

Association of Y chromosome haplogroup I with HIV progression, and HAART outcome

Efe Sezgin · Joanne M. Lind · Sadeep Shrestha · Sher Hendrickson ·
James J. Goedert · Sharyne Donfield · Gregory D. Kirk · John P. Phair ·
Jennifer L. Troyer · Stephen J. O'Brien · Michael W. Smith

Received: 11 November 2008 / Accepted: 25 December 2008 / Published online: 24 January 2009
© US Government 2009

Abstract The host genetic basis of differential outcomes in HIV infection, progression, viral load set point and highly active retroviral therapy (HAART) responses was examined for the common Y haplogroups in European Americans and African Americans. Accelerated progression to acquired immune deficiency syndrome (AIDS) and related death in European Americans among Y chromosome haplogroup I (Y-I) subjects was discovered. Additionally, Y-I haplogroup subjects on HAART took a longer time to HIV-1 viral suppression and were more likely to fail HAART. Both the accelerated progression and longer time to viral suppression results observed in haplogroup Y-I

were significant after false-discovery-rate corrections. A higher frequency of AIDS-defining illnesses was also observed in haplogroup Y-I. These effects were independent of the previously identified autosomal AIDS restriction genes. When the Y-I haplogroup subjects were further subdivided into six I subhaplogroups, no one subhaplogroup accounted for the effects on HIV progression, viral load or HAART response. Adjustment of the analyses for population stratification found significant and concordant haplogroup Y-I results. The Y chromosome haplogroup analyses of HIV infection and progression in African Americans were not significant. Our results suggest that

Electronic supplementary material The online version of this article (doi:[10.1007/s00439-008-0620-7](https://doi.org/10.1007/s00439-008-0620-7)) contains supplementary material, which is available to authorized users.

E. Sezgin · J. M. Lind · S. Shrestha · S. Hendrickson ·
J. L. Troyer · S. J. O'Brien · M. W. Smith
Laboratory of Genomic Diversity,
NCI-Frederick, Frederick, MD 21702, USA

S. Shrestha · J. L. Troyer · M. W. Smith
Basic Research Program, SAIC-Frederick, Inc.,
NCI-Frederick, Frederick, MD 21702, USA

J. J. Goedert
Viral Epidemiology Branch,
National Cancer Institute, Rockville,
MD 20582, USA

S. Donfield
Rho, Incorporated, Chapel Hill,
NC 27517, USA

G. D. Kirk
Department of Epidemiology,
Johns Hopkins Bloomberg School of Public Health,
Baltimore, MD 21205, USA

J. P. Phair
Northwestern University Medical School,
Comprehensive AIDS Center, 6980 North Lake Shore Drive,
Suite 1106, Chicago, IL 60611, USA

Present Address:
J. M. Lind
School of Medicine,
University of Western Sydney, Sydney, Australia

Present Address:
M. W. Smith (✉)
Genetics and Genomics, Advanced Technology Program,
SAIC-Frederick, Inc., National Cancer Institute,
915 Tollhouse Ave., Suite 211, Frederick, MD 21701, USA
e-mail: smithmw@mail.nih.gov

Present Address:
S. Shrestha
Department of Epidemiology and International Health,
School of Public Health, University of Alabama,
Birmingham, AL 35294, USA

one or more loci on the Y chromosome found on haplogroup Y-I have an effect on AIDS progression and treatment responses in European Americans.

Introduction

Genetic variation has been shown to play an important role in acquired immune deficiency syndrome (AIDS), with numerous AIDS restriction genes (ARGs) affecting HIV-1 cell entry, acquired and innate immunity, and cytokine defenses against HIV-1 (O'Brien and Nelson 2004). In the United States, HIV is now one of the leading causes of death in young, predominantly male, adults (Anderson 2001). While there is little evidence to suggest a gender difference in progression to AIDS (Pezzotti et al. 1996), the immune environments between males and females differ in T helper (T_H) lymphocyte response, antibody production and cell-mediated immunity (Gleicher and Barad 2007; Marriott and Huet-Hudson 2006; McCarthy 2000; Whitacre 2001; Whitacre et al. 1999). Proteins encoded by the Y chromosome may play a role in this difference between male and female immune responses, as well as variable immune responses within males (Teuschert et al. 2006; Wesley et al. 2007). In humans, males are more likely to exhibit risky behaviors such as physical aggression, resulting in injuries, and are more likely to travel to new geographical regions (Bosch et al. 2003; Carvajal-Carmona et al. 2000; Carvajal-Carmona et al. 2003; Cavalli-Sforza and Hewlett 1982; Mesa et al. 2000), resulting in more exposure to infectious diseases. These and other circumstances have created unique selective environments that likely specifically impacted Y chromosome evolution and variation (Jobling et al. 1998).

The haploid 60-megabase Y chromosome is obviously unique for genetic studies because only males are carriers. Further, 95% of the Y chromosome that does not undergo recombination contains 78 redundant protein-coding genes that encode 27 distinct proteins (Jobling and Tyler-Smith 2003; Skaletsky et al. 2003). Previous work has defined a Y chromosome phylogeny of 311 haplogroups where 20 major clades are evident (Karafet et al. 2008). Over the past decade attempts have been made to carry out association studies of Y chromosome mutations with differing male phenotypes and some effects related to infertility, testicular cancer and hypertension have been suggested (Krausz et al. 2004).

Previous studies have also examined Y chromosome genes for immune cell expression, immunological function, and the role of the encoded proteins in the transplantation response. We know that of the 11 ubiquitously expressed genes, five (*DDX3Y*, *UTY*, *USP9Y*, *SMCY* and *TMSB4Y*) are highly expressed in immunological tissues (Ditton et al.

2004; Lahn and Page 1997; Skaletsky et al. 2003). All five of these proteins also function as transplantation antigens in graft-versus-host disease (Agulnik et al. 1994; Ivanov et al. 2005; Laurin et al. 2006; Torikai et al. 2004; Vogt et al. 2000; Warren et al. 2000). These results suggest that several Y chromosome genes play a role in immunological monitoring.

Extensive studies of Y chromosome variability have focused on human population structure and migration patterns. Comparisons of the Y chromosome, autosomes, and the mitochondrial DNA along with human history suggests different migration patterns of males and females (Hamilton et al. 2005; Seielstad et al. 1998; Wilder et al. 2004; Wood et al. 2005). Some have argued that, in general, females and males in pre-agricultural societies migrated at equal rates (Wilkins and Marlowe 2006). In contrast, some studies have reported elevated female migration rates (Hamilton et al. 2005; Seielstad et al. 1998) which must be balanced with more recent male directional migration associated with colonization (Bosch et al. 2003; Mesa et al. 2000; Wilkins 2006). Diamond (2005) and colleagues (Dobson and Carper 1996; Wolfe et al. 2007) have argued that infectious diseases played a major role in the European colonization and the associated failure to expand in the tropics. Taken together, there is a unique history of the Y chromosome relative to autosomes and the X chromosome.

We postulated that underlying functional variants in any infection response and resistance could be identified by a Y chromosome major haplogroup analysis. Markers for the ten known major haplogroups of European Americans and African Americans were genotyped in five major HIV-1/AIDS cohorts. The haplogroups were then examined for association with HIV-1 infection, progression, viral load set point and highly active retroviral therapy (HAART) response.

Materials and methods

Study population

We studied 3,727 males (2,292 European Americans, 1,233 African Americans, and 202 individuals from other racial groups) from five HIV-1/AIDS cohorts: Multicenter AIDS Cohort Study (MACS $n = 1,594$) (Phair et al. 1992); San Francisco City Cohort (SFCC $n = 195$) (Buchbinder et al. 1994); Multicenter Hemophilia Cohort Study (MHCS $n = 703$) (Goedert et al. 1989); Hemophilia Growth and Development Study (HGDS $n = 213$) (Hilgartner et al. 1993); and the AIDS Linked to Intravenous Experience (ALIVE $n = 1,022$) (Vlahov et al. 1998). Participants were divided into one of three categories as follows: seronegative (SN, negative result with HIV antibody test, $n = 656$);

seroconverter (SC, date of HIV-1 infection known, $n = 920$); and seroprevalent (SP, HIV-1-positive before study enrollment, $n = 2,151$). SCs were limited to those with 2 years or fewer between first positive and last negative clinic visits. Study subjects were HAART naïve individuals and thus were censored until a conservative HAART initiation cutoff date of 1 January 1996 since HAART became available in 1996. An exception to this was 467 MACS subjects that were only used for HAART analyses (see “HAART analysis” section). Estimated dates of seroconversion were available for SP subjects from their respective cohorts (Buchbinder et al. 1994; Goedert et al. 1989; Phair et al. 1992).

Individual DNA samples were extracted from the immortal lymphoblastoid B cell lines as previously described (Dean et al. 1996). This study was approved by the Protocol Review Office of the Institutional Review Board of the National Cancer Institute. Informed consent was obtained from all individuals at the study sites.

Y chromosome analysis

Each male was genotyped for ten single nucleotide polymorphisms (SNPs) located on the non-recombining region of the Y chromosome (NRY) using the 5′ nuclease assay (TaqMan Assay-by-Design SNP genotyping products, Applied Biosystems) (Lind et al. 2007). Males were then assigned to 1 of 11 major Y chromosome haplogroups (those common in Europeans and Africans) according to the Y Chromosome Consortium (2002). Individuals who had the m89 derived allele but did not belong to the H, I or K haplogroups were joined to form the juxtaposed F* (xI, H, K) group and referred as F* throughout the manuscript. After some promising results were obtained for haplogroup I (Y-I), five additional markers, M253, M227, P37, M26, and M223 (Cinnioglu et al. 2004; Rootsi et al. 2004; Underhill et al. 2001) were genotyped to define the I1a*, I1a4, I1b*, I1b2, and I1c subhaplogroups (the corresponding YCC 2008 (Karafet et al. 2008) names are I1, I1b, I2a, I2a2 and I2b, respectively). The remaining Y-I haplogroup subjects who lacked the derived alleles at these five markers were assigned into the I* (xI1a*, xI1a4, xI1b*, xI1b2, xI1c) haplogroup.

Statistical analyses

Haplogroups with frequency less than 5% were excluded from statistical analysis. Haplogroups compared were I, R and F* in European Americans, and I, R and E in African Americans. Statistical analyses were performed with SAS 9.1 (SAS Institute, Cary, NC, USA). P values reported were nominal and uncorrected for multiple comparisons. The proportion of false positives due to multiple testing

was estimated using the QVALUE (version 1.0) program (Storey and Tibshirani 2003). A q value cutoff of 0.05 was used for significance. Throughout the main body of this paper q values are given when $P \leq 0.05$ and in the display items q values ≤ 0.05 are noted.

Population structure

For 90% of the European-American and 97% of the African-American SCs, 800K autosomal SNP data from Affymetrix 6.0 Genechip platform (Troyer et al. in preparation) with call rates greater than or equal to 95% was applied a principal components analysis (PCA) using the EIGENSOFT (Price et al. 2006) program to examine and adjust for potential population stratification. ANOVA F statistic was performed on the recovered eigenvectors given the Y chromosome haplogroups. Further, the top three most informative eigenvector values for each subject were included in a Cox proportional hazards model to correct for their contribution to the estimated hazard ratios.

HIV infection and progression

Separate analyses were conducted for European-American and African-American subjects. Haplogroup frequencies were compared in HIV-1-negative and HIV-1-positive subjects using the log-likelihood chi-square test. For SC subjects the Kaplan–Meier and Cox proportional hazard model analyses were used to examine the relationship between Y haplogroups and four AIDS-related outcomes: time to CD4 T lymphocyte count of fewer than 200 per cubic millimeter ($CD4 < 200$); AIDS-1993 definition (Centers for Disease Control 1992); AIDS-1987 definition (Centers for Disease Control 1986); and death from AIDS-defining illnesses prior to HAART use. The Cox models were run by stratifying for cohorts in order to account for individual cohort effects. The nine ALIVE European American samples, of which only three had AIDS outcome data, were not included in the stratified analyses to maintain statistical power. The time dependence of AIDS progression was examined initially using graphical analyses and later with partitioned time interval Cox analysis. Genes known to influence HIV-1 progression [*CCR5-Δ32* (Dean et al. 1996), *CCR2-64I* (Smith et al. 1997), *CCR5-P1/P1* promoter (Martin et al. 1998), and *IL10-5′A* (Shin et al. 2000)] were accounted for in the Cox proportional hazard models. More generally, the full set of known ARGs was summarized as a genetic propensity index (O’Brien and Nelson 2004) that also included *HLA-B*27*, *HLA-B*57*, *HLA-B*35*, *HLA-B*35Px*, *KIR3DS1*, *TSG101*, *SDF*, *RANTES*, and HLA class I heterozygosity for each subject for the four AIDS-related outcomes and used as a continuous variable in the Cox analyses. SP subjects who avoided one or more

AIDS outcomes for more than 9 years were included only in the categorical analyses of AIDS progression and outcomes similar to our previous reports (Martin et al. 1998; Shin et al. 2000; Smith et al. 1997).

HIV-1 viral load set point analysis

Viral load measurements for 381 European-American SCs from the MACS cohort were quantified using either the reverse-transcription polymerase chain reaction (RT-PCR, Amplicor; Roche Diagnostics, Nutley, NJ, USA) or the Roche Ultrasensitive Assay described previously (Lyles et al. 2000). Viral load set points were calculated as the mean of all viral measurements for these SCs using the HIV-1 RNA measurements obtained between 6 months to 3 years after the first HIV-seropositive visit and analyzed as \log_{10} HIV-1 RNA copies/mL.

HAART analysis

HAART was defined (Office of AIDS Research Advisory Council 2006) as one of the following: (1) two or more nucleoside reverse transcriptase inhibitors (NRTIs) in combination with at least one protease inhibitors (PI) or one non-nucleoside reverse transcriptase inhibitors (NNRTI); (2) one NRTI in combination with at least one PI and at least one NNRTI; (3) a regimen consisting of ritonavir and saquinavir in combination with one NRTI and no NNRTIs; and (4) an abacavir- or tenofovir- containing regimen of three or more NRTIs without both PIs and NNRTIs. Combinations of zidovudine (AZT) and stavudine (d4T) with either a PI or NNRTI were not considered HAART. The date of HAART initiation was defined as the midpoint between the last visit without HAART and the first visit at which HAART was reported. Only subjects with less than 1 year between these visits were included in the study. Baseline viral load and CD4 counts were taken within 6 months of HAART initiation.

We only used the European-American MACS cohort subjects to detect the suppression of viral load upon HAART. Time to suppression was defined as the time from HAART initiation to the first undetectable viral load visit for each subject. Differential survival rates were examined graphically using Kaplan–Meier analyses and with the hazard ratio using Cox proportional hazards models adjusted for clinical AIDS prior to HAART, baseline viral load, and CD4 cell count, and *CCR5-Δ32*. The measurement of HIV-1 RNA was obtained by RT-PCR (Amplicor HIV Monitor Assay, Roche Diagnostics, Nutley, NJ). We performed separate analysis of subjects below thresholds of <200 copies/mL and <50 copies/mL. Failure of viral suppression (<200 copies/mL) was reported as HAART failure.

Results

In our analyses, we focused on the European-American and African-American samples, and excluded other racial groups due to their small sample size in these cohorts. The Y major haplogroup frequencies of the remaining 3,490 subjects are shown in Table 1.

European Americans

Population structure and stratification analysis

Since specific Y chromosomes track population stratification in a population study, we determined whether any observed signals were due to the Y chromosome itself or due to an autosomal region associated with the Y chromosome because of geographical origins. We examined data for 800K autosomal SNPs and applied the EIGENSOFT-PCA (Price et al. 2006) approach to identify and correct for population stratification in the I, R and F* haplogroup subjects. The known substructure of European Americans (Price et al. 2008) from northwest to southeast Europe and the division within southern Europeans were evident. An ANOVA of the most significant (top three) eigen vector values (principal components) showed that I and R haplogroup subjects were not significantly different from each other (vector-1: $P = 0.60$; vector-2: $P = 0.08$; vector-3: $P = 0.67$). However, the distribution of the F* subjects along the first and second eigen vectors were significantly different compared to the I and R subjects (vector-1: $df = 2$, $P = 0.0001$; vector-2: $df = 2$, $P = 0.001$; vector-3: $df = 2$, $P = 0.91$). For completeness, the progression analyses (of SCs) were corrected for these top three most informative eigen vectors.

HIV-1 infection and AIDS progression

An analysis of susceptibility to infection showed no significant frequency difference when comparing HIV-1-negative and SC subjects in European Americans ($P = 0.38$; Table 1). A survival difference for the AIDS-1987 outcome was observed between the SCs belonging to the three most common (I, R, and F*) European-American haplogroups ($df = 2$, log-rank $P = 0.02$, $q = 0.19$; Fig. 1), where faster disease progression in the Y-I haplogroup subjects was evident. A trend toward faster AIDS progression in the Y-I haplogroup was also suggested by the Cox proportional hazards analyses [Table 2; relative hazard (RH) = 1.35, $P = 0.07$ for AIDS-1987]. A close examination of survival plots indicated acceleration to AIDS in the latter stage of HIV-1 pathogenesis among the Y-I haplogroup subjects. To address the time dependence of AIDS progression of Y-I haplogroup individuals, various partitioned time intervals were examined

Table 1 Frequency (%) of Y chromosome haplogroups in European-American and African-American seroconverters, seroprevalents and seronegatives

Haplogroup	Markers ^a										Sero-status					
											European Americans ^b			African Americans ^c		
	m60	m216	m203	m96	m89	m69	m170	m9	m45	m207	Converter	Prevalent	Negative	Converter	Prevalent	Negative
											<i>n</i> = 625	<i>n</i> = 1411	<i>n</i> = 229	<i>n</i> = 246	<i>n</i> = 582	<i>n</i> = 397
Y* (xB, C, DE, F)	D	C	G	G	G	T	A	C	G	A	0.0	0.1	0.4	1.4	0.5	0.8
B	T	C	G	G	G	T	A	C	G	A	0.0	0.1	0.0	1.8	3.1	2.5
C	D	T	G	G	G	T	A	C	G	A	0.0	0.1	0.0	0.0	0.0	0.0
DE (xE)	D	C	C	G	G	T	A	C	G	A	0.1	0.0	0.0	0.5	0.2	1.5
E	D	C	C	C	G	T	A	C	G	A	5.8	7.5	7.0	76.5	67.8	71.5
F* (xI, H, K)	D	C	G	G	A	T	A	C	G	A	10.9	10.6	12.3	0.9	2.2	2.0
H	D	C	G	G	A	C	A	C	G	A	0.0	0.1	0.0	0.0	0.0	0.0
I	D	C	G	G	A	T	C	C	G	A	16.2	17.8	19.3	4.1	5.2	4.8
K* (xP)	D	C	G	G	A	T	A	G	G	A	1.0	2.2	1.8	0.0	0.9	0.5
P* (xR)	D	C	G	G	A	T	A	G	A	A	0.3	0.8	0.0	0.0	0.5	0.3
R	D	C	G	G	A	T	A	G	A	G	65.6	60.8	59.2	17.5	19.6	16.1

^a The order of markers reflects the hierarchical structure of the Y chromosome phylogeny. Each individual was typed for the non-recombining Y biallelic markers m60(rs2032623); m216(rs2032666); m203(rs2032623); m96(rs9306841); m89(rs2032652); m69(rs2032673); m170(rs2032597); m9(rs3900); m45(rs2032631); m207(rs2032658) with typing results shown for each haplogroup. Following the YCC nomenclature, individuals who had the m89 derived allele but do not belong to the H, I or K haplogroups are joined together to form the juxtaposed F* (xI, H, K) group. Similarly, individuals with the m9 derived allele but not P and individuals with the derived m45 allele but not R are represented as K*(xP) and P*(xR), respectively. Ambiguous genotypes or genotyping failure occurred for 35 males (27 European Americans and 8 African Americans). Our haplogroup assignments, based on YCC 2002 major haplogroup defining markers, are in good agreement with the revised YCC 2008 Y chromosome haplogroup tree (Karafet et al. 2008). The only difference is the joining of I and J haplogroups into the 'IJ' clade on the revised Y chromosome tree. We note that most individuals in the F* grouping are haplogroup J (Karafet et al. 2008)

^b The distribution of haplogroups between seroconverters and seronegatives were not significantly different either considering all haplogroups ($P = 0.34$) or only the major haplogroups of E, R and I ($P = 0.38$)

^c The distribution of haplogroups between seroconverters and seronegatives were not significantly different either considering all haplogroups ($P = 0.40$) or only the major haplogroups of R, I, F*, and E ($P = 0.45$)

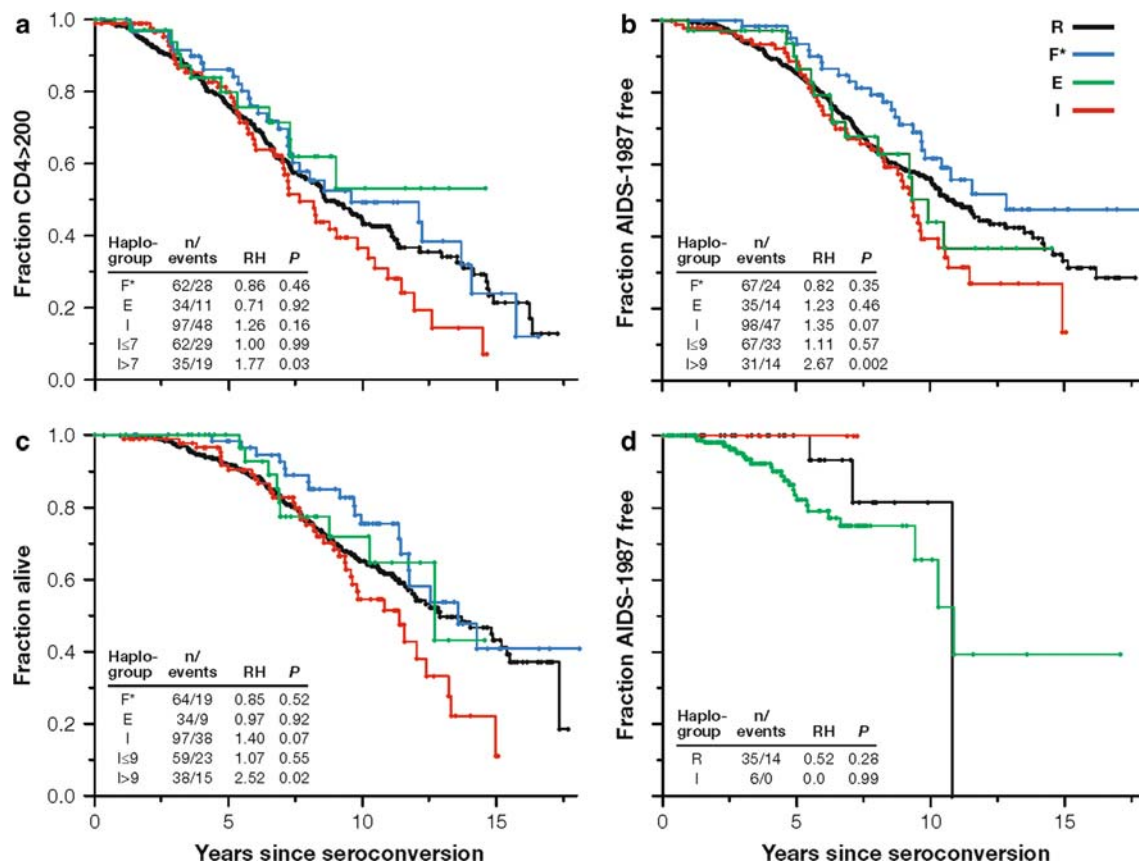


Fig. 1 Survival analyses of most common European-American and African-American Y chromosome haplogroups in all AIDS cohorts. Number of individuals (*n*) and number of events, Cox proportional hazards model *P* values (*P*) and relative hazards (RH) are presented. **a–c** Kaplan–Meier survival curve of European-American individuals for three AIDS-related outcomes. Haplogroup R is the reference

(data not shown). Generally, the most significant Cox models were divided into early progressors (0–7 years after seroconversion) and late/slow progressors (>7 years after seroconversion) for CD4 < 200 and AIDS-1993 outcomes; and at 9 years for the AIDS-1987 and death. Our Cox model interval analyses indicated that Y-I haplogroup subjects depleted CD4 cells more quickly (RH = 2.05, *P* = 0.007, *q* = 0.10) and progressed to AIDS (RH = 2.84, *P* = 0.001, *q* = 0.03) and death (RH = 2.48, *P* = 0.003, *q* = 0.07) faster in the latter phase (Table 2). The faster progression to AIDS-1987 and death (in the MHCS cohort) results were significant after multiple test corrections among the Y-I haplogroup individuals (Table 3). This direction of association was consistent in all the cohorts, but the results were not always as significant (Supplementary Table 1). An additional categorical analysis, including the seroprevalent subjects also showed the same trend (results not shown).

To address whether any one of the Y-I subhaplogroups carried the faster progression signal, we further genotyped the Y-I haplogroup subjects for the most common Y-I subhaplogroup markers (I1a*, I1a4, I1b*, I1b2 and I1c) known

group. Haplogroup E was excluded from further analyses as it had only 35 subjects (5.8% of seroconverters, Table 1) and behaved nearly identically to the reference haplogroup R for the major AIDS outcomes. **d** Kaplan–Meier survival curve of African-American individuals. Most common E haplogroup is used as the reference group

from Europe. Y-I subhaplogroups show geographical frequency differences in Europe most probably representing their places of origin. Subhaplogroups I1a* and I1a4 are mostly found in northern Europe. I1b* and I1b2 are most frequent in Eastern Europe and the Balkans, with I1b2 reaching the highest frequency in Sardinia. Finally, I1c despite its wide range in Europe, is found mostly in north-west Europe (Rootsi et al. 2004). The observed frequencies of I1a*, I1a4, I1b*, I1b2 and I1c were 59.2, 0.3, 11.5, 2.8 and 23.1%, respectively. The remaining 3.1% of the Y-I haplogroup samples that lacked these five subhaplogroup-defining derived alleles were labeled I*. The Cox analyses suggested a relatively faster progression signal for the relatively common I1a* subhaplogroup (RH = 1.42, *P* = 0.09 for AIDS-1987; Table 4) compared to the other subhaplogroups. However, a Kaplan–Meier survival analysis did not find significant differences between the subhaplogroups (*df* = 4, log-rank *P* = 0.72; Supplementary Fig. 1) indicating that the Y-I haplogroup as a whole best explains the acceleration to AIDS outcomes rather than any particular Y-I subhaplogroup. To correct for any population stratification

Table 2 Faster progression of Y-I haplogroup subjects to AIDS outcomes among European-American seroconverters in early and late time intervals

Haplogroup ^a	Time interval, years ^b	CD4 < 200			AIDS-1993			AIDS-1987			Death		
		n/events	RH ^c (95% CI)	P	n/events	RH ^c (95% CI)	P	n/events	RH ^c (95% CI)	P	n/events	RH ^c (95% CI)	P
No genetic propensity adjustment													
R		372/172	1.00	–	372/187	1.00	–	391/169	1.00	–	385/130	1.00	–
F*		62/28	0.93 (0.62–1.40)	0.74	62/30	0.91 (0.62–1.65)	0.65	66/24	0.82 (0.53–1.26)	0.37	64/19	0.85 (0.53–1.39)	0.52
I		97/48	1.29 (0.93–1.78)	0.12	97/49	1.20 (0.87–1.65)	0.26	98/47	1.35 (0.98–1.87)	0.07	97/38	1.40 (0.97–2.01)	0.07
I	0–7	62/29	1.03 (0.68–1.54)	0.91	63/31	1.00 (0.68–1.48)	0.99	67/33	1.10 (0.76–1.61)	0.60	59/23	1.07 (0.68–1.69)	0.78
I	>7	35/19	2.05 (1.22–3.45)	0.007	34/18	1.76 (1.04–2.99)	0.03	31/14	2.84 (1.51–5.34)	0.001 ^d	38/15	2.48 (1.37–4.52)	0.003
With population stratification adjustment													
R		341/157	1.00	–	341/172	1.00	–	356/153	1.00	–	350/118	1.00	–
F*		55/22	0.88 (0.55–1.42)	0.61	55/24	0.86 (0.55–1.34)	0.50	59/21	0.87 (0.54–1.39)	0.55	58/16	0.94 (0.54–1.62)	0.82
I		87/43	1.44 (1.02–2.03)	0.04	87/44	1.32 (0.95–1.86)	0.10	87/44	1.57 (1.12–2.21)	0.009	86/35	1.44 (1.10–2.38)	0.01
I	0–7	59/27	1.13 (0.73–1.72)	0.58	60/29	1.10 (0.73–1.66)	0.64	64/32	1.26 (0.85–1.87)	0.25	57/23	1.35 (0.84–2.15)	0.21
I	>7	28/16	2.53 (1.43–4.47)	0.001 ^d	27/15	2.11 (1.18–3.76)	0.01	23/12	3.77 (1.93–7.33)	0.0001 ^d	29/12	2.50 (1.30–4.81)	0.006
With genetic propensity adjustment													
R		372/172	1.00	–	372/187	1.00	–	391/169	1.00	–	385/130	1.00	–
F*		62/28	0.92 (0.61–1.38)	0.69	62/30	0.89 (0.60–1.32)	0.57	66/24	0.82 (0.53–1.26)	0.36	64/19	0.90 (0.55–1.47)	0.68
I		97/48	1.34 (0.97–1.85)	0.07	97/49	1.26 (0.91–1.73)	0.16	98/47	1.41 (1.02–1.96)	0.04	97/38	1.45 (1.00–2.08)	0.05
I	0–7	62/29	1.03 (0.69–1.55)	0.88	63/31	1.02 (0.69–1.51)	0.93	67/33	1.13 (0.78–1.65)	0.52	59/23	1.08 (0.69–1.71)	0.74
I	>7	35/19	2.40 (1.41–4.06)	0.001 ^d	34/18	2.03 (1.19–3.46)	0.009	31/14	3.45 (1.81–6.60)	0.0002 ^d	38/15	2.80 (1.53–5.12)	0.001 ^d
Propensity index ^e			2.18 (1.74–2.74)	<0.0001		1.86 (1.56–2.23)	<0.0001		1.73 (1.45–2.07)	<0.0001		2.19 (1.67–2.88)	<0.0001

^a Haplogroup E was excluded because only 35 seroconverters were available for analysis^b In the I haplogroup analysis, the times to CD4 < 200 and AIDS-1993 outcomes were divided into two intervals of less than or equal to seven and greater than 7 years. For AIDS-1987 and death, the time intervals were less than or equal to nine and greater than 9 years^c The R haplogroup is the reference group. Relative Hazards results were obtained after stratification by cohort. Nine ALIVE cohort samples were excluded from the stratified analysis to maintain power. Their inclusion did not alter the results reported but decreased their significance^d $P = 0.001$ corresponds to false-discovery-rate $q = 0.031$, $P = 0.0001$ corresponds to $q = 0.017$, and $P = 0.0002$ corresponds to $q = 0.023$ ^e Calculated as in O'Brien and Nelson (2004)

Table 3 Significant associations of Y-I haplogroup with AIDS progression and viral suppression outcomes with corresponding q values (≤ 0.05)

Outcome	Section	Haplogroup	Patient group ^a	Nominal P	q value	RH (95% CI)	Adjustments
Time to CD4 <200 >7 years	Table 4	I1a*	All	0.001	0.031	2.71 (1.51–4.85)	
	Table 2	I	All	0.001	0.031	2.53 (1.43–4.47)	Population stratification
	Table 2	I	All	0.001	0.031	2.40 (1.41–4.06)	Genetic propensity index
Time to AIDS-1987 >9 years	Table S2	I	MACS cohort	0.0007	0.031	5.52 (2.06–14.84)	Population stratification
	Table 4	I1a*	All	0.0001	0.017	4.91 (2.41–9.88)	
	Table 2	I	All	0.0001	0.017	3.77 (1.93–7.33)	Population stratification
	Table 2	I	All	0.0002	0.023	3.45 (1.81–6.60)	Genetic propensity index
	“Results” ^b	I	All	0.0009	0.031	2.93 (1.45–5.06)	CCR5-Δ32
Time to death >9 years	Table 2	I	All	0.001	0.031	2.84 (1.51–5.34)	
	Table S4	I	MHCS cohort	0.0007	0.031	6.30 (2.18–18.21)	Genetic propensity index
Viral suppression post-HAART	Table 2	I	All	0.001	0.031	2.80 (1.53–5.12)	Genetic propensity index
	Table 5	I	MACS cohort	0.001	0.031	0.62 (0.46–0.84)	

The proportion of false positives estimated from 414 P values (325 European American and 89 African American tests). A false-discovery-rate q value cut-off of 0.05 was used for significance. All significant analyses were Cox proportional hazard tests

^a European-American significant tests after q value analysis. No tests of African Americans were false-discovery-rate q value ≤ 0.05

^b See results subsection “Genetic cofactors”

effect, we included the top three eigen values for each SC as a continuous variable in the Cox model. In these adjusted COX models, Y-I haplogroup individuals trended toward faster CD4 depletion and progression to AIDS and death later in HIV infection (Table 2). The faster progression to AIDS-1987 result was significant after multiple test corrections (Table 3). This direction of association was consistent in all the cohorts, but the results were not always as significant (Supplementary Table 2).

We also examined the frequency of Y chromosome haplogroups among subjects with different AIDS-defining illnesses (Supplementary Table 3). An analysis of haplogroup frequencies among cases and controls for AIDS-defining disease revealed a trend toward elevated frequency of Y-I haplogroup cases compared to SCs for all of the eight disease categories examined and a trend toward elevated malignancy development (RH = 2.0, P = 0.009, q = 0.11) and Kaposi sarcoma (RH = 2.30, P = 0.007, q = 0.10) for the Y-I haplogroup (Supplementary Table 3). These findings were consistent with a faster AIDS progression observed in the Y-I haplogroup subjects. A similar analysis of AIDS-defining disease as the first outcome showed seven of the eight disease categories with elevated Y-I haplogroup frequencies. However, neither the results nor the Cox models for specific disease outcomes were significant (results not shown).

Genetic cofactors

We systematically examined and accounted for the effects of well-known ARGs. First, to evaluate the CCR5-Δ32

influence on AIDS progression, we analyzed the Y-I haplogroup association using CCR5-Δ32 in a Cox model as a covariate. Again, the results indicated a significantly faster progression to AIDS-1987 (after 9 years) in the Y-I haplogroup subjects (adjusted RH = 2.93, P = 0.0009, q = 0.03). The Y-I haplogroup also maintained its significant result when CCR2-64I, CCR5-P1/P1 promoter and IL10-5'A variants were included in the Cox model (results not shown). Further, when we included a genetic propensity index which accounts for HLA-B*27, HLA-B*57, HLA-B*35, HLA-B*35Px, KIR3DS1, TSG101, SDF, RANTES and HLA class I heterozygosity effects in the Cox model, the progression difference was still significant for the Y-I haplogroup (Table 2). This progression difference was also observed in individual cohorts, but the results were not always as significant (Supplementary Table 4). Moreover, the Y-I haplogroup association with rapid CD4 cell depletion, and rapid progression to AIDS-1987 and to death remained significant after multiple test corrections (Table 3).

Viral load and HAART

Plasma viral load set points among subjects infected with HIV-1 may be a strong indicator of future disease progression because higher viral load levels lead to faster CD4 cell depletion and AIDS progression (Bruisten et al. 1997; Henrard et al. 1995; Katzenstein et al. 1996; Lyles et al. 2000). An ANOVA analysis of plasma viral load set point among HIV-infected patients showed no significant differences between I, R, and F* haplogroup subjects (df = 2, P = 0.36; Supplementary Table 5). In contrast, among

Table 4 Association of Y-I subhaplogroups with progression to AIDS outcomes among European American seroconverters

Haplogroup	Time interval, years ^a	CD4 < 200			AIDS-1993			Time interval (years) ^a	AIDS-1987			Death		
		n/events	RH ^b (95% CI)	P	n/events	RH ^b (95% CI)	P		n/events	RH ^b (95% CI)	P	n/events	RH ^b (95% CI)	P
No time partitioning														
R		372/172	1.00	–	372/187	1.00	–		391/169	1.00	–	385/130	1.00	–
F*		62/28	0.92 (0.61–1.38)	0.69	62/30	0.89 (0.60–1.32)	0.57		66/24	0.82 (0.53–1.26)	0.36	64/19	0.90 (0.55–1.47)	0.68
IIa*		55/30	1.36 (0.92–2.02)	0.12	55/30	1.27 (0.86–1.87)	0.23		55/29	1.42 (0.95–2.10)	0.09	54/22	1.34 (0.85–2.12)	0.20
I(xIIa*) ^c		42/18	1.32 (0.81–2.15)	0.27	42/19	1.24 (0.77–1.99)	0.38		43/18	1.41 (0.86–2.30)	0.17	43/16	1.61 (0.96–2.72)	0.07
With time partitioning														
IIa*	0–7	33/15	0.91 (0.53–1.55)	0.72	34/16	0.90 (0.54–1.52)	0.70	0–9	36/18	0.98 (0.60–1.60)	0.92	31/13	1.00 (0.56–1.79)	0.99
IIa*	>7	22/15	2.71 (1.51–4.85)	0.001 ^d	21/14	2.36 (1.30–4.27)	0.005	>9	19/11	4.91 (2.41–9.88)	0.0001 ^d	23/9	2.54 (1.21–5.33)	0.01
I(xIIa*) ^c	0–7	29/14	1.21 (0.70–2.11)	0.50	29/15	1.18 (0.69–2.02)	0.54	0–9	32/16	1.38 (0.82–2.32)	0.22	28/10	1.21 (0.63–2.33)	0.57
I(xIIa*) ^c	>7	13/4	1.67 (0.60–4.69)	0.33	13/4	1.38 (0.49–3.83)	0.54	>9	11/2	1.38 (0.33–5.76)	0.66	15/6	3.28 (1.37–7.83)	0.008
Propensity index			2.24 (1.78–2.82)	<0.0001		1.88 (1.58–2.25)	<0.0001			1.76 (1.47–2.10)	<0.0001		2.23 (1.69–2.92)	<0.0001

Kaplan–Meier survival analyses were not significantly different between the subhaplogroups (Supplementary Fig. 1)

^a In the I haplotype analysis, the times to CD4 < 200 and AIDS-1993 outcomes were divided into two intervals of less than or equal to seven and greater than 7 years. For AIDS-1987 and death, the time intervals were less than or equal to nine and greater than 9 years^b The R haplogroup is the reference group. Relative Hazards results were obtained after stratification by cohort. Nine ALIVE cohort samples were excluded from the stratified analysis to maintain power. Their inclusion did not alter the results reported but decreased their significance^c I haplotypes other than IIa* combined^d $P = 0.001$ corresponds to false-discovery-rate $q = 0.031$ and $P = 0.0001$ corresponds to $q = 0.017$

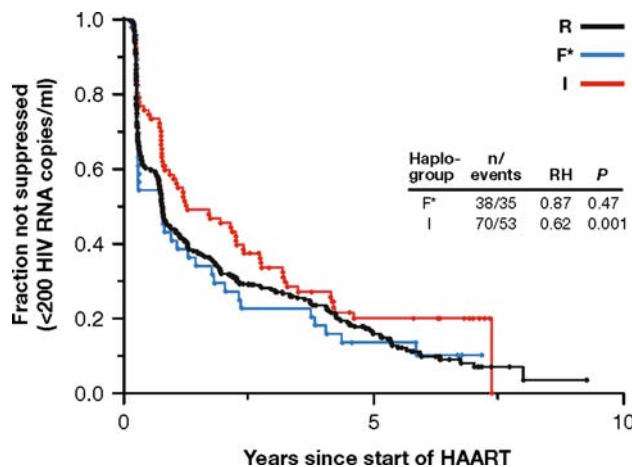


Fig. 2 Kaplan–Meier survival curves for viral suppression of I, R, and F* haplogroups in the MACS (European American) cohort subjects. Number of individuals (*n*) and number of events, Cox proportional hazards model *P* values (*P*) and relative hazards (RH) are presented. Haplogroup R is the reference group. *P* = 0.001 corresponds to false-discovery-rate *q* = 0.03

AIDS patients treated with HAART, Y-I haplogroup-bearing subjects took significantly longer to reach overall viral suppression in a Cox model (<200 HIV RNA copies/mL; RH = 0.62, *P* = 0.001, *q* = 0.03; Fig. 2) and to reach an undetectable viral load (<50 HIV RNA copies/mL; RH = 0.69, *P* = 0.02, *q* = 0.19; Table 5). The slower viral suppression result observed in the Y-I haplogroup was still significant after multiple test corrections (Table 3). The Y-I subhaplogroups had the same trend for AIDS progression, but the significant result was not due to a single subhaplogroup (Table 5). Overall, more individuals with the Y-I haplogroup trended toward HAART failure [OR = 2.42 (95% CI 1.29–4.54), Fishers Exact *P* = 0.009, *q* = 0.11].

Analyses of African Americans

An initial analysis of susceptibility to infection showed no significant frequency difference when comparing HIV-1-negative and SC subjects in African Americans (*P* = 0.45; Table 1). A Kaplan–Meier survival analysis for AIDS-1987 outcome did not demonstrate a significant difference between the SCs in seven African American Y haplogroups (*df* = 6, log-rank *P* = 0.20; Fig. 1). A simplified Kaplan–Meier survival analysis for AIDS-1987 outcome using only the E, R, and I haplogroups also did not demonstrate a significant difference between the three haplogroups (*df* = 2, log-rank *P* = 0.34; Fig. 1). Furthermore, survival analyses using the Cox proportional hazard model for CD4 < 200, AIDS-1993, AIDS-1987, and time-to-death outcomes did not find a significant progression difference between the E, R, and I haplogroups (Supplementary Table 6). Following modified Cox analyses, adjusting for the effects of ARGs via genetic propensity index values (Supplementary Table 7) and population stratification (Supplementary Table 8) also did not find a significant progression difference between the E, R, and I haplogroups. Further analysis of specific outcomes in African Americans failed to show any significant associations with the examined haplogroups.

Discussion

Our study is the first to examine whether major Y haplogroups have an association with HIV infection and progression to AIDS. We genotyped subjects as one of 11 common Y haplogroups in European Americans and African Americans, and observed faster progression to all four AIDS outcomes in European-American Y-I haplogroup subjects in

Table 5 Slower HIV-1 suppression of Y-I subjects among European American haplogroups

Haplogroup	Viral suppression (<200 HIV RNA copies/mL)						Viral suppression (<50 HIV RNA copies/mL)					
	Categorical			Survival			Categorical			Survival		
	<i>N</i>	OR	<i>P</i> ^a	<i>n</i> /events	RH	<i>P</i> ^b	<i>N</i>	OR	<i>P</i> ^a	<i>n</i> /events	RH	<i>P</i> ^b
R	329	1.00	–	253/222	1.00	–	329	1.00	–	253/206	1.00	–
F*	46	1.49	0.54	38/35	0.87	0.47	46	0.85	0.72	38/29	1.00	0.99
I	92	0.59	0.07	70/53	0.62	0.001 ^c	92	0.54	0.02	70/47	0.69	0.02
I1a*	60	0.72	0.18	50/37	0.69	0.06	56	0.70	0.11	50/32	0.72	0.09
I (xI1a*)	32	0.64	0.96	20/16	0.59	0.04	36	0.59	0.44	20/15	0.70	0.15

Categorical events were defined as undetectable viral RNA post-HAART. Survival analyses were adjusted for clinical AIDS, baseline viral load, and baseline CD4 count. All patients were from MACS cohort

^a Fisher's exact test results. Haplogroup R is the reference group

^b Log-likelihood *P* values. Haplogroup R is the reference group

^c False-discovery-rate *q* = 0.031

later years of infection. The early and initial stages of AIDS development indicated by the HIV-1 infection rate and plasma viral load levels, in the I, R, and F* haplogroups were not significantly different from each other. However, the Y-I haplogroup showed significantly longer HIV-1 viral suppression time and higher failure rate in HAART, outcomes that relate to later stages in AIDS progression. Moreover, the significant results after false-discovery-rate corrections for AIDS-1987, AIDS-related death, and HIV-1 viral suppression analyses suggest a potential biological basis for these observations. An independent analysis of AIDS-related illnesses also suggests the increased risk of the Y-I haplogroup in all disease categories. A non-significant opposite trend was observed in African American samples, but results depended on only six individuals. It is hard to draw haplogroup specific conclusions based on the few (<5%) Y-I haplogroup subjects in African Americans.

There are two possible explanations for the Y-I haplogroup effect. First, a locus (or loci) on the Y chromosome is responsible for faster progression to AIDS. As our results do not indicate an infection difference between the haplogroups, this locus might influence the AIDS progression pathway and lead to faster immunosuppression and AIDS outcomes. The relative hazard ratio estimates (1.76–2.84, Table 2) indicate a moderate effect on AIDS progression in the Y-I haplogroup. The alternative explanation for the Y chromosome signal is that we are tracing an autosomal locus in European Americans associated with the Y-I haplogroup. In European Americans, many ARG polymorphisms have been identified and found to influence HIV-1 infection, AIDS progression and mortality (O'Brien and Nelson 2004). A classic example of these protective genes against AIDS infection and progression, the *CCR5-Δ32* variant, shows a north to south gradient in European populations (Libert et al. 1998). Similarly, the Y-I haplogroup is found mostly in northern Europe, and its frequency declines toward southern Europe (Rootsi et al. 2004). This observation may suggest that we see an autosomal loci effect rather than a true Y chromosome effect. However, the effect of the Y-I haplogroup in the latter stage of disease progression is independent of the 13 well-studied autosomal AIDS restriction/susceptibility loci considered in these analyses. Moreover, the European American Y-I subhaplogroups and the population stratification analyses fail to show any specific geographical or population substructure basis for the Y-I haplogroup AIDS progression effect.

Our report of the involvement of Y chromosome major haplogroups in HIV progression and HAART outcomes has both strengths, largely elaborated above, and weaknesses. The genotyping of the known major haplogroups from Europe and Africa precisely differentiated relevant Y chromosome major haplogroups, yet failed to differentiate all of the underlying subhaplogroups. Both progression and

HAART viral suppression signals were seen for the Y-I haplogroup. Therefore, the five known Y-I subhaplogroups were genotyped, but none of them specifically carried the associations. That these two major and different endpoints were both associated with the Y-I haplogroup suggests that an underlying biological process could be at work. A strict Bonferroni correction for the total 414 tests (325 European-American and 89 African-American Cox proportional hazards and viral load tests) presented in figures and tables yields a significance cutoff of $0.05/414 = 1.2 \times 10^{-4}$. Only one of the nominal *P* values across this study was more significant than this stringent correction. However, applying the false-discovery-rate approach, 12 *P* values (0.0001–0.001) yield *q* values (Storey and Tibshirani 2003) ≤ 0.05 (Table 3). Those 12 tests show that the Y-I haplogroup is significantly associated with accelerated AIDS progression, particularly later in infection (after 7–9 years), and with failure to suppress virus during HAART. While these results localize an AIDS progression and HAART treatment signal to the Y chromosome, the underlying causative variant was not identified.

Variation on the human Y chromosome is shaped by mutations ranging from single nucleotide changes to inversions, duplications, and deletions, causing larger structural changes that generate copy number variations (Repping et al. 2006). Some of these structural changes and copy number variations have been linked to diseases such as oligo/azospermia and gonadoblastoma, and shown to be represented frequently at certain Y haplogroups (Krausz et al. 2004; Repping et al. 2003). However, no such mutation or structural change has been documented to be particularly associated with the Y-I haplogroup. The non-recombining nature of the Y chromosome makes the task of discovering the locus involved in HIV progression significantly more difficult than it would be elsewhere in the genome. The Y chromosome can be considered a complete haplotype block; therefore, all the genes on the Y chromosome are positional candidate disease genes, with the primary candidates being those expressed in the immune tissues and involved in the immunological response.

In conclusion, we present a unique association between infectious disease and the Y chromosome haplogroup I (Y-I). The 12 significant false-discovery-rate *q* values for AIDS progression and HAART outcomes observed in the Y-I haplogroup and the independence of the Y chromosome effect from that of the examined autosomal loci suggest a causal variant at a locus on the Y-I haplogroup. Eleven of these 12 associations are with survival and have strong relative hazards ranging from 2.4 to 6.3. So while a “winners curse” may be inflating these estimates, the consistency of these results points to a significant role for the Y chromosome in human health. Further comprehensive genetic and functional studies of the Y chromosome should

speed the discovery of the hidden variation related to HIV-1/AIDS.

Acknowledgments We thank the patients and staff of all the participating cohorts in the study. San Francisco City Cohort samples were provided by Susan Buchbinder. We thank Drs. George W. Nelson, James Lautenberger, Cheryl Winkler, Taras Oleksyk and Yvette Berthier-Schaad for helpful discussions. We are also grateful to Ann Truelove, Viktoriya Grinberg, Randy Johnson, Bailey Kessing, Michael Malasky, and Mary Thompson for their assistance. 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. The project described in this manuscript has been funded in whole or in part with federal funds from the National Cancer Institute, National Institutes of Health, under contract N01-CO-12400. The HGDS is supported by 1 R01 HD41224, NIH/NCICHD.

References

- Agulnik AI, Mitchell MJ, Lerner JL, Woods DR, Bishop CE (1994) A mouse Y chromosome gene encoded by a region essential for spermatogenesis and expression of male-specific minor histocompatibility antigens. *Hum Mol Genet* 3:873–878
- Anderson RN (2001) Deaths: leading Causes for 1999. National Vital Statistics Report, vol 49. National Centre for Health Statistics, Hyattsville
- Bosch E, Calafell F, Rosser ZH, Norby S, Lynnerup N, Hurler ME, Jobling MA (2003) High level of male-biased Scandinavian admixture in Greenlandic Inuit shown by Y-chromosomal analysis. *Hum Genet* 112:353–363
- Bruisten SM, Frissen PH, Van Swieten P, Harrigan PR, Kinghorn I, Larder B, Weigel HM, De Vries E, Regez RM, Henrichs JH, Koot M, Huisman JG (1997) Prospective longitudinal analysis of viral load and surrogate markers in relation to clinical progression in HIV type 1-infected persons. *AIDS Res Hum Retroviruses* 13:327–335
- Buchbinder SP, Katz MH, Hessel NA, Omalley PM, Holmberg SD (1994) Long-term HIV-1 infection without immunological progression. *AIDS* 8:1123–1128
- Carvajal-Carmona LG, Soto ID, Pineda N, Ortiz-Barrientos D, Duque C, Ospina-Duque J, McCarthy M, Montoya P, Alvarez VM, Bedoya G, Ruiz-Linares A (2000) Strong Amerind/white sex bias and a possible Sephardic contribution among the founders of a population in northwest Colombia. *Am J Hum Genet* 67:1287–1295
- Carvajal-Carmona LG, Ophoff R, Service S, Hartiala J, Molina J, Leon P, Ospina J, Bedoya G, Freimer N, Ruiz-Linares A (2003) Genetic demography of Antioquia (Colombia) and the Central Valley of Costa Rica. *Hum Genet* 112:534–541
- Cavalli-Sforza LL, Hewlett B (1982) Exploration and mating range in African Pygmies. *Ann Hum Genet* 46:257–270
- Centers for Disease Control (1986) Classification system for human T-lymphotropic virus type III/lymphadenopathy-associated virus infections. *MMWR Morb Mortal Wkly Rep* 35:334–339
- Centers for Disease Control (1992) 1993 revised classification system for HIV infection and expanded surveillance case definition for AIDS among adolescents and adults. *MMWR Recomm Rep* 41:1–19
- Chromosome Consortium Y (2002) A nomenclature system for the tree of human Y-chromosomal binary haplogroups. *Genome Res* 12:339–348
- Cinnioglu C, King R, Kivisild T, Kalfoglu E, Atasoy S, Cavalleri GL, Lillie AS, Roseman CC, Lin AA, Prince K, Oefner PJ, Shen PD, Semino O, Cavalli-Sforza LL, Underhill PA (2004) Excavating Y-chromosome haplotype strata in Anatolia. *Hum Genet* 114:127–148
- 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 (1996) Genetic restriction of HIV-1 infection and progression to AIDS by a deletion allele of the CKR5 structural gene. Hemophilia Growth and Development Study, Multicenter AIDS Cohort Study, Multicenter Hemophilia Cohort Study, San Francisco City Cohort, ALIVE Study. *Science* 273:1856–1862
- Diamond JM (2005) Guns, germs, and steel: the fates of human societies. Norton, New York
- Ditton HJ, Zimmer J, Kamp C, Rajpert-De Meyts E, Vogt PH (2004) The AZFa gene DBY (DDX3Y) is widely transcribed but the protein is limited to the male germ cells by translation control. *Hum Mol Genet* 13:2333–2341
- Dobson AP, Carper ER (1996) Infectious diseases and human population history—throughout history the establishment of disease has been a side effect of the growth of civilization. *Bioscience* 46:115–126
- Gleicher N, Barad DH (2007) Gender as risk factor for autoimmune diseases. *J Autoimmun* 28:1–6
- Goedert JJ, Kessler CM, Aledort LM, Biggar RJ, Andes WA, White GC, Drummond JE, Vaidya K, Mann DL, Eyster ME, Ragni MV, Lederman MM, Cohen AR, Bray GL, Rosenberg PS, Friedman RM, Hilgartner MW, Blattner WA, Kroner B, Gail MH (1989) A prospective study of human immunodeficiency virus type-1 infection and the development of AIDS in subjects with hemophilia. *N Engl J Med* 321:1141–1148
- Hamilton G, Stoneking M, Excoffier L (2005) Molecular analysis reveals tighter social regulation of immigration in patrilineal populations than in matrilineal populations. *Proc Natl Acad Sci USA* 102:7476–7480
- Henrard DR, Phillips JF, Muenz LR, Blattner WA, Wiesner D, Eyster ME, Goedert JJ (1995) Natural history of HIV-1 cell-free viremia. *JAMA* 274:554–558
- Hilgartner MW, Donfield SM, Willoughby A, Contant CF, Evatt BL, Gomperts ED, Hoots WK, Jason J, Loveland KA, McKinlay SM, Stehens JA (1993) Hemophilia growth and development study—design, methods, and entry data. *Am J Pediatr Hematol Oncol* 15:208–218
- Ivanov R, Hol S, Aarts T, Hagenbeek A, Slager EH, Ebeling S (2005) UTY-specific TCR-transfer generates potential graft-versus-leukemia effector T cells. *Br J Haematol* 129:392–402
- Jobling MA, Tyler-Smith C (2003) The human Y chromosome: an evolutionary marker comes of age. *Nat Rev Genet* 4:598–612
- Jobling MA, Williams GA, Schiebel GA, Pandya GA, McElreavey GA, Salas GA, Rappold GA, Affara NA, Tyler-Smith C (1998) A selective difference between human Y-chromosomal DNA haplotypes. *Curr Biol* 8:1391–1394
- Karafet TM, Mendez FL, Meilerman MB, Underhill PA, Zegura SL, Hammer MF (2008) New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. *Genome Res* 18:830–838
- Katzenstein TL, Pedersen C, Nielsen C, Lundgren JD, Jakobsen PH, Gerstoft J (1996) Longitudinal serum HIV RNA quantification: correlation to viral phenotype at seroconversion and clinical outcome. *AIDS* 10:167–173
- Krausz C, Quintana-Murci L, Forti G (2004) Y chromosome polymorphisms in medicine. *Ann Med* 36:573–583
- Lahn BT, Page DC (1997) Functional coherence of the human Y chromosome. *Science* 278:675–680
- Laurin D, Spierings E, van der Veken LT, Hamrouni A, Falkenburg JH, Souillet G, Vermeulen C, Farre A, Galambun C, Rigal D, Bertrand Y, Goulmy E, Eljaafari A (2006) Minor histocompatibility

- antigen DDX3Y induces HLA-DQ5-restricted T cell responses with limited TCR-Vbeta usage both in vivo and in vitro. *Biol Blood Marrow Transplant* 12:1114–1124
- Libert F, Cochaux P, Beckman G, Samson M, Aksenova M, Cao A, Czeizel A, Claustres M, de la Rua C, Ferrari M, Ferrec C, Glover G, Grinde B, Guran S, Kucinskas V, Lavinha J, Mercier B, Ogur G, Peltonen L, Rosatelli C, Schwartz M, Spitsyn V, Timar L, Beckman L, Parmentier M, Vassart G (1998) The *deltacr5* mutation conferring protection against HIV-1 in Caucasian populations has a single and recent origin in Northeastern Europe. *Hum Mol Genet* 7:399–406
- Lind JM, Hutcheson-Dilks HB, Williams SM, Moore JH, Essex M, Ruiz-Pesini E, Wallace DC, Tishkoff SA, O'Brien SJ, Smith MW (2007) Elevated male European and female African contributions to the genomes of African American individuals. *Hum Genet* 120:713–722
- Lyles RH, Munoz A, Yamashita TE, Bazmi H, Detels R, Rinaldo CR, Margolick JB, Phair JP, Mellors JW (2000) Natural history of human immunodeficiency virus type 1 viremia after seroconversion and proximal to AIDS in a large cohort of homosexual men. Multicenter AIDS Cohort Study. *J Infect Dis* 181:872–880
- Marriott I, Huet-Hudson YM (2006) Sexual dimorphism in innate immune responses to infectious organisms. *Immunol Res* 34:177–192
- 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 (1998) Genetic acceleration of AIDS progression by a promoter variant of CCR5. *Science* 282:1907–1911
- McCarthy M (2000) The “gender gap” in autoimmune disease. *Lancet* 356:1088
- Mesa NR, Mondragon MC, Soto ID, Parra MV, Duque C, Ortiz-Barrientos D, Garcia LF, Velez ID, Bravo ML, Munera JG, Bedoya G, Bortolini MC, Ruiz-Linares A (2000) Autosomal, mtDNA, and Y-chromosome diversity in Amerinds: pre- and post-Columbian patterns of gene flow in South America. *Am J Hum Genet* 67:1277–1286
- O'Brien SJ, Nelson GW (2004) Human genes that limit AIDS. *Nat Genet* 36:565–574
- Office of AIDS Research Advisory Council (2006) Guidelines for the use of antiretroviral agents in HIV-1-infected adults and adolescents. <http://aidsinfo.nih.gov/ContentFiles/AdultandAdolescentGL05042006050.pdf>
- Pezzotti P, Phillips AN, Dorrucci M, Lepri AC, Galai N, Vlahov D, Rezza G (1996) Category of exposure to HIV and age in the progression to AIDS: longitudinal study of 1199 people with known dates of seroconversion. HIV Italian Seroconversion Study Group. *Br Med J* 313:583–586
- Phair J, Jacobson L, Detels R, Rinaldo C, Saah A, Schragger L, Munoz A (1992) Acquired immune deficiency syndrome occurring within 5 years of infection with human immunodeficiency virus type-1: the Multicenter Aids Cohort Study. *J Acquir Immune Defic Syndr Hum Retrovirol* 5:490–496
- Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D (2006) Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 38:904–909
- Price AL, Butler J, Patterson N, Capelli C, Pascali VL, Scarnicci F, Ruiz-Linares A, Groop L, Saetta AA, Korkolopoulou P, Seligsohn U, Waliszewska A, Schirmer C, Ardlie K, Ramos A, Nemesh J, Arbeitman L, Goldstein DB, Reich D, Hirschhorn JN (2008) Discerning the ancestry of European Americans in genetic association studies. *PLoS Genet* 4:e236
- Repping S, Skaletsky H, Brown L, van Daalen SK, Korver CM, Pyntikova T, Kuroda-Kawaguchi T, de Vries JW, Oates RD, Silber S, van der Veen F, Page DC, Rozen S (2003) Polymorphism for a 1.6-Mb deletion of the human Y chromosome persists through balance between recurrent mutation and haploid selection. *Nat Genet* 35:247–251
- Repping S, van Daalen SK, Brown LG, Korver CM, Lange J, Marszalek JD, Pyntikova T, van der Veen F, Skaletsky H, Page DC, Rozen S (2006) High mutation rates have driven extensive structural polymorphism among human Y chromosomes. *Nat Genet* 38:463–467
- Rootsi S, Magri C, Kivisild T, Benuzzi G, Help H, Bermisheva M, Kutuev I, Barac L, Pericic M, Balanovsky O, Pshenichnov A, Dion D, Grobei M, Zhivotovsky LA, Battaglia V, Achilli A, Al-Zahery N, Parik J, King R, Cinnioglu C, Khusnutdinova E, Rudan P, Balanovska E, Scheffrahn W, Simonescu M, Brehm A, Goncalves R, Rosa A, Moisan JP, Chaventre A, Ferak V, Furedi S, Oefner PJ, Shen PD, Beckman L, Mikerezi I, Terzic R, Primorac D, Cambon-Thomsen A, Krumina A, Torroni A, Underhill PA, Santachiara-Benerecetti AS, Vilems R, Semino O (2004) Phylogeography of Y-chromosome haplogroup I reveals distinct domains of prehistoric gene flow in Europe. *Am J Hum Genet* 75:128–137
- Seielstad MT, Minch E, Cavalli-Sforza LL (1998) Genetic evidence for a higher female migration rate in humans. *Nat Genet* 20:278–280
- 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 (2000) Genetic restriction of HIV-1 pathogenesis to AIDS by promoter alleles of IL10. *Proc Natl Acad Sci USA* 97:14467–14472
- Skaletsky H, Kuroda-Kawaguchi T, Minx PJ, Cordum HS, Hillier L, Brown LG, Repping S, Pyntikova T, Ali J, Bieri T, Chinwalla A, Delehaunty A, Delehaunty K, Du H, Fewell G, Fulton L, Fulton R, Graves T, Hou SF, Latrielle P, Leonard S, Mardis E, Maupin R, McPherson J, Miner T, Nash W, Nguyen C, Ozersky P, Pepin K, Rock S, Rohlfing T, Scott K, Schultz B, Strong C, Tin-Wollam A, Yang SP, Waterston RH, Wilson RK, Rozen S, Page DC (2003) The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature* 423:825–837
- 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 (1997) 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* 277:959–965
- Storey JD, Tibshirani R (2003) Statistical significance for genomewide studies. *Proc Natl Acad Sci USA* 100:9440–9445
- Teuschert C, Noubade R, Spach K, McElvany B, Bunn JY, Fillmore PD, Zachary JF, Blankenhorn EP (2006) Evidence that the Y chromosome influences autoimmune disease in male and female mice. *Proc Natl Acad Sci USA* 103:8024–8029
- Torikai H, Akatsuka Y, Miyazaki M, Warren EH 3rd, Oba T, Tsujimura K, Motoyoshi K, Morishima Y, Koderia Y, Kuzushima K, Takahashi T (2004) A novel HLA-A*3303-restricted minor histocompatibility antigen encoded by an unconventional open reading frame of human TMSB4Y gene. *J Immunol* 173:7046–7054
- Underhill PA, Passarino G, Lin AA, Shen P, Lahr MM, Foley RA, Oefner PJ, Cavalli-Sforza LL (2001) The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations. *Ann Hum Genet* 65:43–62
- Vlahov D, Graham N, Hoover D, Flynn C, Bartlett JG, Margolick JB, Lyles CM, Nelson KE, Smith D, Holmberg S, Farzadegan H (1998) Prognostic indicators for AIDS and infectious disease death in HIV-infected injection drug users: plasma viral load and CD4(+) cell count. *JAMA* 279:35–40

- Vogt MH, de Paus RA, Voogt PJ, Willemze R, Falkenburg JH (2000) DFFRY codes for a new human male-specific minor transplantation antigen involved in bone marrow graft rejection. *Blood* 95:1100–1105
- Warren EH, Gavin MA, Simpson E, Chandler P, Page DC, Distech C, Stankey KA, Greenberg PD, Riddell SR (2000) The human UTY gene encodes a novel HLA-B8-restricted H-Y antigen. *J Immunol* 164:2807–2814
- Wesley JD, Tessmer MS, Paget C, Trottein F, Brossay L (2007) A Y chromosome-linked factor impairs NK T development. *J Immunol* 179:3480–3487
- Whitacre CC (2001) Sex differences in autoimmune disease. *Nat Immunol* 2:777–780
- Whitacre CC, Reingold SC, O'Looney PA (1999) A gender gap in autoimmunity. *Science* 283:1277–1278
- Wilder JA, Kingan SB, Mobasher Z, Pilkington MM, Hammer MF (2004) Global patterns of human mitochondrial DNA and Y-chromosome structure are not influenced by higher migration rates of females versus males. *Nat Genet* 36:1122–1125
- Wilkins JF (2006) Unraveling male and female histories from human genetic data. *Curr Opin Genet Dev* 16:611–617
- Wilkins JF, Marlowe FW (2006) Sex-biased migration in humans: what should we expect from genetic data? *Bioessays* 28:290–300
- Wolfe ND, Dunavan CP, Diamond J (2007) Origins of major human infectious diseases. *Nature* 447:279–283
- Wood ET, Stover DA, Ehret C, Destro-Bisol G, Spedini G, McLeod H, Louie L, Bamshad M, Strassmann BI, Soodyall H, Hammer MF (2005) Contrasting patterns of Y chromosome and mtDNA variation in Africa: evidence for sex-biased demographic processes. *Eur J Hum Genet* 13:867–876