubMed: 16815734]
56. Javanbakht H, et al. Effects of human TRIM5alpha polymorphisms on antiretroviral function and susceptibility to human immunodeficiency virus infection. *Virology.* 2006; 354:15–27. [PubMed: 16887163]
57. Nakajima T, et al. Impact of novel TRIM5alpha variants, Gly110Arg and G176del, on the anti-HIV-1 activity and the susceptibility to HIV-1 infection. *AIDS.* 2009; 23:2091–2100. [PubMed: 19710594]
58. Sawyer SL, et al. High-frequency persistence of an impaired allele of the retroviral defense gene TRIM5alpha in humans. *Curr Biol.* 2006; 16:95–100. [PubMed: 16401428]
59. Goldschmidt V, et al. Role of common human TRIM5alpha variants in HIV-1 disease progression. *Retrovirology.* 2006; 3:54. [PubMed: 16925802]
60. van Manen D, et al. The effect of Trim5 polymorphisms on the clinical course of HIV-1 infection. *PLoS pathogens.* 2008; 4:e18. [PubMed: 18248091]
61. Speelman EC, et al. Genetic association of the antiviral restriction factor TRIM5alpha with human immunodeficiency virus type 1 infection. *J Virol.* 2006; 80:2463–2471. [PubMed: 16474153]
62. Kaumanns P, et al. Human TRIM5alpha mediated restriction of different HIV-1 subtypes and Lv2 sensitive and insensitive HIV-2 variants. *Retrovirology.* 2006; 3:79. [PubMed: 17087820]
63. Sewram S, et al. Human TRIM5alpha expression levels and reduced susceptibility to HIV-1 infection. *J Infect Dis.* 2009; 199:1657–1663. [PubMed: 19388851]
64. An P, et al. Regulatory polymorphisms in the cyclophilin A gene, PPIA, accelerate progression to AIDS. *PLoS pathogens.* 2007; 3:e88. [PubMed: 17590083]
65. Bleiber G, et al. Use of a combined ex vivo/in vivo population approach for screening of human genes involved in the human immunodeficiency virus type 1 life cycle for variants influencing disease progression. *J Virol.* 2005; 79:12674–12680. [PubMed: 16188970]
66. Rits MA, et al. Polymorphisms in the regulatory region of the Cyclophilin A gene influence the susceptibility for HIV-1 infection. *PloS one.* 2008; 3:e3975. [PubMed: 19092998]
67. Li J, et al. Discovery of dual inhibitors targeting both HIV-1 capsid and human cyclophilin A to inhibit the assembly and uncoating of the viral capsid. *Bioorganic & medicinal chemistry.* 2009; 17:3177–3188. [PubMed: 19328002]
68. Limou S, et al. Genomewide association study of an AIDS-nonprogression cohort emphasizes the role played by HLA genes (ANRS Genomewide Association Study 02). *J Infect Dis.* 2009; 199:419–426. [PubMed: 19115949]
69. Le Clerc S, et al. Genomewide association study of a rapid progression cohort identifies new susceptibility alleles for AIDS (ANRS Genomewide Association Study 03). *J Infect Dis.* 2009; 200:1194–1201. [PubMed: 19754311]
70. Dalmaso C, et al. Distinct genetic loci control plasma HIV-RNA and cellular HIV-DNA levels in HIV-1 infection: the ANRS Genome Wide Association 01 study. *PloS one.* 2008; 3:e3907. [PubMed: 19107206]

71. Fellay J, et al. A whole-genome association study of major determinants for host control of HIV-1. *Science*. 2007; 317:944–947. [PubMed: 17641165]
72. Herbeck JT, et al. Multi-Stage Genome-Wide Association Study of HIV-1 Disease Progression to Clinical AIDS. *J Infect Dis*. In press.
73. Kopp JB, Winkler C. HIV-associated nephropathy in African Americans. *Kidney international*. 2003;S43–49. [PubMed: 12864874]
74. Kopp JB, et al. MYH9 is a major-effect risk gene for focal segmental glomerulosclerosis. *Nature genetics*. 2008; 40:1175–1184. [PubMed: 18794856]
75. Brass AL, et al. Identification of host proteins required for HIV infection through a functional genomic screen. *Science*. 2008; 319:921–926. [PubMed: 18187620]
76. Yeung ML, et al. A genome-wide short hairpin RNA screening of jurkat T-cells for human proteins contributing to productive HIV-1 replication. *J Biol Chem*. 2009; 284:19463–19473. [PubMed: 19460752]
77. Konig R, et al. Global analysis of host-pathogen interactions that regulate early-stage HIV-1 replication. *Cell*. 2008; 135:49–60. [PubMed: 18854154]
78. Zhou H, et al. Genome-scale RNAi screen for host factors required for HIV replication. *Cell host & microbe*. 2008; 4:495–504. [PubMed: 18976975]
79. Nguyen DG, et al. Identification of novel therapeutic targets for HIV infection through functional genomic cDNA screening. *Virology*. 2007; 362:16–25. [PubMed: 17257639]
80. Imbeault M, et al. Microarray study reveals that HIV-1 induces rapid type-I interferon-dependent p53 mRNA up-regulation in human primary CD4+ T cells. *Retrovirology*. 2009; 6:5. [PubMed: 19146679]
81. Chan EY, et al. Dynamic host energetics and cytoskeletal proteomes in human immunodeficiency virus type 1-infected human primary CD4 cells: analysis by multiplexed label-free mass spectrometry. *J Virol*. 2009; 83:9283–9295. [PubMed: 19587052]
82. Nathans R, et al. Cellular microRNA and P bodies modulate host-HIV-1 interactions. *Mol Cell*. 2009; 34:696–709. [PubMed: 19560422]
83. Goldstein DB. Common genetic variation and human traits. *N Engl J Med*. 2009; 360:1696–1698. [PubMed: 19369660]
84. Fauci AS, et al. Immunopathogenic mechanisms of HIV infection. *Annals of internal medicine*. 1996; 124:654–663. [PubMed: 8607594]
85. Bushman FD, et al. Host cell factors in HIV replication: meta-analysis of genome-wide studies. *PLoS pathogens*. 2009; 5:e1000437. [PubMed: 19478882]
86. Bashirova AA, et al. Consistent effects of TSG101 genetic variability on multiple outcomes of exposure to human immunodeficiency virus type 1. *J Virol*. 2006; 80:6757–6763. [PubMed: 16809281]
87. Catano G, et al. HIV-1 disease-influencing effects associated with ZNRD1, HCP5 and HLA-C alleles are attributable mainly to either HLA-A10 or HLA-B*57 alleles. *PloS one*. 2008; 3:e3636. [PubMed: 18982067]
88. Deeks SG, Walker BD. Human immunodeficiency virus controllers: mechanisms of durable virus control in the absence of antiretroviral therapy. *Immunity*. 2007; 27:406–416. [PubMed: 17892849]
89. Loeuillet C, et al. In vitro whole-genome analysis identifies a susceptibility locus for HIV-1. *PLoS biology*. 2008; 6:e32. [PubMed: 18288889]
90. van Manen D, et al. Association of HLA-C and HCP5 gene regions with the clinical course of HIV-1 infection. *AIDS*. 2009; 23:19–28. [PubMed: 19050382]
91. Lewden C, et al. Comparison of Early CD4 T-Cell Count in HIV-1 Seroconverters in Cote d'Ivoire and France: The ANRS PRIMO-CI and SEROCO Cohorts. *J Acquir Immune Defic Syndr*. 2009
92. Oh DY, et al. CCR5Delta32 genotypes in a German HIV-1 seroconverter cohort and report of HIV-1 infection in a CCR5Delta32 homozygous individual. *PloS one*. 2008; 3:e2747. [PubMed: 18648518]

93. van Loggerenberg F, et al. Establishing a cohort at high risk of HIV infection in South Africa: challenges and experiences of the CAPRISA 002 acute infection study. *PloS one*. 2008; 3:e1954. [PubMed: 18414658]
94. Land AM, et al. Human immunodeficiency virus (HIV) type 1 proviral hypermutation correlates with CD4 count in HIV-infected women from Kenya. *J Virol*. 2008; 82:8172–8182. [PubMed: 18550667]
95. Tang J, et al. Human leukocyte antigen class I genotypes in relation to heterosexual HIV type 1 transmission within discordant couples. *J Immunol*. 2008; 181:2626–2635. [PubMed: 18684953]
96. Singh KK, et al. Associations of chemokine receptor polymorphisms With HIV-1 mother-to-child transmission in sub-Saharan Africa: possible modulation of genetic effects by antiretrovirals. *J Acquir Immune Defic Syndr*. 2008; 49:259–265. [PubMed: 18845960]
97. Smith MW, et al. Contrasting genetic influence of CCR2 and CCR5 variants on HIV-1 infection and disease progression. Hemophilia Growth and Development Study (HGDS), Multicenter AIDS Cohort Study (MACS), Multicenter Hemophilia Cohort Study (MHCS), San Francisco City Cohort (SFCC), ALIVE Study. *Science*. 1997; 277:959–965. [PubMed: 9252328]
98. Duggal P, et al. Genetic influence of CXCR6 chemokine receptor alleles on PCP-mediated AIDS progression among African Americans. *Genes and immunity*. 2003; 4:245–250. [PubMed: 12761559]
99. Garcia-Morujia C, et al. Molecular phenotype of CXCL12beta 3 UTR G801A polymorphism (rs1801157) associated to HIV-1 disease progression. *Current HIV research*. 2009; 7:384–389. [PubMed: 19601773]
100. Winkler C, et al. Genetic restriction of AIDS pathogenesis by an SDF-1 chemokine gene variant. ALIVE Study, Hemophilia Growth and Development Study (HGDS), Multicenter AIDS Cohort Study (MACS), Multicenter Hemophilia Cohort Study (MHCS), San Francisco City Cohort (SFCC). *Science*. 1998; 279:389–393. [PubMed: 9430590]
101. He W, et al. Reply to: “Experimental aspects of copy number variant assays at CCL3L1”. *Nat Med*. 2009; 15:1117–1120. [PubMed: 19812563]
102. Martin MP, et al. Association of DC-SIGN promoter polymorphism with increased risk for parenteral, but not mucosal, acquisition of human immunodeficiency virus type 1 infection. *J Virol*. 2004; 78:14053–14056. [PubMed: 15564514]
103. Shin HD, et al. Genetic restriction of HIV-1 pathogenesis to AIDS by promoter alleles of IL10. *Proc Natl Acad Sci U S A*. 2000; 97:14467–14472. [PubMed: 11121048]
104. An P, et al. A tumor necrosis factor-alpha-inducible promoter variant of interferon-gamma accelerates CD4+ T cell depletion in human immunodeficiency virus-1-infected individuals. *J Infect Dis*. 2003; 188:228–231. [PubMed: 12854077]
105. Kanari Y, et al. Genotypes at chromosome 22q12-13 are associated with HIV-1-exposed but uninfected status in Italians. *AIDS*. 2005; 19:1015–1024. [PubMed: 15958832]
106. Koning FA, et al. Defining APOBEC3 expression patterns in human tissues and hematopoietic cell subsets. *J Virol*. 2009; 83:9474–9485. [PubMed: 19587057]
107. Harari A, et al. Polymorphisms and splice variants influence the antiretroviral activity of human APOBEC3H. *J Virol*. 2009; 83:295–303. [PubMed: 18945781]

Box 1**Approaches to identify host genetic factors affecting HIV-1 acquisition and progression**

- i. HIV-1 acquisition and transmission: The identification of host genetic factors influencing HIV infection is confounding by non-genetic factors such as route of transmission, risk behavior, concurrent infection with sexually transmitted diseases, and infectiousness (viral load) of the transmitting partner. This has hampered the identification of genetic factors predicting transmission and acquisition. (Figure 2):
- ii. Disease progression (Figure 1): With a seroconverter (i.e., known time of HIV infection) cohort, the rate of progression to specific AIDS endpoints can be assessed with Kaplan-Meier survival analysis and Cox proportional hazards models using time to CD4<200 or <350, AIDS-defining condition or death as endpoints. When the date of infection is unknown, KM survival and Cox models are not robust. General linear or mixed effects models can be used to assess the effects of genotype on CD4+ T cell trajectory if longitudinal data is available.
- iii. Control of HIV replication: Viral set point during the asymptomatic phase is a predictor of HIV progression and a commonly used outcome variable (Figure 1). Other studies have used the general linear or mixed effects models to assess the HIV RNA trajectory.
- iv. Extreme phenotypes: Extremes of the distribution in terms of AIDS-free survival, CD4⁺ cell, or control of viremia can be enriched for genetic variation associated with the extreme phenotype and require fewer individuals for gene detection. However, caution is required in comparing the extreme tails of the distribution to each other since, for example, factors promoting rapid progression can be quite different than those involved associated with long term AIDS-free survival (analogous to comparing infant mortality to centenarians). Different design strategies are to compare the genotype distribution in the extreme group to a normal population or to a middle group showing median or average progression (Table S1 online).
- v. The major cause of failure to replicate is that the original association was a false positive or had a smaller than originally reported effect size—the winner's curse.
- vi. Candidate gene studies were used to identify CCR5, chemokine ligands, HLA, and KIRS in HIV/AIDS, though they are known for high rate of false positives. Replication in independent, well-powered populations is the gold standard of validation. On the other hand, the curse of GWAS is high rate of false negatives. The stringent threshold for statistical significance (5 or 1×10^{-8}) required to correct for multiple comparisons in GWAS avoid false positive associations, but comes at the expense of false negatives. Separating noise from true signals in genome-wide studies will continue to be a challenge, but hopefully this barrier will be breached through international collaborations to increase power.

The term *winner's curse* is also used in a related but distinct sense in science, in particular genome-wide association studies. In studies involving many tests on one sample of the full population, the consequent stringent standards for significance make it likely that the first person to report a significant test (the *winner*) will also report an effect size much larger than is likely to be seen in subsequent replication studies.

Box 2**Strength and weakness of the approaches to identify host genetic factors****i. Candidate gene association study****Strength**

- Cost-effective method for direct genetic interrogation
- Prior evidence of biological relevance
- Hypothesis-driven test
- Provides estimates of attributable and relative risks

Weakness

- Subject to false positive bias (winner's curse)
- Limited to what is already known

ii. GWAS (genome-wide association study)**Strength**

- Unbiased, Identifies previously undetected variation
- Provides estimates of attributable and relative risks
- Stringent standards for genome-wide significance reduce false positive associations
- Robust for common alleles (>5%), which tend to have small to moderate effect sizes

Weakness

- High noise to signal ratio: statistically impossible to separate false from true negatives
- Expensive and requires large sample sizes
- Does not identify rare mutations

iii. Genome-wide RNA interference (RNAi) screen: Systematically for genes controlling specific biological processes *in vivo* by silencing each of individual endogenous mRNAs with double-stranded (ds) RNA.**Strength**

- Genetic variation not required for gene detection
- Unbiased

Weakness

- Off-target silencing leading to low replication among studies
- Cannot identify extracellular factors
- No quantification of population variation or risk

iv. Whole-genome re-sequencing: Next generation, high throughput sequencing**Strength**

Weakness

- Unbiased detection of all genome or transcriptome sequence variation
- Bioinformatics complex
- Massively parallel coverage to avoid false detection of rare variation

Figure 1.

A typical natural course of HIV infection. (i) Individual infected with HIV-1; (ii) Acute phase lasting for 6-12 weeks with flu-like symptoms, peak viral load and drop in CD4⁺ T cells; (iii) Chronic asymptomatic phase last on average for 7-10 years; following the acute phase, viral replication reaches a steady level known as 'set point'; (iv) AIDS onset is associated with increasing viral replication and declining CD4⁺ cells decline to < 200/mm³; time to the AIDS onset vary in individuals ranging from as little as 2 years to more than 15 years after seroconversion. The figure was modified from Ref[84].

Figure 2.

Study groups and phenotypes in HIV-1 genetic association analysis. Case-control and cohort studies are used to evaluate genetic impact on HIV-1 infection or disease progression by assessing an AIDS-predicting phenotype (setpoint viral load), time from infection to a specific outcome (e.g. CD4⁺ cell, AIDS and Death), or case-control comparison between fast, median and non-progressors.

Figure 3. Host factors that interact with HIV

HIV-1 co-opts host cellular factors for its life cycle: cell entry, uncoating, reverse transcription to viral DNA, nucleus entry, integration into host DNA, transcription of viral RNA, assembling and budding of new virions. Cellular factors such as CCR5, CD4 and CypA promote and other factors such as TRIM5 and APOBEC3G impede HIV replication.

Figure 4. Replication among GWAS, silencing RNA screen, and candidate gene studies
 Genes identified from a GWAS [71], gene silencing screens [75-78] and numerous candidate gene studies (Table 1) were compared for overlapping genes. ZNRD1 was identified in two of the siRNA screens [75, 78] and in the Fellay GWAS [71]. HCP5 tracks *HLA-B*57*, a well-defined, protective *HLA* allele. *HLA-C* was uniquely revealed by GWAS. This gene would not be identified by siRNA screens targeting intracellular factors. Genes from silencing RNA studies that were identified in two out of four studies are indicated [75-78, 85]. Genes in parenthesis are also listed in the NCBI HIV-1 interaction database (<http://www.ncbi.nlm.nih.gov/RefSeq/HIVInteractions/>). Tsg101 was identified by candidate association [86] and in one siRNA screen [78]. Not all genes identified by uniquely by candidate gene analysis (Table 1) are listed.

Figure 5. Interdisciplinary approaches to identify host factors affecting HIV-1/AIDS

The complexity of host-HIV interactions and defenses require a multi-pronged, holistic approach to fully understand the biological systems that promote and impede HIV replication and pathogenesis. The integration of knowledge gained from the emerging technologies in genomics, transcriptomics, proteomics and epigenetics should lead to novel strategies for vaccine development and identify multiple, vulnerable targets for therapeutic intervention.

Table 1
Published host genetic factors of HIV/AIDS

Gene	SNP rs# and name	Mechanism	Effect	Refs
HIV entry Chemokine receptors				
<i>CCR5</i>	rs33332	Truncated HIV-1 coreceptor	32/ 32: Prevent HIV acquisition	[2]
			32/+ : Delay AIDS	[2]
	rs1799987 Promoter SNP 59029A *	Upregulate CCR5 expression	Accelerate AIDS	[21]
<i>CCR2</i>	rs1799864 V64I	LD with <i>CCR5</i> ?	Delay AIDS	[97]
<i>CXCR6</i>	rs2234355 E3K		Increased survival time after PCP diagnosis	[98]
Chemokines				
<i>SDF1</i>	rs1801157 3 A (3 UTR)	Increased levels of CXCL12 mRNA and enhanced mRNA stability compared to 3 G [99]	Delay AIDS	[100]
<i>CCL5</i>	rs2280789 In1.1C (intronic)	Down-regulate <i>CCL5</i>	Accelerate AIDS, risk to HIV acquisition in AA, Thai [32], infection in Chinese [33], German [31],	[28, 30, 34, 65]
	Haplotype R3+R5		Accelerate AIDS in EA[28] and under HAART	[28]
	rs2107538 -403		Increased risk of infection	[28]
	rs1800825 -28	Up-regulation	Delay progression in Japanese [29], Thai [32]	[29]
	3 222C		Increased risk of infection	[28]
<i>CCL3L1</i>	Copy number polymorphism at 17q11.2	Correlate with <i>CCL3L1</i> level	Increasing copy number may be associated with favorable outcomes	[37, 39, 40, 63, 101]
<i>CCL2-CCL17-CCL11</i>	Hap 7 (31 kb) at 17q 11.2-q12	Immune modifiers?	Prevent HIV acquisition	[1]
<i>CCL18-CCL3-CCL4</i>	rs1719153 rs1719134 47 kb haplotype at 17q12		Accelerate AIDS	[1]
<i>DC-SIGN</i>	rs4804803 -336G	Lower DC-SIGN expression on dendritic cells	Increased risk of HIV acquisition in EA	[102]
Post-entry cellular viral cofactor				
<i>CUL5</i>	rs7117111 SNP5	Unknown	Accelerate AIDS in AA	[53]
	rs11212495 SNP6	Transcription factor binding	Accelerate AIDS in AA	[53]
	rs7103534 SNP4	Unknown	Delay progression in EA	[53]

Gene	SNP rs# and name	Mechanism	Effect	Refs
<i>PPIA/CypA</i>	rs8177826 1604G	Binding of transcription factors	Accelerate AIDS in EA and AA	[64]
	rs6850 1650G	Increase HIV infectivity	Increased risk of HIV acquisition	[64, 65]
<i>Tsg101</i>	Haplotype C		Delay progression	[86]
<i>APOBEC3G</i>	rs8477832 H186R		Accelerate AIDS in AA, high viral load in Africans	[44, 46, 93]
	rs3736685 197193C	Intronic, in LD with H186R	Accelerate AIDS in AA	[44]
	rs2294367 199376C	Intronic	Accelerate AIDS in EA	[44]
	C40693T	Intronic	Increased risk of HIV acquisition	[43]
<i>APOBEC3B</i>	3B/ 3B	Deletion of <i>APOBEC3B</i> gene	Increased risk of HIV acquisition	[52]
<i>TRIM5</i>	rs16934386 Promoter SNP2 rs7127617 Promoter SNP 3	Regulate TRIM5 expression in T-cell?	Increased risk of HIV acquisition	[56]
	rs10838525 R136Q	Better anti-HIV-1 activity	Prevent HIV acquisition	[56]
	rs3740996 H43Y		Prevent HIV acquisition	[56]
Cytokines				
<i>IL10</i> (Th2 cytokine)		Down-regulates IL10	Accelerate AIDS	[103]
<i>IFNG</i> (Th1 cytokine)	rs2069709 -179T	Aberrant IFNG regulation	Accelerate AIDS	[104]
<i>IRF-1</i>	rs17848395 619A rs17848424 6516G	Down-regulation and lowered response to IFN- γ	Prevent HIV acquisition	[16]
<i>CXCR1</i> (IL8 receptor)	Haplotype Ha	Modulation of CD4 and CXCR4 expression	Delay progression	[17]
HLA system				
<i>HLA</i>	Class I A,B,C homozygotes	Reduced epitope recognition repertoire	Accelerate AIDS	[9, 95]
	<i>HLA</i> class I Concordance	Narrow epitope recognition repertoire	Increased risk of sexual and mother-to-infant transmission	[12]
	B *35-Px	Weak epitope binding helps HIV immune escape	Accelerate AIDS	[10]
	B *27	Hinders HIV immune escape	Delay AIDS	[7]
	B *57	Hinders HIV immune escape	Delay AIDS	[7]
	rs9264942, 5' upstream of <i>HLA-C</i>	Regulate <i>HLA-C</i> expression	Decease viral set point	[71]
<i>HCP5</i>	rs2395029-G	LD with B *57	Decease viral set point	[71]

Gene	SNP rs# and name	Mechanism	Effect	Refs
<i>KIR</i>	<i>KIR3DS1</i> + <i>HLA</i> Bw4-80I	Signaling the NK cells to kill the HIV infected cells	Delay AIDS	[7]
	<i>KIR3DS1</i> in absence of ligand	Poor regulation of NK cell activity	Accelerate AIDS	[7]
Others				
<i>ZNRD1</i>	rs9261174 5' region of <i>ZNRD1</i> , 1Mb upstream of MHC	Regulate <i>ZNRD1</i> expression	Delay AIDS	[71]
<i>Ly6</i>	rs2572886	Higher cellular susceptibility to HIV-1 in B cells and in primary T cells	Higher viral load, fast CD4 ⁺ T-cell decrease	[89]
Chr. 22q13	Microsatellite D22S423 at 22q13	Controlling virus-neutralizing antibody?	Prevent HIV acquisition	[105]
<i>Prox1</i>	rs17762192-C	Regulate <i>IFNG</i> expression in T-cells?	Delay progression	[72]

aNote. AA, Africa American; EA, European American.

* 59029A (−2459 relative to the translational site, or 303 based on AF031237), representative of P1 of Martin's [21], or HHE+HHG*1 of Gonzalez's [95].

Table 2

Anti-HIV-1 Profile of APOBEC3 proteins

Gene	Anti-HIV-1	mRNA expression in CD4 ⁺ T-cell	Resistance to HIV-Vif
APOBEC3A	No	low, IFN- α inducible	No
APOBEC3B	Strong	Low	Complete
APOBEC3C	Weak	No	No
APOBEC3D/E	Weak	No	No
APOBEC3F	Strong	High	Partial
APOBEC3G	Strong	High	No
APOBEC3H	Strong/weak by variants	High	Yes/no by variants

Data from Chiu and Greene [41], Koning et al. [106] and Harari et al.[107].

Table 3

In vitro and *in vivo* effects of *TRIM5* variants^a

Variants	MAF in Africa American	<i>In vitro</i> HIV restriction	<i>In vivo</i> Genetic association
Rs16934386	0.08	Not known	Risk for HIV acquisition [56]
H43Y	0.09	Loss of function [56, 58]	Accelerates progression[57][60] Delays progression (p>0.05) AA [56]
G110E	0.005	Loss of function [57]	Risk for HIV acquisition [57]
V112F	0.01	Loss of function [56]	No effect [56, 57]
R136Q	0.13	Slight gain of function[56]	Restricts HIV acquisition[56] and delays progression[60] Risk for HIV acquisition [61]
G249D	0.26	No change [56]	No observed effect [56]
H419Y	0.01	No change [56]	No observed effect [56]

^a Abbreviation: MAF, Minor allele frequency.

Table 4

Summary of GWAS results in association with HIV-1/AIDS^a

Cohort	Number of Patients	Phenotype	Platform	SNP identified	Nearest gene	Chromosome	Effect	P value	Refs
Euro-CHAVI	486 Caucasian SC and 140 for replication	Setpoint viral load; CD4 cell count below 350	Illumina HumanHap550 BeadChip	rs2395029 rs9264942 rs3869068	<i>HCP5</i> <i>HLA-C</i> <i>ZNRD1</i>	6p21.33 6p21.33 6p21.33	Lower VL Lower VL CD4+ loss	$P=9.36 \times 10^{-12}$ $P=3.77 \times 10^{-9}$ $P=3.89 \times 10^{-7}$	Fellay et al. 2007 [71]
ANRS 01 PRIMO	605 French SC and 45 controllers.	Plasma HIV-RNA and cellular HIV-DNA levels during primary infection	Illumina HumanHap300 BeadChips	rs2395029 rs10484554 rs6503919 rs2575735	<i>HCP5</i> <i>HLA-C</i> <i>DDX40</i> and <i>YPEL2</i> Syndecan2	6p21.33 6p21.33 chr17 chr8	=-0.92 =0.53 =-0.21 =-0.18	FDR<25%, $P=10^{-4} \sim 10^{-9}$	Dalmasso et al., 2008 [70]
ANRS 02	275 French nonprogressors and 626 HIV-positive patients for replication	Nonprogression vs. regular progression	Illumina HumanHap300 BeadChips	Illumina HumanHap300 BeadChips	<i>HCP5</i> <i>HLA-C</i> <i>ZNRD1</i>	6p21.33 6p21.33 6p21.33	OR=3.47 NA NA	$P=6.8 \times 10^{-9}$ $P=3.77 \times 10^{-04}$ $P=3.78 \times 10^{-03}$	Limou et al. 2009 [68]
ANRS 03	85 French rapid progressors and 275 non-progressors	Rapid progressors vs. non-progressors	Illumina HumanHap300 BeadChips	rs4118325 rs1522232 rs10800098 rs1020064	<i>PRMT6</i> <i>SOX5</i> <i>RXRG</i> <i>TGFBP1</i>	chr1 chr12 chr1 chr2	OR=0.24 OR=0.45 OR=3.29 OR=0.34	FDR<25%: $P=10^{-6} \sim 10^{-7}$	Le Clerc et al., 2009 [69]
Swiss HIV cohort	15 CEPH pedigrees 805 HIV-positive patients	In vitro GWAS--cellular permissiveness to HIV; viral load and CD4 cell dynamics	Linkage analysis with 2500 SNPs in CEPH	rs2572886	<i>LY6</i>	8q24	Higher VL	0.009	Loeuillet et al., 2008 [89]

^a Abbreviations: VL, viral load; SC, seroconverter

Table 5
Design of HIV genome-wide siRNA and shRNA screens^a

Study	Cell types	siRNA treatment time	HIV-1 strain	Reading time post-infection	Measurement of infection	Number of HIV dependent factors	Refs
siRNA, Brass et al.	HeLa	72 hrs	HXB2	48 hrs; 48 hrs infection of new cells	Reduced P24 $\geq 2SD$ of plate mean or -gal reporter expression	281(237 filtered)	[75]
siRNA Zhou et al.	HeLa	24 hrs	HXB2	48 and 96 hrs	-gal reporter expression decrease or increase $>2SD$ compared to control	311 (232 filtered)	[78]
siRNA Konig et al.	HEK293T	48 hrs	VSV-G-Pseudotyped HIV-1 pNL4-3	24 hrs	Reduced HIV luciferase reporter expression by $>45\%$	295	[77]
shRNA Yeung	Jurkat T-cells	4 weeks	NL4-3 (X4)	4 weeks	> 2 -22-fold enrichment of mRNA from the survival cells over control for lytic replication	252	[76]
cDNA Nguyen et al	HeLa	12 hrs	HXB2	96 hrs	Tat-mediated luciferase	315	[79]
Transcriptomic, Imbeault et al.	Human primary CD4 ⁺ T lymphocytes		HIV NL4-3				[15, 80]
Proteomic, Chan, et al.	Primary CD4		HIV LAI				[81]
miRNA microarray, Nathans et al.	H9 T lymphocytes human T lymphocytes		HIV NL4-3				[82]

^a HXB2 is an X4 strain that does not express Nef or Vpu and is truncated for Vpr; shRNA, short hairpin RNA Screening; VSV, vesicular stomatitis virus G protein.

Table 6
Major cellular pathways identified in genomic-wide screens

Pathways	Selected cellular Tactors	Refs
HIV transcription required	NF- B AKT, CycT1, CDK9, ERCC3, PLOR2L, GTF2A1	[75, 78, 85]
Regulate transcription	Mediator complex genes, (8 genes including Med28)	[75, 78] [85]
Nuclear import/export, Golgi proteins, Transport,	Nup98, Nup153, Nup155, Nup210, RANBP2, RGPD8, GCC1, GOLM1, GOSR2, TNPO3	[75, 77, 78, 85]
Microtubule	MAP4, MID11P1, CAV2	[75, 77, 78, 85]
RNA binding/splicing	HNRPF, SF3B2, SFRS2	[75, 77, 78, 85]
Proteasome	PSMD4, PSME2, PSMA2	[75, 78, 85]
Reverse transcription	BMP2K, NEIL3, SERPINB6	[78]