Published in final edited form as:

Int J Immunogenet. 2016 December ; 43(6): 369–375. doi:10.1111/iji.12292.

HLA Class I and II Imputation in a Multiracial Population

Mark H. Kuniholm¹, Xianhong Xie¹, Kathryn Anastos^{1,3}, Xiaonan Xue¹, Laura Reimers², Audrey L. French⁴, Stephen J. Gange⁵, Seble G. Kassaye⁶, Andrea Kovacs⁷, Tao Wang¹, Bradley E. Aouizerat^{8,9,*}, and Howard D. Strickler^{1,*}

¹Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, NY

²Department of Obstetrics & Gynecology and Women's Health, Albert Einstein College of Medicine, Bronx, NY

³Department of Medicine, Montefiore Medical Center, Bronx, NY

⁴Ruth M. Rothstein CORE Center, Stroger Hospital of Cook County, Chicago, IL

⁵Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD

⁶Department of Medicine, Georgetown University Medical Center, Washington, DC

⁷Department of Pediatrics, University of Southern California, Los Angeles, CA

⁸Bluestone Center for Clinical Research, New York University, New York, NY

⁹Department of Oral and Maxillofacial Surgery, New York University, New York, NY

SUMMARY

Human leukocyte antigen (HLA) genes play a central role in response to pathogens and in autoimmunity. Research to understand effects of HLA genes on health has been limited because HLA genotyping protocols are labor intensive and expensive. Recently, algorithms to impute HLA genotype data using genome-wide association study (GWAS) data have been published. However, imputation accuracy for most of these algorithms was based primarily on training datasets of European ancestry individuals. We considered performance of two HLA-dedicated imputation algorithms - SNP2HLA and HIBAG - in a multiracial population of n=1,587 women with HLA genotyping data by gold-standard methods. We first compared accuracy – defined as the percentage of correctly predicted alleles – of HLA-B and HLA-C imputation using SNP2HLA and HIBAG using a breakdown of the dataset into an 80% training group and a 20% testing group. Estimates of accuracy for HIBAG were either the same or better than those for SNP2HLA. We then conducted a more thorough test of HIBAG imputation accuracy using 5 independent 10-fold cross-validation procedures with delineation of ancestry groups using ancestry informative markers. Overall accuracy for HIBAG was 89%. Accuracy by HLA gene was: 93% for HLA-A,

CORRESPONDING AUTHOR: Mark H. Kuniholm, PhD, Department of Epidemiology & Population Health, Albert Einstein College of Medicine, Belfer Building, Room 1308C, 1300 Morris Park Ave. Bronx, NY 10461. Phone: (718) 430-4129. Fax: (718) 430-8780. mark.kuniholm@einstein.yu.edu.

*Both senior authors contributed equally

CONFLICTS OF INTEREST AND SOURCE OF FUNDING

All authors, no conflicts of interest declared.

84% for HLA-B, 94% for HLA-C, 83% for HLA-DQA1, 91% for HLA-DQB1, and 88% for HLA-DRB1. Accuracy was highest in the African ancestry group (the largest group) and lowest in the Hispanic group (the smallest group). Despite suboptimal imputation accuracy for some HLA gene/ancestry group combinations, the HIBAG algorithm has the advantage of providing posterior estimates of accuracy which enable the investigator to analyze subsets of the population with high predicted (e.g., >95%) imputation accuracy.

Keywords

human leukocyte antigen; HLA; imputation; HIV; multiracial; African American

INTRODUCTION

Classical human leukocyte antigen (HLA) class I and II molecules play a central role in host response to pathogens and in autoimmunity. The genes that encode classical HLA molecules are located on chromosome 6 and are among the most polymorphic in the human genome – an important factor in the ability of the immune system to respond to a wide variety of pathogens. HLA genes are clinically important and significant associations in the HLA region have been reported for a diverse range of infectious diseases including HIV-1 (Pereyra et al. 2010), hepatitis B and C (Duggal et al. 2013; Kamatani et al. 2009). HLA genes are also of interest for their role in autoimmune conditions, including Graves' disease (Chu et al. 2011), pediatric asthma (Noguchi et al. 2011), multiple sclerosis (Hafler et al. 2007), rheumatoid arthritis (Wellcome Trust Case Control Consortium. 2007), and type 1 diabetes (Wellcome Trust Case Control Consortium. 2007).

Despite increasing appreciation of the clinical importance of HLA gene variation, relatively few laboratories have the expertise to conduct high-resolution HLA genotyping and HLA genotyping protocols (e.g., by the polymerase chain reaction-sequence specific oligonucleotide (PCR-SSO) probe method) are labor intensive and expensive (Trowsdale & Knight 2013; Jia et al. 2013). However, many cohorts and other clinical studies have genome-wide association study (GWAS) data. Given this availability, a natural question is whether these GWAS data can be used to impute or otherwise infer classical HLA allele data. Until recently, the answer was generally no. A few HLA alleles are in almost perfect linkage disequilibrium (LD) with specific single nucleotide polymorphisms (SNPs - e.g., rs2395029 is in high LD with HLA-B*57:01 in European ancestry (EA) populations (Colombo et al. 2008)) but there was no method to infer complete classical HLA data.

In 2008 Leslie et al. published the first algorithm to impute complete classical HLA allele data from GWAS data (Leslie et al. 2008). This seminal algorithm is based on the principle of identity by descent (IBD) and can be used to impute HLA alleles at either two-digit or four-digit resolution. Four-digit resolution is preferred for most research applications because these [four-digit] data correspond to unique HLA molecules as compared to two-digit groups of similar HLA molecules. The algorithm proposed by Leslie et al. requires a large training dataset in which both classical HLA alleles and SNPs have been typed. Using HapMap data (de Bakker et al. 2006) for training and a validation dataset of individuals of

United Kingdom origin, Leslie et al. reported an HLA imputation accuracy of >90% at four-digit resolution for most class I and II HLA genes (Leslie et al. 2008).

Since the 2008 report by Leslie et al. new algorithms to impute classical HLA data from GWAS data have been proposed (Setty et al. 2011; Xie et al. 2010; Zheng et al. 2014) along with improvements to the original Leslie et al. algorithm (Dilthey et al. 2013; Dilthey et al. 2011). Other groups used general genotype imputation methods such as BEAGLE (Browning & Browning 2009) to impute HLA data from GWAS data (Pereyra et al. 2010; Raychaudhuri et al. 2012; Jia et al. 2013). Imputation accuracy in most of these studies was based on training datasets of EA individuals and/or racially homogeneous HapMap populations (Xie et al. 2010; Setty et al. 2011; Dilthey et al. 2011; Jia et al. 2013). However, two studies reported validation of their imputation algorithms using training datasets from diverse (though mainly European or EA) populations (Dilthey et al. 2013; Zheng et al. 2014). However, the use of heterogeneous populations is increasingly common in genetic association studies that include an evaluation of HLA regions and the performance of the HLA imputation methods is unknown.

The goal of the described analyses is to determine how selected HLA imputation methods perform when applied to majority multiracial (admixed) populations. We present our experience conducting HLA imputation in the Women's Interagency HIV Study (WIHS), a cohort study of HIV-positive and HIV-negative women of diverse racial/ethnic backgrounds (i.e., 57% African American, 25% Hispanic, 14% White, 1% Asian, 1% Native American, 2% Other).

METHODS

Study Population

Characteristics of the WIHS study population have been described previously (Bacon et al. 2005). Briefly, HIV-positive and HIV-negative women were recruited from similar risk settings at six US sites (Bronx, Brooklyn, Washington D.C., Chicago, San Francisco and Los Angeles) during 1994-1995, 2001-2002, and 2011-2012. In 2013 the WIHS expanded to include new sites in the Southern US. The procedures outlined in this manuscript are restricted to women enrolled at the original six sites because only these women have HLA and GWAS data.

HLA Typing

Four-digit resolution HLA data were typed in n=1,587 WIHS women, as described previously (Kuniholm et al. 2010; Kuniholm et al. 2011). Protocols for HLA genotyping have been standardized by the International Histocompatibility Working Group (<http://www.ihwg.org>). Briefly, genomic DNA was prepared from subjects' lymphoblastoid B cell lines or from peripheral blood lymphocytes. HLA class I genes (i.e., HLA-A, -B, and -C) were amplified using locus-specific PCR primers flanking exons 2 and 3, the polymorphic segments of the class I genes. The 1 kb PCR products were blotted on nylon membranes and hybridized with a panel of sequence-specific oligonucleotide (SSO) probes. The HLA alleles were assigned by the hybridization patterns of the SSO probes, according to known HLA

sequences. Ambiguous SSO probing was resolved by sequencing analysis, as previously described (Gao et al. 2001). HLA class II typing was conducted using high resolution SSO typing for HLA-DQA, -DQB, and -DRB1 loci, using the polymorphic exon 2. DRB genotyping involved a two-step procedure. First, the broad serological DR types were determined using a pair of DRB generic primers and a panel of SSO probes. Allele-level DRB typing was then achieved by using group-specific primers to amplify the DRB alleles determined in the generic typing followed by SSO hybridization. For DQA and DQB, locus-specific PCR were performed followed by SSO hybridization.

Because HLA class I and HLA-DRB1 genes are currently the most relevant for studies of pathogens that are common in WIHS (i.e., HIV, human papillomavirus and hepatitis C virus) women were typed for these genes (i.e., HLA-A: n=1,567 – 46 unique alleles, HLA-B: n=1,537 – 99 unique alleles, HLA-C: n=1,497 – 37 unique alleles, HLA-DRB1: n=1,415 – 50 unique alleles). The majority of these women also had HLA-DQA1 (n=1,034 – 13 unique alleles) and HLA-DQB1 (n=1,036 – 18 unique alleles) typing.

SNP Typing

SNPs were genotyped using the Illumina HumanOmni2.5-quad beadchip (Illumina, San Diego) for all WIHS women enrolled in the 1994-1995 and 2001-2002 recruitment waves who provided informed consent for genetic testing (n=3,353). Excluded from the analytic dataset were SNPs with a genotype call rate of <95% and SNPs that failed our in-house quality control criteria. Specifically, SNPs with at least 2 discordant genotypes among greater than 20 duplicate samples were excluded.

SNPs flanking (i.e., approximately 100kb surrounding) the classical HLA genes typed in WIHS were used for imputation (i.e., 379 SNPs spanning HLA-A (chr6: 29,810,001-30,013,000), 725 SNPs spanning HLA-B/C (chr6: 31,136,000-31,424,000), 848 SNPs spanning HLA-DQ/DR (chr6: 32,307,000-32,734,000)). Mapping positions were defined using NCBI build 37, hg19. Ancestry informative marker (AIM) SNPs (n=185, Supplementary Table 1) from the Illumina beadchip were selected from across the genome using Helix Tree (Golden Helix, Bozeman, MT) to calculate principal components (PCs) of ancestry.

Selection of an HLA Imputation Algorithm

We considered imputation of classical HLA genotype data using two HLA-dedicated algorithms: SNP2HLA (Jia et al. 2013) (<https://www.broadinstitute.org/mpg/snp2hla/>) and HIBAG (Zheng et al. 2014) (<http://www.biostat.washington.edu/~bsweir/HIBAG/>). To compare whether SNP2HLA or HIBAG might be best suited to impute HLA alleles in the WIHS, we conducted an initial sub-study using SNPs from the HLA-B/C region (chr6: 31,136,000-31,424,000) and HLA-B/C alleles generated by PCR-SSO. We selected the HLA-B/C region for this sub-study because of its central importance in HIV/AIDS research. Specifically, we randomly split women with both HLA-B/C allele and HLA-B/C SNP data into two groups: an 80% training group and a 20% testing group (a commonly used split for training/testing studies). We also split subsets of women who self-reported as African American and White into 80%/20% training/testing groups. SNP2HLA was run using

default software parameters on the 80% training data which was then used to impute alleles in the testing data. Similarly, HIBAG was run using default software parameters except specifying the number of classifiers=100 (to maximize or stabilize prediction accuracies (Zheng et al. 2014)) on the 80% training data; this was then used to impute alleles in the testing data.

Imputation accuracy statistics of SNP2HLA and HIBAG for HLA-B and -C in the 20% testing data are shown in Table 1. Specifically, we considered two metrics of accuracy: (i) the percentage of total alleles correctly predicted, and (ii) the percentage of women with both alleles correctly predicted. The estimates of accuracy for HIBAG were either the same or better than those for SNP2HLA across both metrics and in each self-reported racial/ethnic group. For example, among women in all racial/ethnic groups HIBAG accurately predicted 87% of HLA-B alleles and 94% of HLA-C alleles. For SNP2HLA the estimates were 86% for HLA-B and 92% for HLA-C. HIBAG also predicted with much greater accuracy than SNP2HLA the percent of White women with both HLA-B alleles correctly predicted (74% vs. 26% for HIBAG and SNP2HLA, respectively).

We also considered percent agreement between SNP2HLA and HIBAG at HLA-B/-C loci. As seen in Supplementary Table 2, 83% of HLA-B alleles were correctly predicted by both algorithms, 10% of HLA-B alleles were incorrectly predicted by both algorithms and 7% of allele predictions were inconsistent. For HLA-C, 92% of alleles were correctly predicted by both algorithms, 5% of alleles were incorrectly predicted by both algorithms and 3% of allele predictions were inconsistent (Supplementary Table 2).

Finally, we considered whether accuracy of SNP2HLA and HIBAG predictions might vary by the specific allele being imputed. To do this, we calculated imputation accuracy for the 5 most common and 5 rarest HLA-B/-C alleles, reasoning that allele vs. allele comparisons are best made within strata of allele frequency. Allele vs. allele differences in imputation accuracy were minor for the most common HLA-B/-C alleles (e.g., imputation accuracy using HIBAG for the most common HLA-B alleles was 91%-100%). HIBAG accurately predicted the rarest HLA-B/-C alleles better than SNP2HLA (Supplementary Table 3). Based on the data shown in Table 1 and Supplementary Tables 2 and 3, we made the decision to use HIBAG rather than SNP2HLA for a more thorough test of HLA imputation accuracy in WIHS women.

Definition of Ancestry Groups

For our HIBAG HLA imputation procedures we delineated ancestry groups by visual inspection of PCs rather than by self-report. Specifically, we defined ancestry as European ancestry (EA): $PC1 < -0.017$ and $PC2 < 0.002$; African ancestry (AA): $PC1 > 0$; Hispanic ancestry (HA): $PC1 < -0.008$ and $PC2 > 0.002$; Other ancestry (OA): all other women (Figure 1).

Procedures for HLA Imputation

We used HIBAG with default software parameters except for specification of the number of classifiers=100 (to maximize or stabilize prediction accuracies (Zheng et al. 2014)) to impute missing HLA alleles for WIHS women. Subsets of WIHS women with HLA alleles

measured by PCR-SSO served as the training data for the HIBAG algorithm and HLA imputation was conducted within strata of genetically-defined ancestry.

To estimate the accuracy of HLA imputation by HIBAG, we used a 10-fold cross validation procedure. Specifically, within ancestry groups we randomly divided the training data (i.e., the subsets of WIHS women with gold-standard HLA data by PCR-SSO) into ten equally sized groups. Nine of these groups were used as training data while the tenth was used as the testing data to estimate accuracy. This process was repeated 10 times so that each of the 10 random groups was included in the training dataset nine times and served as the testing data one time. Overall accuracy was based on average results of 5 independent 10-fold cross-validation procedures. Cross-validation was computationally intensive with each 10-fold cross-validation procedure taking approximately two weeks to complete on a shared university computer cluster.

RESULTS

The accuracy of HLA imputation using the HIBAG algorithm was determined for 1,587 women with gold-standard HLA allele data (HLA-A: n=1,567; HLA-B: n=1,537; HLA-C: n=1,497; HLA-DRB1: n=1,415; HLA-DQA1: n=1,034; HLA-DQB1: n=1,036) and 1,952 SNPs in the HLA region (see Methods). Overall accuracy – the percentage of correctly predicted alleles based on 5 independent 10-fold cross-validation procedures – was 89% (Table 2). Accuracy by HLA gene was: 93% for HLA-A, 84% for HLA-B, 94% for HLA-C, 83% for HLA-DQA1, 91% for HLA-DQB1, and 88% for HLA-DRB1.

Imputation accuracy using the HIBAG algorithm varied between ancestry groups as defined by principal components of ancestry informative markers - the percentage of correctly predicted alleles was highest in the AA group (the largest group) and lowest in the HA group (the smallest group - Table 2). However, for the least polymorphic HLA genes considered in this study (HLA-DQA1 and HLA-DQB1), percentage of correctly predicted alleles was similar for EA, AA and HA ancestry groups (80%-83% and 90%-93% for HLA-DQA1 and HLA-DQB1, respectively). In contrast, for the most polymorphic HLA class I and II genes (HLA-B and HLA-DRB1), imputation accuracy varied widely between EA, AA and HA ancestry groups (66%-88% and 69%-91% for HLA-B and HLA-DRB1, respectively).

When imputation accuracy was calculated as the percentage of women for whom both alleles were correctly predicted, overall accuracy was 82% (Table 2). Wide variation in imputation accuracy between ancestry groups was also observed for HLA-B and HLA-DRB1 when imputation accuracy was calculated as the percentage of women for whom both alleles were correctly predicted (45%-78% and 48%-84% for HLA-B and HLA-DRB1, respectively).

DISCUSSION

In a large cohort of HIV-positive and HIV-negative women of diverse racial/ethnic backgrounds, the HIBAG algorithm accurately imputed 89% of classical HLA alleles. However, there is, to our knowledge, no consensus in the scientific literature regarding which specific HLA imputation algorithm is superior in terms of performance and we cannot

exclude the possibility that algorithms not considered in the current study could prove more accurate. Furthermore, there was wide variation in imputation accuracy among ancestry groups for certain HLA genes (e.g., 66%-88% for HLA-B), which may reflect limitations of the HIBAG algorithm. Nevertheless, the current study provides independent validation that HLA imputation algorithms can provide accurate predictions of HLA alleles in large multiracial populations when gold-standard HLA allele data are available for a subset of study participants.

Our observation that HIBAG was slightly more accurate for prediction of HLA-B and HLA-C alleles than SNP2HLA is consistent with the findings of a recent study that conducted HLA class II imputation in African Americans (Levin et al. 2014). The authors of that study found HIBAG to be the most accurate of three HLA imputation algorithms that they considered (i.e., HIBAG, BEAGLE and HLA*IMP:02 (Dilthey et al. 2013)) when a representative sample of individuals with both HLA allele and GWAS data was used as training data. We attempted to use a prior version of the latter software (HLA*IMP:01 (Dilthey et al. 2011)) but could not impute HLA allele data because the software required uploads of SNPs which were not genotyped on the Illumina HumanOmni2.5-quad chip. HIBAG employs bootstrap aggregation known as bagging to impute HLA alleles whereas SNP2HLA relies on an internal BEAGLE run to impute HLA alleles. The authors of the HIBAG algorithm observed that HIBAG is robust to departures from HWE (Zheng et al. 2014) and it is possible that differences in how HIBAG and SNP2HLA handle departures from HWE accounted for the slight differences in accuracy between HIBAG and SNP2HLA in our dataset.

Accuracy of HIBAG allele imputation for certain highly polymorphic HLA genes (e.g., HLA-B and HLA-DRB1) was highly dependent on the specific ancestry group considered. Specifically, accuracy of imputation at these genes was highest for the AA group (which had the largest sample size) and lowest for the HA group (which had the lowest sample size). The authors of the HIBAG algorithm observed a similar relationship between imputation accuracy and sample – size in their study imputation accuracy was on average highest in the EA group (n=2,668) and lowest in the AA group (n=173) (Zheng et al. 2014). Taken together, these data suggest that imputed allele data in small studies should be interpreted with caution – particularly if HLA-B and/or HLA-DRB1 are of primary interest. On the other hand, in our dataset there was little difference in imputation accuracy across ancestry groups for the least polymorphic HLA genes (HLA-DQA1 and HLA-DQB1). However, imputation accuracy for HLA-A also did not vary dramatically (87%-95%) across ancestry groups – HLA-A is the second most polymorphic class I gene and more research is needed to better understand relationships between sample size, the extent of HLA gene polymorphism, and imputation accuracy.

Misclassification of explanatory variables in epidemiologic studies, as could occur through imperfect imputation of HLA alleles, generally attenuates statistical power but does not induce bias. In certain circumstances however a bias could occur. For example, if the outcome (disease) phenotype of a study is imperfectly measured in one but not other racial groups in a multiracial population a bias could occur because accuracy of HLA imputation

may also vary by race. Careful attention to study design and stratified analyses may help to mitigate bias in such circumstances.

Limitations must be considered in the interpretation of these results. First, we limited the scope of our comparisons to two HLA-dedicated imputation algorithms, SNP2HLA and HIBAG. Second, we evaluated the performance of these HLA-dedicated imputation algorithms in a single multiracial cohort – the majority of study participants were African American. Finally, our dataset did not include HLA-DP alleles and we did not conduct imputation for this gene.

We conclude that accurate imputation of high-resolution HLA class I and II alleles can be conducted in admixed populations provided that gold-standard HLA data are available for a subset of study participants. Sample size of the admixed study population correlated strongly with accuracy of imputation. HIBAG predicted HLA-B and HLA-C alleles more accurately than SNP2HLA. An additional advantage of HIBAG is that posterior prediction probabilities are calculated for each study participant which enables the investigator to analyze subsets of the population with high predicted (e.g., >95%) imputation accuracy. We consider that such a strategy is potentially valid because in our assessment of imputation accuracy for specific alleles (Supplementary Table 3), we observed only minor differences in allele vs. allele prediction accuracy. Nevertheless, it is possible that such a strategy could induce a selection bias (i.e., alleles with lower imputation accuracy could be systematically excluded from the dataset) and therefore caution should be exercised when making such exclusions. If these and other issues are carefully considered, imputation of HLA data in multiracial populations may lead to significant new knowledge about health and disease.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This project was supported in part by R01AI057006 and R01CA085178 (to H.D.S.). Additional data in this manuscript were collected by the Women's Interagency HIV Study (WIHS). The contents of this publication are solely the responsibility of the authors and do not represent the official views of the National Institutes of Health (NIH). WIHS (Principal Investigators): UAB-MS WIHS (Michael Saag, Mirjam-Colette Kempf, and Deborah Konkle-Parker), U01-AI-103401; Atlanta WIHS (Ighovwerha Ofotokun and Gina Wingood), U01-AI-103408; Bronx WIHS (Kathryn Anastos), U01-AI-035004; Brooklyn WIHS (Howard Minkoff and Deborah Gustafson), U01-AI-031834; Chicago WIHS (Mardge Cohen and Audrey French), U01-AI-034993; Metropolitan Washington WIHS (Mary Young and Seble Kassaye), U01-AI-034994; Miami WIHS (Margaret Fischl and Lisa Metsch), U01-AI-103397; UNC WIHS (Adaora Adimora), U01-AI-103390; Connie Wofsy Women's HIV Study, Northern California (Ruth Greenblatt, Bradley Aouizerat, and Phyllis Tien), U01-AI-034989; WIHS Data Management and Analysis Center (Stephen Gange and Elizabeth Golub), U01-AI-042590; Southern California WIHS (Joel Milam), U01-HD-032632 (WIHS I – WIHS IV). The WIHS is funded primarily by the National Institute of Allergy and Infectious Diseases (NIAID), with additional co-funding from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD), the National Cancer Institute (NCI), the National Institute on Drug Abuse (NIDA), and the National Institute on Mental Health (NIMH). Targeted supplemental funding for specific projects is also provided by the National Institute of Dental and Craniofacial Research (NIDCR), the National Institute on Alcohol Abuse and Alcoholism (NIAAA), the National Institute on Deafness and other Communication Disorders (NIDCD), and the NIH Office of Research on Women's Health. WIHS data collection is also supported by UL1-TR000004 (UCSF CTSA) and UL1-TR000454 (Atlanta CTSA).

References

1. Bacon MC, von W V, Alden C, Sharp G, Robison E, Hessol N, et al. The Women's Interagency HIV Study: an observational cohort brings clinical sciences to the bench. *Clinical and Vaccine Immunology*. 2005; 12:1013.
2. Browning BL, Browning SR. A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. *American Journal of Human Genetics*. 2009; 84:210. [PubMed: 19200528]
3. Chu X, Pan CM, Zhao SX, Liang J, Gao GQ, Zhang XM, et al. A genome-wide association study identifies two new risk loci for Graves' disease. *Nature Genetics*. 2011; 43:897. [PubMed: 21841780]
4. Colombo S, Rauch A, Rotger M, Fellay J, Martinez R, Fux C, et al. The HCP5 single-nucleotide polymorphism: a simple screening tool for prediction of hypersensitivity reaction to abacavir. *Journal of Infectious Diseases*. 2008; 198:864. [PubMed: 18684101]
5. de Bakker PI, McVean G, Sabeti PC, Miretti MM, Green T, Marchini J, et al. A high-resolution HLA and SNP haplotype map for disease association studies in the extended human MHC. *Nature Genetics*. 2006; 38:1166. [PubMed: 16998491]
6. Dilthey A, Leslie S, Moutsianas L, Shen J, Cox C, Nelson MR, et al. Multi-population classical HLA type imputation. *PLoS Computational Biology*. 2013; 9:e1002877. [PubMed: 23459081]
7. Dilthey AT, Moutsianas L, Leslie S, McVean G. HLA*IMP--an integrated framework for imputing classical HLA alleles from SNP genotypes. *Bioinformatics*. 2011; 27:968. [PubMed: 21300701]
8. Duggal P, Thio CL, Wojcik GL, Goedert JJ, Mangia A, Latanich R, et al. Genome-wide association study of spontaneous resolution of hepatitis C virus infection: data from multiple cohorts. *Annals of Internal Medicine*. 2013; 158:235. [PubMed: 23420232]
9. Gao X, Nelson GW, Karacki P, Martin MP, Phair J, Kaslow R, et al. Effect of a single amino acid change in MHC class I molecules on the rate of progression to AIDS. *New England Journal of Medicine*. 2001; 344:1668. [PubMed: 11386265]
10. Hafler DA, Compston A, Sawcer S, Lander ES, Daly MJ, De Jager PL, et al. Risk alleles for multiple sclerosis identified by a genomewide study. *New England Journal of Medicine*. 2007; 357:851. [PubMed: 17660530]
11. Jia X, Han B, Onengut-Gumuscu S, Chen WM, Concannon PJ, Rich SS, et al. Imputing amino acid polymorphisms in human leukocyte antigens. *PLoS One*. 2013; 8:e64683. [PubMed: 23762245]
12. Kamatani Y, Watanapokayakit S, Ochi H, Kawaguchi T, Takahashi A, Hosono N, et al. A genome-wide association study identifies variants in the HLA-DP locus associated with chronic hepatitis B in Asians. *Nature Genetics*. 2009; 41:591. [PubMed: 19349983]
13. Kuniholm MH, Gao X, Xue X, Kovacs A, Anastos K, Marti D, et al. Human leukocyte antigen genotype and risk of HIV disease progression before and after initiation of antiretroviral therapy. *Journal of Virology*. 2011; 85:10826. [PubMed: 21849458]
14. Kuniholm MH, Kovacs A, Gao X, Xue X, Marti D, Thio CL, et al. Specific human leukocyte antigen class I and II alleles associated with hepatitis C virus viremia. *Hepatology*. 2010; 51:1514. [PubMed: 20169624]
15. Leslie S, Donnelly P, McVean G. A statistical method for predicting classical HLA alleles from SNP data. *American Journal of Human Genetics*. 2008; 82:48. [PubMed: 18179884]
16. Levin AM, Adrianto I, Datta I, Iannuzzi MC, Trudeau S, McKeigue P, et al. Performance of HLA allele prediction methods in African Americans for class II genes HLA-DRB1, -DQB1, and -DPB1. *BMC Genetics*. 2014; 15:72. [PubMed: 24935557]
17. Noguchi E, Sakamoto H, Hirota T, Ochiai K, Imoto Y, Sakashita M, et al. Genome-wide association study identifies HLA-DP as a susceptibility gene for pediatric asthma in Asian populations. *PLoS Genetics*. 2011; 7:e1002170. [PubMed: 21814517]
18. Pereyra F, Jia X, McLaren PJ, Telenti A, de Bakker PI, Walker BD, et al. The major genetic determinants of HIV-1 control affect HLA class I peptide presentation. *Science*. 2010; 330:1551. [PubMed: 21051598]

19. Raychaudhuri S, Sandor C, Stahl EA, Freudenberg J, Lee HS, Jia X, et al. Five amino acids in three HLA proteins explain most of the association between MHC and seropositive rheumatoid arthritis. *Nature Genetics*. 2012; 44:291. [PubMed: 22286218]
20. Setty MN, Gusev A, Pe'er I. HLA type inference via haplotypes identical by descent. *J Comput Biol*. 2011; 18:483. [PubMed: 21385049]
21. Trowsdale J, Knight JC. Major histocompatibility complex genomics and human disease. *Annual Review Genomics and Human Genetics*. 2013; 14:301.
22. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*. 2007; 447:661. [PubMed: 17554300]
23. Xie M, Li J, Jiang T. Accurate HLA type inference using a weighted similarity graph. *BMC Bioinformatics*. 2010; 11(Suppl 11):S10.
24. Zheng X, Shen J, Cox C, Wakefield JC, Ehm MG, Nelson MR, et al. HIBAG--HLA genotype imputation with attribute bagging. *Pharmacogenomics Journal*. 2014; 14:192. [PubMed: 23712092]

Figure 1. Racial/ethnic ancestry groups for WIHS women based on self-report and the first two principal components (PCs) of ancestry informative marker data.

Table 1

Accuracy of imputed HLA-B and -C alleles based on SNP2HLA and HIBAG software packages in a 20% test sample of women^a

All women		HLA-B	HLA-C
SNP2HLA	% of correctly predicted total alleles	392/458 (86%)	393/426 (92%)
	% women with both alleles correctly predicted	170/229 (74%)	186/213 (87%)
HIBAG	% of correctly predicted total alleles	399/458 (87%)	401/426 (94%)
	% women with both alleles correctly predicted	177/229 (77%)	194/213 (91%)
African American women			
SNP2HLA	% of correctly predicted total alleles	248/288 (86%)	241/268 (90%)
	% women with both alleles correctly predicted	107/144 (74%)	112/134 (84%)
HIBAG	% of correctly predicted total alleles	253/288 (88%)	249/268 (93%)
	% women with both alleles correctly predicted	111/144 (77%)	120/134 (90%)
White women			
SNP2HLA	% of correctly predicted total alleles	40/70 (57%)	54/64 (84%)
	% women with both alleles correctly predicted	9/35 (26%)	24/32 (75%)
HIBAG	% of correctly predicted total alleles	59/70 (84%)	59/64 (92%)
	% women with both alleles correctly predicted	26/35 (74%)	28/32 (88%)

^aWomen with both HLA-B/C allele and HLA-B/C SNP data were randomly split into two groups: an 80% training group and a 20% testing group. Accuracy estimates are computed for the 20% testing group.

Table 2

Accuracy of HLA class I and II allele imputation by HIBAG software using GWAS (Illumina HumanOmni2.5) and gold-standard HLA allele (PCR-SSO)
data^{a,b}

	Ancestry Group ^c	HLA-A	HLA-B	HLA-C	HLA-DQA1	HLA-DQB1	HLA-DRB1	Total
% of correctly predicted total alleles	EA	453/476 (95%)	385/462 (83%)	418/448 (93%)	210/262 (80%)	243/262 (93%)	372/428 (87%)	2081/2338 (89%)
	AA	1674/1782 (94%)	1534/1748 (88%)	1615/1696 (95%)	992/1200 (83%)	1089/1204 (90%)	1475/1622 (91%)	8379/9252 (91%)
	HA	169/194 (87%)	123/186 (66%)	161/186 (87%)	79/96 (82%)	86/96 (90%)	107/154 (69%)	725/912 (79%)
	OA	225/258 (87%)	192/254 (76%)	229/250 (92%)	175/196 (89%)	177/196 (90%)	201/236 (85%)	1199/1390 (86%)
	Total	2521/2710 (93%)	2234/2650 (84%)	2423/2580 (94%)	1456/1754 (83%)	1595/1758 (91%)	2155/2440 (88%)	12384/13892 (89%)
Median % women with both alleles correctly predicted	EA	216/238 (91%)	161/231 (70%)	198/224 (88%)	92/131 (70%)	113/131 (86%)	164/214 (77%)	944/1169 (81%)
	AA	791/891 (89%)	682/874 (78%)	784/848 (92%)	463/600 (77%)	511/602 (85%)	679/811 (84%)	3910/4626 (85%)
	HA	75/97 (77%)	42/93 (45%)	71/93 (76%)	37/48 (77%)	39/48 (81%)	37/77 (48%)	301/456 (66%)
	OA	99/129 (77%)	80/127 (63%)	108/125 (86%)	82/98 (84%)	79/98 (81%)	88/118 (75%)	536/695 (77%)
	Total	1181/1355 (87%)	965/1325 (73%)	1161/1290 (90%)	674/877 (77%)	742/879 (84%)	968/1220 (79%)	5691/6946 (82%)

EA: European ancestry; AA: African ancestry; HA: Hispanic ancestry; OA: Other ancestry

^a GWAS data include 1,952 genotyped SNPs in the HLA region and gold-standard PCR-SSO allele typing was conducted for n=1,587 women (see Methods)

^b Estimates are based on the average of five 10-fold cross-validation procedures

^c Ancestry groups were defined according to principal components (PCs) of ancestry informative markers (see Methods)