Supplemental Data

AJHG*, Volume *86

Quantitative Trait Loci for CD4:CD8 Lymphocyte Ratio Are Associated with Risk of Type 1 Diabetes and HIV-1 Immune Control

Manuel A.R. Ferreira, Massimo Mangino, Chanson J. Brumme, Zhen Zhen Zhao, Sarah E. Medland, Margaret J. Wright, Dale R. Nyholt, Scott Gordon, Megan Campbell, Brian P. McEvoy, Anjali Henders, David M. Evans, Jerry S. Lanchbury, Florencia Pereyra, International HIV Controllers Study, Bruce D. Walker, David W. Haas, Nicole Soranzo, Tim D. Spector, Paul I.W. de Bakker, Ian H. Frazer, Grant W. Montgomery, and Nicholas G. Martin

Outline

Supplemental Figures

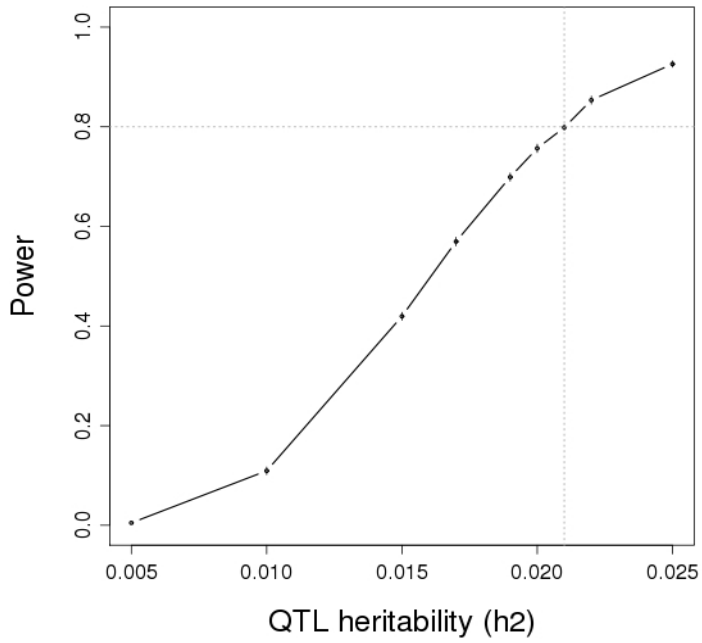

Figure S1. Power ($\alpha = 8.3x10^{-9}$) provided by the Australian adolescent GWAS dataset to detect a quantitative trait locus (QTL) explaining 0.5% to 2.5% of the total phenotypic variance (h^2). Power was estimated as the proportion of 1,000 datasets simulated with Merlin[1] under the null hypothesis of no association with a $P < 8.3x10^{-9}$. Datasets were simulated while preserving the original missingness patterns and assuming a trait with a total heritability of 75%, and a sibling correlation of 0.4. Eighty percent power (horizontal line) was achieved for a QTL heritability of 2.1% (vertical line).

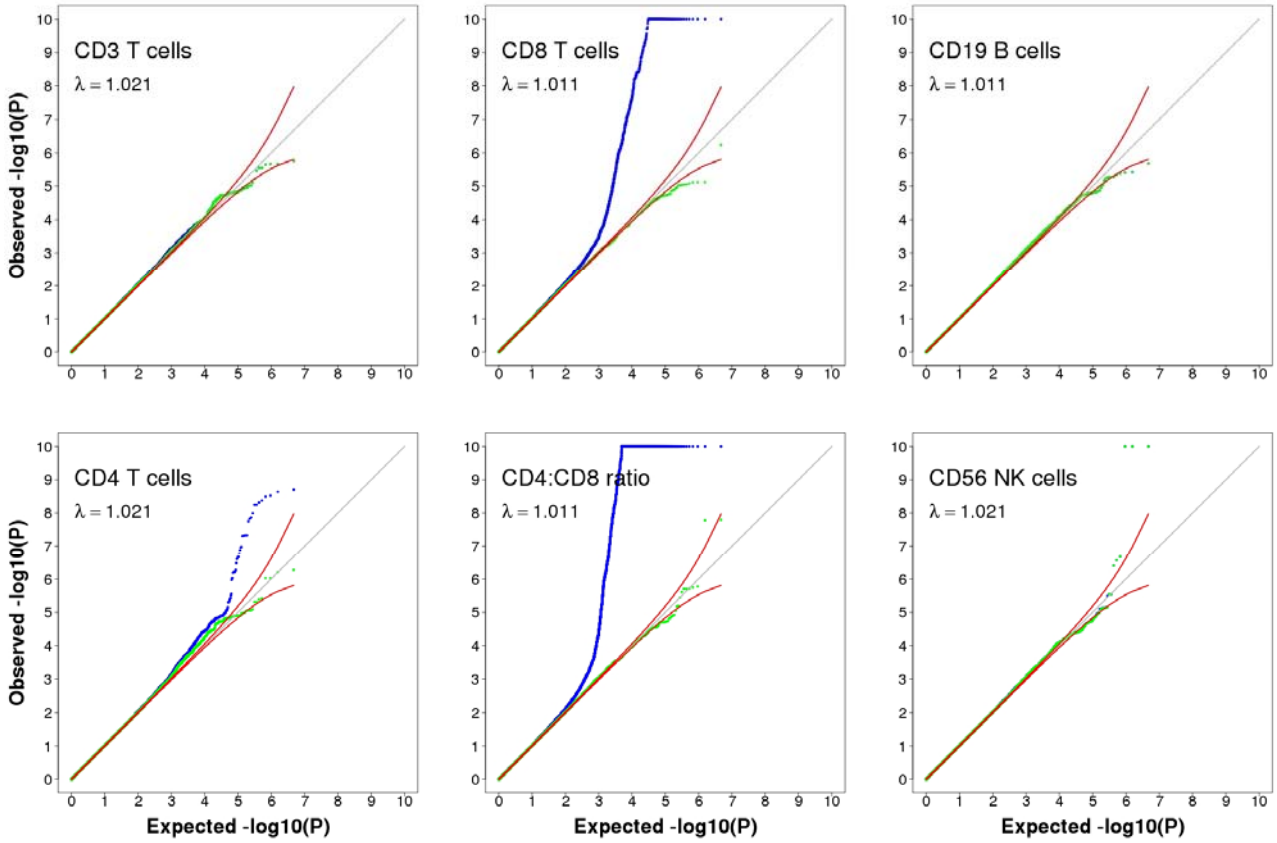

Figure S2. Quantile-quantile (QQ) plots for the six lymphocyte traits tested in the Australian GWAS panel, before (blue) and after (green) excluding data from the MHC region (chromosome 6, from 25 to 35 Mb). The upper and lower boundaries of the 95% confidence bands are represented by the red lines. The corresponding genomic inflation factor (λ) is also shown for each trait. Observed *P*-values were truncated at 10^{-10}.

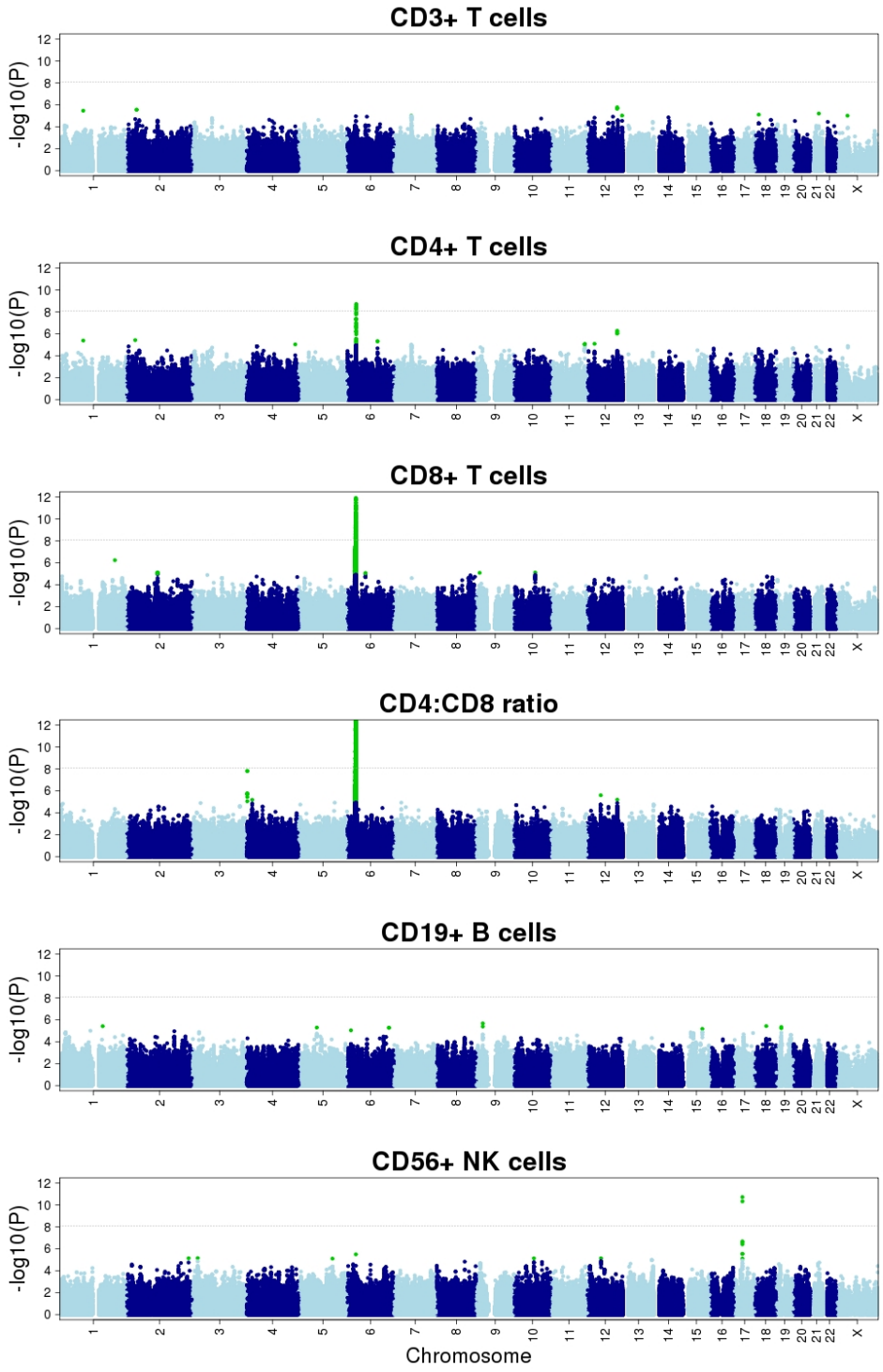

Figure S3. Genome-wide association results for the six lymphocyte phenotypes tested. Results (–log10*P*) are shown in chromosomal order for individual SNPs genotyped in 1,089 Australian families. Horizontal line indicates a *P*-value of 8.3×10^{-9}. SNPs with a *P*-value < 10^{-5} are highlighted in green. NK: natural killer. Observed *P*-values were truncated at 10^{-12}.

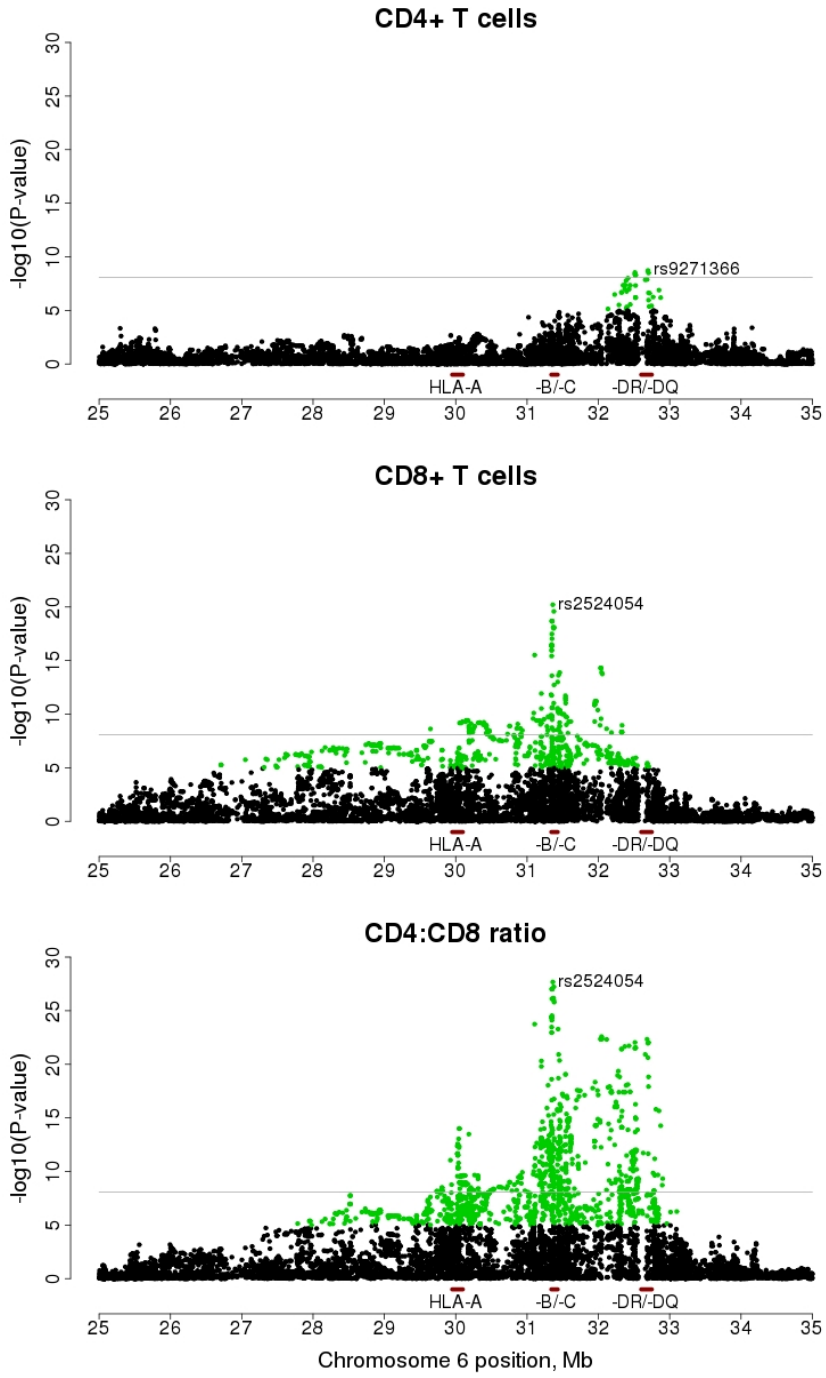

Figure S4. Association results between the three major T cell phenotypes tested and the MHC region. The most associated SNP for each trait is shown, as well as the location of the main HLA class I (*HLA-A*, *-B* and *-C*) and class II (*HLA-DR* and *-DQ*) gene clusters. Horizontal line indicates a *P*-value of 8.3×10^{-9}. SNPs with a *P*-value < 1×10^{-5} are highlighted in green.

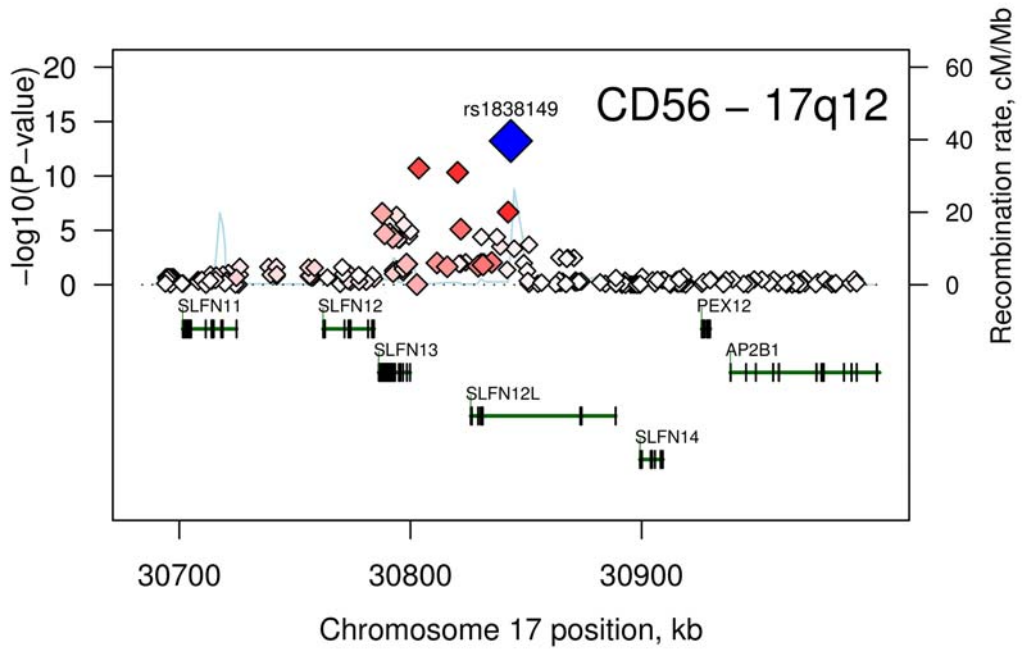

Figure S5. Regional association plot between 17q12 variants and CD56+ NK cell levels. The most associated SNP (rs1838149, imputed) is shown in blue, and the color of the remaining markers reflects the linkage disequilibrium (r^2) with the top SNP (increasing red hue associated with increasing r^2). The most significant genotyped SNP in this region was rs9916629 ($P = 4.7\text{x}10^{-11}$). The recombination rate (second *y*-axis) is plotted in light blue and is based on the CEU HapMap population. Exons for each gene are represented by vertical bars, based on all isoforms available from the March 2006 UCSC genome browser assembly.

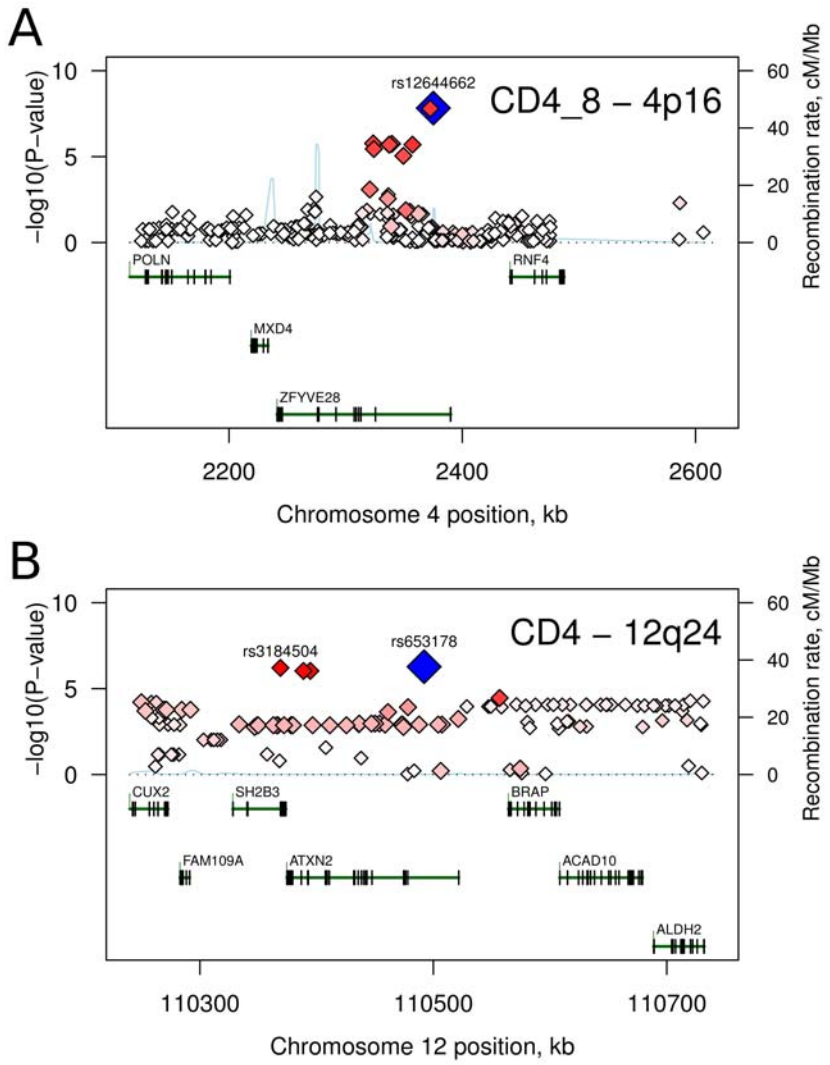

Figure S6. Regional association plots between (A) 4p16 variants and CD4:CD8 T cell ratio and (B) 12q24 variants and CD4 T cell levels. The most associated SNP in each panel is shown in blue, and the color of the remaining markers reflects the linkage disequilibrium (r^2) with the top SNP (increasing red hue associated with increasing r^2). The recombination rate (second y-axis) is plotted in light blue and is based on the CEU HapMap population. Exons for each gene are represented by vertical bars, based on all isoforms available from the March 2006 UCSC genome browser assembly.

Supplemental Tables

Table S1. Characteristics of the study participants for the three panels studied that were ascertained from the general population

	Australian GWAS	**Australian replication**	**TwinsUK replication**
	N individuals		
Genotyped (& clinical)[a]	3,995 (2,538)	978 (592)	396 (396)
	Demographics[b]		
Female sex, %	52	48	100
Mean age (range), years	15 (10-37)	14 (10-22)	50 (19-80)
	Mean trait levels (SD, range)		
CD3, 10^9 cells / L	1.76 (0.43 ,0.34-3.46)	1.80 (0.46, 0.58-4.03)	-
CD4, 10^9 cells / L	1.03 (0.27, 0.21-2.53)	1.04 (0.30, 0.20-2.80)	0.87 (0.33, 0.39-2.38)
CD8, 10^9 cells / L	0.61 (0.20, 0.08-1.66)	0.64 (0.22, 0.14-1.69)	0.46 (0.18, 0.12-1.26)
CD4:CD8 ratio	1.79 (0.52, 0.43-4.09)	1.73 (0.55, 0.32-3.93)	2.00 (0.84, 0.46-5.70)
CD19, 10^9 cells / L	0.47 (0.17, 0.06-1.31)	0.51 (0.20, 0.13-1.67)	-
CD56, 10^9 cells / L	0.23 (0.11, 0.03-0.81)	0.24 (0.12, 0.06-0.73)	-
	Core family configuration N families (N individuals) analysed[c]		
2 parents, 4+ offspring	42 (256)	3 (18)	0
2 parents, 3 offspring	154 (771)	31 (155)	0
2 parents, 2 offspring	519 (2,076)	83 (332)	0
2 parents, 1 offspring	14 (42)	12 (36)	0
1 parent, 4 offspring	1 (5)	9 (45)	0
1 parent, 3 offspring	2 (8)	19 (76)	0
1 parent, 2 offspring	14 (42)	86 (258)	0
1 parent, 1 offspring	1 (2)	14 (28)	0
0 parents, 4+ offspring	17 (71)	0	0
0 parents, 3 offspring	77 (233)	2 (6)	0
0 parents, 2 offspring	229 (465)	11 (22)	118 (236)
0 parents, 1 offspring	19 (24)	2 (2)	160 (160)
Total	1,089 (3,995)	272 (978)	278 (396)

[a] For the Australian GWAS and replication panels, parents were not phenotyped for hematology traits but 1,457 (GWAS) and 386 (replication) were genotyped to improve genotype error detection and the power of association analysis.

[b] Figures are based on individuals that were both genotyped and clinically tested.

[c] Counts consider founders in a family if genotyped and non-founders if both genotyped and clinically tested. For a small number of Australian families, data were included from relatives other than those part of the core family configuration (eg. half-sibs). For this reason, the total number of individuals may differ from that expected from the main family configuration and number of families with data.

Table S2. Heritabilities[a] (main diagonal) and cross-trait phenotypic correlations[b] (below main diagonal) for the six lymphocyte phenotypes tested in the Australian GWAS panel

	CD3	CD4	CD8	CD19	CD56	CD4:CD8
CD3+ T cells	79%					
CD4+ T cells	0.88	80%				
CD8+ T cells	0.82	0.51	82%			
CD19+ B cells	0.51	0.51	0.36	85%		
CD56+ NK cells	0.23	0.18	0.23	0.15	76%	
CD4:CD8 ratio	-0.11	0.33	-0.62	0.07	-0.07	84%

[a] Heritabilities were estimated with Merlin[2] after excluding phenotypic outliers, adjusting for age and sex, and normalising each trait with an inverse normal transformation. For more detailed heritability analyses on this dataset please see Evans et al[3].

[b] Cross-trait phenotypic correlations were estimated on a subset of 1,106 unrelated individuals.

Table S3. Breakdown of SNP and sample filtering during QC of the Australian GWAS panel

		SNPs	Individuals
N at start of QC		599,011	4,406
N dropped at QC step:			
1	BeadStudio GenCall score < 0.7	47,418	-
2	Individuals with call rate < 0.95	-	1
3	SNPs with call rate < 0.95	539	-
4	SNPs with HWE failure $P < 10^{-6}$	813	-
5	SNPs with MAF < 0.01	20,520	-
6	Individuals from families with pedigree errors[a]	-	21
7	SNPs with Mendel failure rate in > 0.05 families	0	-
8	Individuals from families with Mendel failure rate in > 0.05 SNPs	-	0
9	Population stratification gross outliers[b]	-	88
10	Individuals from families with no available lymphocyte data		301
N at end of QC		529,721	3,995

[a] We compared self-reported with genotype-inferred family relationships, the latter based on genome-wide IBS sharing. Forty-eight families with pedigree errors were identified; 21 samples from these families were excluded to correct errors which could not be resolved.

[b] We excluded 88 individuals identified as outliers from populations of European descent through the estimation of genetic ancestry using EIGENSTRAT[4] and data from eleven populations of the HapMap 3 and five Northern European populations genotyped by the GenomeEUtwin consortium.

Table S4. Association analysis of the MHC region in the Australian GWAS panel before and after conditioning on the the most associated variants for CD4:CD8 ratio

Analysis	Variant most associated with CD4:CD8 ratio			CD4:CD8 ratio				CD4 T cell levels				CD8 T cell levels			
	SNP, allele	Nearest gene	Freq	Effect[a]	SE	h^2	*P*	Effect	SE	h^2	*P*	Effect	SE	h^2	*P*
Unconditional	rs2524054, A	*HLA-B*	0.32	0.37	0.03	5.7	2.1×10^{-28}	0.04	0.03	0.1	0.26	-0.31	0.03	4.1	6.2×10^{-21}
Conditional on:															
rs2524054	rs9270986, A	*HLA-DRB*	0.18	0.28	0.04	2.4	4.4×10^{-12}	0.23	0.04	1.4	8.3×10^{-8}	-0.08	0.04	0.2	0.06
rs2524054 & rs9270986	rs2523471, G	*MICA*	0.10	-0.27	0.05	1.4	1.5×10^{-7}	-0.05	0.05	0.1	0.34	0.20	0.05	0.8	8.6×10^{-5}
rs2524054, rs9270986 & rs2523471	rs2523946, T	*HLA-A*	0.44	-0.16	0.03	1.4	1.4×10^{-7}	-0.09	0.03	0.4	0.006	0.08	0.03	0.3	0.01
rs2524054, rs9270986, rs2523471 & rs2523946	rs1431403, C	*HLA-DP*	0.30	0.16	0.03	1.2	9.1×10^{-7}	0.13	0.04	0.7	0.0002	-0.04	0.03	0.1	0.26

[a] Effect corresponds to standard deviation units for the transformed phenotype.

Table S5. Gene expression results from Dixon et al[5] for two peak MHC variants associated with CD4:CD8 ratio (cf. **Table S4**)

Probe ID[a]	Related gene	Beta	h^2	*P*
	MHC class I peak (rs6457374, C)[b]			
209698_at	*CCHCR1*	0.63	16.0	1.9×10^{-12}
AVG_CCHCR1	*CCHCR1*	0.54	11.7	1.6×10^{-9}
211911_x_at	*HLA-B*	-0.53	11.3	2.5×10^{-9}
	MHC class II peak (rs9267992, G)			
209480_at	*HLA-DQB1*	0.94	20.1	8.6×10^{-18}
238900_at	*HLA-DRB1*	0.88	17.8	1.9×10^{-15}
211654_x_at	*HLA-DQB1*	0.88	18.0	2.3×10^{-15}
212999_x_at	*HLA-DQB1*	0.86	16.8	1.6×10^{-14}
213831_at	*HLA-DQA1*	0.81	15.1	1.1×10^{-13}
236203_at	*HLA-DQA1*	0.76	13.2	9.2×10^{-12}
204670_x_at	*HLA-DRB1*	-0.68	10.6	9.1×10^{-10}

[a] The table shows all gene expression probes with >10% phenotypic variance (h^2) explained by rs6457374 or rs9267992. Data were extracted using the mRNA by SNP Browser v1.0.1[5].

[b] The variants most associated with CD4:CD8 ratio (MHC class I: rs2524054; class II: rs9270986) were not present in the mRNA by SNP Browser v1.0.1 database and so results are reported for the best available proxy for each (class I: rs6457374, $r^2 = 0.96$; class II: rs9267992, $r^2 = 0.92$).

Table S6. Two-locus (MHC class I x MHC class II) genotypic means for CD4:CD8 ratio, and single-locus genotypic means for CD4:CD8, CD4 and CD8 cell levels, based on a subset of 1,058 unrelated individuals

Two-locus analysis[a]			**MHC class II** (rs9270986)			**CD4:CD8**[b]	**CD4**	**CD8**
			AA	AC	CC			
MHC class I (rs2524054)	AA	N	12	51	47	110	112	112
		Mean	1.92	2.22	2.10	2.13	1.06	0.52
		SD	0.28	0.55	0.56	0.54	0.26	0.18
	AC	N	8	160	265	433	433	433
		Mean	1.93	1.98	1.73	1.83	1.05	0.61
		SD	0.69	0.56	0.47	0.52	0.26	0.19
	CC	N	9	92	414	515	518	517
		Mean	1.89	1.83	1.62	1.67	1.04	0.65
		SD	0.52	0.49	0.44	0.45	0.28	0.20
CD4:CD8[b]		N	29	303	726			
		Mean	1.91	1.97	1.69			
		SD	0.48	0.55	0.47			
CD4		N	29	305	729			
		Mean	1.15	1.10	1.02			
		SD	0.26	0.29	0.26			
CD8		N	29	305	728			
		Mean	0.63	0.58	0.63			
		SD	0.22	0.18	0.20			

[a] A 4-df interaction test (for the parameters rs2524054_additive*rs9270986_additive, rs2524054_additive*rs9270986_dominant, rs2524054_dominant*rs9270986_additive and rs2524054_dominant*rs9270986_dominant) was not significant ($P = 0.29$).

[b] Significant dominance deviation for the effect of rs2524054 ($P = 0.037$) and rs9270986 ($P = 0.002$) on CD4:CD8 ratio.

Table S7. Association results between the two MHC variants with major effects on basal lymphocyte levels and disease

	MHC class I (rs2524054, A)				**MHC class II** (rs9270986, A)			
	Allele frequency		OR	*P*	Allele frequency		OR	*P*
	Cases	Controls			Cases	Controls		
WTCCC disease cohort:[a]								
Type 1 diabetes	0.34	0.28	1.33	$3.3x10^{-9}$	0.01	0.15	0.04	$1.0x10^{-125}$
Rheumatoid arthritis	0.26	0.28	0.92	0.08	0.10	0.15	0.60	$2.7x10^{-15}$
Crohn's disease	0.26	0.28	0.87	0.01	0.13	0.15	0.81	0.001
Type-2 diabetes	0.26	0.28	0.9	0.02	0.14	0.15	0.92	0.16
Coronary artery disease	0.26	0.28	0.91	0.05	0.15	0.15	0.96	0.46
Hypertension	0.26	0.28	0.90	0.03	0.13	0.15	0.85	0.01
Bipolar disease	0.28	0.28	0.97	0.58	0.17	0.15	1.11	0.07
HIV+ cohort:[b]								
Host control of HIV	0.13	0.26	0.32	$9.1x10^{-11}$	0.10	0.14	0.72	0.036
HLA alleles tagged[c] (r^2)	HLA-C*0702 (.67), HLA-B*0702 (.50)				HLA-DQB1*0602 (1.00), HLA-DRB1*1501 (1.00), HLA-DQA1*0102 (0.95)			

[a] The MHC variant rs2524054 was not present in the Affymetrix 500K chip used by the WTCCC[6] and so was imputed using PLINK[7] as described previously[8]. The SNP was imputed with high confidence (information score 0.94).

[b] The HIV cohort comprised 445 HIV controllers (cases) of European descent which were compared to 733 individuals with progressive HIV disease (controls), also of European descent. HIV Controllers are individuals who naturally control HIV plasma viral load to extremely low or undetectable levels without medications. The rs2524054 and rs9270986 SNPs were imputed in all samples using PLINK with high confidence (information score >0.92). Both SNPs were tested for association to HIV controller status using a logistic regression model, correcting for population structure by including EIGENSTRAT principal components as covariates.

[c] The classic HLA alleles were imputed in the HapMap (release 2) CEU population using PLINK and the set of proxies for each HLA allele identified by de Bakker et al[9]. The linkage disequilibrium (r^2) was then estimated between the HLA alleles and both rs2524054 and rs9270986, and the best tagged HLA alleles identified.

Table S8. Characteristics of the HIV study participants[a]

	HIV controllers	HIV progressors
Genotyped (& clinical)	400 (293-400)	733 (0)
Female sex, %	16	9
Mean age (range), years	47 (23-80)	39 (17-77)
	Mean lymphocyte levels (SD, range):	
CD4, 10^9 cells / L	0.75 (0.29, 0.12-2.17)	-
CD8, 10^9 cells / L	0.92 (0.41, 0.20-2.97)	-
CD4:CD8 ratio	0.94 (0.47, 0.14-2.94)	-

[a]A cohort of 445 HIV-1 controllers and 733 HIV-1 chronically infected individuals ("progressors") of European descent was ascertained through The International HIV Controllers Study. HIV-1 controllers are individuals who naturally control virus replication to levels below 2000 copies/mL in the absence of antiretroviral therapy. The HIV-1 progressors were from AIDS Clinical Trials Group (ACTG) protocols A384, A5095 and A5142 who provided DNA under ACTG protocol A5128[10]. Clinical data were collected longitudinally over the course of routine HIV monitoring visits and were provided by the collaborating physicians and research centers. Patients are typically seen every three months. A subset of the HIV controllers had longitudinal CD4 (N = 400), CD8 and derived CD4:CD8 ratio (N = 293) measurements available (median 8 measurements per person; IQR 4-14). The average of all measurements was taken. Measurements > 5 SD above the mean were excluded and an inverse-normal transformation was applied. Genotyping was performed on the Illumina HumanHap650Y and 1M Duo platforms in two phases. Only samples with a genotyping call rate ≥ 95% were included. SNPs having a call rate < 0.95, Hardy-Weinberg $P < 10^{-6}$, minor allele frequency < 0.01, non-random missingness $P < 10^{-8}$ and differential missingness between cases and controls $P < 10^{-3}$ were excluded. To test for European ancestry, multidimensional scaling plots were drawn using all HapMap phase 3 reference populations. Samples that were gross outliers from the CEU and TSI populations were excluded.

Table S9. Association between the two peak MHC variants[a] and lymphocyte levels in HIV Controllers of European descent

	MHC class I (rs2524054, A)							**MHC class II** (rs9270986, A)						
	N	Allele Frequency	Power[b]	Effect[c]	SE	h^2	*P*	N	Allele Frequency	Power	Effect	SE	h^2	P
CD4:CD8 ratio	293	0.18	0.99	0.08	0.16	0.0	0.64	293	0.16	0.76	-0.19	0.13	0.0	0.17
CD4 levels	400	0.16	-	0.22	0.14	0.1	0.11	400	0.13	0.53	-0.10	0.14	0.0	0.48
CD8 levels	292	0.18	0.94	0.24	0.16	0.1	0.14	292	0.16	-	0.03	0.14	0.0	0.81

[a] SNPs rs2524054 and rs9270986 were imputed in all samples using PLINK with HapMap release 22 CEU founder haplotypes as reference. Both SNPs were imputed with high confidence (information score 0.95 and 1.02, respectively). Association tests for rs2524054 and rs9270986 with CD4 count, CD8 count and CD4:CD8 ratio were performed in PLINK using linear regression. In order to control for the effects of population stratification the first ten EIGENSTRAT principal components were included as covariates in the model. We also adjusted for the effects of age, sex, duration of HIV infection and degree of viremia (undetectable vs. HIV plasma virus load <2000 copies/mL), but the results remained essentially unchanged when these covariates were excluded. In a separate genome-wide association study we are currently testing the hypothesis that host genetic factors are associated with durable suppression of virus replication in 445 HIV controllers as compared to 733 HIV progressors of European descent (www.hivcontrollers.org). Results of the complete study will be presented in greater detail elsewhere.

[b] Theoretical power ($\alpha = 0.05$) to detect the effect size estimated in the Australian GWAS panel (cf **Table 1** in the main text and **Table S4**).

[c] Effect corresponds to standard deviation units for the transformed phenotype.

Members of the International HIV Controllers Study

Providers

Brian Agan, Shanu Agarwal, Brady Allen, Sherly Altidor, Eric Altschuler, Sujata Ambardar, Kathryn Anastos, Ben Anderson, Ushan Andrady, Teresa Angell, Jonathan Appelbaum, Kathy Arnold, Julio Arroyo, Ernie Arthur, Daniel Barbaro, Irene Barnett, Tom Barrett, William Barrie, Vincent Barthel, Mary E. Bartlett, Simon Barton, Pat Basden, Nesli Basgoz, Nicholas Bellos, Judith Berger, Annette Bernard, Nicole Bernard, Donald Blair, Stanley Bodner, Robert Bolan, Emilie Boudreaux, James Braun, Jon Brndjar, Edward Brooks, Kathy Brown, Stephen Brown, Susan Bryant, Rena Bryson, Susan Buchbinder, Diane Bullock, Jedidah Burack, Mary Busalacchi, Larry Bush, Virginia Cafaro, John Campbell, Beth Cardosi, Rob Carlson, Kevin Carmichael, Pat Carr, Mary Carrington, Chris Cavacuiti, Greg Celestin, Michele Chaborek, John Chau, Nancy Chez, Paul Cimoch, Daniel E. Cohen, Bruce Cohn, Lillian Cohn, Brian Conway, David Cooper, Julie Cooper, Brian Cornellson, Anastacia Cowles, David Cox, Roxanne Cox-Iyamu, George Cuchural, Michael Daily, Jennifer Daly, Son Dang, Benjamin Davis, Kristine Davis, P. De Koning, Melissa DeCardi Hladek, Mark Dechter, Steven Deeks, Joseph DeFoto, Edwin DeJesus, Deborah T. Dennis, Matthew Dolan, Joan Duggan, Michael Dunn, Karen Egenes, Todd Ellerin, Kathy Elliott, Joe Eron, John Fangman, Helen Ferlazzo, Yvette Figueroa, Terri Fiola, Stuart Fisk, Maureen Flaherty, Anita Fleenor-Ford, Alice Forster, Todd Fralich, Elliot Francke, Renee Frankel, Kenneth Freedberg, Neel French, Jonathan D. Fuchs, John Fuller, Jonna I. Gaberman, Joel Gallant, Rajesh Gandhi, Anurdha Ganesan, Efrain Garcia, Adrian Gardner, Patricia Garver, Joseph Gastaldo, Joseph Gathe, Cyril Gaultier, Wondwoosen Gebre, Marcy Gelman, Frank Gilman, John Gilmore, Boris Gimbarevsky, Jordan Glaser, Paul Goepfert, Gary Good, Michael Gottlieb, Claudia Goulston, William Grimm, Richard Groger, Jacques Guiteau, T. Douglas Gurley, Stuart Haber, Braden Hale, Roberta Halpenny, Theresa Hardin, W. David Hardy, Todd Hargan, Richard Harrigan, Janna Haston, Trevor Hawkins, Walter Hellinger, Robert Henderson, Tom Henderson, Keith Henry, Susan Hermann, Salvador Hernandez, Claire Hicks, Robert Hoffman, Chris Holland, James Horton, Scott Hoskinson, Ricky K. Hsu, Janice Huber, Gregory Huhn, Rene Humphrey, Mark Hupert, Rocio Hurtado, Mark Illeman, Dennis Israelski, Robert Jellinger, Heiko Jessen, Mina John, Kay Johnson, Mary Jones, Wilbert Jordan, Carol Kauffman, Anna Keba, Homayoon Khanlou, Rob Killian, Arthur Kim, David Kim, Ford Kinder, Jeff Kirchner, Mary Ann Kish, Kathryn Kocurek, Laura Kogelman, E. Milu Kojic, Todd Kortuis, Julia Kovacs, Stacey Kreiswirth, Greg Kuldanek, Wayne Kurisu, Philip Lackey, Ahmadi Ladan, Massimiliano Lanzafame, Michael Lederman, David Lee, Ed Lee, Jenny Lee, Marah Lee, James LeGrand, Janice Lemoine, Steve Levin, Jay Levy, Michael Liguori, Nina Lin, Susan Little, Richard Livine, Josep M. Llibre, Al Lopez, Mona Loutfy, Dawn Loy, Barbara Lubarda, Thomas Magee, Harvey Maksvytis, Alan Man, Stacey Mancuso, Vincent Marconi, Martin Markowitz, Rui Marques, Jr. Martin, William Martin, Gal Mayer, Kenneth Mayer, Cindy McCrone, Cheryl McDonald, Theresa McGhee, Dawn McIntyre, P. Scott McLean, Michael McNett, Hazel Mead, Katherine Mechling, Michael Melia, Craig Metroka, Dirk Meyer-Olson, Brent Michael, Andy Miller, John Montana, Kate Montgomery, Anthony Morice, Karam Mounzer, Michael Munoz, Iris Nagin, Ron Nahass, Thumbi Ndung'u, Sandra Nelson, Craig Nielsen, David Norene, Susan Novales, David O'Connor, Laura O'Donnell, Edward Oldfield, Susan Olender, Bisola Ojikutu, William Owens, Todd Panzer, Brad Parks, Julie Pattison, Andy Pavlatos, Aaron Perlmutter, Michael Norman Pierce, Deborah Pollak, Emile Powe, Richard Pressnell, Lawrence Price, Laurie Proia, Richard Prokesch, Peter Prutch, Heather Pujet, Michael Rausch, Jayashree Ravinshankar, Dalilah Restrepo, Frank Rhame, Constance Richards, Douglas D. Richman, Gregory Robbins, Richard Rose III, Daniel Rosenthal, Polly Ross, David Rubin, Elease Rumbaugh, Will Sanchez, Veeraf Sanjana, Stephen Santiago, Rene Santos, Paul Sax, Debra Scherman, Wolfgang Schmidt, Shannon Schrader, Hanneke Schuitemaker, Anita Scribner, Phillip Sestak, Peter Shalit, William Shay, Jeffrey Sherman, Vivian Shirvani, James Sims, Jay Sizemore, Ross Slotten, Joseph Smiddy, Marcia Sokol-Anderson, Rachel Somogi, Renee Sordjan, James Sosman, Paul Stabile, Jack Stapleton, Sharee Starrett, Kaleo Staszkow, Ben Stearn, Hans-Jurgen Stellbrink, Lisa Sterman, Tinus Steyn, David Stone, Valerie Stone, Ha Ta, Giuseppe Tambussi, Randy Taplitz, Karen Tashima, Daniel Tavari, Ellen Tedaldi, Amalio Telenti, Marianne Tierney, Carolyn Tillquist, Lorraine Tosiello, Cécile Tremblay, Virginia Jeanne Triant, Marc Tribble, P. Trinh, Susan Tuscher, Anthony J. Vaccaro, Emilia Valadas, Isabel Vecino, Theresa Wagner, Charles Walworth, Chingchai Wanidworanun, Douglas Ward, Daniel Warner, Terry Watson, Karl Weber, Terry Weber, Duncan Webster, Winkler Weinberg, Amy Weintrob, Steve Weis, Scott Wenger, Gary Werntz, David Wheeler, David White, Ed Wilkins, Alan Winston, Clifford Wlodaver, Eva Wolf, David Wright, Otto Yang, Sigal Yawetz, David L. Yurdin, Brandon Zabukovic, Kimon Zachary, Beth Zeeman

Ragon Institute of MGH, MIT and Harvard

Sample management and processing: Brett Baker*, Brian Block*, Conrad Braaten, Aisha Tiala Darrah, Brooke LaTour, Jeffrey Lian, Cheryl McClurg, Alicja Piechocka-Trocha, Jacqueline Proudfoot, Alissa Rothchild*, Ildiko Toth. *Recruitment*: Laurie Barefoot, Suzane Bazner, Emily Cutrell, Mary Johnston*, Paul Lemay, Kristin Moss, Olakunle Oladehin, Rachel Rosenberg*, Jake Tinsley, Alice Tsao, Florencia Pereyra (leader). *Investigators*: Marylyn Addo, Todd Allen, Galit Alter, Marcus Altfeld, Sylvie Legall, Florencia Pereyra, Bruce Walker, Xu Yu

Broad Institute of MIT and Harvard

Genotyping: Noël Burtt (Project Manager), Amy Camargo, Gabriel Crawford, Andrew Crenshaw, Leela Davies,

Stacey Gabriel, Lauren Gianniny*, Candace Guiducci. *Analysis*: Chanson Brumme, Paul de Bakker, Xiaoming Jia, Paul McLaren

AIDS Clinical Trials Group (ACTG)
David Haas, Daniel Kuritzkes, Roy Gulick, Heather Ribaudo

Project management
Clinical cohort: Florencia Pereyra, Steve Deeks. *Genetics*: Paul de Bakker, Mary Carrington. *Virology*: Todd Allen, Matthew Henn. *Immunology*: Marcus Altfeld, Sylvie Le Gall. *Bioinformatics*: David Heckerman, Chanson Brumme. *Overall leadership*: Bruce Walker

* Former members

Supplemental References

1. Chen WM, Abecasis GR (2007) Family-based association tests for genomewide association scans. Am J Hum Genet 81:913-926
2. Abecasis GR, Cherny SS, Cookson WO, Cardon LR (2002) Merlin--rapid analysis of dense genetic maps using sparse gene flow trees. Nat Genet 30:97-101
3. Evans DM, Frazer IH, Martin NG (1999) Genetic and environmental causes of variation in basal levels of blood cells. Twin Res 2:250-257.
4. 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
5. Dixon AL, Liang L, Moffatt MF, Chen W, Heath S, Wong KC, Taylor J, Burnett E, Gut I, Farrall M, et al. (2007) A genome-wide association study of global gene expression. Nat Genet 39:1202-1207
6. Wellcome Trust Cases Control Consortium (2007) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447:661-678
7. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, et al. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81:559-575
8. Ferreira MA, O'Donovan MC, Meng YA, Jones IR, Ruderfer DM, Jones L, Fan J, Kirov G, Perlis RH, Green EK, et al. (2008) Collaborative genome-wide association analysis supports a role for ANK3 and CACNA1C in bipolar disorder. Nat Genet 40:1056-1058
9. de Bakker PI, McVean G, Sabeti PC, Miretti MM, Green T, Marchini J, Ke X, Monsuur AJ, Whittaker P, Delgado M, et al. (2006) A high-resolution HLA and SNP haplotype map for disease association studies in the extended human MHC. Nat Genet 38:1166-1172
10. Haas DW, Wilkinson GR, Kuritzkes DR, Richman DD, Nicotera J, Mahon LF, Sutcliffe C, Siminski S, Andersen J, Coughlin K, et al. (2003) A multi-investigator/institutional DNA bank for AIDS-related human genetic studies: AACTG Protocol A5128. HIV Clin Trials 4:287-300