Supplementary Materials

Rotger et al.

Table S1A: Demographic characteristics of study participants.

	VNP (n=6)	RP (n=66)	EC (n=9)	CP (n=5)
Male gender, n(%)	5 (83)	54 (82)	5 (56)	3 (60)
White ethnicity, n(%)	5 (83)	56 (89)	9 (100)	5 (100)
Mode of transmission, n (%)				
Heterosexual	1 (17)	17 (26)	2 (22)	1 (20)
MSM	5 (83)	29 (44)	3 (33)	1 (20)
Intravenous drug use	0 (0)	14 (21)	3 (33)	3 (60)
Other	0 (0)	6 (9)	1 (88)	0 (0)
Median (IQR) age, years	38 (34-39)	37 (34-40)	31 (27-36)	36 (30-38)
Set point HIV RNA, log_{10} cp/ml	5.4 (5.1-5.5)	4.7 (4.3-5.2)	<1.7 (<1.7-2)	3.9 (3.3-4.7)
Viral subtype (n)	B(5), F(1)	B(51), A(2), 01_AE(1). 02_AG(1), BF(1), BD(1), C(2), G(1), NA(6)	B(3), 11_CPX(1), NA(5)	B(2), 01_AG(1), NA(2)

[#] IQR: interquartile range; MSM: Men who have sex with men, CP= chronic progressors

Table S1B: Viral load and CD4 T cell count at transcriptome day.

Median (IQR)

	samples	VL at sample date	CD4 at sample date	VL setpoint
RP	27	4.7 (4.3-5.1)	341 (280-469)	4.7 (4.3-5.2)
EC	9	0 (0-0)	858 (655-1140)	0 (0-2)
CP	5	4 (3.4-4.7)	932 (345-1289)	3.9 (3.3-4.7)
VNP	5	5.1 (4.9-5.1)	556 (490.5-923)	5.3 (5-5.5)

Immunogenetics.

The carriage frequencies of HLA-A, -B and -C alleles in individuals with rapid HIV disease progression were compared to the allele frequencies of 1609 participants of the Swiss HIV Cohort Study (SHCS) (**Figure S1**). HLA alleles previously associated with lower HIV RNA levels and/or slower HIV disease progression in Caucasians were classified as *protective,* those with unfavorable effects as *risk* alleles, as summarized by R. Kaslow (http://retroconference.org/2010/Abstracts/39871.htm). Protective alleles were clearly underrepresented in RP compared to the overall HIV-infected population. The proportion of individuals carrying at least one protective HLA allele was significantly lower in RP compared to the general HIV infected population (17.2% vs 30.7%, p=0.02). The depletion of protective alleles was particularly evident for the two strongest protective markers *HLA-B*57:01* and *HLA-B*27:05*. There was an increase in the prevalence of *HLA-B*58:01* in RP, this may be explained by the low frequency of the allele, or by the fact that the original description of this allele as protective was in individuals with HIV-1 clade C infection. Risk alleles were more common in RP compared to the general population, however this difference was not statistically significant (23.3% vs 14.9%, P=0.2). Among VNP, one individual carried *HLA-B*57:01*; a comparative analyses of HLA allele frequencies for VNP was not possible due to the rarity of this profile (**Supplementary Table S2**).

In contrast to HLA alleles, there was no depletion of protective KIR alleles or KIR/HLA combinations in RP. The distribution of protective and risk alleles in RP was very similar to the overall HIV-infected population (**Figure S2**). The only significant difference was found in the prevalence of the *KIR3DL1-B*57* combination due to the absence of *HLA*B57:01* in RP.

Table S2: *HLA* alleles of viremic non progressors.

ID	HLA_AI	HLA_AII	HLA_BI	HLA_BII	HLA_CI	HLA_CII
VNP1	1101	3201	1801	4403	1203	1601
VNP2	0201	0301	0702	4001	0304	0702
VNP3	2902	3004	3501	4501	0401	0602
VNP4	0201	2402	1501	3906	0303	0702
VNP5	0201	3301	NA	NA	NA	NA
VNP6	0101	0201	4402	5701	0501	0701

NA= not available

Table S3: Results for suggestive associations from genome-wide association analysis of rapid progressors. Reported frequencies and OR are with respect to the A1 (minor) allele. Positions are relative to Human genome build 18. (Separate file in attachment)

Table S4: Results for SNPs previously reported to be associated with rapid progression. Reported frequencies and OR are with respect to the A1 (minor) allele.

SNP	Chr	A1	A2	Case freq	Control freq	OR previous	OR current	P previous	P current
rs4118325	1	A	G	0.18	0.17	0.24	1.12	6.09E-07	6.39E-01
rs1522232	12	T	C	0.46	0.48	0.45	1.17	1.80E-06	3.87E-01
rs1360517	9	A	G	0.07	0.07	3.09	0.99	3.27E-06	9.67E-01
rs10800098	1	A	G	0.05	0.06	3.29	0.99	3.86E-06	9.80E-01
rs10494056	1	A	C	0.20	0.17	0.27	1.34	4.29E-06	2.01E-01
rs12351740	9	T	C	0.05	0.04	3.46	0.97	6.63E-06	9.45E-01
rs1020064	2	T	G	0.21	0.22	0.34	0.94	7.04E-06	7.65E-01

Table S5: Differentially expressed genes in CD8+ T cells of rapid progressors. (Separate file in attachment)

Table S6: Sets of differentially expressed genes based on a large microarray dataset (Bosinger et al. J Clin Invest 2009) detailing longitudinal SIV infection in rhesus macaques, which develop disease, and sooty mangabeys, a non-pathogenic, natural host species. (Separate file in attachment)

Figure S1A: *HLA* allele frequencies in individuals with rapid progression (n=66) and the general HIV infected population (n=1609). Only *HLA* alleles with allele frequencies ≥1% are shown. Panel A, HLA-A; Panel B, HLA-B; Panel C, HLA-C.

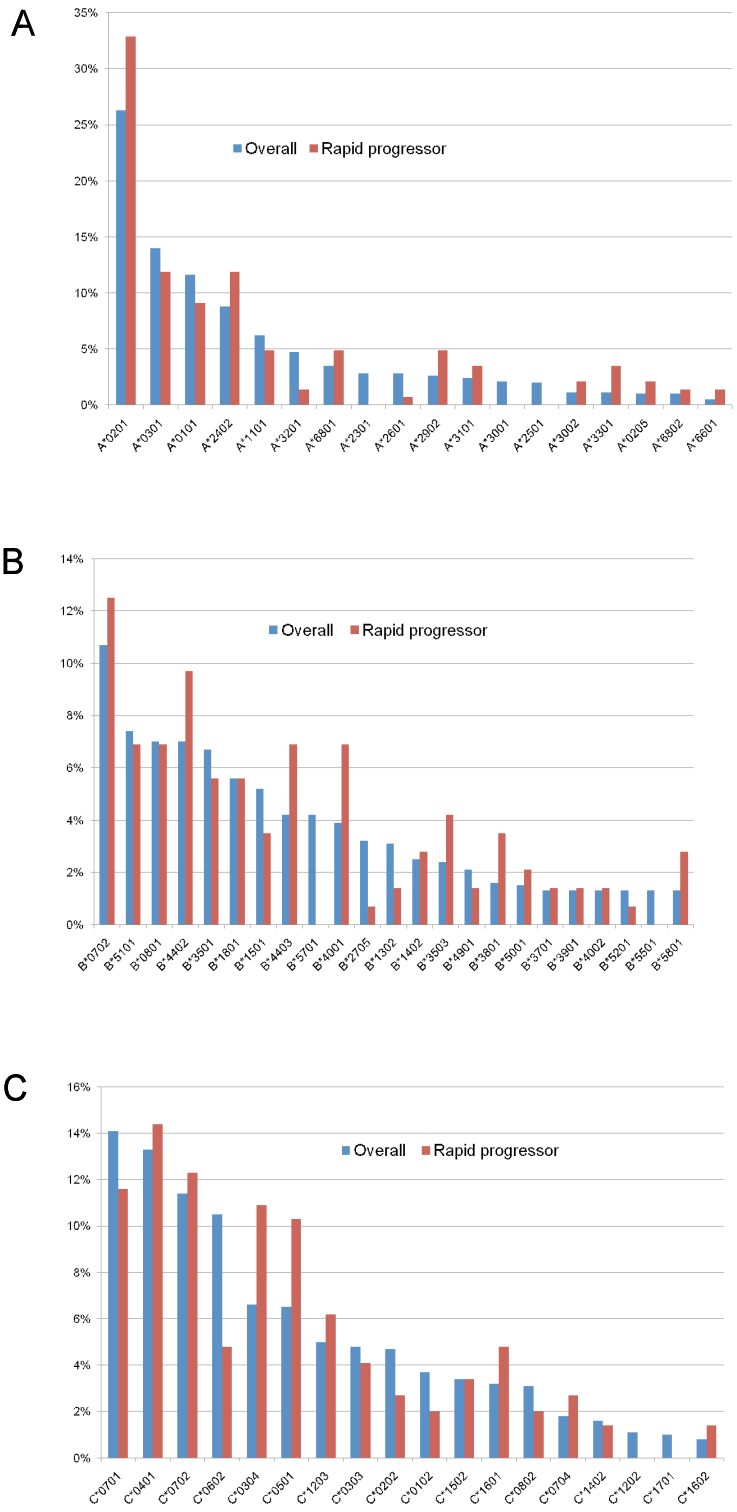


Figure S1B: Frequency of selected protective and risk HLA alleles in individuals with rapid progression (n=66) and the general HIV infected population (n=1609). The indicated effect (protective vs risk alleles) is based on published references. Only significant P-values are shown.

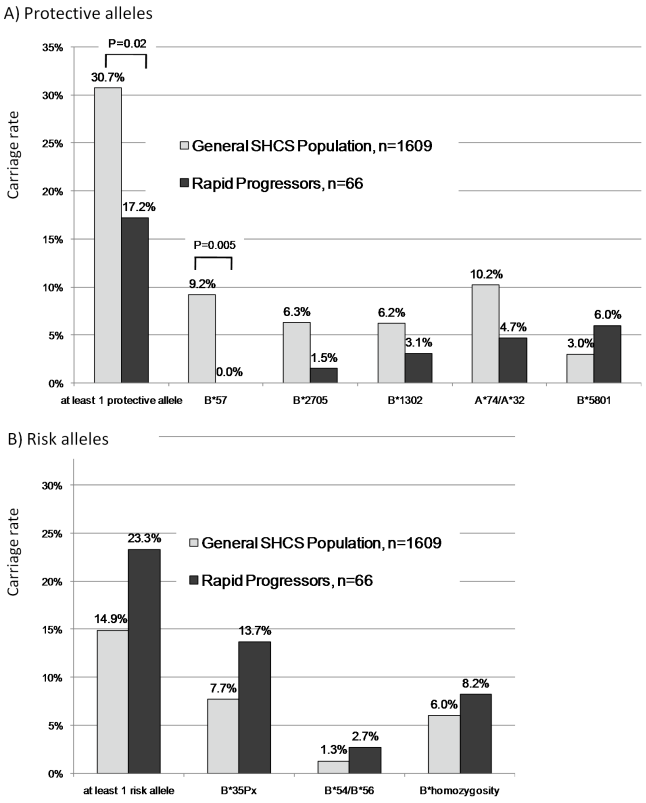


Figure S2: *KIR* allele frequencies and HLA ligand pairs in individuals with rapid progression, and in the general HIV infected population. The indicated effect (protective vs risk alleles) is based on published references. Only significant P-values are shown.

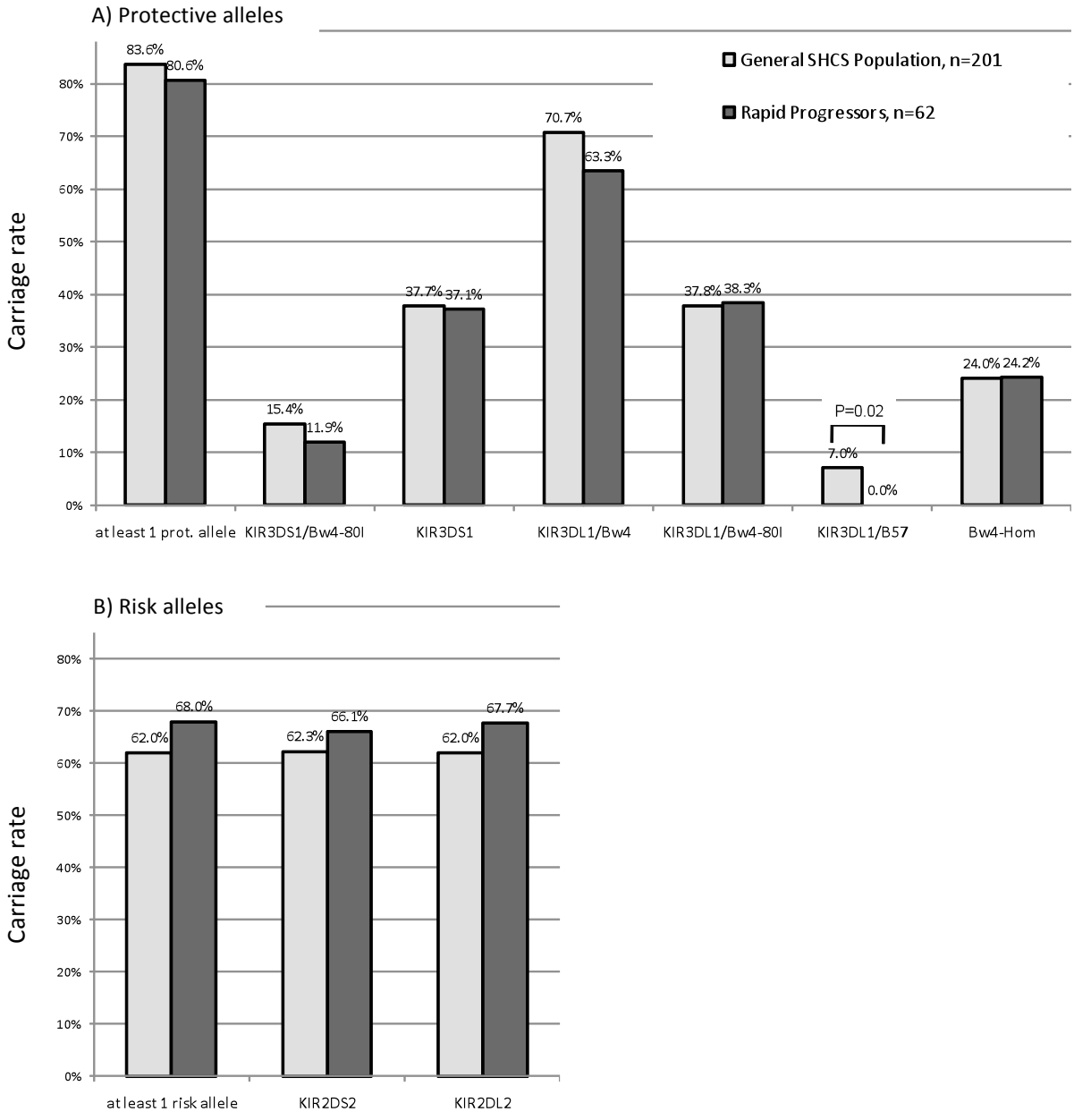
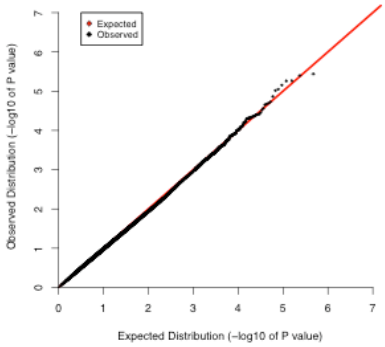


Figure S3: Genome-wide association results comparing rapid progressors to controls. Top panel shows the p-value distribution across the data set. The data fall along the null distribution, showing no evidence for confounding due to population structure (lambda = 1). Bottom panel shows Manhattan plot of p-values genome-wide. No SNPs pass the genome-wide threshold of $p<5x10^{-8}$ (dotted line).



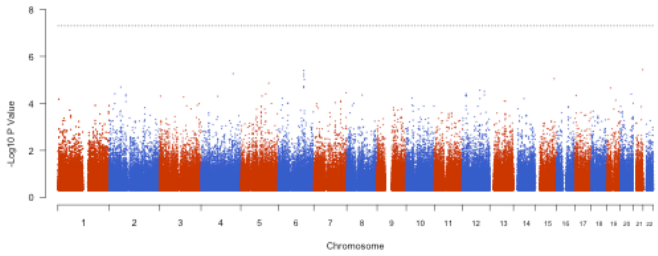


Figure S4. Predicted interaction networks of genes differentially upregulated in CD8+ T cells during HIV-1 infection. Genes (n=180) upregulated in RP compared to elite/viremic controllers (FDR adjusted p-value<0.05) are shown: links have been predicted using STRING (http://string.embl.de/). Interactions are depicted according to the strength of evidence obtained from direct (physical) and indirect (functional) associations (the thickness of the connecting blue line indicates the strength of confidence). Information is derived from four sources: genomic context, high-throughput experiments, conserved co-expression, and previous knowledge from literature.

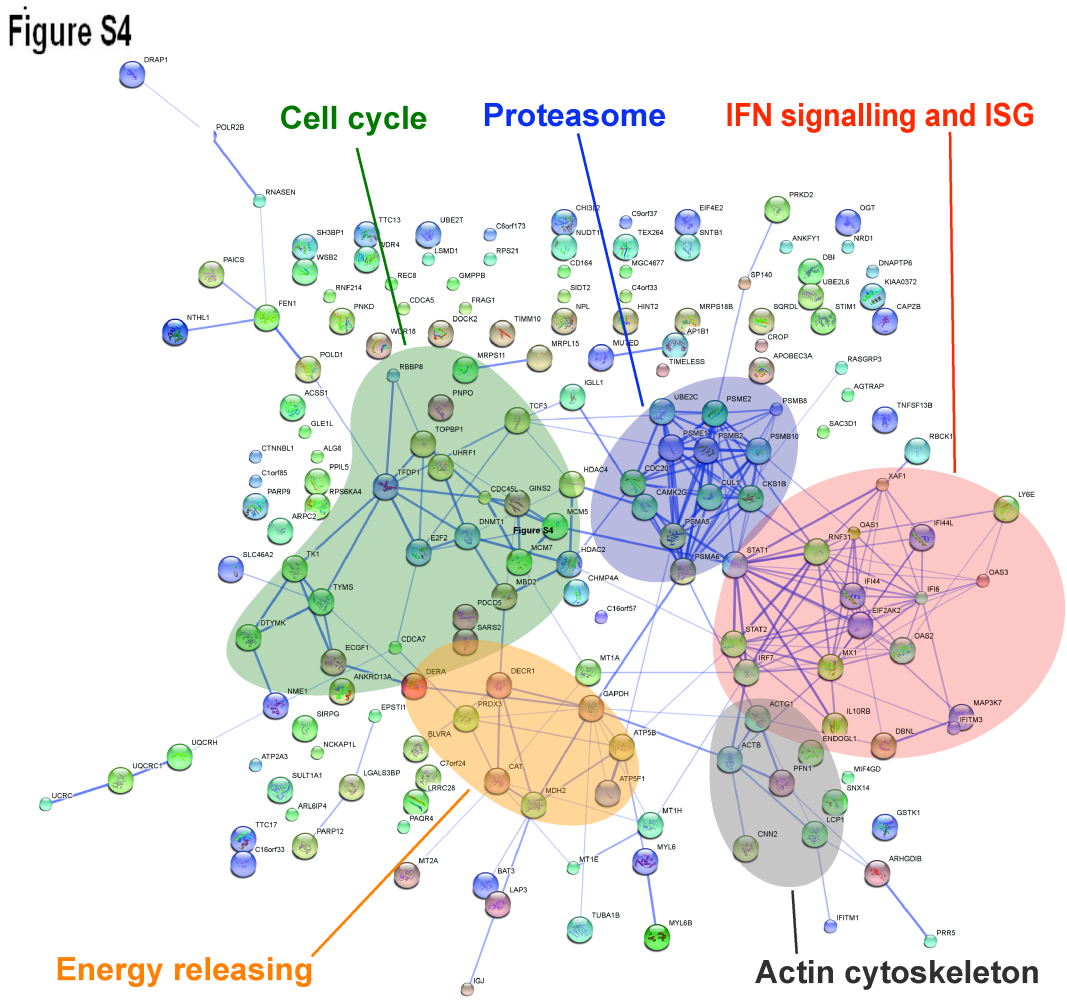


Figure S5. Gene set enrichment analysis. The upper (human CD8+ T cells) and lower panels (human CD4 T cells) represent the enrichment analysis in RP and VNP datasets of query sets of genes differentially expressed in rhesus macaques (RM) or sooty mangabeys (SM) in the course of SIV infection. The explored query (primate) sets included interferon stimulated genes (ISGs) and immune activation (IA) genes differentially upregulated in RM, and genes differentially upregulated in SM during chronic infection, and a random set of genes as control. The vertical dotted line denotes the "tipping point" in the ranked list; genes to right (red dots) are preferentially ranked among RPs; genes on the left (blue) are preferentially ranked among VNPs.