



The Major Genetic Determinants of HIV-1 Control Affect HLA Class I Peptide Presentation

The International HIV Controllers Study

Science **330**, 1551 (2010);

DOI: 10.1126/science.1195271

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by [clicking here](#).

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines [here](#).

The following resources related to this article are available online at www.sciencemag.org (this information is current as of December 9, 2010):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:

<http://www.sciencemag.org/content/330/6010/1551.full.html>

Supporting Online Material can be found at:

<http://www.sciencemag.org/content/suppl/2010/11/02/science.1195271.DC1.html>

This article **cites 31 articles**, 15 of which can be accessed free:

<http://www.sciencemag.org/content/330/6010/1551.full.html#ref-list-1>

This article has been **cited by** 2 articles hosted by HighWire Press; see:

<http://www.sciencemag.org/content/330/6010/1551.full.html#related-urls>

This article appears in the following **subject collections**:

Genetics

<http://www.sciencemag.org/cgi/collection/genetics>

RXLR effector genes (15, 23). Moreover, *Hpa* effectors generally were not located in syntenic locations relative to *Phytophthora* genomes, except for three families of effectors, which have unusually high levels of sequence conservation (Fig. 2).

As obligate biotrophs, downy mildews may have lost some metabolic pathways. We identified several potential metabolic defects in *Hpa* compared with *P. sojae* and *P. ramorum* (fig. S9). For example, genes for nitrate and nitrite reductases, a nitrate transporter, and sulfite reductase were missing (fig. S10 and table S3), which is also a feature of the genomes of obligately parasitic powdery mildew fungi (24). *Hpa* also lacks genes required for synthesis of arachidonic acid and polyamine oxidases.

Flagellated zoospores are produced by many oomycetes (25). Contrastingly, several downy mildew lineages germinate by extending infective germ tubes from nonmotile conidiospores, although evidence exists for a rare zoospore stage in some otherwise conidial downy mildews (26, 27). To conclusively determine whether spore motility has been lost from the *Hpa* lineage, we searched the *Hpa* genome for 90 flagella-associated genes using *Chlamydomonas* sequences and their *Phytophthora* orthologs (28). No matches were detected in *Hpa* for any of these. Similarly, many *Phytophthora* adhesion-related genes are reduced in number or absent from *Hpa*, consistent with the lack of adherent cysts that normally develop from zoospores during infection.

Analysis of *Hpa* gene space revealed genomic signatures of major alterations in pathogenic strategy, metabolism, and development that occurred

during the evolution of obligate biotrophy from a facultative, hemibiotrophic ancestor. Interestingly, some features of *Hpa* gene space (large numbers of secreted effectors, reduction in degradative enzymes, and loss of N and S assimilation) are mirrored in genomes of biotrophic fungi (24, 29, 30). These similarities indicate that convergent adaptations occurred during the independent evolution of biotrophy in fungal and oomycete lineages.

References and Notes

1. E. B. Holub, *Eur. J. Plant Pathol.* **122**, 91 (2008).
2. M. E. Coates, J. L. Beynon, *Annu. Rev. Phytopathol.* **48**, 329 (2010).
3. J. Clark, P. Spencer-Phillips, in *Encyclopedia of Microbiology* (Academic Press, 2000), vol. 2, pp. 117–129.
4. X. Giresse, S. Ahmed, S. Richard-Cervera, F. Delmotte, *J. Phytopathol.* **158**, 321 (2010).
5. M. Göker, H. Voglmayr, A. Riethmüller, F. Oberwinkler, *Fungal Genet. Biol.* **44**, 105 (2007).
6. R. Panstruga, P. N. Dodds, *Science* **324**, 748 (2009).
7. M. Thines, *PLoS ONE* **4**, e4790 (2009).
8. B. J. Haas et al., *Nature* **461**, 393 (2009).
9. B. M. Tyler et al., *Science* **313**, 1261 (2006).
10. Materials and methods are available as supporting material on Science Online.
11. P. D. Bittner-Eddy, R. L. Allen, A. P. Rehmany, P. Birch, J. L. Beynon, *Mol. Plant Pathol.* **4**, 501 (2003).
12. M. Gijzen, T. Nürnberger, *Phytochemistry* **67**, 1800 (2006).
13. S. Kamoun, *Annu. Rev. Phytopathol.* **44**, 41 (2006).
14. E. Gaulin et al., *Plant Cell* **18**, 1766 (2006).
15. R. H. Jiang, S. Tripathy, F. Govers, B. M. Tyler, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 4874 (2008).
16. S. C. Whisson et al., *Nature* **450**, 115 (2007).
17. D. Dou et al., *Plant Cell* **20**, 1930 (2008).
18. S. Kale et al., *Cell* **142**, 284 (2010).
19. D. Dou et al., *Plant Cell* **20**, 1118 (2008).
20. K. H. Sohn, R. Lei, A. Nemri, J. D. Jones, *Plant Cell* **19**, 4077 (2007).

21. R. L. Allen et al., *Science* **306**, 1957 (2004).
22. A. P. Rehmany et al., *Plant Cell* **17**, 1839 (2005).
23. J. Win et al., *Plant Cell* **19**, 2349 (2007).
24. P. Spanu, *Science* **330**, 1543 (2010).
25. A. R. Hardham, G. J. Hyde, *Adv. Bot. Res.* **24**, 353 (1997).
26. D. G. Milbrath, *J. Agric. Sci.* **23**, 989 (1923).
27. V. Skalicky, *Preslia* **38**, 117 (1966).
28. G. J. Pazour, N. Agrin, J. Leszyk, G. B. Witman, *J. Cell Biol.* **170**, 103 (2005).
29. J. Kämper et al., *Nature* **444**, 97 (2006).
30. F. Martin et al., *Nature* **452**, 88 (2008).
31. We thank E. Holub for providing the Emoy2 isolate, D. Greenshields and N. Bruce for technical assistance, A. Heck and M. Slijper for analysis of secreted *Hpa* proteins, R. Hubley for creating repeat modeller libraries, and participants in the 2007 Annotation Jamboree and in the 2008 and 2009 Oomycete Bioinformatics Training Workshops for sequence annotations. This research was supported by grants EF-0412213, IOS-0744875, IOS-0924861, and MCB-0639226 from the U.S. NSF and 2004-35600-15055 and 2007-35319-18100 from the U.S. Department of Agriculture National Institute of Food and Agriculture to B.M.T. and J.M.M.; Biotechnology and Biological Sciences Research Council (BBSRC) BB/C509123/1, BB/E024815/1, and Engineering and Physical Sciences Research Council/BBSRC Systems Biology DTC student EP/F500025/1 to J.B.; Gatsby GAT2545 and BBSRC BB/F0161901, BB/E024882/1, and BBSRC CASE studentship T12144 to J.D.G.J. Other support is detailed in the supporting online material. Genome browsers are maintained at the Virginia Bioinformatics Institute (vmd.vbi.vt.edu) and the Sainsbury Laboratories (gbrowse2.tsl.ac.uk/cgi-bin/gb2/gbrowse/hpa_emoy2_publication).

Supporting Online Material

www.sciencemag.org/cgi/content/full/330/6010/1549/DC1
Materials and Methods

Figs. S1 to S10

Tables S1 to S3

References

16 July 2010; accepted 25 October 2010
10.1126/science.1195203

The Major Genetic Determinants of HIV-1 Control Affect HLA Class I Peptide Presentation

The International HIV Controllers Study*†

Infectious and inflammatory diseases have repeatedly shown strong genetic associations within the major histocompatibility complex (MHC); however, the basis for these associations remains elusive. To define host genetic effects on the outcome of a chronic viral infection, we performed genome-wide association analysis in a multiethnic cohort of HIV-1 controllers and progressors, and we analyzed the effects of individual amino acids within the classical human leukocyte antigen (HLA) proteins. We identified >300 genome-wide significant single-nucleotide polymorphisms (SNPs) within the MHC and none elsewhere. Specific amino acids in the *HLA-B* peptide binding groove, as well as an independent *HLA-C* effect, explain the SNP associations and reconcile both protective and risk *HLA* alleles. These results implicate the nature of the HLA–viral peptide interaction as the major factor modulating durable control of HIV infection.

HIV infection is characterized by acute viremia, often in excess of 5 million viral particles per milliliter of plasma, followed by an average 100-fold or greater decline to a relatively stable plasma virus load set point (1). In the absence of antiretroviral therapy, the

level of viremia is associated with the rate of CD4⁺ T cell decline and progression to AIDS. There is substantial interperson variability in the virus load set point, with most individuals having stable levels exceeding 10,000 RNA copies/ml. Yet a small number of people demonstrate sus-

tained ability to control HIV replication without therapy. Such individuals, referred to as HIV controllers, typically maintain stable CD4⁺ cell counts, do not develop clinical disease, and are less likely to transmit HIV to others (2).

To determine the genetic basis for this rare phenomenon, we established a multinational consortium (www.hivcontrollers.org) to recruit HIV-1 controllers, who are defined by at least three measurements of plasma virus load (VL) < 2000 RNA copies/ml over at least a 12-month period in the absence of antiviral therapy. We performed a genome-wide association study (GWAS) in the HIV controllers (median VL, CD4 count, and disease duration of 241 copies/ml, 699 cells/mm³, and 10 years, respectively) and treatment-naïve chronically infected individuals with advanced disease (median VL and CD4 count of 61,698 copies/ml and 224 cells/mm³, respectively) enrolled in antiviral treatment studies led by the AIDS Clinical Trials Group. After quality control and imputation on the basis of HapMap

*All authors with their contributions and affiliations appear at the end of this paper.

†To whom correspondence should be addressed. E-mail: bwalker@partners.org (B.D.W.); pdebakker@rics.bwh.harvard.edu (P.I.W.d.B.)

Phase 3 (3), we obtained data on 1,384,048 single-nucleotide polymorphisms (SNPs) in 974 controllers (cases) and 2648 progressors (controls) from multiple populations (table S1).

After stratification into European, African American, and Hispanic ethnic groups (fig. S1), we tested each SNP for association using logistic regression, including the major principal components as covariates to correct for population substructure (4). In the largest group, comprising 1712 individuals of European ancestry, we identified 313 SNPs with genome-wide signif-

icance, defined by $P < 5 \times 10^{-8}$ due to correction for multiple comparisons (table S2). All SNPs that reached genome-wide significance were located in the major histocompatibility complex (MHC) region on chromosome 6 (Fig. 1A). We obtained similar results for the other two ethnic groups and in a meta-analysis of all participants (fig. S2). We also performed a genome-wide analysis to test the influence of local chromosomal ancestry in the African American sample (4), but we detected no signal outside the MHC (figs. S3 and S4). The impact of the MHC was further un-

derscored when we specifically tested published associations related to HIV disease progression outside the MHC. Only variants in the *CCR5-CCR2* locus—namely, *CCR5*Δ32 deletion polymorphism (5), C927T in *CCR5* (6), and Val⁶⁴→Ile⁶⁴ in *CCR2* (7)—replicate with nominal statistical significance in our study (Fig. 1B and table S3).

Closer examination of the significant SNPs within the MHC showed that they are located within a 3-Mb region concentrated around class I *human leukocyte antigen (HLA)* genes (fig. S5), but extensive linkage disequilibrium (LD) makes precise assignment of causal variants challenging (8). Therefore, we used stepwise regression to define independent markers associated with host control. From the initial set of 313 SNPs that reached genome-wide significance in the European sample, for which the greatest numbers of participants were available, we found only four independent markers of association (Table 1). rs9264942, located 35 kb upstream of *HLA-C* and a putative variant associated with *HLA-C* expression levels [odds ratio (OR) = 2.9, $P = 2.8 \times 10^{-35}$, where an OR > 1 indicates a protective effect], and rs2395029, a proxy for *HLA-B*57:01* (OR = 5.3, $P = 9.7 \times 10^{-26}$), had been previously reported to be associated with virus load set point after acute infection (9). We also defined rs4418214, a noncoding SNP near *MICA* (OR = 4.4, $P = 1.4 \times 10^{-34}$), and rs3131018 in *PSORSIC3*, a gene implicated in psoriasis (OR = 2.1, $P = 4.2 \times 10^{-16}$). These four SNPs explain 19% of the observed variance of host control in the European sample; together with those in *CCR5*, these SNPs explain 23%, using Nagelkerke's approximation (Fig. 1C) (10).

In the smaller African American sample, we observed 33 SNPs with genome-wide significance, four of which were identified as independent markers, but all differed from those in the European sample (Table 1). This suggests that shared causal variants are tagged by different SNPs in these two populations or that the mechanism of control differs with ethnicity. Only rs2523608 was previously identified, in a recent study of virus load set point in African Americans (11). Despite no evidence for historical recombi-

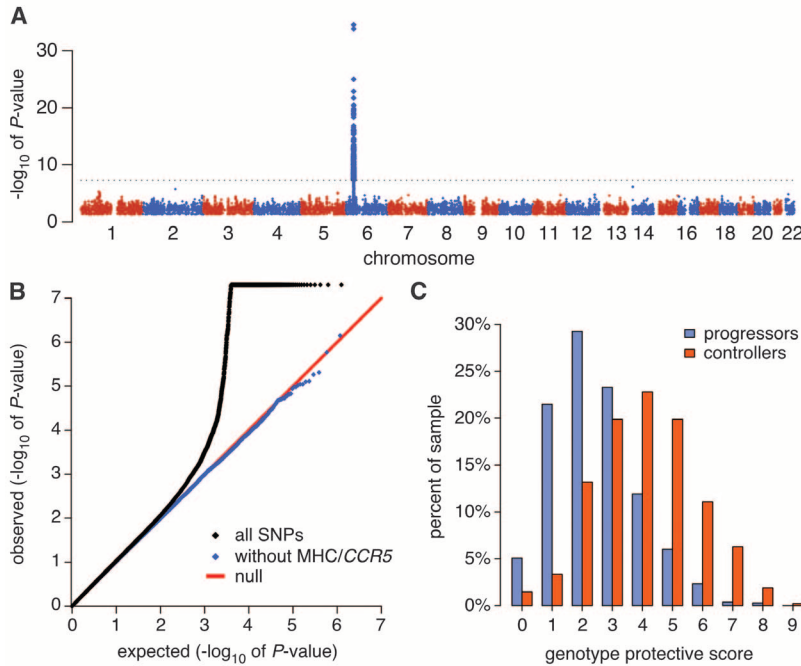


Fig. 1. Genome-wide association results in the European sample. (A) Manhattan plot of 1.3 million autosomal SNPs. Only SNPs in the MHC on chromosome 6 reach genome-wide significance, indicated by the horizontal dotted line ($P < 5 \times 10^{-8}$). Red and blue colors alternate between chromosomes. (B) Quantile-quantile plot of the association results with (black) and without (blue) SNPs in the extended MHC and the *CCR5-CCR2* locus, indicating that the detectable effect is entirely attributable to these two loci. The red line denotes the expected distribution under the null hypothesis of no effect. (C) Distribution of the genotype protective score, defined as the total number of alleles associated with host control at the four independent SNPs in the MHC and the variants at *CCR5-CCR2*, showing marked differences in controllers (orange) and progressors (blue). In aggregate, these variants explain 23% of the observed variance of durable host control.

Table 1. Association results for the independent SNPs in the MHC identified with stepwise regression in the European and African American samples. The odds ratio and frequency is given for the A1 allele, where OR >

1 indicates a protective effect. Odds ratios and *P* values were computed for univariate and multivariate regression models. C, cytosine; G, guanine; T, thymine; A, adenine.

SNP	A1	A2	Frequency in controllers	Frequency in progressors	Univariate		Multivariate	
					OR	<i>P</i> value	OR	<i>P</i> value
<i>European</i>								
rs9264942	C	T	0.595	0.336	2.9	2.8×10^{-35}	2.1	6.3×10^{-16}
rs4418214	C	T	0.240	0.075	4.4	1.4×10^{-34}	1.8	4.9×10^{-4}
rs2395029	G	T	0.139	0.032	5.3	9.7×10^{-26}	2.1	3.5×10^{-4}
rs3131018	C	A	0.777	0.625	2.1	4.2×10^{-16}	1.5	1.2×10^{-5}
<i>African American</i>								
rs2523608	G	A	0.522	0.326	2.6	8.9×10^{-20}	2.3	3.7×10^{-15}
rs2255221	T	G	0.264	0.137	2.7	3.5×10^{-14}	1.9	2.1×10^{-6}
rs2523590	C	T	0.300	0.164	2.4	1.7×10^{-13}	2.3	1.2×10^{-12}
rs9262632	G	A	0.097	0.034	3.1	1.0×10^{-8}	2.2	2.8×10^{-4}

nation ($D' = 1$), this SNP is only weakly correlated ($r^2 < 0.1$) with *HLA-B*57:03*, the class I allele most strongly associated with durable control of HIV in populations of African ancestry (11–13). In the Hispanic sample, which was much smaller, the most significant SNP was rs2523590, 2 kb upstream of *HLA-B*, also identified in the African American sample described here.

Given the localization of significant SNPs entirely to the HLA class I region, as well as previous studies showing *HLA* alleles to affect disease progression (13–20), we next sought to evaluate whether these SNP and HLA associations might be due to specific amino acids within HLA. Because *HLA* types were available for only a portion of the entire cohort, we developed a method to impute classical *HLA* alleles and their corresponding amino acid sequences (4) on the basis of haplotype patterns in an independent data set collected by the Type 1 Diabetes Genetics Consortium (T1DGC) (21). This data set contains genotype data for 639 SNPs in the MHC that overlap with genotyped SNPs in our GWAS and classical *HLA* types for class I and II loci at four-digit resolution in 2767 unrelated individuals of European descent.

We imputed *HLA* types in the European sample of our study and validated the imputations by comparing to empirical four-digit *HLA* typing data collected for class I loci in a subset ($n = 371$) of the HIV controllers. The quality of the imputations was such that the imputed and true frequencies for all *HLA* alleles in this subset were in near-perfect agreement (Fig. 2A) ($r^2 = 0.99$). Furthermore, the positive predictive value was 95.2% and the sensitivity was 95.2% at two-digit resolution (92.7 and 95.6%, respectively, at four-digit resolution) for *HLA* alleles with frequency $>2\%$ (Fig. 2B). This indicates that the performance of the imputation was generally excellent for common alleles, consistent with previous work (22). We used *HLA* allele imputations in all participants (even those with *HLA* types defined by sequencing) for association analyses to avoid systematic bias between cases and controls. Lower imputation quality would only decrease power, not increase the false-positive rate, because cases and controls would be equally affected.

We tested all *HLA* alleles for association via logistic regression, adjusting for the same covariates used in SNP analysis (tables S4 and S5). The most significant HLA association is *B*57:01* (OR = 5.5, $P = 1.4 \times 10^{-26}$), which explains the proxy association of rs2395029 in *HCP5*. With the use of stepwise regression modeling in the European sample of controllers and progressors, we were able to implicate *B*57:01*, *B*27:05*, *B*14/Cw*08:02*, *B*52*, and *A*25* as protective alleles and *B*35* and *Cw*07* as risk alleles. These associations are consistent with earlier studies that highlighted a role for HLA class I loci (13–20), and particularly *HLA-B* alleles in control of HIV, which indicated that the imputations are robust. Collectively explaining 19% of the variance of host control, these *HLA* allele associations are consistent with the effects of the four independent SNPs.

Virus-infected cells are recognized by CD8⁺ T cells after presentation of short viral peptides within the binding groove of HLA class I, and HIV-specific CD8⁺ T cells are strongly associated with control (23). We thus evaluated whether the SNP associations identified in the GWAS, and the HLA associations derived from imputation, might be due to specific amino acid positions within the HLA molecules, particularly those involved in the interaction between the viral peptide and the HLA class I molecule. Using the official DNA sequences defined for known *HLA* alleles (24), we encoded all variable amino acid positions within the coding regions of the *HLA* genes in each of the previously *HLA*-typed 2767 individuals in the T1DGC reference panel, and we used this data set to impute the amino acids in the cases and controls (4). Among a total of 372 polymorphic amino acid positions in class I and II HLA proteins, 286 are biallelic like a typical nonsynonymous coding SNP. The remaining 86 positions accommodate more than two amino acids; position 97 is the most diverse in HLA-B with six possible amino acids observed in European populations.

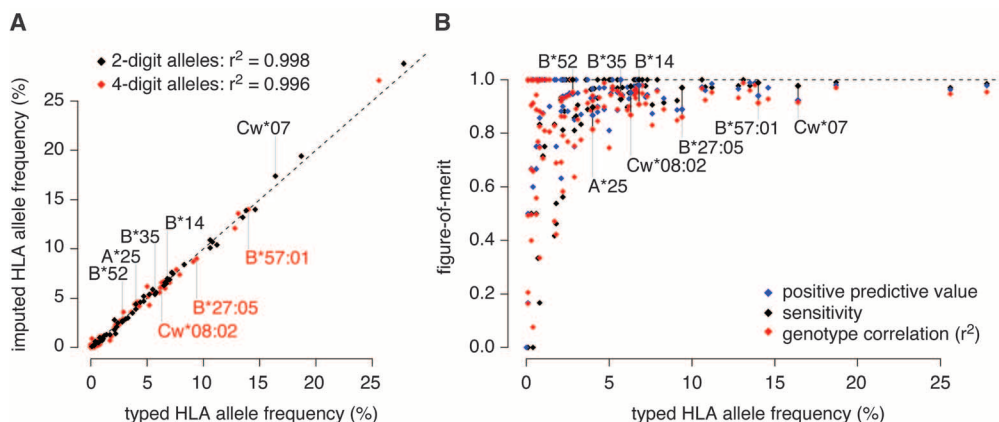
After imputing these amino acids in the European sample, we used logistic regression to test all positions for association with host control (fig. S6 and table S6). Notably, position 97 in HLA-B

was more significant (omnibus $P = 4 \times 10^{-45}$) than any single SNP in the GWAS, and three amino acid positions (67, 70, and 97), all in HLA-B, showed much stronger associations than any single classical *HLA* allele, including *B*57:01* (Fig. 3A). Moreover, allelic variants at these positions were associated with substantial frequency differences between cases and controls (Fig. 3B). These results indicate that the effect of *HLA-B* on disease outcome could be mediated, at least in part, by these positions. These three amino acid positions are located in the peptide binding groove, which suggests that conformational differences in peptide presentation at these sites contribute to the protective or susceptible nature of the various *HLA-B* allotypes. Although both innate and adaptive mechanisms could be at play, the hypothesis that HLA affects peptide presentation and subsequent T cell functionality is supported by experimental data showing substantial functional differences between CTL targeting identical epitopes but restricted by different *HLA* alleles (25).

We next performed stepwise regression modeling and identified six residues as independent markers associated with durable control of HIV. These include Arg⁹⁷, Cys⁶⁷, Gly⁶², and Glu⁶³, all in HLA-B; Ser⁷⁷ in HLA-A; and Met³⁰⁴ in HLA-C, which collectively explain 20% of the observed variance (similar to the variance explained by the seven classical *HLA* alleles described above). With the exception of Met³⁰⁴ in the transmembrane domain of HLA-C, these residues are all located in the MHC class I peptide binding groove, again suggesting that the binding pocket—and, by inference, the conformational presentation of class I-restricted epitopes—plays a key role in host control.

Having identified these amino acid positions as strong candidates to account for the SNP and HLA association signals in this study, we next investigated their effects on protection or risk, revealing allelic variants at these positions linked to both extremes (Table 2). HLA-B position 97 (omnibus $P = 4 \times 10^{-45}$), located at the base of the C pocket, has important conformational properties for peptide binding (26). Position 97 has six allelic variants: Protective haplotypes *B*57:01*, *B*27:05*, and *B*14* are uniquely defined by Val⁹⁷

Fig. 2. Imputation quality of classical *HLA* alleles in the European sample. **(A)** Concordance between imputed (y -axis) and observed (x -axis) frequencies of classical *HLA* types in 371 HIV-1 controllers with four-digit *HLA* types obtained through Sanger sequencing. **(B)** Positive predictive value, sensitivity, and genotype correlation (r^2) with typed alleles as a function of the observed frequency.



(3% frequency in controls), Asn⁹⁷ (4%) and Trp⁹⁷ (3%), respectively; the other amino acids at this position (Ser, Thr, Arg) segregate on a diverse set of haplotypes. Ser⁹⁷ (27% frequency) lies on

risk haplotypes *Cw*07*, *B*07*, and others, whereas Thr⁹⁷ (11%) lies on protective *B*52* (and others). Arg⁹⁷ is the most common amino acid (51%) and is carried by risk allele *B*35*, among others.

The importance of this amino acid position to host control is underscored by conditional analyses revealing significance when we adjust incrementally for Val⁹⁷ (omnibus test for position

Fig. 3. Associations at amino acids in *HLA-B* in the European sample. **(A)** Association results for all variable amino acid positions, as calculated by the omnibus test. Colors denote conventional pocket positions. *P* values for significant classical *HLA-B* alleles are shown for comparison. **(B)** Marked allele frequency differences between controllers and progressors for amino acids at positions 67, 70, and 97. Numbers above the bars indicate odds ratios (values >1 indicate a protective effect). **(C)** Associations between allelic variants at amino acid positions 67, 70, and 97 and quantitative virus load set point in the independent Swiss HIV cohort study. Effect estimates (beta coefficients from a linear-regression model) are given in log₁₀ units of virus load set point. *P* values refer to the omnibus test for association at each position. Error bars indicate the standard error of the beta coefficient.

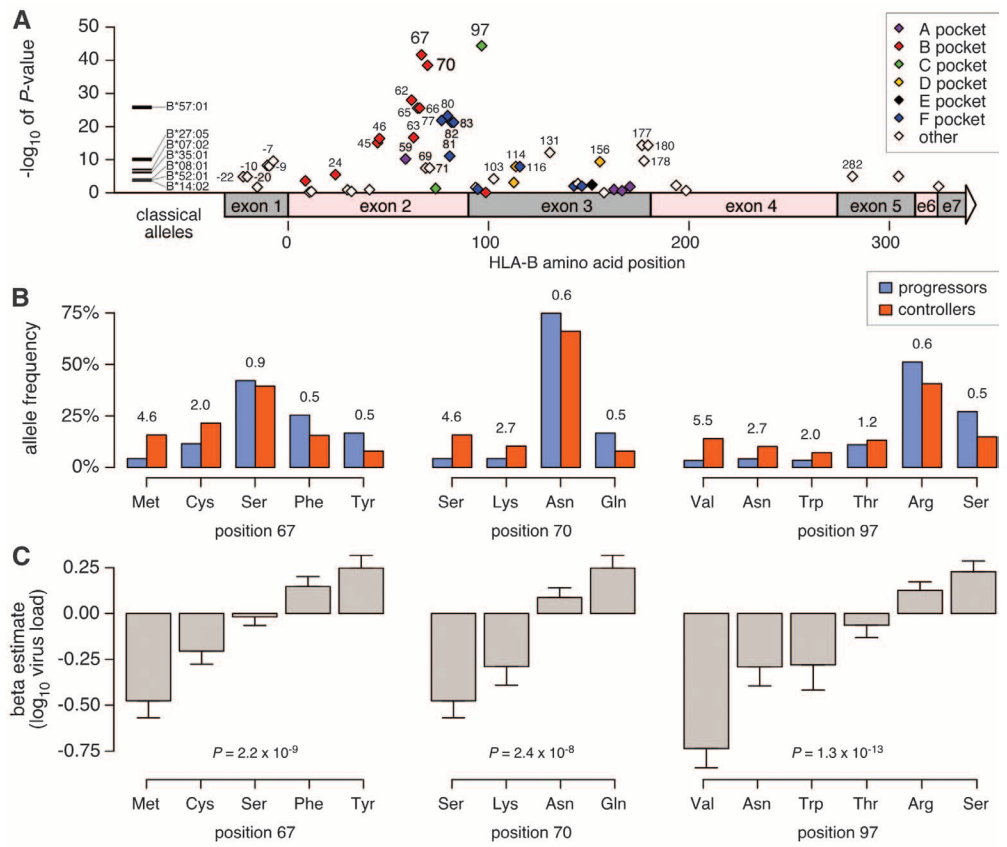


Table 2. Haplotypes defined by the four independent SNPs, classical *HLA* alleles, and amino acids associated with host control in the European sample. Haplotypes are ordered by the estimated odds ratio, where the most common haplotype was taken as reference (OR = 1). *P* values are

for each haplotype tested against all other haplotypes. Only haplotypes with >1% frequency are listed, accounting for >85% of haplotype diversity. *HLA-A* alleles were excluded to limit the number of haplotypes. See (33).

rs3131018	<i>HLA-C</i>		rs9264942	<i>HLA-B</i>					rs4418214	rs2395029	Frequency	OR	<i>P</i> value	
	Classical	304		Classical	62	63	67	70						97
C		M	C	<i>B*57:01</i>	G	E	M	S	V	C	G	0.060	7.05	1.5E-26
C		M	C	<i>B*52:01</i>	R	E	S	N	T	T	T	0.011	6.32	4.2E-05
C		V	C	<i>B*27:05</i>	R	E	C	K	N	C	T	0.051	3.41	1.3E-10
C		M	C		R	E	S	N	T	T	T	0.024	2.78	1.3E-03
C	<i>Cw*08:02</i>	M	C	<i>B*14:02</i>	R	N	C	N	W	T	T	0.030	2.58	6.0E-03
C		M	C		R	N	S	N	R	T	T	0.021	2.16	4.2E-02
C		V	C		R	N	F	N	T	T	T	0.021	2.02	4.6E-01
C		M	C		R	N	C	N	R	T	T	0.025	1.58	1.7E-01
C		V	C		R	E	S	N	S	T	T	0.012	1.50	4.5E-01
A		M	C		R	E	S	N	R	T	T	0.067	1.38	6.5E-01
C		M	T		R	N	F	N	T	T	T	0.020	1.29	8.9E-01
A		M	C		R	N	S	N	R	T	T	0.016	1.03	1.7E-01
C		V	T		R	E	S	N	R	T	T	0.168	(reference)	1.6E-03
C		M	C		R	E	S	N	R	T	T	0.022	0.98	4.4E-01
A		V	T		R	E	S	N	R	T	T	0.018	0.87	6.0E-02
C	<i>Cw*07:01</i>	V	T		R	N	S	N	R	T	T	0.016	0.80	9.5E-02
C	<i>Cw*07:01</i>	V	T	<i>B*08:01</i>	R	N	F	N	S	T	T	0.085	0.79	6.0E-05
C		V	T		R	N	Y	Q	T	T	T	0.018	0.67	3.3E-02
A	<i>Cw*07:02</i>	V	T	<i>B*07:02</i>	R	N	Y	Q	S	T	T	0.116	0.65	3.2E-08
A		V	T	<i>B*35:01</i>	R	N	F	N	R	T	T	0.050	0.51	4.3E-06
A		V	T		R	N	F	N	R	T	T	0.017	0.29	4.1E-05

Downloaded from www.sciencemag.org on December 9, 2010

97, $P = 3 \times 10^{-20}$), Asn⁹⁷ ($P = 2 \times 10^{-9}$), and Trp⁹⁷ ($P = 7 \times 10^{-5}$). Thus, at a single position within the peptide binding groove (position 97, C-pocket), discrete amino acids are associated with opposite disease outcomes, even after controlling for *B*57* and *B*27*, alleles associated with host control.

We also found similar discordant associations for alleles at positions 67, 63, and 62 (Table 2), all of which line the $\alpha 1$ helix along the peptide binding groove and help shape the B-pocket (Fig. 4). At position 67 (omnibus $P = 2 \times 10^{-42}$), risk haplotypes *B*35* and *B*07* carry aromatic residues Phe⁶⁷ and Tyr⁶⁷, respectively, whereas protective *B*57:01*, *B*27:05*, and *B*14* alleles carry sulfur-containing residues Met⁶⁷ or Cys⁶⁷. Position 62 ($P = 5 \times 10^{-27}$) is biallelic (Arg/Gly) with the Gly⁶² allele segregating with protective alleles *B*57:01* and *B*58* (<1% frequency, OR = 1.7, $P = 0.2$). Adjacent position 63 ($P = 9 \times 10^{-16}$) is also biallelic (Glu/Asn) with Glu⁶³ appearing in complete LD ($D' = 1$) with *B*57:01*, *B*27:05*, and *B*52*. In contrast, at this position the risk alleles *B*07* (14% frequency, OR = 0.5, $P = 1 \times 10^{-7}$) and *B*35* both carry Asn⁶³. Position 70 (omnibus $P = 3 \times 10^{-39}$) accommodates four alleles that are tightly coupled with positions 67 and 97: Ser⁷⁰ appears exclusively with Met⁶⁷ (which defines *B*57* and *B*58*), Gln⁷⁰ with Tyr⁶⁷, and Lys⁷⁰ with Asn⁹⁷ (*B*27*). Hence, these data create a consistent and parsimonious model that can explain the associations of classical *HLA-B* alleles by specific amino acids lining the binding groove (and residues tightly coupled to them), which are expected to have an impact on the three-dimensional structure of the peptide-MHC complex.

To further investigate the role of individual amino acid positions in *HLA-B*, we implemented a permutation procedure to assess how consistent the above observations are with a null model in which there is no relation between amino acids at a particular position and host control (4). The

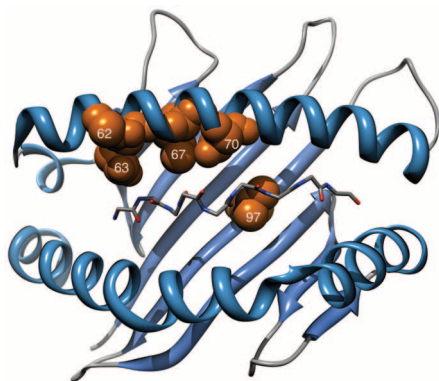


Fig. 4. Three-dimensional ribbon representation of the HLA-B protein based on Protein Data Bank entry 2bvp (30), highlighting amino acid positions 62, 63, 67, 70, and 97 lining the peptide binding pocket. The peptide backbone of the epitope is also displayed. This figure was prepared with UCSF Chimera (32).

results of this procedure provided evidence that multiple amino acid positions in the peptide binding groove are indeed associated with host control (table S7), including positions 62, 63, 67, 70, and 97, thus providing a structural basis for the effect of *HLA-B* on host control (Fig. 4).

Within HLA-A position 77, which lies on the α helix contributing to the F-pocket, we identified a weaker but still significant association (omnibus $P = 3 \times 10^{-6}$). Ser⁷⁷ (6% frequency, OR = 2.0, $P = 2 \times 10^{-6}$) is carried by only two *HLA-A* alleles (joint $r^2 = 1$): *A*25* (2.4% frequency, OR = 2.6, $P = 1 \times 10^{-5}$) and *A*32* (3.2%, OR = 1.6, $P = 0.02$). Given its location and earlier association evidence for the A10 supertype (27), *HLA-A* could play a role in host control, although the evidence is not as strong as for *HLA-B*.

The signals within *HLA-C* are less straightforward to interpret. Position 304 is a biallelic variant (Val/Met) located in the transmembrane domain (Met³⁰⁴, 28% frequency, OR = 2.3, $P = 7 \times 10^{-23}$). Met³⁰⁴ is in moderate LD ($r^2 = 0.5$) with rs9264942, which is known to be associated with *HLA-C* expression levels (28). Addition of this SNP to a multivariate model of all six amino acids is marginally significant ($P = 0.013$) but eliminates the effect of Met³⁰⁴ ($P = 0.06$). Similarly, addition of rs9264942 to a multivariate model of all seven independent classical *HLA* alleles is also significant ($P = 2 \times 10^{-4}$) but eliminates the effect of *Cw*07* ($P = 0.08$). These observations make it difficult to determine the extent to which epitope presentation in the *HLA-C* peptide binding pocket is important for host control. Thus, rs9264942 could be a proxy for not only many protective and risk *HLA* alleles (predominantly at *HLA-B*), but also for an independent effect on *HLA-C* gene expression, differentially affecting the response to HIV (29).

We next evaluated associations for the SNPs in the MHC, classical *HLA* alleles, and amino acids in a second independent cohort of untreated HIV-infected persons from Switzerland (fig. S7 and tables S8 and S9) (4), in whom virus load set point was measured as a quantitative trait. Allelic variants at positions 67, 70, and 97 were also associated with highly significant differences in virus load set point in this second cohort (Fig. 3C). The effect estimates of all variable amino acids in *HLA-B* ($r^2 > 0.9$) and, to a lesser degree, those in *HLA-C* ($r^2 > 0.8$) in that cohort are in excellent agreement (figs. S8 and S9). As before, position 97 in *HLA-B* is the most significant association (omnibus $P = 1 \times 10^{-13}$). The *HLA-A* associations (*A*25* or Ser⁷⁷) did not replicate, which reduces the likelihood that *HLA-A* plays a major role in host control.

In the African American sample (fig. S10), the most significant *HLA* allele association was observed for two-digit *B*57* (OR = 5.1, $P = 1.7 \times 10^{-21}$) and four-digit *B*57:03* (OR = 5.1, $P = 2.8 \times 10^{-17}$; tables S10 and S11), consistent with previous studies (11–13). Position 97 in *HLA-B* (omnibus $P = 2 \times 10^{-25}$) is again the most significant amino acid (table S12). The consistency

of these results demonstrates that imputation and association testing at amino acid resolution in multiple ethnicities can resolve disparate SNP associations in the MHC and help with fine-mapping of classical HLA associations.

Altogether, these results link the major genetic impact of host control of HIV-1 to specific amino acids involved in the presentation of viral peptides on infected cells. Moreover, they reconcile previously reported SNP and HLA associations with host control and lack of control to specific amino acid positions within the MHC class I peptide binding groove. Although variation in the entire HLA protein is involved in the differential response to HIV across *HLA* allotypes, the major genetic effects are condensed to the positions highlighted in this study, indicating a structural basis for the HLA association with disease progression that is probably mediated by the conformation of the peptide within the class I binding groove. The most significant residue, position 97 in the floor of the peptide binding groove of *HLA-B*, is associated with the extremes of viral load, depending on the expressed amino acid. This residue has been shown to have important conformational properties that affect epitope-contacting residues within the binding groove (26, 30) and has also been implicated in HLA protein folding and cell-surface expression (31).

Although the main focus of this study was on common sequence variation, it remains an open question as to the role of variants outside the MHC and the contribution of epistatic effects and epigenetic regulation. Additional factors also contribute to immune control of HIV, including fitness-altering mutations, immunoregulatory networks, T cell help, thymic selection, and innate effector mechanisms such as killer cell immunoglobulin-like receptor recognition (23), some of which are influenced by the peptide-HLA class I complex. However, the combination and location of the significant amino acids defined here are most consistent with the genetic associations observed being modulated by HLA class I restricted CD8⁺ T cells. These results implicate the nature of the HLA-viral peptide interaction as the major genetic factor modulating durable control of HIV infection and provide the basis for future studies of the impact of HLA-peptide conformation on immune cell induction and function.

References and Notes

1. A. J. McMichael, P. Borrow, G. D. Tomaras, N. Goonetilleke, B. F. Haynes, *Nat. Rev. Immunol.* **10**, 11 (2010).
2. S. G. Deeks, B. D. Walker, *Immunity* **27**, 406 (2007).
3. The International HapMap 3 Consortium, *Nature* **467**, 52 (2010).
4. See supporting online material on Science Online for detailed background on the analyses that we performed.
5. M. Dean et al., *Science* **273**, 1856 (1996).
6. M. P. Martin et al., *Science* **282**, 1907 (1998).
7. M. W. Smith et al., *Science* **277**, 959 (1997).
8. P. I. W. de Bakker et al., *Nat. Genet.* **38**, 1166 (2006).
9. J. Fellay et al., *Science* **317**, 944 (2007); 10.1126/science.1143767.

10. N. J. D. Nagelkerke, *Biometrika* **78**, 691 (1991).
11. K. Pelak et al., *J. Infect. Dis.* **201**, 1141 (2010).
12. F. Pereyra et al., *J. Infect. Dis.* **197**, 563 (2008).
13. C. Costello et al., *AIDS* **13**, 1990 (1999).
14. M. R. Klein et al., *J. Infect. Dis.* **169**, 1244 (1994).
15. R. A. Kaslow et al., *Nat. Med.* **2**, 405 (1996).
16. M. Carrington et al., *Science* **283**, 1748 (1999).
17. S. A. Migueles et al., *Proc. Natl. Acad. Sci. U.S.A.* **97**, 2709 (2000).
18. P. O. Flores-Villanueva et al., *Proc. Natl. Acad. Sci. U.S.A.* **98**, 5140 (2001).
19. M. Carrington, S. J. O'Brien, *Annu. Rev. Med.* **54**, 535 (2003).
20. P. Kiepiela et al., *Nature* **432**, 769 (2004).
21. W. M. Brown et al., *Diabetes Obes. Metab.* **11** (suppl. 1), 2 (2009).
22. S. Leslie, P. Donnelly, G. McVean, *Am. J. Hum. Genet.* **82**, 48 (2008).
23. H. W. Virgin, B. D. Walker, *Nature* **464**, 224 (2010).
24. J. Robinson et al., *Nucleic Acids Res.* **37**, D1013 (2009).
25. A. Leslie et al., *J. Immunol.* **177**, 4699 (2006).
26. T. Fagerberg, J. C. Cerottini, O. Michielin, *J. Mol. Biol.* **356**, 521 (2006).
27. G. Catano et al., *PLoS ONE* **3**, e3636 (2008).
28. B. E. Stranger et al., *Nat. Genet.* **39**, 1217 (2007).
29. R. Thomas et al., *Nat. Genet.* **41**, 1290 (2009).
30. G. B. Stewart-Jones et al., *J. Immunol.* **175**, 2459 (2005).
31. M. A. Blanco-Gelaz et al., *Int. Immunol.* **18**, 211 (2006).
32. E. F. Pettersen et al., *J. Comput. Chem.* **25**, 1605 (2004).
33. Single-letter abbreviations for the amino acid residues referred to in Table 2 are as follows: E, Glu; F, Phe; G, Gly; K, Lys; M, Met; N, Asn; Q, Gln; R, Arg; S, Ser; V, Val; W, Trp; and Y, Tyr.
34. This work was made possible through a generous donation from the Mark and Lisa Schwartz Foundation and a subsequent award from the Collaboration for AIDS Vaccine Discovery of the Bill and Melinda Gates Foundation. This work was also supported in part by the Harvard University Center for AIDS Research (grant P-30-AI060354); University of California San Francisco (UCSF) Center for AIDS Research (grant P-30 AI27763); UCSF Clinical and Translational Science Institute (grant UL1 RR024131); Center for AIDS Research Network of Integrated Clinical Systems (grant R24 AI067039); and NIH grants AI28568 and AI030914 (B.D.W.); AI087145 and K24AI069994 (S.G.D.); AI069513, AI34835, AI069432, AI069423, AI069477, AI069501, AI069474, AI069428, AI69467, AI069415, AI32782, AI27661, AI25859, AI28568, AI30914, AI069495, AI069471, AI069532, AI069452, AI069450, AI069556, AI069484, AI069472, AI34853, AI069465, AI069511, AI38844, AI069424, AI069434, AI46370, AI68634, AI069502, AI069419, AI068636, and RR024975 (AIDS Clinical Trials Group); and AI077505 and MH071205 (D.W.H.). The Swiss HIV Cohort Study is supported by the Swiss National Science Foundation (SNF grants 33CS00-108787 and 310000-110012). S. Ripke acknowledges support from NIH/National Institute of Mental Health (grant MH085520). This project has been funded in whole or in part with funds from National Cancer Institute/NIH (grant HHSN261200800001E to M. Carrington). The content of this publication does not necessarily reflect the views or policies of the U.S. Department of Health and Human Services, nor does the mention of trade names, commercial products, or organizations imply endorsement by the U.S. government. This research was supported in part by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research.

Author Contributions

Writing team: Florencia Pereyra,^{1,2,†} Xiaoming Jia,^{3,†} Paul J. McLaren,^{4,5,†} Amalio Telenti,⁶ Paul I.W. de Bakker,^{4,5,7,8,†} (co-chair), Bruce D. Walker,^{1,9,†} (co-chair)
Analysis team: Xiaoming Jia,³ Paul J. McLaren,^{4,5} (project leaders), Stephan Ripke,^{4,10} Chanson J. Brumme,¹ Sara L. Pulit,^{4,5} Amalio Telenti,⁶ Mary Carrington,^{1,11} Carl M. Kadie,¹² Jonathan M.

Carlson,¹³ David Heckerman,¹³ Paul I.W. de Bakker,^{4,5,7,8,†} (chair)
Study design: Florencia Pereyra,^{1,2} Paul I.W. de Bakker,^{4,5,7,8,†} Robert R. Graham,¹⁴ Robert M. Plenge,^{4,15} Steven G. Deeks,¹⁶ Bruce D. Walker,^{1,9,†} (chair)
SNP genotyping, HLA typing, and sample management: Lauren Gianniny,⁴ Gabriel Crawford,⁴ Jordan Sullivan,⁴ Elena Gonzalez,⁴ Leela Davies,⁴ Amy Camargo,⁴ Jamie M. Moore,⁴ Nicole Beattie,⁴ Supriya Gupta,⁴ Andrew Crenshaw,⁴ Noël P. Burt,⁴ Candace Guiducci,⁴ Namrata Gupta,⁴ Mary Carrington,^{1,11} Xiaojiang Gao,¹¹ Ying Qi,¹¹ Yuko Yuki¹¹
HIV controllers recruitment and sample management: Florencia Pereyra^{1,2} (project leader), Alicia Piechocka-Trocha,¹ Emily Cutrell,¹ Rachel Rosenberg,¹ Kristin L. Moss,¹ Paul Lemay,¹ Jessica O'Leary,¹ Todd Schaefer,¹ Pranshu Verma,¹ Ildiko Toth,¹ Brian Block,¹ Brett Baker,¹ Alissa Rothchild,¹ Jeffrey Lian,¹ Jacqueline Proudfoot,¹ Donna Marie L. Alvinio,¹ Seanna Vine,¹ Marylyn M. Addo,¹ Todd M. Allen,¹ Marcus Altfeld,¹ Matthew R. Henn,⁴ Sylvie Le Gall,¹ Hendrik Streeck,¹ Bruce D. Walker^{1,9,†} (chair)
AIDS Clinical Trials Group: David W. Haas,¹⁷ Daniel R. Kuritzkes,² Gregory K. Robbins,¹⁸ Robert W. Shafer,¹⁹ Roy M. Gulick,²⁰ Cecilia M. Shikuma,²¹ Richard Haubrich,²² Sharon Riddler,²³ Paul E. Sax,² Eric S. Daar,²⁴ Heather J. Ribaud,²⁵
HIV controllers referral team: Brian Agan,²⁶ Shanu Agarwal,²⁷ Richard L. Ahern,¹⁸ Brady L. Allen,²⁸ Sheryl Altidor,²⁹ Eric L. Altschuler,³⁰ Sujata Ambardar,³¹ Kathryn Anastos,³² Ben Anderson,³³ Val Anderson,³⁴ Ushan Andrady,³⁴ Diana Antoniskis,³⁵ David Bangsberg,^{1,18} Daniel Barbaro,³⁶ William Barrie,³⁷ J. Bartczak,³⁸ Simon Barton,³⁹ Patricia Basden,⁴⁰ Nesli Baszoglu,¹⁸ Suzanne Bazner,¹ Nicholaos C. Bellos,⁴¹ Anne M. Benson,⁴⁰ Judith Berger,⁴² Nicole F. Bernard,⁴³ Annette M. Bernard,⁴⁴ Christopher Birch,¹ Stanley J. Bodner,⁴⁵ Robert K. Bolan,⁴⁶ Emilie T. Boudreaux,⁴⁷ Meg Bradley,¹ James F. Braun,⁴⁸ Jon E. Brndjar,⁴⁹ Stephen J. Brown,⁵⁰ Katherine Brown,⁵¹ Sheldon T. Brown,⁵² Jeddiah Burack,⁵³ Larry M. Bush,⁵⁴ Virginia Cafaro,⁵⁵ Omololaj Campbell,¹⁸ John Campbell,⁵⁶ Robert H. Carlson,⁵⁷ J. Kevin Carmichael,⁵⁸ Kathleen K. Casey,⁵⁹ Chris Cavacuiti,⁶⁰ Gregory Celestin,⁶¹ Steven T. Chambers,⁶² Nancy Chez,⁶³ Lisa M. Church,⁶⁴ Paul J. Cimocho,⁶⁵ Daniel Cohen,⁶⁶ Lillian E. Cohn,⁶⁷ Brian Conway,⁶⁸ David A. Cooper,⁶⁹ Brian Cornelison,⁶⁰ David T. Cox,⁷⁰ Michael V. Cristofano,⁷¹ George Cuchural Jr.,⁷² Julie L. Czartoski,⁷³ Joseph M. Dahman,⁷⁴ Jennifer S. Daly,⁷⁵ Benjamin T. Davis,¹⁸ Kristine Davis,⁷⁶ Sheila M. Davod,¹⁸ Steven G. Deeks,¹⁶ Edwin DeJesus,⁷⁷ Craig A. Dietz,⁷⁸ Eleanor Dunham,⁶⁴ Michael E. Dunn,⁷⁹ Todd B. Ellerin,⁸⁰ Joseph J. Eron,⁸¹ John J.W. Fangman,⁸² Claire E. Farel,² Helen Ferlazzo,⁸³ Sarah Fidler,⁸⁴ Anita Fleenor-Ford,⁸⁵ Renee Frankel,⁸⁶ Kenneth A. Freedberg,¹⁸ Neel K. French,⁸⁷ Jonathan D. Fuchs,⁸⁸ Jon D. Fuller,⁸⁹ Jonna Gaberman,⁹⁰ Joel E. Gallant,⁹¹ Rajesh T. Gandhi,¹⁸ Efrain Garcia,⁹² Donald Garmon,⁹³ Joseph C. Gathe Jr.,⁹⁴ Cyril R. Gaultier,⁹⁵ Wondwoosen Gebre,⁹⁶ Frank D. Gilman,⁹⁷ Ian Gilson,⁹⁸ Paul A. Goepfert,⁹⁹ Michael S. Gottlieb,¹⁰⁰ Claudia Goulet,¹⁰¹ Richard K. Groger,¹⁰² T. Douglas Gurley,¹⁰³ Stuart Haber,¹⁰⁴ Robin Hardwicke,¹⁰⁵ W. David Hardy,²⁴ P. Richard Harrigan,¹⁰⁶ Trevor N. Hawkins,¹⁰⁷ Sonya Heath,⁹⁹ Frederick M. Hecht,¹⁶ W. Keith Henry,¹⁰⁸ Melissa Hladek,¹⁰⁹ Robert P. Hoffman,¹¹⁰ James M. Horton,¹¹¹ Ricky K. Hsu,¹¹² Gregory D. Huhn,¹¹³ Peter Hunt,¹⁶ Mark J. Hupert,¹⁶ Mark L. Illeman,¹¹⁴ Hans Jaeger,¹¹⁵ Robert M. Jellinger,¹¹⁶ Mina John,¹¹⁷ Jennifer A. Johnson,¹¹⁸ Theresa A. McGhee,¹⁴² Heather Johnson,³⁶ Kay Johnson,¹¹⁸ Jennifer Joly,⁶⁴ Wilbert C. Jordan,¹¹⁹ Carol A. Kauffman,¹²⁰ Homayoun Khanlou,¹²¹ Robert K. Killian,¹²² Arthur Y. Kim,¹²³ David D. Kim,¹²³ Clifford A. Kinter,¹²⁴ Jeffrey T. Kirchner,¹²⁵ Laura Kogelman,¹²⁶ Erna Milunka Kojic,¹²⁷ P. Todd Korhuis,¹²⁸ Wayne Korusu,⁹⁷ Douglas S. Kwon,¹ Melissa LaMar,⁹³ Harry Lampiris,¹⁶ Massimiliano Lanzafame,¹²⁹ Michael M. Lederman,¹³⁰ David M. Lee,²⁸ Jean M.L. Lee,⁷³ Marah J. Lee,¹³¹ Edward T.Y. Lee,¹³² Janice Lemoine,¹³³ Jay A. Levy,¹⁶ Josep M. Llibre,¹³⁴ Michael A. Liguori,¹¹² Susan J. Little,²² Anne Y. Liu,² Alvaro J. Lopez,¹³⁵ Mono R. Loutfy,¹³⁶ Dawn Loy,¹³⁷ Debbie Y. Mohammed,³⁰ Alan Man,³⁵ Michael K. Mansour,¹⁸ Vincent C. Marconi,¹³⁸ Martin Markowitz,¹³⁹ Rui Marques,¹⁴⁰ Jeffrey N. Martin,¹⁶ Harold L. Martin Jr.,¹⁴¹ Kenneth Hugh Mayer,⁶⁶ M. Juliana McElrath,⁷³ Theresa A. McGhee,¹⁴² Barbara H. McGovern,¹²⁶ Katherine McGowan,² Dawn McIntyre,⁵⁹ Gavin X. McLeod,¹⁴³ Prema Menezes,⁸¹ Greg Mesa,¹⁴⁴ Craig E. Metroka,²⁹ Dirk Meyer-Olson,¹⁴⁵ Andy O. Miller,¹⁴⁶ Kate Montgomery,¹⁴⁷ Karam C. Mounzer,¹⁴⁸ Ellen H. Nagami,¹ Iris Nagin,¹⁴⁹ Ronald G. Nahass,¹⁵⁰ Margaret O. Nelson,¹⁸ Craig

Nielsen,¹⁵¹ David L. Norene,¹⁵² David H. O'Connor,¹⁵³ Bisola O. Ojikutu,¹⁸ Jason Okulicz,¹⁵⁴ Olakunle O. Oladehin,¹⁸ Edward C. Oldfield III,¹⁵⁵ Susan A. Olander,¹⁵⁶ Mario Ostrowski,¹⁵⁶ William F. Owen Jr.,¹⁵⁷ Eunice Pae,¹ Jeffrey Parsonnet,¹⁵⁸ Andrew M. Pavlatos,¹⁵⁹ Aaron M. Perlmutter,¹⁶⁰ Michael N. Pierce,²¹⁸ Jonathan M. Pincus,¹⁶¹ Leandro Pisani,¹⁶² Lawrence Jay Price,¹⁶³ Laurie Proia,¹⁶⁴ Richard C. Prokesch,¹³⁷ Heather Calderon Pujet,¹⁶⁵ Moti Ramgopal,¹⁶⁶ Almas Rathod,¹ Michael Rausch,¹⁶⁷ J. Ravishankar,¹⁶⁸ Frank S. Rhame,¹⁶⁹ Constance Shamyarira Richards,¹⁷⁰ Douglas D. Richmond,²² Gregory K. Robbins,¹⁸ Berta Rodes,¹⁷¹ Milagros Rodriguez,¹⁶² Richard C. Ross II,¹⁷² Eric S. Rosenberg,¹⁸ Daniel Rosenthal,¹⁷³ Polly E. Ross,¹⁷⁴ David S. Rubin,¹⁷⁵ Elasee Rumbaugh,³⁵ Luis Saenz,¹⁶² Michelle R. Salvaggio,¹⁷⁶ William C. Sanchez,¹⁷⁷ Veerav M. Sanjana,¹⁷⁸ Steven Santiago,¹⁶² Wolfgang Schmidt,¹⁷⁹ Hanneke Schuitemaker,¹⁸⁰ Philip M. Sestak,¹⁸¹ Peter Shalit,¹⁸² William Shay,¹⁰⁴ Vivian N. Shirvani,¹⁸³ Vanessa I. Silebi,¹⁸⁴ James M. Sizemore Jr.,¹⁸⁵ Paul R. Skolnik,⁸⁹ Marcia Sokol-Anderson,¹⁸⁶ James M. Sosman,¹⁵³ Paul Stabile,¹⁸⁷ Jack T. Stapleton,¹⁸⁸ Sheree Starrett,¹⁸⁹ Francine Stein,⁸³ Hans-Jürgen Stellbrink,¹⁹⁰ F. Lisa Sterman,¹⁹¹ Valerie E. Stone,¹⁸ David R. Stone,¹⁹² Giuseppe Tambussi,¹⁹³ Randy A. Taplitz,²² Ellen M. Tedaldi,¹⁹⁴ Amalio Telenti,⁶ William Theisen,² Richard Torres,¹⁹⁵ Lorraine Tosiello,¹⁹⁶ Cecile Tremblay,¹⁹⁷ Marc A. Tribble,¹⁹⁸ Phuong D. Trinh,¹⁹⁹ Alice Tsao,¹ Peggy Ueda,¹ Anthony Vaccaro,²⁰⁰ Emilia Valadas,²⁰¹ Thanes J. Vanig,²⁰² Isabel Vecino,²⁰³ Vilma M. Vega,¹³⁷ Wenoah Veikley,¹⁰⁷ Barbara H. Wade,²⁰⁴ Charles Walworth,⁶⁵ Chingchai Wanidworanun,²⁰⁵ Douglas J. Ward,²⁰⁶ Daniel A. Warner,²⁰⁷ Robert D. Weber,²⁰⁸ Duncan Webster,²⁰⁹ Steve Weis,²⁰³ David A. Wheeler,²¹⁰ David J. White,²¹¹ Ed Wilkins,²¹² Alan Winston,⁸⁴ Clifford G. Wlodaver,²¹³ Angélique van't Wout,¹⁸⁰ David P. Wright,²¹⁴ Otto O. Yang,²⁴ David L. Yurdin,²¹⁵ Brandon W. Zabukovic,²¹⁶ Kimon C. Zachary,¹⁸ Beth Zeeman,¹ Meng Zhao²¹⁷

¹Ragon Institute of Massachusetts General Hospital, Massachusetts Institute of Technology (MIT) and Harvard, Boston, MA, USA. ²Department of Medicine, Division of Infectious Disease, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA. ³Harvard-MIT Division of Health Sciences and Technology, Boston, MA, USA. ⁴Broad Institute of Harvard and MIT, Cambridge, MA, USA. ⁵Department of Medicine, Division of Genetics, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA. ⁶Institute of Microbiology, University of Lausanne, Lausanne, Switzerland. ⁷Department of Medical Genetics, Division of Biomedical Genetics, University Medical Center Utrecht, Netherlands. ⁸Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Netherlands. ⁹Howard Hughes Medical Institute, Chevy Chase, MD, USA. ¹⁰Department of Medicine, Center for Human Genetic Research, MGH, Harvard Medical School, Boston, MA, USA. ¹¹Cancer and Inflammation Program, Laboratory of Experimental Immunology, SAIC-Frederick, NCI-Frederick, Frederick, MD, USA. ¹²Microsoft Research, Redmond, WA, USA. ¹³Microsoft Research, Los Angeles, CA, USA. ¹⁴Genentech, South San Francisco, CA, USA. ¹⁵Department of Medicine, Division of Rheumatology, Immunology and Allergy, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA. ¹⁶University of California San Francisco, San Francisco, CA, USA. ¹⁷Vanderbilt University School of Medicine, Nashville, TN, USA. ¹⁸MGH, Harvard Medical School, Boston, MA, USA. ¹⁹Stanford University, Palo Alto, CA, USA. ²⁰Weill Medical College of Cornell University, New York, NY, USA. ²¹Hawaii Center for AIDS, John A. Burns School of Medicine, University of Hawaii, Honolulu, HI, USA. ²²University of California San Diego, San Diego, CA, USA. ²³University of Pittsburgh, Pittsburgh, PA, USA. ²⁴University of California Los Angeles, Los Angeles, CA, USA. ²⁵Department of Biostatistics, Harvard School of Public Health, Boston, MA, USA. ²⁶Infectious Disease Clinical Research Program, Uniformed Services University of the Health Sciences, Bethesda, MD, USA. ²⁷Summa Health System, Akron, OH, USA. ²⁸Uptown Physicians Group, Dallas, TX, USA. ²⁹St. Luke's Roosevelt Hospital, New York, NY, USA. ³⁰New Jersey Medical School, University Hospital, Newark, NJ, USA. ³¹Infectious Disease Physicians, Annandale, VA, USA. ³²Montefiore Medical Center, Albert Einstein College of Medicine, Bronx, NY, USA. ³³St. Leonards Medical Centre, St. Leonards, Australia. ³⁴Syby Gwynedd Hospital, Gwynedd, UK. ³⁵Kaiser Permanente, Portland, OR, USA. ³⁶Tarrant County Infectious Disease Associates, Fort Worth, TX, USA. ³⁷Private Practice of William Barrie,

- Toronto, Canada. ³⁸Rowan Tree Medical, Wilton Manors, FL, USA. ³⁹Chelsea and Westminster Hospital, St. Stephen's Centre, London, UK. ⁴⁰Beaver Street Family Practice, Flagstaff, AZ, USA. ⁴¹Southwest Infectious Disease Associates, Dallas, TX, USA. ⁴²St. Barnabas Hospital, Bronx, NY, USA. ⁴³Research Institute, McGill University Health Centre, Montreal General Hospital, Montreal, Canada. ⁴⁴Thacker, Thompson and Bernard, Atlanta, GA, USA. ⁴⁵Vanderbilt University School of Medicine, Hermitage, TN, USA. ⁴⁶LA Gay and Lesbian Center, Los Angeles, CA, USA. ⁴⁷Louisiana State University Health Sciences Center, University Medical Center East Clinic, Lafayette, LA, USA. ⁴⁸Physicians' Research Network, Callen-Lorde Community Health Center, New York, NY, USA. ⁴⁹Brndjar Medical Associates, Allentown, PA, USA. ⁵⁰AIDS Research Alliance, Los Angeles, CA, USA. ⁵¹David Powell Community Health Center, Austin, TX, USA. ⁵²James J. Peters VA Medical Center, Bronx, NY, USA. ⁵³Sunrise Medical Group, Brooklyn, NY, USA. ⁵⁴University of Miami-Miller School of Medicine, Lake Worth, FL, USA. ⁵⁵WellSpring Medical Group, San Francisco, CA, USA. ⁵⁶Moses Cone Health System, Greensboro, NC, USA. ⁵⁷HealthPartners Infectious Disease, St Paul, MN, USA. ⁵⁸El Rio Special Immunology Associates, Tuscon, AZ, USA. ⁵⁹Jersey Shore University Medical Center, Neptune, NJ, USA. ⁶⁰St. Michaels Hospital, Toronto, Canada. ⁶¹The Brooklyn Hospital Center, PATH Center, Brooklyn, NY, USA. ⁶²University of Otago, Christchurch, New Zealand. ⁶³H.E.L.P./Project Samaritan, Bronx, NY, USA. ⁶⁴David E. Rogers Center for HIV/AIDS Care, Southampton, NY, USA. ⁶⁵Center for Special Immunology, Fountain Valley, CA, USA. ⁶⁶Fenway Community Health, Boston, MA, USA. ⁶⁷9th Street Internal Medicine Associates, Philadelphia, PA, USA. ⁶⁸University of British Columbia, Vancouver, Canada. ⁶⁹National Centre in HIV Epidemiology and Clinical Research, Sydney, Australia. ⁷⁰Metro Infectious Disease Consultants, Indianapolis, IN, USA. ⁷¹John H. Stroger Hospital of Cook County, Chicago, IL, USA. ⁷²New England Quality Care Alliance, Braintree, MA, USA. ⁷³Fred Hutchinson Cancer Research Center, Seattle, WA, USA. ⁷⁴Desert AIDS Project, Palm Springs, CA, USA. ⁷⁵University of Massachusetts Memorial Medical Center, Worcester, MA, USA. ⁷⁶University of Iowa Hospitals and Clinics, Iowa City, IA, USA. ⁷⁷Orlando Immunology Center, Orlando, FL, USA. ⁷⁸The Kansas City Free Health Clinic, Kansas City, MO, USA. ⁷⁹Private Practice of Michael E. Dunn, M.D., Tampa, FL, USA. ⁸⁰South Shore Hospital, Weymouth, MA, USA. ⁸¹University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. ⁸²AIDS Resource Center of Wisconsin, Milwaukee, WI, USA. ⁸³Visiting Nurse Association of Central New Jersey, Community Health Center, Asbury Park, NJ, USA. ⁸⁴Imperial College, London, UK. ⁸⁵Heartland Clinic, Paducah, KY, USA. ⁸⁶Morristown Memorial Hospital, Morristown, NJ, USA. ⁸⁷Private Practice of Neel K. French, M.D., Chicago, IL, USA. ⁸⁸San Francisco Department of Public Health, San Francisco, CA, USA. ⁸⁹Boston University Medical Center, Boston, MA, USA. ⁹⁰Baystate Medical Center, Springfield, MA, USA. ⁹¹Johns Hopkins University School of Medicine, Baltimore, MD, USA. ⁹²Private Practice of Efrain Garcia, M.D., Miami, FL, USA. ⁹³The Rockefeller University, New York, NY, USA. ⁹⁴Private Practice of Joseph C. Gathe Jr., M.D., Houston, TX, USA. ⁹⁵Tower Infectious Disease, Los Angeles, CA, USA. ⁹⁶Nassau University Medical Center, East Meadow, NY, USA. ⁹⁷Sharp Rees Stealy Medical Center, San Diego, CA, USA. ⁹⁸Medical College of Wisconsin, Milwaukee, WI, USA. ⁹⁹University of Alabama, Birmingham, Birmingham, AL, USA. ¹⁰⁰Synergy Hematology and Oncology, Los Angeles, CA, USA. ¹⁰¹University of Utah, Salt Lake City, UT, USA. ¹⁰²South Dayton Acute Care Consultants, Dayton, OH, USA. ¹⁰³T. Douglas Gurley, M.D., Atlanta, GA, USA. ¹⁰⁴St. Vincent's Hospital, New York, NY, USA. ¹⁰⁵University of Texas Health Science Center, Houston, TX, USA. ¹⁰⁶British Columbia Centre for Excellence in HIV/AIDS, Vancouver, Canada. ¹⁰⁷Southwest C.A.R.E. Center, Santa Fe, NM, USA. ¹⁰⁸Hennepin County Medical Center, Minneapolis, MN, USA. ¹⁰⁹The Catholic University of America, School of Nursing, Washington, DC, USA. ¹¹⁰Mercy Medical Center, Springfield, MA, USA. ¹¹¹CMC Myers Park Medical Center, Charlotte, NC, USA. ¹¹²New York University Medical Center, New York, NY, USA. ¹¹³The Ruth M. Rothshon Care Center, Chicago, IL, USA. ¹¹⁴Feldman Medical Group, San Francisco, CA, USA. ¹¹⁵HIV Research and Clinical Care Centre, Munich, Germany. ¹¹⁶Albany Medical College, Albany, NY, USA. ¹¹⁷Murdoch University, Murdoch, Australia. ¹¹⁸University of Cincinnati, Cincinnati, OH, USA. ¹¹⁹OASIS Clinic, Los Angeles, CA, USA. ¹²⁰VA Ann Arbor Healthcare System, Ann Arbor, MI, USA. ¹²¹AIDS Healthcare Foundation, Los Angeles, CA, USA. ¹²²Capitol Hill Medical, Seattle, WA, USA. ¹²³Astor Medical Group, New York, NY, USA. ¹²⁴The Kinder Medical Group, Miami, FL, USA. ¹²⁵Lancaster General Hospital, Lancaster, PA, USA. ¹²⁶Tufts Medical Center, Boston, MA, USA. ¹²⁷Alpert Medical School of Brown University, Providence, RI, USA. ¹²⁸Oregon Health and Science University, Portland, OR, USA. ¹²⁹G.B. Rossi Hospital, Verona, Italy. ¹³⁰Case Western Reserve University, Cleveland, OH, USA. ¹³¹LifeWay, Fort Lauderdale, FL, USA. ¹³²Saint Claire Medical Associate, Toronto, Canada. ¹³³Greater Lawrence Family Health Center, Lawrence, MA, USA. ¹³⁴Hospital Universitari Germans Trias i Pujol, Barcelona, Spain. ¹³⁵Infectious Disease Consultants, Tucker, GA, USA. ¹³⁶University of Toronto, Toronto, Canada. ¹³⁷Infectious Disease Associates, Sarasota, FL, USA. ¹³⁸Emory University, School of Medicine, Atlanta, GA, USA. ¹³⁹Aaron Diamond AIDS Research Center, Rockefeller University, New York, NY, USA. ¹⁴⁰Deruico Doencas Infecciosas, Porto, Portugal. ¹⁴¹Park Nicollet Clinic, St. Louis, MN, USA. ¹⁴²Absolute Care, Atlanta, GA, USA. ¹⁴³College of Physicians and Surgeons, Columbia University, New York, NY, USA. ¹⁴⁴Highland Medical Associates, Hendersonville, NC, USA. ¹⁴⁵Medizinische Hochschule, Abteilung Klinische Immunologie, Hannover, Germany. ¹⁴⁶Hospital for Special Surgery, New York, NY, USA. ¹⁴⁷Family Practice Specialists, Phoenix, AZ, USA. ¹⁴⁸Philadelphia FIGHT, Philadelphia, PA, USA. ¹⁴⁹Lower East Side Service Center, New York, NY, USA. ¹⁵⁰Infectious Diseases Care, Hillsborough, NJ, USA. ¹⁵¹University of Colorado, Denver, Aurora, CO, USA. ¹⁵²Sutter Medical Group, Sacramento, CA, USA. ¹⁵³University of Wisconsin in Madison, Madison, WI, USA. ¹⁵⁴Brooke Army Medical Center, San Antonio, TX, USA. ¹⁵⁵Eastern Virginia Medical School, Norfolk, VA, USA. ¹⁵⁶Columbia University Medical Center, New York, NY, USA. ¹⁵⁷CA Pacific Medical Center, San Francisco, CA, USA. ¹⁵⁸Dartmouth-Hitchcock Medical Center, Lebanon, NH, USA. ¹⁵⁹St. Joseph Hospital, Chicago, IL, USA. ¹⁶⁰Aaron M. Perlmutter, M.D., Beverly Hills, CA, USA. ¹⁶¹Codman Square Health Center, Dorchester, MA, USA. ¹⁶²CARE Resource, Miami, FL, USA. ¹⁶³Castro-Mission Health Center, San Francisco, CA, USA. ¹⁶⁴Rush Medical College, Chicago, IL, USA. ¹⁶⁵Boulder Community Hospital, Boulder, CO, USA. ¹⁶⁶Midway Immunology and Research Center, Fort Pierce, FL, USA. ¹⁶⁷Aerztezentrum Nollendorferplatz, Berlin, Germany. ¹⁶⁸State University of New York Downstate Medical Center, Brooklyn, NY, USA. ¹⁶⁹Clinic 42, Minneapolis, MN, USA. ¹⁷⁰King Edward Memorial Hospital, Paget, Bermuda. ¹⁷¹Fundacion para la Investigacion Biomedica del Hospital Carlos III, Madrid, Spain. ¹⁷²Summit Medical Group, Knoxville, TN, USA. ¹⁷³Medical Consultants of South Florida, Coral Springs, FL, USA. ¹⁷⁴Western North Carolina Community Health Services, Asheville, NC, USA. ¹⁷⁵New York Hospital Medical Center of Queens, Flushing, NY, USA. ¹⁷⁶University of Oklahoma Health Sciences Center, Oklahoma City, OK, USA. ¹⁷⁷Georgetown University Medical Center, Washington, DC, USA. ¹⁷⁸Village Care Health Center, New York, NY, USA. ¹⁷⁹Aerzteforum Seestrassen, Berlin, Germany. ¹⁸⁰Academic Medical Center Amsterdam, Amsterdam, Netherlands. ¹⁸¹St. Paul's Hospital, Vancouver, Canada. ¹⁸²Swedish Medical Center, Seattle, WA, USA. ¹⁸³Cedars-Sinai Medical Center, Los Angeles, CA, USA. ¹⁸⁴Mercy Hospital, Miami, FL, USA. ¹⁸⁵University of Tennessee, Chattanooga, TN, USA. ¹⁸⁶St. Louis University, St. Louis, MO, USA. ¹⁸⁷William F. Ryan Community Health Center, New York, NY, USA. ¹⁸⁸The University of Iowa, Iowa City, IA, USA. ¹⁸⁹Rivington House and Village Care, New York, NY, USA. ¹⁹⁰Infektionsmedizinisches Centrum Hamburg, Hamburg, Germany. ¹⁹¹California Pacific Medical Center, San Francisco, CA, USA. ¹⁹²Lemuel Shattuck Hospital, Boston, MA, USA. ¹⁹³Fondazione San Raffaele Del Monte Tabor, Milan, Italy. ¹⁹⁴Temple University School of Medicine, Philadelphia, PA, USA. ¹⁹⁵Yale University School of Medicine, Bridgeport, CT, USA. ¹⁹⁶Jersey City Medical Center, Jersey City, NJ, USA. ¹⁹⁷University of Montreal, Montreal, Canada. ¹⁹⁸Baylor University Medical Center, Dallas, TX, USA. ¹⁹⁹Montgomery Infectious Disease Associates, Silver Spring, MD, USA. ²⁰⁰Northwestern University, Chicago, IL, USA. ²⁰¹Hospital de Santa Maria, Faculdade de Medicina de Lisboa, Lisbon, Portugal. ²⁰²Spectrum Medical Group, Phoenix, AZ, USA. ²⁰³University of North Texas Health Science Center, Fort Worth, TX, USA. ²⁰⁴Infectious Diseases Associates of Northwest Florida, Pensacola, FL, USA. ²⁰⁵Private Practice of Chingchai Wanidworanun, M.D., Arlington, VA, USA. ²⁰⁶Dupont Circle Physician's Group, Washington, DC, USA. ²⁰⁷Consultative Medicine, Daytona Beach, FL, USA. ²⁰⁸Infectious Disease Specialists, Colorado Springs, CO, USA. ²⁰⁹Saint John Regional Hospital, Saint John, Canada. ²¹⁰Clinical Alliance for Research and Education—Infectious Diseases, Annandale, VA, USA. ²¹¹Hawthorn House, Birmingham Heartlands Hospital, Birmingham, UK. ²¹²North Manchester General Hospital, Manchester, UK. ²¹³Private Practice of Clifford Wlodaver, Midwest City, OK, USA. ²¹⁴Central Texas Clinical Research, Austin, TX, USA. ²¹⁵Primary Health Care, Des Moines, IA, USA. ²¹⁶Memorial Neighborhood Health Center Central Clinic, South Bend, IN, USA. ²¹⁷United Health Services Hospitals, Binghamton, NY, USA. ²¹⁸All Med and Rehabilitation of New York, Bronx, NY, USA.

†These authors contributed equally to this work.

Supporting Online Material

www.sciencemag.org/cgi/content/full/science.1195271/DC1
Materials and Methods
Figs. S1 to S12
Tables S1 to S13
References

19 July 2010; accepted 26 October 2010
Published online 4 November 2010;
10.1126/science.1195271