Genome **Biology**

REVIEW

Host genomic influences on HIV/AIDS

Stephen J O'Brien[1]* and Sher L Hendrickson[2]

Abstract

The AIDS era has seen multiple advances in the power of genetics research; scores of host genetic protective factors have been nominated and several have translated to the bedside. We discuss how genomics may inform HIV/AIDS prevention, treatment and eradication.

Introduction

Thirty-one years have passed since reports of severely immune-compromised gay men in San Francisco and New York first drew attention to the AIDS epidemic [1,2]. It took 3 years to identify the cause as human immuno-deficiency virus (HIV)-1, a lentivirus that subverts the lymphoid system to become viral factories producing 1-2 billion virions per day. HIV-1 infects and destroys CD4-bearing lymphocytes, decreasing the victim's ability to overcome normally innocuous agents such as *Pneumocystis, Candida*, cytomegalovirus and a score of other opportunistic infections (Box 1). Left untreated, these infections and certain rare cancers contribute further to the immune system collapse and to the demise of the patients. An effective vaccine is yet to be developed, and anti-retroviral drugs are not ideal since they do not work well for all patients, can generate adverse side effects and do not eliminate the virus from an infected patient. Since the first reports of AIDS in 1981, the world has seen more than 60 million cases of HIV infections, and 30 million victims have died as a consequence, 90% of them in the developing world [3].

The genomic era began eleven years ago with the release of the draft sequence of the human genome. A complete human gene set and DNA sequence was available to help combat the more inscrutable deadly chronic and infectious diseases [4], and this has enabled the investigation of HIV/AIDS using the new tools of genomics. AIDS exhibits abundant epidemiological heterogeneity, much of which could be attributed to host genetic factors. Beginning in the early 1980s, epidemiologists began collecting longitudinal cohorts of at-risk AIDS populations to describe that heterogeneity. Many collaborated with geneticists, who used population genetics-based association analyses to uncover genes with natural variants that exerted an effect on HIV infection, on the dynamics of AIDS progression, and on the results of highly active anti-retroviral therapy (HAART). This review will attempt to update the status of discovered host genetic influences on stages of AIDS beyond the informative reviews that have appeared to date [5-10]. We will highlight those host gene influences that are credible, replicated and implicated in AIDS disease, and we will discuss their clinical relevance in stemming the spread of HIV .

Candidate AIDS restriction genes: the case of CCR5-Δ32

Before the first of ten HIV-AIDS genome-wide association studies (GWAS) appeared [11], single nucleotide polymorphism (SNP) variants in numerous candidate genes were suggested as being associated with HIV or AIDS in patients. Several hundred candidate gene SNPs were tested for association with AIDS progression and for HIV transmission. Over 35 plausible AIDS restriction genes (ARGs) have been identified (Table 1). Their discovery, replication and functional interpretation have been reviewed in depth, and these reviews should be consulted to appreciate the details uncovered around each ARG discovery [5-10].

The first and perhaps most provocative ARG involved the description of *CCR5-Δ32,* a 32 base pair deletion frameshift mutation that truncates C-C chemokine receptor 5 (CCR5), the HIV entry receptor on lymphoid cells [12-14] (see Box 2 for a discussion of the origin of this mutation). Epidemiological studies showed that individuals homozygous for *CCR5-Δ32* had a 100-fold reduction in HIV infection incidence: the genotype seems to confer near complete protection from HIV infection (Figure 1). Individuals in large cohort studies that are heterozygous for *CCR5-Δ32*, although susceptible to infection, consistently have a delayed onset of AIDS by two to four years, likely due to the diminishing onset of

*Correspondence: lgdchief@gmail.com
[1]Theodosius Dobzhansky Center for Genome Bioinformatics, 41A Sredniy Avenue, St Petersburg State University, St Petersburg, 199034, Russia
Full list of author information is available at the end of the article

BioMed Central

Box 1. The stages of AIDS progression
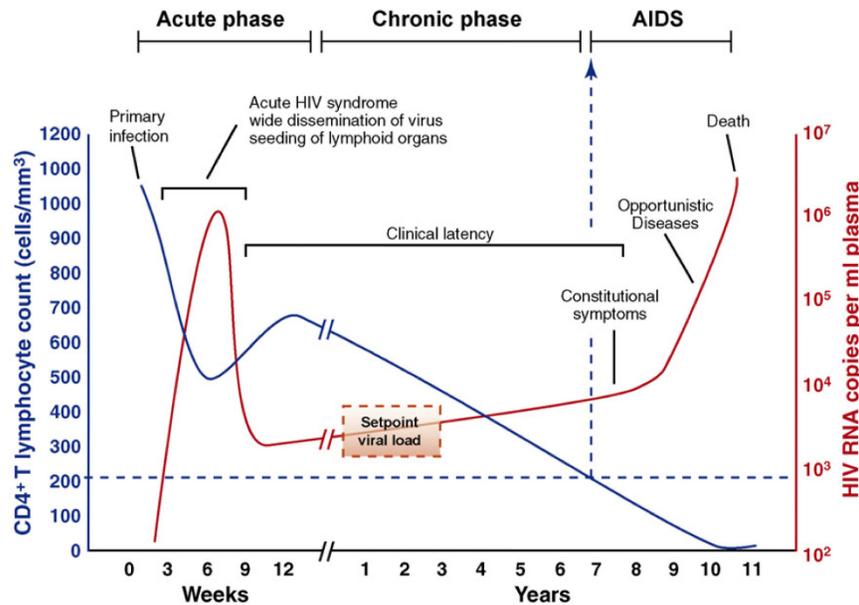
AIDS is one of the most intensively studied diseases in history, providing not only a rich description of the steps in steady progression to immune collapse and death, but also discrete stages that can be investigated for host genetic influence. The process begins when a person becomes exposed to HIV-1, which can occur through sexual contact with an infected partner, intravenous injection with contaminated needles used to deliver addictive drugs, or receipt of contaminated blood or blood products (for example, clotting Factor VIII and IX in the case of hemophiliacs). Once an individual becomes infected, HIV-1 seeks out and overtakes CD4-bearing T-lymphocytes to produce a struggle between viral production (tracked as the viral load – the virus titer in circulating blood) and immune defenses that struggle to control viral proliferation in the beginning weeks. With more than 1 billion new HIV copies produced each day by an acutely infected person, and with new mutational variants arising at a rate of one every second virus copy, immune control is no mean feat. Nonetheless. in most infected people the viral load settles to a chronic plateau level, which varies among individuals from 1×10^3 to 1×10^5 virions/ml (known as the viral set point). Over time the virus gradually destroys the CD4 lymphocytes and the blood count drops to below 200 $CD4^+$ cells/mm^3. The time in which CD4 cell depletion occurs can vary from within a year of infection to more than 20 years later, with the mean time at approximately 10 years post infection. Once the CD4 cell counts drop, the patients succumb to a variety of normally innocuous infections (*Candida*, *Pneumocystis carinii*, cytomegalovirus, herpes, tuberculosis and others) and to rare cancers such as Kaposi's sarcoma and B cell lymphomas. If left untreated, these infections and cancers almost invariably lead to death.

In 1996, a powerful triple drug anti-retroviral therapy was initiated and offered a milestone in AIDS treatment and prognosis. In most patients receiving this treatment, known as highly active anti-retroviral therapy (HAART), the viral load drops to below detectable levels. As promising as these treatments became, there were caveats. Many people (20 to 40%) do not tolerate the drugs and develop adverse reactions. The treatment does not clear HIV from many tissues, and suspension of treatment therefore invariably results in viral load rebound in the blood. Last, in spite of viral replication repression, there is still slow viral-induced pathology and so many patients continue, albeit more slowly, to progress to AIDS and death.

Genetic researchers noted epidemiological differences in four principal stages: (1) HIV acquisition of infection; (2) the rate of progression to AIDS among HIV-infected individuals; (3) the nature of the AIDS-defining condition displayed; and (4) the outcome of HAART, viral load suppression, adverse events and AIDS progression post-treatment [5-10] (Table 1). Case–control and survivor analyses (in the form of Cox proportional hazards models) parsed into alternative SNP alleles or genotypes allow for statistical comparison for differences in patient populations that exhibit better or worse outcome for each stage, leading to gene/SNP association.

Source: An and Winkler [5], reproduced with permission.

B-cell lymphoma [12,15]. Further, AIDS patients on HAART live longer when they are heterozygous for the *CCR5-Δ32* mutation, due to faster HIV viral load suppression, and progress to AIDS more slowly than *CCR5*$^{+/+}$ patients on HAART (Figure 1c) [16]. It is clear that latent HIV continues its pathogenesis in these patients, even in the presence of effective anti-retroviral treatment. We therefore need to develop better AIDS therapies.

CCR5 in clinical practice
Individuals homozygous for *CCR5-Δ32* seem to be relatively healthy into advanced age, suggesting that

Table 1. Candidate AIDS restriction genes

Year	Gene symbol	Allele	Mode	Effect	Citation(s)
1996	CCR5	Δ32	Recessive	Prevents infection	[12-14]
1996	CCR5	Δ32	Dominant	Delays AIDS	[12]
1997	CCR2	64I	Dominant	Delays AIDS	[80]
1998	CCR5P	P1	Recessive	Accelerates AIDS	[81]
1998	SDF1	3'A	Recessive	Delays AIDS	[82]
1999	CCR5	Δ32	Dominant	Prevents lymphoma	[15]
1999	HLA	A,B,C,'Homozy'	Co-dominant	Accelerates AIDS	[83]
2000	IL10	5'A	Dominant	Limits infection	[84]
	IL10	5'A	Dominant	Accelerates AIDS	[84]
2001	HLA	B*35Px	Co-dominant	Accelerates AIDS	[85]
2002	RANTES	−403A	Dominant	Accelerates AIDS	[86]
		In1.1C	Co-dominant	Accelerates AIDS	[86]
2002	KIR	3DS1	Epistatic (Bw4-801)	Delays AIDS	[28,29]
2003	EOTAXIN-MCP1	Hap7	Dominant	Enhances infection	[87]
2003	HLA	B*57	Co-dominant	Delays AIDS	[88]
2003	IFNG	179T	Dominant	Accelerates AIDS	[89]
2003	CXCR6	E3K	Dominant	Accelerates PCP	[90]
2004	APOBEC3G	H186R	Recessive	Accelerates AIDS	[91]
2004	DCSIGN	−336T	Dominant	Decreases infection	[92]
2006	HLA	B27	Co-dominant	Delays AIDS	[93]
2006	TSG101	Hap2	Dominant	Accelerates AIDS	[94]
2006	TRIM5	Hap4	Dominant	Increases infection	[95]
2007	CUL5	Hapl	Co-dominant	Accelerates CD4 loss	[96]
2007	PP1A (cyclophilin A)	SNP-4	Dominant	Accelerates AIDS	[97]
2007	HLA	Bw4	Dominant	Reduces HIV Transmission	[98]
2008	MYH9			End Stage Renal Disease	[31]
2008	MYH9			HIV FSGN	[30]
2008	mtDNA	Hap-J, U5a	Dominant	Accelerates AIDS	[56]
2008	mtDNA	Hap-H	Dominant	Increases Lipoatrophy post HAART	[55]
2008	mtDNA	Hap-J	Dominant	Delays CMV-NRD	[99]
2009	HCP5	T>G; rs2395029	Dominant	HIV set point	[33]
2009	HLA	rs9264942	Dominant	HIV set point	[33]
2009	PROX1	Hap-CGT	Recessive	Delays AIDS progression	[37]
2009	APOBEC 3B	ΔV_	Recessive	Increases infection	[100]
2010	PECI	G	Dominant	Accelerates AIDS	[52]
2010	ACSM4	A	Co-dominant	Delays AIDS	[52]
2010	NCOR2	T	Dominant	Increases infection	[57]
2010	IDH1	C	Dominant	Prevents infection	[57]

All these genes were discovered or validated using the NCI-Laboratory of Genomic Diversity. Eight AIDS cohorts comprising more than 10,000 study participants were used in these studies from 1985 to 2011 [5,10,79,101].

Box 2. The origins of the *CCR5-Δ32* mutation in human populations

The *CCR5-Δ32* allele occurs at a relatively high frequency in Europe (about 10%), but is completely absent in native African and Asian people. This distribution, plus the high incidence of amino acid-altering mutations (about 80%) among all SNP variants in the *CCR5* locus, lent support to the idea that *CCR5-Δ32* was likely the object of strong selective pressures favoring *CCR5-Δ32* frequency elevation in European ancestors [18,71,72]. The extended haplotype of SNPs around *CCR5-Δ32* is about 1,000 kb, 100 times longer than the average haplotype size in Europeans, suggesting a recent mutational origin of *CCR5-Δ32* and possibly a selective advantage of *CCR5-Δ32* carriers [72-74].

But what was the selective pressure? It was likely a global deadly epidemic that *CCR5-Δ32* carriers were resistant to, but it could not have been HIV/AIDS since this epidemic emerged in the early 20th century, too recent to account for the high frequency of *CCR5-Δ32* in European populations. An evolutionary computation based upon the length of the extended *CCR5* haplotype was used to estimate that the date of the latest selective elevation of *CCR5-Δ32* was about 700 years ago [72]. This calculation raised the prospect that survivors of the Black Death (in which bubonic plague killed some 25,000,000 Europeans in 1348–1353 CE) had a role in favoring the elevation of *CCR5-Δ32*. A functional demonstration that cultured cells from *Ccr5$^{-/-}$* knockout mice (equivalent to *CCR5-Δ32/Δ32* homozygous people) show a 30-fold reduction in sensitivity to infection by *Yersinia pestis*, the plague bacillus, compared with *Ccr5$^{+/+}$* normal mice [75] lends credence to this provocative explanation. A more detailed narrative of the still unproven plague hypothesis is presented in [76].
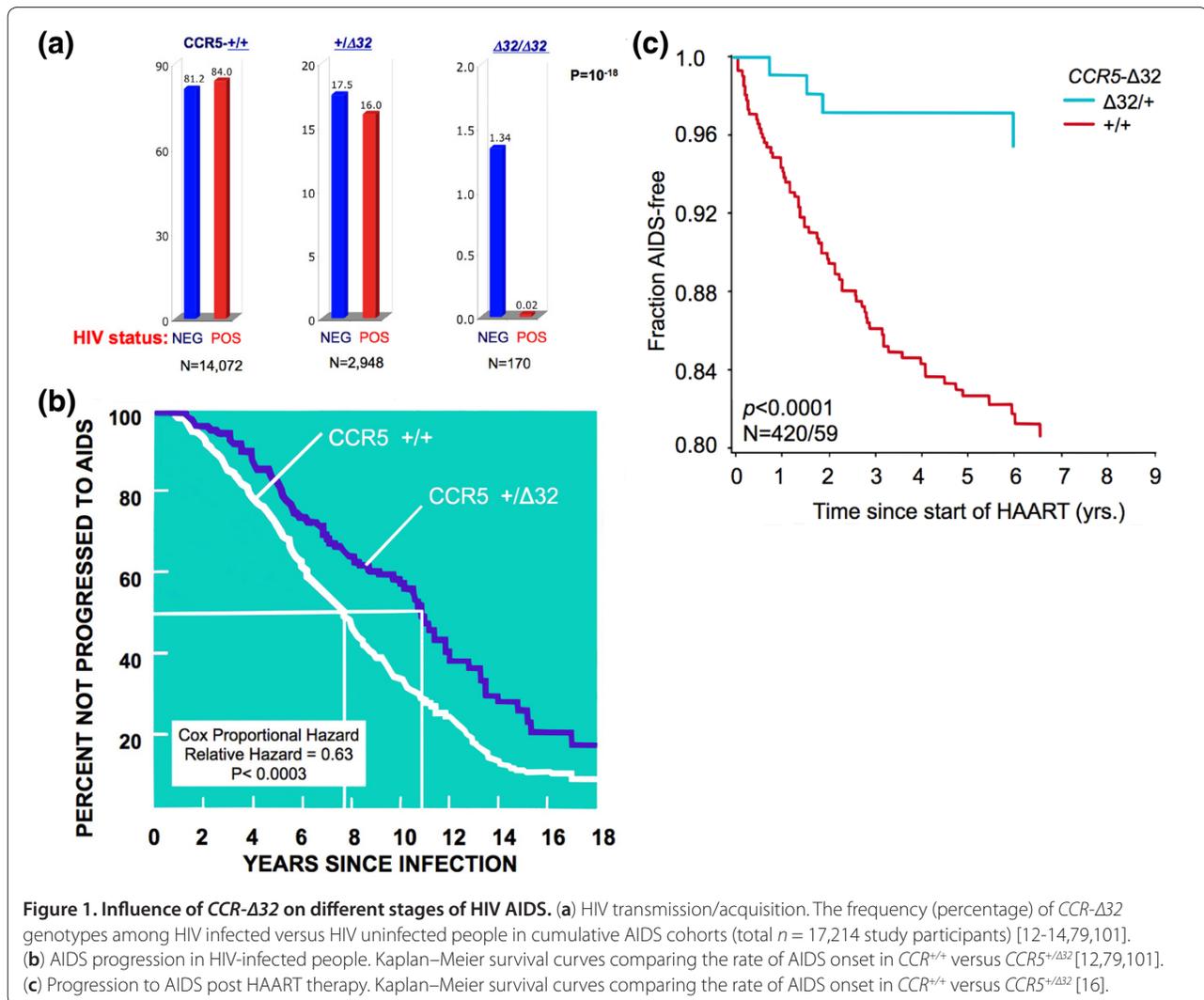
There are alternative interpretations as to the cause of *CCR5-Δ32*-selective elevation (some have proposed that anthrax, small pox, Ebola and other contagions played a selective role in the rise of *CCR5-Δ32* in medieval Europe) and even for Black Death etiology. Whether the direction of the allele frequency cline (gradient) across Europe tracks the plague progression or whether plague mortality was large enough to elevate *CCCR5-Δ32* sufficiently have been raised in critiques of the plague/*CCR5-Δ32* hypothesis. Nonetheless, the functional connection of a CCR5 receptor requirement for *Yersinia* infection, the coalescent date of 700 years for *CCR5-Δ32* haplotype origins, the enormous toll of the Black Death (an estimated 25 million Europeans – one-third of the population – perished by the Black Death) and the multiple plague episodes across Europe tracing from the Justinian plague (541 to 542 CE) to the 17th-century Great Plague of London (1664 to 1666 CE) all lend support to plague being a principal selective pressure on *CCR5-Δ32* [18,72,76].

CCR5 function is dispensable in the human genome, a notion supported by the redundancy of other human chemokine receptors that also interact with specific chemokine ligands of CCR5 [17]. The knowledge that CCR5 is required for HIV-1 infection of lymphoid cells in exposed people, plus the suspicion that CCR5 function is dispensable for good health [18] (see Box 3), stimulated pharmaceutical development of a dozen new anti-AIDS treatments termed HIV entry inhibitors. Entry inhibitors were designed to interfere with HIV–CCR5 cell receptor binding as a block to HIV pathogenesis [19-21]. Two powerful HIV entry inhibitors, enfuviritude (T20) and maraviroc, were approved for AIDS patient treatment by the Food and Drug Administration (FDA) in 2003 and 2007, respectively. The rapid development of this new avenue for AIDS therapy was clearly encouraged by the characterization of HIV resistance in *CCR5-Δ32* homozygotes (the 30-plus other HAART drugs all target HIV-encoded enzymes: reverse transcriptase – targeted by both nucleotide reverse transcriptase inhibitors and non-nucleoside reverse transcriptase inhibitors – and the HIV-encoded protease).

A remarkable example of *CCR5-Δ32* clinical translation to the bedside involves the treatment of Timothy Leigh Brown, dubbed the 'Berlin patient' [22]. Brown was a 25-year-old American living in Berlin in 1995 when he was diagnosed with HIV infection. Despite receiving sporadic anti-HIV medication, his condition deteriorated and he was diagnosed with acute myelogenous leukemia in 2006. His AIDS clinician, Gero Hutter, prescribed a stem cell transplant of healthy HIV-negative donor lymphocytes following x-irradiation of his lymphoid system [23]. Hutter searched the European bone marrow databases to find an HLA-matched donor who was homozygous for *CCR5-Δ32*, so as to block HIV transmission to the donor graft cells. After the stem cell transplant in February 2007, Brown's anti-retroviral treatment was suspended to allow the graft to grow, though monthly blood tests were initiated to detect the inevitable latent HIV rebound. But rebound never occurred. Three years later, Hutter and his team reported that Brown had consistently tested negative for circulating plasma HIV since the transplant [23]. Subsequent rather invasive biopsies of multiple tissues (plasma, gut, brain, testis, liver) using hyper-sensitive virus assays were also negative, indicating that his body had been cleared of replicating HIV, the first such event recorded in history [22-24]. Though no one is certain that each and every one of the scores of human tissues that can become infected with HIV are clear in Mr Brown's body, as of now – five years later – his case is being heralded as the single documented clearance or apparent cure of HIV infection.

Bone marrow and stem cell transplants are of limited promise as a universal therapy. They are expensive (circa US$250,000) and not very safe, primarily because even HLA-matched donors are immunologically allogeneic (that is, non-self). The immune system of the recipient must be destroyed, deadly graft versus host reaction against the host is common, and a dangerous transplantation therapy must be followed by a lifelong

Figure 1. Influence of *CCR-Δ32* on different stages of HIV AIDS. (a) HIV transmission/acquisition. The frequency (percentage) of *CCR-Δ32* genotypes among HIV infected versus HIV uninfected people in cumulative AIDS cohorts (total *n* = 17,214 study participants) [12-14,79,101]. **(b)** AIDS progression in HIV-infected people. Kaplan–Meier survival curves comparing the rate of AIDS onset in *CCR⁺/⁺* versus *CCR5⁺/Δ32* [12,79,101]. **(c)** Progression to AIDS post HAART therapy. Kaplan–Meier survival curves comparing the rate of AIDS onset in *CCR⁺/⁺* versus *CCR5⁺/Δ32* [16].

regimen of anti-rejection therapies. A promising alternative to this may be immunologically syngeneic (self) grafts with *CCR5* knockout strategies. Carl June and Bruce Levine at University of Pennsylvania have developed such a protocol using zinc finger targeted knockout of normal *CCR5* genes in HIV-infected patients [25]. Twelve such patients have been infused with their own-engineered *CCR5⁻/⁻* lymphocytes. The results are promising to date, with donor cell reconstitution observed and few safety issues encountered, and several patients remained plasma HIV-negative several months post transplant infusion.

News of the Berlin patient's amazing recovery travelled quickly in the AIDS research community and beyond. Attention was also paid to the fascinating secondary observation that the Berlin patient has never developed graft versus host disease (GVHD), a deadly reaction that afflicts 30–50% of bone marrow and stem transplant recipients. Transplant researchers at University of Pennsylvania reasoned that perhaps the *CCR5-Δ32* genotype of the Berlin patient's transplant donor might have prevented his GVHD; this would affirm a role for intact CCR5 in GVHD. The availability of safe and FDA approved CCR5 antagonist drugs for AIDS, such as maraviroc, prompted a Phase I and II clinical trial to see if maraviroc-treated transplant patients might avoid GVHD, as with the Berlin patient [26]. A group of 38 cancer patients treated with maraviroc post-transplant demonstrated a dramatic reduction in GVHD without any major adverse side effects. If confirmed by a larger Phase III trial, CCR5 may therefore be implicated as a key player in cancer treatments as well as AIDS therapies. It should be noted that maraviroc is considered to be a safe drug and that *CCR5-Δ32/32* homozygotes are generally healthy. If such bold new therapies do prove efficacious, they will stand upon the findings of the *CCR5-Δ32/Δ32* homozygotes' genetic resistance to HIV acquisition and the apparent success with the Berlin patient.

Box 3. The mixed blessing of inheriting a *CCR5-Δ32/Δ32* genotype

The evidence that *CCR5-Δ32* was protective against AIDS, and also the object of historic selective events likely from infectious agents other than HIV in the distant past, has stimulated a search for other viruses that may co-opt CCR5 molecules as entry portals to human cells. One fascinating study using West Nile Virus (WNV) raises some troubling issues [77,78]. WNV, a mosquito-borne flavivirus that was first introduced into New York City in 1999, induces a fatal encephalitis in 20% of human infections. The agent has caused 17,000 infections and 700 encephalitis cases to date, and so far there is no effective vaccine or treatment. Philip Murphy's group at the National Institutes of Health (NIH) were searching for other CCR5–virus interactions when they examined the WNV infection response of *Ccr5*$^{-/-}$ knockout mice compared with *Ccr5*$^{+/+}$ mice [77]. WNV upregulated the expression of *CCR5* in the central nervous system of infected mice and this was associated with a massive infusion of CD4$^+$, CD8$^+$, NK1.1 lymphocytes and macrophages, all expressing CCR5, into the central nervous system (CNS). Survival and recovery in normal mice was about 60%. However, when *Ccr5*$^{-/-}$ knockout mice (equivalent to *CCR5-Δ32* homozygotes) were infected there was little immune reaction and the infection was uniformly (100%) fatal, implicating CCR5 as a critical innate host defense against WNV in mice. Murphy's group mounted a large genetic epidemiological association study of WNV-infected people and discovered that *CCR5-Δ32/Δ32* homozygotes were elevated 4- to 13-fold among the fatal encephalitis cases relative to the survivors [78]. Clearly the intact *CCR5* gene function exerts a critical defense in the face of WNV. The alternative AIDS resistance and WNV hypersensitivity poses a mixed blessing for homozygous carriers of *CCR5-Δ32* and a stark reminder that genetic adaptation can come with a cost in a changing microbial environment.

AIDS resistance genes beyond *CCR5*

Each of the genes listed in Table 1 has the potential to have a detailed narrative such as that described for *CCR5-Δ32*. For example, *HLA* allelic resistance and susceptibility to AIDS progression is well accepted, and at least one *HLA* genotype, *HLAB*5701*, seems to pre-dispose carriers to a severe immunological disorder (fever, gastrointestinal and respiratory distress) when they are treated with the nucleoside reverse transcriptase HIV inhibitor abacavir [27]. *HLA* genotyping is now a prerequisite to abacavir prescription. Epistatic inter-actions of HLA class I protein products and the killer immunoglobulin cluster KIR ligands (chr19] are power-ful, but have not yet informed clinical practice [28,29].

SNP variants within loci that exert powerful anti-HIV restrictions *in vitro* and against other viruses (*TRIM5, APOBEC, TSG101, CUL5*) have shown modest effects on AIDS progression in association studies (Table 1). Unlike *CCR5-Δ32*, which confers a complete loss of function mutation, the consequence of most other ARG SNP variants are limited by quantitative influences on gene action. An important predictive influence on HIV-asso-ciated nephropathy, particularly among African Americans, appears to be associated with multi-SNP haplotypes within the *MYH9* (non-muscle myosin IIA) gene [30,31]. The risk alleles are quite common in African Americans but rare in European Caucasians, which explains a 20-fold higher risk for HIV-associated nephropathy in African Americans. Taken together, combining func-tional data that connects ARG variants directly to HIV pathology, robust statistical gene SNP associations and clinical translation has improved our understanding of diagnosis, pathogenesis, treatment and ablation of AIDS.

HIV/AIDS GWAS

Unlike association studies of candidate genes, GWAS scan the entire genome with hundreds of thousands (if not millions) of common human SNPs to discover genome regions that statistically associate with explicit AIDS endpoints in large cohort studies (Box 1). Fellay *et al.* [11] reported the first AIDS GWAS in 2007. Their study tracked the HIV set point (mean plasma viral RNA level over several months once the immune system has settled to a steady state level after initial spikes in virus upon HIV infection; see Box 1) in 486 European AIDS patients from the Euro-Center for HIV/AIDS Vaccine Immunology (CHAVI) cohort that were genotyped with an Illumina 550,000 SNP array. Fellay *et al.* discovered three genome-wide significant (after Bonferroni correc-tion for multiple SNP tests; <9.3 × 10^{-8}) SNPs nested within the *HLA* complex, including a proxy-SNP in linkage disequilibrium (LD) *HLA-B*5701*, a known pro-tective influence on AIDS progression. Since then ten additional GWAS using different cohorts, and different AIDS stages (phenotypes), arrays and numbers of SNPs, have appeared [32-41]. The combined GWAS have consistently fingered *HLA* SNPs, but agreed on little else.

Additional GWAS-based gene discoveries outside *HLA* appear robust, but they have seldom included the candidate genes listed in Table 1, nor have the implicated gene associations been replicated by independent studies. Further, the fraction of the epidemiological variance of AIDS progression that individual or combined GWAS discovered genes can explain falls between 10% and 20% for AIDS progression and far less for other stages of AIDS (for example, HIV acquisition, HAART outcomes, and so on) [10,11,40]. Rediscovering *HLA* and the low level of explanative variance were disappointing to the AIDS community given the promise of GWAS for resolv-ing gene determinants in other complex diseases. The advantages of GWAS over candidate gene studies are well known: (1) they can reveal genes we may not have con-sidered since they survey the entire genome; (2) the patterns of LD allow for haplotype-based closing in on

the causal or operative SNP variant; (3) copy number variants can be resolved by SNP genotyping arrays; and (4) subsets of gene families can be explored using SNP array data (see below). The disadvantages of GWAS are that there are no *a priori* assumptions about a possible connection to AIDS as there are for candidate genes; therefore, SNP association discovery must be complemented with additional function-based implication.

The large number of SNPs interrogated requires statistical correction for multiple testing, and to avoid false positives the significance threshold must be stringent. Yet mixed among the below-threshold SNPs are a few gems, such as those in *CCR5* and *KIR*, whose presence in pathways associated with AIDS pathogenesis make them too good to miss (for instance, many ARGs in Table 1 are not revealed by GWAS). Finally, GWAS cannot resolve rare causal variants (f <5%).
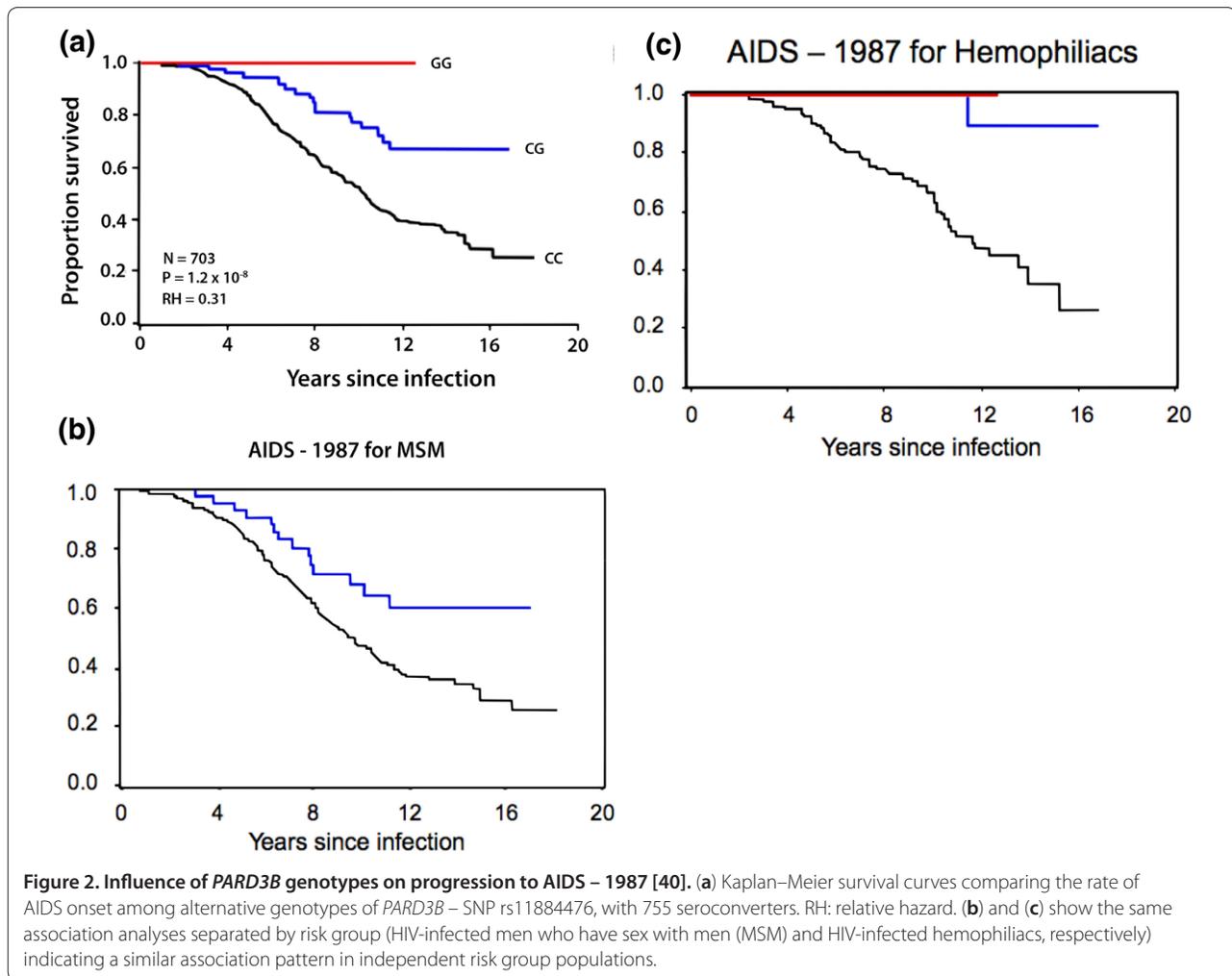
Goldstein has suggested that rare variants may account for undiscovered genetic influences on AIDS and other complex diseases [42,43], but these genetic variants would not be uncovered by GWAS. Although it is impossible to predict future discoveries with confidence, the frequent occurrence of deadly infectious disease episodes in human history has clearly provided abundant selective pressures (possibly raising allele frequencies for *CCR5-Δ32*, *HB-S*, *DUFFY*, *HLA* and others) elevating many disease immune defense alleles above the 'rare allele' frequency. We also suggest that the study design of some of the previous AIDS GWAS might be improved by assessing clinical endpoint/phenotypes less variable than viral load, set point and CD4 cell trajectory. Although viral load/set point are valid predictors of AIDS progression on a course scale, in our cohorts the signal suffers somewhat from noise, and is also less continuous than survival analyses (using Cox proportional hazards analyses for detecting ARGs that influence AIDS progression). Better approaches for resolving the true association signals that fall below genome-wide significance threshold are needed if we are to maximize the potential of AIDS GWAS, and will likely require the evaluation of gene associations from a variety of perspectives.

One recent example of a potentially 'better approach' is the recent GWAS that associated several SNPs within the large 1.1 Mbp *PARD3B* gene (chr2), a locus that regulates cell polarity in epithelial cells [40,44]. Nine of the 15 lowest *P*-values in the GWAS fell in a short 5 kb LD block across a single *PARD3B* exon. One of the nine SNPs reached genome-wide significance ($P = 3.4 \times 10^{-9}$), while the others fell just below that threshold. The basis for the association (Figure 2) demonstrates that the rare genotype (GG) for the SNP almost never progress to AIDS, while people that are heterozygous for this SNP progress markedly slower than those homozygous for the common allele (CC). Support for the gene association came from

six angles: (1) a nearly identical effect was observed among two distinct risk groups (hemophiliacs and gay men; Figure 2b,c); (2) the SNP association result was replicated, albeit with a weaker statistical strength ($P = 0.03$), in independent European cohorts; (3) one of the associated SNPs is a predicted exon splicing enhancer; (4) a quantitative measure of *PARD3B* splice products in lymphoblastoid B-cell lines with alternative *PARD3B* genotypes revealed a dominant 30 to 40% reduction in *PARD3B* isoform transcripts; (5) in the GWAS, *PARD3B* alleles explained 4.3% of the epidemiological variance of AIDS progression, higher than any of the other candidate genes in Table 1 similarly assessed in the same cohorts; (6) *PARD3B* gene products are known to interact with SMAD proteins, signal transduction modulators that mediate gene signaling, including through a direct interaction with HIV regulatory proteins [45]. Taken together, the cumulative replication, HIV1 regulatory connection, transcription influence and association data on *PARD3B* offers a plausible new genetic influence of AIDS progression.

Beyond GWAS: new approaches
The *PARD3B* study was intentionally limited to a single endpoint or phenotype (time of progression as defined by the AIDS 1987 Centers for Disease Control) to avoid statistical penalties for multiple tests. Yet many different gene association tests are possible with the rich clinical and genetic data we have collected across the years on these retrospective AIDS cohorts; indeed, many of these tests were performed and considered supportive in the previous candidate gene studies [7,10] (Table 1). In spite of the aversion to multiple test statistical correction penalties, there is information than can be gleaned from detailed clinical data that should be explored. Hutcheson *et al.* [7] described useful heat plots that allow one to inspect the pattern of genetic association across many tests, as well as across small regions of the genome in strong LD. The combination of multiple non-independent statistical signals is a challenge to interpret, but in cases of known genes, such as *CCR5-Δ32*, *HLA* and *IL10*, the patterns are illuminating. The heat plots also provide a derivative approach to public data sharing of GWAS results without the problem of violating the informed consent and privacy constraints set out by Institutional Review Boards, which often restrict open access to human cohort data [46]. For example, posting an unabridged table of GWAS results online (odds ratios, *P*-values and confidence intervals), together with two-dimensional and three-dimensional visualization heat plots, would allow researchers to freely inspect the results of a published GWAS [40], in any genomic region they may discover in a different cohort, for independent replication purposes – without viewing patients' confidential clinical or genotypic information.

Figure 2. Influence of *PARD3B* genotypes on progression to AIDS – 1987 [40]. (**a**) Kaplan–Meier survival curves comparing the rate of AIDS onset among alternative genotypes of *PARD3B* – SNP rs11884476, with 755 seroconverters. RH: relative hazard. (**b**) and (**c**) show the same association analyses separated by risk group (HIV-infected men who have sex with men (MSM) and HIV-infected hemophiliacs, respectively) indicating a similar association pattern in independent risk group populations.

Now that ten AIDS GWAS have been completed [11,30-41], it is important to share and compare these data to maximize the discoveries and minimize Type 1 (false positive) errors. Meta-analyses such as the International HIV Acquisition Consortium and the HIV Controllers Consortium are important advances that should be encouraged [35,47]. Further, most of the published studies with any statistical power involve European Caucasian patients. Newer cohorts in Africa and Asia would increase the chance of discovering adaptive variants that may have arisen in those ethnicities in recent centuries. Such cohort developments are ongoing with the CHAVI consortium [48], the Botswana Harvard Partnership [49] and the Chinese AIDS collaboration [50]. It is also important to expand interrogation of critical stages of HIV/AIDS by exome and whole genome sequence analyses to capture SNP variants that are not well represented in available genotyping arrays [51]. Whole genome sequencing will offer an unabridged assessment of gene variants in these cohorts and will

likely reveal additional host factors that regulate AIDS pathogenesis.

A compromise approach between the specificity of candidate gene screens and the stringency of GWAS statistical thresholds is to explore certain candidate gene subsets represented within GWAS. This strategy limits the considered SNPs to specific subsets of candidate genes and lowers the multiple SNP test threshold for genome-wide significance. Hendrickson *et al.* [52] used this approach to explore the association of a group of human genes that are involved in mitochondrial metabolism, named nuclear encoded mitochondrial proteins (NEMPs) [53]. Mitochondrial metabolism is relevant because AIDS pathogenesis clearly involves mitochondrial deterioration and cell death, because mitochondrial toxicity has been linked to several adverse events of HAART, and because mitochondrial haplotypes appear to be associated with differential survival to AIDS in longitudinal cohorts [54-56].

Two tested NEMP-encoding genes displayed significant (corrected for all SNPs considered) genotypic influences

on AIDS progression: peroxisomal D3,D2-enoyl-CoA isomerase (*PECI*) on chromosome 6 and acyl-CoA synthetase medium-chain family member 4 (*ACSM4*) on chromosome 12 (Table 1). The gene subset approach allowed the detection of gene associations that would have been discounted in an agnostic full GWAS since their *P*-values of association significance ($P = 6.5 \times 10^{-6}$ and 3.6×10^{-6} for *PECI* and *ACSM4*, respectively) fall below a full GWAS *P*-value threshold.

Another GWAS subset explored was that of genes required for HIV infection of cultured human cells, termed HIV dependency factors (HDFs) [57]. In 2008, four siRNA-based *in vitro* studies were published that collectively identified several hundred HDF genes [58-61]. Chinn *et al.* [57] identified 6,380 SNPs that occurred within 278 HDF genes and used a subset of the *PARD3B* GWAS study data [40] to test them for gene associations with HIV infection, AIDS progression and AIDS-defining sequelae. SNP variants in two genes, *NCOR2* (also known as *SMRT*) and *IDH1*, showed significant associations with HIV infection, consistent with their selection from HDF genes. *IDH1* was particularly notable, first since its HDF status was one of a few replicated in independent HDF studies [58,59], and second because *IDH1* can function as both a tumor suppressor (loss of isocitrate dehydrogenase 1 (IDH1] function) and an oncogene in human glioblastomas and other neoplasms [62-64].

The zinc finger ribbon domain 1 (*ZNRD1*) gene (located within the *HLA* locus) was one of the most strongly associated variants with HIV set point and disease progression in several GWAS [11,33]. *ZNRD1* also appeared as a candidate gene in one HDF screen [58], while Chinn *et al.* [57] implicated this gene as being associated with both HIV acquisition and HIV progression. These multiple indications of AIDS influence using different approaches (GWAS, set point and survival analyses) together make a strong case for the influence of the *ZNRD1* genotype on HIV/AIDS.

Conclusions

Thirty years of genomics investigation of AIDS has produced some important findings that have improved understanding of the disease and informed attempts to stem its deadly pace. Many host genes that affect AIDS have been postulated through candidate gene associations, GWAS, HDF analyses, molecular virology and other approaches. Many, but not all, of the implicated genes have been independently replicated by various approaches designed to overcome the statistical challenge posed by multiple tests (for example, *CCR5-Δ32, HLA, KIR, IDH1, ZFNR1* and others). Most important, at least five discoveries have led to bedside translation: (1) development of *CCR5-Δ32*-based HIV entry inhibitor therapy [20,21]; (2) HLA-based pharmacogenomics

screening for *B*5701*-mediated abacavir adverse events [27]; (3) clearing of HIV from the Berlin patient by a stem cell transplant from a *CCR5-Δ32/Δ32* homozygous donor led to gene therapy trials of autologous *CCR5-Δ32* knockout in AIDS patients [22-24]; (4) use of the CCR5 antagonist, maraviroc, as GVHD prophylaxis in cancer patients receiving marrow and stem cell transplants [26]; and (5) *IDH1* discoveries have stimulated the search for drugs that block the accumulation of 2-hydroxyglutarate for cancer [62-64]. These cancer-targeted pharmaceuticals may well equally inform HIV/AIDS therapies in the future.

The disappointments of the ARG searches are several. The total number of ARGs described account for less than 10 to 15% of the variance in AIDS progression and far less for other AIDS-associated measures: namely HIV acquisition, HAART outcomes or specific AIDS-defining conditions. This means that most of the determinants (whether genetic or non-genetic) are yet to be determined. The stringent statistical threshold for GWAS may force us to discount real variants that do not pass the genome-wide significance level; we need to develop a solution to this conundrum. In addition, genotyping arrays used to date, which contain between 500,000 and 1,000,000 SNPs, cover at best 10 to 15% of the variants in genes and gene regions that we want to interrogate, so we depend upon LD to detect the causal variants through inference. We are missing much of the common variation, especially that found outside of genes (such as in the recently described 4 million regulatory regions [65]), and nearly 100% of the rare variants. Finally, because present GWAS largely target European Caucasian ethnic groups, variants accumulated in other ethnicities have not been thoroughly screened.

The next generation of ARG GWAS will address and remedy these shortcomings as follows. Combined meta-analyses, such as the International HIV Acquisition Consortium [35,47], will help affirm the statistically uncertain discoveries that are robust. Subset analyses such as the HDF and NEMP groups described in this review will help here. In addition, better strategies for data sharing of large GWAS and association studies will go a long way to encourage independent inspection, replication and composite interpretations. Bigger genotyping arrays (for example, Illumina 2.5M Infinium arrays) will increase coverage greatly. By far the most comprehensive approach, whole genome sequencing, will accomplish this ever better, but cost is still the limiting factor for this strategy. Further, the development of large HIV/AIDS cohorts in Africa and Asia suitable for GWAS and genome sequence analyses will be welcome additions to the screening of variants for important clinically relevant ARGs [48-50].

Finally, genomics is but one of several players in AIDS research and there are some extremely promising

developments outside our field that must be mentioned. First, it is now well demonstrated that circumcision is a powerful protective aspect of HIV acquisition that is being implemented widely in the developing world to decrease the spread of HIV [66,67]. Second, three recent clinical trials have demonstrated a powerful near absolute protection against HIV infection among high-risk gay men and heterosexuals who took a single daily combination anti-retroviral pill (such pre-emptive treatment strategies are termed pre-exposure prophylaxis (PrEP)) [68,69]. In a large global cohort of 2,499 men who have sex with men, those treated with Truvada (a combination of two anti-retroviral agents, tenofovir and emtricitabine) showed a 48% reduction in HIV acquisition; this jumped to 95% protection when only those whose drug-taking compliance was 100% were considered [68]. Similar results were observed in an African trial where PrEP-treated heterosexual at-risk women were studied [69]. Anti-retroviral microbicides are also making progress: two recent trials demonstrated 80% reduction in transmission among compliant study participants [70]. Last, in what may be his greatest legacy, former US President George W. Bush established the President's Emergency Plan for AIDS Relief (PEPFAR) in 2002, offering US$15 billion in federal aid for treatment of HIV-infected people across the developing world. He added US$48 billion to the program in 2008. In 2002 there were less than 300,000 people on HAART in the developing world but today, thanks to PEPFAR, there are 6.6 million people on HAART. President Obama has continued PEPFAR funding and there is now tangible hope for slowing the spread of this contagion through education, treatment and research. By interdisciplinary application of important advances, the real dream is to one day see a generation of people that are free from AIDS.

Abbreviations
ARG: AIDS restriction gene; CCR5: C-C chemokine receptor; CHAVI: Center for HIV/AIDS Vaccine Immunology; GVHD: graft versus host disease; GWAS: genome-wide association study; HAART: highly active anti-retroviral therapy; HDF: HIV dependency factor; LD: linkage disequilibrium; NEMP: nuclear encoded mitochondrial protein; PEPFAR: President's Emergency Plan for AIDS Relief; PrEP: pre-exposure prophylaxis; SNP: single nucleotide polymorphism; WNV, West Nile virus.

Author details
[1]Theodosius Dobzhansky Center for Genome Bioinformatics, 41A Sredniy Avenue, St Petersburg State University, St Petersburg, 199034, Russia. [2]Department of Biology, Shepherd University, P.O. Box 5000, Shepherdstown, West Virginia 25443-5000, USA.

References
1. Pneumocystis pneumonia – Los Angeles. *MMWR Morb Mortal Wkly Rep* 1981, **30**:250–252.
2. Kaposi's sarcoma and pneumocystis pneumonia among homosexual men – New York City and California. *MMWR Morb Mortal Wkly Rep* 1981, **30**:305–308.
3. World Health Organization. Global Health Observatory Data Repository. Life Expectancy [http://apps.who.int/ghodata/?vid=710%5D]
4. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, *et al.*: Initial sequencing and analysis of the human genome. *Nature* 2001, **409**:860–921.
5. An P, Winkler CA: Host genes associated with HIV/AIDS: advances in gene discovery. *Trends Genet* 2010, **26**:119–31.
6. Carrington M, O'Brien SJ: The influence of HLA genotype on AIDS. *Ann Rev Med* 2003, **54**:535–551.
7. Hutcheson H, Lautenberger J, Nelson G, Pontuis JJ, Kessing BD, Smith MW, Johonson R, Stephens R, Phair J, Goedert J, Donfield S, O'Brien SJ: Detecting AIDS restriction genes: from candidate genes to genome-wide association discovery. *Vaccine* 2008, **26**:2951–2965.
8. Fellay J, Shianna KV, Telenti A, Goldstein DB: Host genetics and HIV-1: The final phase? *PLoS Pathog* 2010, **6**:e1001033.
9. Lederman MM, Alter G, Daskalakis DC, Rodriguez B, Sieg SF, Hardy G, Cho M, Anthony D, Harding C, Weinberg A, Silverman RH, Douek DC, Margolis L, Goldstein DB, Carrington M, Goedert JJ: Determinants of protection among high risk HIV seronegative persons – an overview. *J Infect Dis* 2010, **202**:333–338.
10. O'Brien SJ, Nelson GW: Human genes that limit AIDS. *Nat Genet* 2004, **36**:565–574.
11. Fellay J, Shianna KV, Ge D, Colombo S, Ledergerber B, Weale M, Zhang K, Gumbs C, Castagna A, Cossarizza A, Cozzi-Lepri A, De Luca A, Easterbrook P, Francioli P, Mallal S, Martinez-Picado J, Miro JM, Obel N, Smith JP, Wyniger J, Descombes P, Antonarakis SE, Letvin NL, McMichael AJ, Haynes BF, Telenti A, Goldstein DB: A whole-genome association study of major determinants for host control of HIV-1. *Science* 2007, **317**:944–947.
12. Dean M, Carrington M, Winkler C, Huttley GA, Smith MW, Allikmets R, Goedert JJ, Buchbinder SP, Vittinghoff E, Gomperts E, Donfield S, Vlahov D, Kaslow R, Saah A, Rinaldo C, Detels R, O'Brien SJ: Genetic restriction of HIV-1 infection and progression to AIDS by a deletion allele of the CKR5 structural gene. *Science* 1996, **273**:1856–1862.
13. Samson M, Libert F, Doranz BJ, Rucker J, Liesnard C, Farber CM, Saragosti S, Lapoumeroulie C, Cognaux J, Forceille C, Muyldermans G, Verhofstede C, Burtonboy G, Georges M, Imai T, Rana S, Yi Y, Smyth RJ, Collman RG, Doms RW, Vassart G, Parmentier M: Resistance to HIV-1 infection in Caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene. *Nature* 1996, **382**:722–725.
14. Liu R, Paxton WA, Choe S, Ceradini D, Martin SR, Horuk R, MacDonald ME, Stuhlmann H, Koup RA, Landau NR: Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. *Cell* 1996, **86**:367–377.
15. Dean M, Jacobson LP, McFarlane G, Margolick JB, Jenkins FJ, Howard OMZ, Dong HF, Goedert JJ, Buchbinder S, Gomperts E, Vlahov D, Oppenheim JJ, O'Brien SJ, Carrington M: Reduced risk of AIDS lymphoma in individuals heterozygous for the CCR5 delta32 mutation. *Cancer Res* 1996, **59**:3561–3564.
16. Hendrickson SL, Jacobson LP, Nelson G, Phair JP, Johnson R, Margolick JB, Detels R, Rinaldo C, O'Brien SJ: Host genetic influences on highly active antiretroviral therapy efficacy and AIDS-free survival. *J Acquir Immune Defic Syndr* 2008, **48**:263–271.
17. Murphy PM: International Union of Pharmacology. XXX. Update on chemokine receptor nomenclature. *Pharmacol Rev* 2002, **54**:227–229.
18. O'Brien SJ, Dean M: In search of AIDS-resistance genes. *Sci Am* 1997, **277**:44–51.
19. Manfredi R, Sabbatani S: A novel antiretroviral class (fusion inhibitors) in the management of HIV infection. Present features and future perspectives of enfuvirtide (T-20). *Curr Med Chem J* 2006, **13**:2369–2384.
20. Aquaro S, D'Arrigo R, Svicher V, Perri GD, Caputo SL, Visco-Comandini U, Santoro M, Bertoli A, Mazzotta F, Bonora S, Tozzi V, Bellagamba R, Zaccarelli M, Narciso P, Antinori A, Perno CF: Specific mutations in HIV-1 gp41 are associated with immunological success in HIV-1-infected patients receiving enfuvirtide treatment. *J Antimicrob Chemother* 2006, **58**:714–722.
21. Dorr P, Westby M, Dobbs S, Griffin P, Irvine B, Macartney M, Mori J, Rickett G, Smith-Burchnell C, Napier C, Webster R, Armour D, Price D, Stammen B, Wood A, Perros M: Maraviroc (UK-427,857), a potent, orally bioavailable, and selective small-molecule inhibitor of chemokine receptor CCR5 with broad-spectrum anti-human immunodeficiency virus type 1 activity.

Antimicrob Agents Chemother 2005, **49**:4721–4732.

22. Cohen J: **The emerging race to cure HIV infections**. *Science* 2011, **332**:784–789.

23. Hutter G, Nowak D, Mossner M, Ganepola S, Mussig A, Allers A, Schneider T, Hofmann J, Kucherer C, Blau O, Blau I, Hofmann WK, Thiel E: **Long-term control of HIV by CCR5 delta32/delta32 stem-cell transplantation.** *New Engl J Med* 2009, **360**:692–698.

24. Allers K, Hütter G, Hofmann J, Loddenkemper C, Rieger K, Thiel E, Schneider T: **Evidence for the cure of HIV infection by CCR5Δ32/Δ32 stem cell transplantation.** *Blood* 2011, **117**:2791–2799.

25. June C, Levine B: **Blocking HIV's attack.** *Sci Am* 2012, **3**:54–59.

26. Reshef R, Luger SM, Hexner EO, Loren AW, Frey NV, Nasta SD, Goldstein SC, Stadtmauer EA, Smith J, Bailey S, Mick R, Heitjan DF, Emerson SG, Hoxie JA, Vonderheide RH, Porter DL: **Blockade of lymphocyte chemotaxis in visceral graft-versus-host disease**. *N Engl J Med* 2012, **367**:135–145.

27. Mallal S, Phillips E, Carosi G, Molina JM, Workman C, Tomažič J, Jägel-Guedes E, Rugina S, Kozyrev O, Flores Cid J, Hay P, Nolan D, Hughes S, Hughes A, Ryan S, Fitch N, Thorborn D, Benbow A: **HLA-B*5701 screening for hypersensitivity to abacavir.** *New Engl J Med* 2008, **358**:568–579.

28. Martin MP, Gao X, Lee HL, Nelson G, Detels R, Goedert JJ, Buchbinder S, Hoots K, Vlahov D, Trowsdale J, Wilson M, O'Brien SJ, Carrington M: **Epistatic interaction between KIR3DS1 and HLA-B delays the progression to AIDS.** *Nat Genet* 2002, **31**:429–434.

29. Martin MP, Qi Y, Gao XJ, Brown EE, Phair J, Goedert JJ, Buchbinder S, Kirk G D, O'Brien SJ, Parham P, McVica EW, Carrington M: **Innate partnership of HLA-B and KIR3DL1 subtypes against HIV-1.** *Nat Genet* 2007, **39**:733–740.

30. Kopp JB, Smith MW, Nelson GW, Johnson RC, Freedman BI, Bowden DW, Oleksyk T, McKenzie LM, Kajiyama H, Ahuja TS, Berns JS, Briggs W, Cho ME, Dart RA, Kimmel PL, Korbet SM, Michel DM, Mokrzycki MH, Schelling JR, Simon E, Trachtman H, Vlahov D, Winkler CA: **MYH9 is a major-effect risk gene for focal segmental glomerulosclerosis.** *Nat Genet* 2008, **40**:1175–84.

31. Kao WH, Klag MJ, Meoni LA, Reich D, Berthier-Schaad Y, Li M, Coresh J, Patterson N, Tandon A, Powe NR, Fink NE, Sadler JH, Weir MR, Abboud HE, Adler SG, Divers J, Iyengar SK, Freedman BI, Kimmel PL, Knowler WC, Kohn OF, Kramp K, Leehey DJ, Nicholas SB, Pahl MV, Schelling JR, Sedor JR, Thornley-Brown D, Winkler CA, Smith MW, Parekh RS, Family Investigation of Nephropathy and Diabetes Research Group: **MYH9 is associated with nondiabetic end-stage renal disease in African Americans.** *Nat Gen* 2008, **40**:1185–1192.

32. Dalmasso C, Carpentier W, Meyer L, Rouzioux C, Goujard C, Chaix ML, Lambotte O, Avettand-Fenoel V, Le Clerc S, de Senneville LD, Deveau C,Boufassa F, Debré P, Delfraissy JF, Broet P, Theodorou I, ANRS Genome Wide Association 01: **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.

33. Fellay J, Ge D, Shianna KV, Colombo S, Ledergerber B, Cirulli ET, Urban TJ, Zhang K, Gumbs CE, Smith JP, Castagna A, Cozzi-Lepri A, De Luca A, Easterbrook P, Günthard HF, Mallal S, Mussini C, Dalmau J, Martinez-Picado J, Miro JM, Obel N, Wolinsky SM, Martinson JJ, Detels R, Margolick JB, Jacobson LP, Descombes P, Antonarakis SE, Beckmann JS, O'Brien SJ, *et al.*: **Common genetic variation and the control of HIV-1 in humans.** *PloS Genet* 2009, **5**:e1000791.

34. Le Clerc S, Limou S, Coulonges C, Carpentier W, Dina C, Taing L, Delaneau O, Labib T, Sladek R; ANRS Genomic Group, Deveau C, Guillemain H,Ratsimandresy R, Montes M, Spadoni JL, Therwath A, Schächter F, Matsuda F, Gut I, Lelièvre JD, Lévy Y, Froguel P, Delfraissy JF, Hercberg S, Zagury JF: **Genome wide 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.

35. International HIV Controllers Study, Pereyra F, Jia X, McLaren PJ, Telenti A, de Bakker PI, Walker BD, Ripke S, Brumme CJ, Pulit SL, Carrington M, Kadie CM, Carlson JM, Heckerman D, Graham RR, Plenge RM, Deeks SG, Gianniny L, Crawford G, Sullivan J, Gonzalez E, Davies L, Camargo A, Moore JM, Beattie N, Gupta S, Crenshaw A, Burtt NP, Guiducci C, Gupta N, *et al.*: **The major genetic determinants of HIV-1 control affect HLA class I peptide presentation.** *Science* 2010, **330**:1551–1557.

36. Limou S, Le Clerc S, Coulonges C, Carpentier W, Dina C, Delaneau O, Labib T, Taing L, Sladek R, Deveau C, Ratsimandresy R, Montes M, Spadoni JL,Lelièvre JD, Lévy Y, Therwath A, Schächter F, Matsuda F, Gut I, Froguel P, Delfraissy JF, Hercberg S, Zagury JF; ANRS Genomic Group: **Genome wide 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.

37. Herbeck JT, Gottlieb GS, Winkler CA, Nelson GW, An P, Maust BS, Wong KG, Troyer JL, Goedert JJ, Kessing BD, Detels R, Wolinsky SM, Martinson J, Buchbinder S, Kirk GD, Jacobson LP, Margolick JB, Kaslow RA, O'Brien SJ, Mullins JI: **Multistage genome wide association study identifies a locus at 1q41 associated with rate of HIV-1 disease progression to clinical AIDS.** *J Infect Dis* 2010, **201**:618–626.

38. Limou S, Coulonges C, Herbeck JT, van Manen D, An P, Le Clerc S, Delaneau O, Diop G, Taing L, Montes M, van't Wout AB, Gottlieb GS, Therwath A, Rouzioux C, Delfraissy JF, Lelièvre JD, Lévy Y, Hercberg S, Dina C, Phair J, Donfield S, Goedert JJ, Buchbinder S, Estaquier J, Schächter F, Gut I, Froguel P, Mullins JI, Schuitemaker H, Winkler C, Zagury JF: **Multiple-cohort genetic association study reveals CXCR6 as a new chemokine receptor involved in long-term nonprogression to AIDS.** *J Infect Dis* 2010, **202**:908–915.

39. Pelak K, Goldstein DB, Walley NM, Fellay J, Ge D, Shianna KV, Gumbs C, Gao X, Maia JM, Cronin KD, Hussain SK, Carrington M, Michael NL, Weintrob AC, Infectious Disease Clinical Research Program HIV Working Group, National Institute of Allergy and Infectious Diseases Center for HIV/AIDS Vaccine Immunology (CHAVI): **Host determinants of HIV-1 control in African Americans.** *J Infect Dis* 2010, **201**:1141–1149.

40. Troyer JL, Nelson GW, Lautenberger JA, Chinn L, McIntosh C, Johnson RC, Sezgin E, Kessing B, Malasky M, Hendrickson SL, Li G, Pontius J, Tang M, An P, Winkler CA, Limou S, Le Clerc S, Delaneau O, Zagury JF, Schuitemaker H, van Manen D, Bream JH, Gomperts ED, Buchbinder S, Goedert JJ, Kirk GD, O'Brien SJ: **Genome-wide association study implicates PARD3B-based AIDS restriction.** *J Infect Dis* 2011, **203**:1491–502.

41. Limou S, Delaneau O, van Manen D, An P, Sezgin E, Le Clerc S, Coulonges C, Troyer JL, Veldink JH, van den Berg LH, Spadoni JL, Taing L, Labib T, Montes M, Delfraissy JF, Schachter F, O'Brien SJ, Buchbinder S, van Natta ML, Jabs DA, Froguel P, Schuitemaker H, Winkler CA, Zagury JF: **Multicohort genomewide association study reveals a new signal of protection against HIV-1 acquisition.** *J Infect Dis* 2012, **205**:1155–1162.

42. Goldstein DB: **Common genetic variation and human traits.** *New Engl J Med* 2009, **360**:1696–1698.

43. Cirulli ET, Goldstein DB: **Uncovering the roles of rare variants in common disease through whole-genome sequencing.** *Nat Rev Genet* 2010, **11**:415–425.

44. Kohjima M, Noda Y, Takeya R, Saito N, Takeuchi K, Sumimoto H: **PAR3beta, a novel homologue of the cell polarity protein PAR3, localizes to tight junctions.** *Biochem Biophys Res Commun* 2002, **299**:641–646.

45. Abraham S, Sawaya BE, Safak M, Batuman O, Khalili K, Amini S: **Regulation of MCP-1 gene transcription by Smads and HIV-1 Tat inhuman glial cells.** *Virology* 2003, **309**:196–202.

46. O'Brien SJ: **Stewardship of human biospecimens, DNA, genotype, and clinical data in the GWAS era.** *Ann Rev Genomics Hum Genet* 2009, **10**:193–209.

47. McLaren PJ, Zagury JF, Fellay J: **Joining forces in HIV host genomics: the International HIV Acquisition Consortium**. *CROI* 2012, abstract #295.

48. Petrovski S, Fellay J, Shianna KV, Carpenetti N, Kumwenda J, Kamanga G, Kamwendo DD, Letvin NL, McMichael AJ, Haynes BF, Cohen MS, Goldstein DB, Center for HIV/AIDS Vaccine Immunology: **Common human genetic variants and HIV-1 susceptibility: a genome-wide survey in a homogeneous African population**. *AIDS* 2011, **25**:513–518.

49. Wester CW, Eden SK, Shepherd BE, Bussmann H, Novitsky V, Samuels DC, Hendrickson SL, Winkler CA, O'Brien SJ, Essex M, D'Aquila RT, Degruttola V, Marlink RG: **Risk factors for symptomatic hyperlactatemia and lactic acidosis among combination antiretroviral therapy-treated adults in Botswana: results from a clinical trial.** *AIDS Res Hum Retroviruses* 2012, **28**:759–765.

50. Yu XF, Liu W, Chen J, Kong W, Liu B, Zhu Q, Liang F, McCutchan F, Piyasirisilp S, Lai S. **Maintaining low HIV type 1 env genetic diversity among injection drug users infected with a B/C recombinant and CRF01_AE HIV type 1 in southern China.** *AIDS Res Hum Retroviruses* 2002, **18**:167–170.

51. Cirulli ET, Singh A, Shianna KV, Ge D, Smith JP, Maia JM, Heinzen EL, Goedert JJ, Goldstein DB, Center for HIV/AIDS Vaccine Immunology (CHAVI): **Screening the human exome: a comparison of whole genome and whole transcriptome sequencing.** *Genome Biol* 2010, **11**:R57.

52. Hendrickson, SL, Lautenberger, J, Chinn, L, Phair, J, Goedert, JJ, Vlahov, D, Donfield, S. Buchbinder, SP, Troyer, J, O'Brien, SJ: **Genetic variants in nuclear-encoded mitochondrial genes influence AIDS progression.** *PLoS One* 2010,

5:e12862.

53. Pagliarini DJ, Calvo SE, Chang B, Sheth SA, Vafai SB, Ong SE, Walford GA, Sugiana C, Boneh A, Chen WK, Hill DE, Vidal M, Evans JG, Thorburn DR,Carr SA, Mootha VK: **A mitochondrial protein compendium elucidates complex I disease biology.** *Cell* 2008, **134:**112–123.

54. LewisW, Dalakas MC: **Mitochondrial toxicity of antiviral drugs.** *Nat Med* 1995, **1:**417–422.

55. Hendrickson SL, Kingsley LA, Ruiz-Pesini E, Poole JC, Jacobson LP, Palella FJ, Bream JH, Wallace DC, O'Brien SJ: **Mitochondrial DNA haplogroups influence lipoatrophy after highly active antiretroviral therapy.** *J Acquir Immune Defic Syndr* 2009, **51:**111–116.

56. Hendrickson SL, Hutcheson HB, Ruiz-Pesini E, Poole JC, Lautenberger J, Sezgin E, Kingsley L, Goedert JJ, Vlahov D, Donfield S, Wallace DC, O'Brien SJ: **Mitochondrial DNA haplogroups influence AIDS progression.** *AIDS* 2008, **22:**2429–2439.

57. Chinn LW, Tang M, Kessing BD, Lautenberger JA, Troyer JL, Malasky MJ, McIntosh C, Kirk GD, Wolinsky SM, Buchbinder SP, Gomperts ED, Goedert JJ, O'Brien SJ: **Genetic associations of variants in genes encoding HIV-dependency factors required for HIV-1 infection.** *J Infect Dis* 2010, **202:**1836–1845.

58. Brass AL, Dykxhoorn DM, Benita Y, Yan N, Engelman A, Xavier RJ, Lieberman J, Elledge SJ: **Identification of host proteins required for HIV infection through a functional genomic screen.** *Science* 2008, **319:**921–926.

59. König R, Zhou Y, Elleder D, Diamond TL, Bonamy GM, Irelan JT, Chiang CY, Tu BP, De Jesus PD, Lilley CE, Seidel S, Opaluch AM, Caldwell JS,Weitzman MD, Kuhen KL, Bandyopadhyay S, Ideker T, Orth AP, Miraglia LJ, Bushman FD, Young JA, Chanda SK: **Global analysis of host-pathogen interactions that regulate early-stage HIV-1 replication.** *Cell* 2008, **135:**49–60.

60. Zhou H, Xu M, Huang Q, Gates AT, Zhang XD, Castle JC, Stec E, Ferrer M, Strulovici B, Hazuda DJ, Espeseth AS: **Genome-scale RNAi screen for host factors required for HIV replication.** *Cell Host Microbe* 2008, **4:**495–504.

61. Yeung ML, Houzet L, Yedavalli VS, Jeang KT: **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.

62. Parsons DW, Jones S, Zhang X, Lin JC, Leary RJ, Angenendt P, Mankoo P, Carter H, Siu IM, Gallia GL, Olivi A, McLendon R, Rasheed BA, Keir S, Nikolskaya I, Nikolsky Y, Busam DA, Tekleab H, Diaz LA Jr, Hartigan J, Smith DR, Strausberg RL, Marie SK, Shinjo SM, Yan H, Riggins GJ, Bigner DD, Karchin R, Papadopoulos N, Parmigiani G, *et al.*: **An integrated genomic analysis of human glioblastoma multiforme.** *Science* 2008, **321:**1807–1812.

63. Gross S, Cairns RA, Minden MD, Driggers EM, Bittinger MA, Jang HG, Sasaki M, Jin S, Schenkein DP, Su SM, Dang L, Fantin VR, Mak TW: **Cancer-associated metabolite 2-hydroxyglutarate accumulates in acute myelogenous leukemia with isocitrate dehydrogenase 1 and 2 mutations.** *J Exp Med* 2010, **207:**339–344.

64. Garber K: **Oncometabolite? IDH1 discoveries raise possibility of new metabolism targets in brain cancers and leukemia.** *J Nat Cancer Inst* **102:**926–928.

65 ENCODE Project Consortium, Bernstein BE, Birney E, Dunham I, Green ED, Gunter C, Snyder M: **An integrated encyclopedia of DNA elements in the human genome.** *Nature* 2012, **489:**57–74.

66. Bailey RC, Moses S, Parker CB, Agot K, Maclean I, Krieger JN, Williams CF, Campbell RT, Ndinya-Achola JO: **Male circumcision for HIV prevention in young men in Kisumu, Kenya: a randomised controlled trial.** *Lancet* 2007, **369:**643–656.

67. Gray RH, Kigozi G, Serwadda D, Makumbi F, Watya S, Nalugoda F, Kiwanuka N, Moulton LH, Chaudhary MA, Chen MZ, Sewankambo NK, Wabwire-Mangen F, Bacon MC, Williams CF, Opendi P, Reynolds SJ, Laeyendecker O, Quinn TC, Wawer MJ: **Male circumcision for HIV prevention in men in Rakai, Uganda: a randomised trial.** *Lancet* 2007, **369:**657–666.

68. Grant RM, Lama JR, Anderson PL, McMahan V, Liu AY, Vargas L, Goicochea P, Casapía M, Guanira-Carranza JV, Ramirez-Cardich ME, Montoya-Herrera O, Fernández T, Veloso VG, Buchbinder SP, Chariyalertsak S, Schechter M, Bekker LG, Mayer KH, Kallás EG, Amico KR: **Preexposure chemoprophylaxis for HIV prevention in men who have sex with men.** *New Engl J Med* 2010, **363:**2587–2599.

69. **Safety and Effectiveness of Tenofovir 1% Gel, Tenofovir Disproxil Fumarate, and Emtricitabine/Tenofovir Disoproxil Fumarate Tablets in Preventing HIV in Women. Clinical trial** [http://clinicaltrials.gov/ct2/show/NCT00705679]

70. Abdool Karim Q, Abdool Karim SS, Frohlich JA, Grobler AC, Baxter C, Mansoor LE, Kharsany AB, Sibeko S, Mlisana KP, Omar Z, Gengiah TN, Maarschalk S, Arulappan N, Mlotshwa M, Morris L, Taylor D, CAPRISA 004 Trial Group: **Effectiveness and safety of tenofovir gel, an antiretroviral microbicide, for the prevention of HIV infection in women.** *Science* 2010, **329:**1168–1174.

71. Carrington M, Kissner T, Gerrard B, Ivanov S, O'Brien SJ, Dean, M: **Novel alleles of the chemokine-receptor gene CCR5.** *Am J Hum Genet* 1997, **61:**1261–1267.

72. Stephens JC, **Dating the origin of the CCR5-Δ32 AIDS resistance gene allele by the coalescence of haplotypes**. *Am J Hum Genet* 1998, **62:**1507–1515.

73. Bamshad MJ, Mummidi S, Gonzalez E, Ahuja SS, Dunn DM, Watkins WS, Wooding S, Stone AC, Jorde LB, Weiss RB, Ahuja SK: **A strong signature of balancing selection in the 5' cis-regulatory region of CCR5.** *Proceed Nat Acad Sci U S A* 2002, **99:**10539–10544.

74. Sabeti PC, Walsh E, Schaffner SF, Varilly P, Fry B, Cullen M, Mikkelsen T, Roy J, Patterson N, Cooper R, Altshuler D, Lander ES: **The case for selection at CCR5Δ32.** *PloS Biol* 2005, **3:**1963–1969.

75. Elvin SJ, Williamson ED, Scott JC, Smith JN, Pérez De Lema G, Chilla S, Clapham P, Pfeffer K, Schlöndorff D, Luckow B: **Evolutionary genetics: Ambiguous role of CCR5 in *Y. pestis* infection.** *Nature* 2004, **430:**417–424.

76. O'Brien SJ: *Tears of the Cheetah and Other Tales from the Genetic Frontier.* St Martin's Press: New York; 2003.

77. Glass WG, Lim JK, Cholera R, Pletnev AG, Gao JL, Murphy PM: **Chemokine receptor CCR5 promotes leukocyte trafficking to the brain and survival in West Nile virus infection.** *J Exp Med* 2005, **202:**1087–98.

78. Glass WG, McDermott DH, Lim JK, Lekhong S, Yu SF, Frank WA, Pape J, Cheshier RC, Murphy PM: **CCR5 deficiency increases risk of symptomatic West Nile virus infection.** *J Exp Med* 2006, **203:**35–40.

79. O'Brien SJ, Moore J: **The effect of genetic variation in chemokines and their receptors on HIV transmission and progression to AIDS.** *Immunol Rev* 2000, **177:**99–111.

80. Smith MW, Dean M, Carrington M, Winkler C, Huttley GA, Lomb DA, Goedert JJ, O'Brien TR, Jacobson LP, Kaslow R, Buchbinder S, Vittinghoff E, Vlahov D, Hoots K, Hilgartner MW, O'Brien SJ: **Contrasting genetic influence of CCR2 and CCR5 variants on HIV-1 infection and disease progression.** *Science* 1997, **277:**959–965.

81. Martin MP, Dean M, Smith MW, Winkler C, Gerrard B, Michael NL, Lee B, Doms RW, Margolick J, Buchbinder S, Goedert JJ, O'Brien TR, Hilgartner MW, Vlahov D, O'Brien SJ, Carrington M: **Genetic acceleration of AIDS progression by a promoter variant of CCR5.** *Science* 1998, **282:**1907–1911.

82. Winkler C, Modi W, Smith MW, Nelson GW, Wu X, Carrington M, Dean M, Honjo T, Tashiro K, Yabe D, Buchbinder S, Vittinghoff E, Goedert JJ, O'Brien TR, Jacobson LP, Detels R, Donfield S, Willoughby A, Gomperts E, Vlahov D, Phair J, O'Brien SJ: **Genetic restriction of AIDS pathogenesis by an SDF-1 chemokine gene variant.** *Science* 1998, **279:**389–393.

83. Carrington M, Nelson GW, Martin MP, Kissner T, Vlahov D, Goedert JJ, Kaslow R, Buchbinder S, Hoots K, O'Brien SJ: **HLA and HIV-1:** **Heterozygote advantage and B*35-Cw*04 disadvantage.** *Science* 1999, **283:**1748–1752.

84. Shin HD, Winkler C, Stephens JC, Bream J, Young H, Goedert JJ, O'Brien TR, Vlahov D, Buchbinder S, Giorgi J, Rinaldo C, Donfield S, Willoughby A, O'Brien SJ, Smith MW: **Genetic restriction of HIV-1 pathogenesis to AIDS by promoter alleles of IL10.** *Proc Natl Acad Sci U S A* 2000, **97:**14467–14472

85. Gao X, Nelson GW, Karacki P, Martin MP, Phair J, Kaslow R, Goedert JJ, Buchbinder S, Hoots K, Vlahov D, O'Brien SJ, Carrington M: **Effect of a single amino acid change in MHC class I molecules on the rate of progression to AIDS.** *New Engl J Med* 2001, **344:**1668–1675.

86. An P, Nelson GW, Wang L, Donfield S, Goedert JJ, Phair J, Vlahov D, Buchbinder S, Farrar WL, Modi W, O'Brien SJ, Winkler CA: **Modulating influence on HIV/AIDS by interacting RANTES gene variants.** *Proc Natl Acad Sci U S A* 2002, **99:**10002–10007.

87. Modi WS, Goedert JJ, Strathdee S, Buchbinder S, Detels R, Donfield S, O'Brien SJ, Winkler C: **MCP-1-MCP-3-Eotaxin gene cluster influences HIV-1 transmission.** *AIDS* 2003, **17:**2357–2365.

88. Carrington M, O'Brien SJ: **The influence of HLA genotype on AIDS.** *Ann Rev Med* 2003, **54:**535–551.

89. An P, Vlahov D, Margolick JB, Phair J, O'Brien TR, Lautenberger J, O'Brien SJ, Winkler CA: **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.

90. Duggal P, An P, Beaty TH, Strathdee SA, Farzadegan H, Markham RB, Johnson L, O'Brien SJ, Vlahov D, Winkler CA: **Genetic influence of CXCR6 chemokine receptor alleles on PCP-mediated AIDS progression among African-**

Americans. *Genes Immun* 2003, **4**:245–250.

91. An P, Bleiber G, Duggal P, Nelson G, May M, Mangeat B, Alobwede I, Trono D, Vlahov D, Donfield S, Goedert JJ, Phair J, Buchbinder S, O'Brien SJ, Telenti A, Winkler CA: **APOBEC3G genetic variants and their influence on the progression to AIDS**. J Virol 2004, **78**:11070–11076.

92. Martin MP, Lederman MM, Hutcheson HB, Goedert JJ, Nelson GW, van Kooyk Y, Detels R, Buchbinder S, Hoots K, Vlahov D, O'Brien SJ, Carrington M: **Association of DC SIGN promoter polymorphisms with increased risk for parenteral but not mucosal acquisition of HIV-1 infection.** *J Virol* 2003, **78**:14053–14056.

93. Gao X, Bashirova A, Iversen AK, Phair J, Goedert JJ, Buchbinder S, Hoots K, Vlahov D, Altfeld M, O'Brien SJ, Carrington M: **AIDS restriction HLA allotypes target distinct intervals of HIV-1 pathogenesis**. *Nat Med* 2006, **11**:1290–1292.

94. Bashirova AA, Bleiber G, Qi Y, Hutcheson H, Yamashita T, Johnson RC, Cheng J, Alter G, Goedert JJ, Buchbinder S, Hoots K, Vlahov D, May M, Maldarelli F, Jacobson L, O'Brien SJ, Telenti A, Carrington M: **Consistent effects of TSG101 genetic variability on multiple outcomes of exposure to human immunodeficiency virus type 1.** *J Virol* 2006, **80**:6757–6763.

95. Javanbakht H, An P, Gold B, Petersen DC, O'Huigin C, Nelson GW, O'Brien SJ, Kirk GD, Detels R, Buchbinder S, Donfield S, Shulenin S, Song B, Perron MJ, Stremlau M, Sodroski J, Dean M, Winkler C: **Effects of human TRIM5alpha polymorphisms on antiretroviral function and susceptibility to human immunodeficiency virus infection.** *Virology* 2006, **354**:15–27.

96. An P, Duggal P, Wang LH, O'Brien SJ, Donfield S, Goedert JJ, Phair J, Buchbinder S, Kirk GD, Winkler CA: **Polymorphisms of CUL5 area ssociated with CD4+ T cell loss in HIV-1 infected individuals**. *PLoS Genet* 2007, **3**:e19.

97. An P, Wang LH, Hutcheson-Dilks H, Nelson G, Donfield S, Goedert JJ, Rinaldo CR, Buchbinder S, Kirk GD, O'Brien SJ, Winkler CA: **Regulatory polymorphisms in the cyclophilin A gene, PPIA accelerate progression to AIDS**. *PLoS Pathog* 2007, **3**:849–857.

98. Welzel TM, Gao X, Pfeiffer RM, Martin MP, O'Brien SJ, Goedert JJ, Carrington M, O'Brien TR: **HLA Bw4 alleles and HIV-1 transmission in heterosexual couples.** *AIDS* 2007, **21**:225–229.

99. Hendrickson SL, Jabs DA, Van Natta M, Lewis RA, Wallace DC, O'Brien SJ: **Mitochondrial haplogroups are associated with risk of neuroretinal disorder in HIV-positive patients.** *J Acquir Immune Defic Syndr* 2010, **53**:451–455.

100. An P, Johnson R, Phair J, Kirk GD, Yu XF, Donfield S, Buchbinder S, Goedert JJ, Winkler CA: **APOBEC3B** deletion and risk of HIV-1 acquisition. *J Infect Dis* 2008, **200**:1054–1058.

101. O'Brien SJ, Nelson GW, Winkler CA, Smith MW: **Polygenic and multifactorial disease gene association in man: lessons from AIDS**. *Ann Rev Genet* 2000, **34**:563–591.