

## ORIGINAL ARTICLE

# Evaluating genetic susceptibility to *Staphylococcus aureus* bacteremia in African Americans using admixture mapping

DD Cyr<sup>1</sup>, AS Allen<sup>1,2</sup>, G-J Du<sup>3</sup>, F Ruffin<sup>4</sup>, C Adams<sup>4</sup>, JT Thaden<sup>4</sup>, SA Maskarinec<sup>4</sup>, M Souli<sup>4,5</sup>, S Guo<sup>6</sup>, DM Dykxhoorn<sup>6</sup>, WK Scott<sup>6</sup> and VG Fowler Jr<sup>1,4</sup>

The incidence of *Staphylococcus aureus* bacteremia (SAB) is significantly higher in African American (AA) than in European-descended populations. We used admixture mapping (AM) to test the hypothesis that genomic variations with different frequencies in European and African ancestral genomes influence susceptibility to SAB in AAs. A total of 565 adult AAs (390 cases with SAB; 175 age-matched controls) were genotyped for AM analysis. A case-only admixture score and a mixed  $\chi^2$ (1df) score (MIX) to jointly evaluate both single-nucleotide polymorphism (SNP) and admixture association ( $P < 5.00e-08$ ) were computed using MIXSCORE. In addition, a permutation scheme was implemented to derive multiplicity adjusted  $P$ -values (genome-wide 0.05 significance threshold:  $P < 9.46e-05$ ). After empirical multiplicity adjustment, one region on chromosome 6 (52 SNPs,  $P = 4.56e-05$ ) in the HLA class II region was found to exhibit a genome-wide statistically significant increase in European ancestry. This region encodes genes involved in HLA-mediated immune response and these results provide additional evidence for genetic variation influencing HLA-mediated immunity, modulating susceptibility to SAB.

*Genes and Immunity* (2017) 18, 95–99; doi:10.1038/gene.2017.6; published online 23 March 2017

## INTRODUCTION

*Staphylococcus aureus* infections are a leading cause of disease in humans. A growing body of evidence suggests that genetic variation can influence susceptibility to infection with *S. aureus*.<sup>1</sup> Genetic variation is associated with susceptibility to *S. aureus* infection in inbred mice,<sup>2,3</sup> sheep<sup>4</sup> and cattle.<sup>5</sup> In humans, first-degree relatives of patients previously hospitalized with *S. aureus* bacteremia (SAB) were themselves significantly more likely to develop SAB than the population as a whole.<sup>6</sup> Our recent genome-wide study of >50 000 white subjects found that two single-nucleotide polymorphisms (SNPs) in the human leukocyte antigen (HLA) class II region on chromosome 6 were associated with susceptibility to *S. aureus* infection at a genome-wide significant level.<sup>7</sup>

African Americans (AA) represent a complex mixture of West African and European ancestry, with an average 20% of each AA genome inherited from European ancestors.<sup>8</sup> Such admixture is traditionally regarded as a liability in genetic association studies and is typically avoided by studying genetically homogeneous populations. This, in turn, has contributed to under-representation of AAs in large-scale genetic association studies in general<sup>9</sup> and in *S. aureus* susceptibility studies in particular. For example, all three of the genome-wide association studies (GWASs) performed to evaluate genetic susceptibility to *S. aureus* were conducted exclusively in individuals of European descent.<sup>7,10,11</sup> This is problematic, as rates of invasive *S. aureus* are significantly higher among AAs than European-descended populations.<sup>12,13</sup>

Admixture mapping (AM) is an innovative strategy to overcome the issue of genetic admixture in human genotyping studies. For

diseases where associated genetic variants differ substantially in frequency between ancestral populations, admixed individuals with disease will have an overrepresentation of ancestry from the population with the higher proportion of risk alleles.<sup>9</sup> For this reason, AM can be more powerful than standard genetic association methods in cases where there is substantial genetic heterogeneity at a disease locus.<sup>9</sup> Diseases for which risk varies substantially between ancestral populations make particularly good targets for AM studies. For example, AM has been used successfully in hypertension,<sup>14</sup> prostate cancer,<sup>15,16</sup> peripheral artery disease<sup>17</sup> and type II diabetes.<sup>18</sup>

In the current study, we evaluated the role of genetic variation on susceptibility to SAB in AAs. We used AM to test the hypothesis that, at particular sites in the genome, cases will be enriched for a given local ancestry relative to that expected given their average, genome-wide, admixture proportion. Such sites could represent the presence of risk or absence of protective alleles in the population showing enrichment.

## RESULTS

A total of 390 cases and 175 controls were included in the analysis. Compared with AA controls, SAB cases were more likely to be diabetic (48.4 vs 33.7%,  $P < 0.001$ ) and hemodialysis dependent (42.3 vs 5.7%;  $P < 0.001$ ), less likely to have a neoplasm (10.6 vs 23.4%,  $P < 0.001$ ) and less likely to have had a recent surgical procedure performed (20.7 vs 36.0%,  $P < 0.001$ ) (Table 1). Genome-wide ancestry for each sample was defined as the average of local ancestry estimates across the genome. Genome-

<sup>1</sup>Duke Clinical Research Institute, Duke University Medical Center, Durham, NC, USA; <sup>2</sup>Department of Biostatistics and Bioinformatics, Duke University School of Medicine, Durham, NC, USA; <sup>3</sup>Duke Center for Genomic and Computational Biology, Durham, NC, USA; <sup>4</sup>Division of Infectious Diseases, Duke University Medical Center, Durham, NC, USA; <sup>5</sup>School of Medicine, National and Kapodistrian University of Athens, Chaidari, Greece and <sup>6</sup>Dr John T. Macdonald Foundation Department of Human Genetics and John P. Hussman Institute for Human Genomics, Miller School of Medicine, University of Miami, Miami, FL, USA. Correspondence: Dr VG Fowler Jr, Division of Infectious Diseases, Duke University Medical Center, Room 185 Hanes Building 315 Trent Drive, Durham, NC 27710, USA.

E-mail: vance.fowler@duke.edu

Received 2 December 2016; revised 14 February 2017; accepted 16 February 2017; published online 23 March 2017

**Table 1.** Characteristics of the admixture mapping population

Characteristic	Cases (N = 389)	Controls (N = 175)	P-value
Age (years), n (%)			0.051
10–19	3 (0.77%)	0 (0.00%)	
20–29	23 (5.91%)	15 (8.57%)	
30–39	30 (7.71%)	18 (10.29%)	
40–49	65 (16.71%)	23 (13.14%)	
50–59	112 (28.79%)	35 (20.00%)	
60–69	90 (23.14%)	44 (25.14%)	
70–79	46 (11.83%)	26 (14.86%)	
80–89	15 (3.86%)	14 (8.00%)	
> 89	5 (1.29%)	0 (0.00%)	
Gender, n (%)			0.491
Female	199 (51.16%)	95 (54.29%)	
Male	190 (48.84%)	80 (45.71%)	
Injection drug use, n (%)			0.177
No	362 (95.77%)	142 (98.61%)	
Yes	16 (4.23%)	2 (1.39%)	
Diabetic, n (%)			0.001
No	195 (51.59%)	116 (66.29%)	
Yes	183 (48.41%)	59 (33.71%)	
Hemodialysis dependence, n (%)			< 0.001
No	218 (57.67%)	165 (94.29%)	
Yes	160 (42.33%)	10 (5.71%)	
Neoplasm, n (%)			< 0.001
No	338 (89.42%)	134 (76.57%)	
Yes	40 (10.58%)	41 (23.43%)	
Rheumatoid arthritis, n (%)			0.477
No	369 (98.40%)	142 (97.26%)	
Yes	6 (1.60%)	4 (2.74%)	
Corticosteroid use past 30 days, n (%)			0.207
No	303 (80.16%)	124 (84.93%)	
Yes	75 (19.84%)	22 (15.07%)	
Surgery past 30 days, n (%)			< 0.001
No	299 (79.31%)	112 (64.00%)	
Yes	78 (20.69%)	63 (36.00%)	
Permanently implanted foreign body, n (%)			0.354
No	128 (34.04%)	56 (38.36%)	
Yes	248 (65.96%)	90 (61.64%)	

The denominator for each summary is the number of non-missing values for the respective characteristic.

wide European ancestry in cases and controls were similar (cases 20.4%, s.d. 1.5%; controls 20.9%, s.d. 2.4%) (Table 2).

No SNPs were found to exhibit genome-wide statistically significant evidence of case-control association (accounting for admixture) using the MIX statistic at the traditional GWAS multiple-comparison corrected threshold of  $P < 5.00e-08$ . For the admixture score (ADM) test of admixture, strong correlations between test statistics due to large blocks of local ancestry make conventional thresholds inappropriate. Therefore, we implemented a permutation scheme to derive empirical multiplicity adjusted  $P$ -values (described in Materials and Methods section and in complete detail in the accompanying Supplementary Materials). Using this permutation scheme, the empirical threshold used to declare statistically significant increased admixture is  $9.46e-05$ , which corresponds to the 5th percentile of the permutation distribution of the maximum ADM test statistic.

Using the empirical threshold value, one region on chromosome 6 in the HLA class II region (52 SNPs; from physical position 32377284 to 32660943 (hg19); see Supplementary Materials for list) exhibited increased admixture association at a level of genome-wide significance ( $P = 4.56e-05 < 9.46e-05$ ) (Figure 1).

Despite the significant evidence that local admixture proportions differ from the genome-wide average, visually there is no notable difference in the estimated proportion of European ancestry in cases and controls in the region on chromosome 6 (Figure 2).

**Table 2.** Genome-wide average proportions of European ancestry

	N	Mean (s.d.)	Median (Q1, Q3)
Overall	565	0.206 (0.014)	0.204 (0.196, 0.214)
Cases	390	0.204 (0.015)	0.204 (0.194, 0.213)
Controls	175	0.209 (0.024)	0.209 (0.191, 0.226)

## DISCUSSION

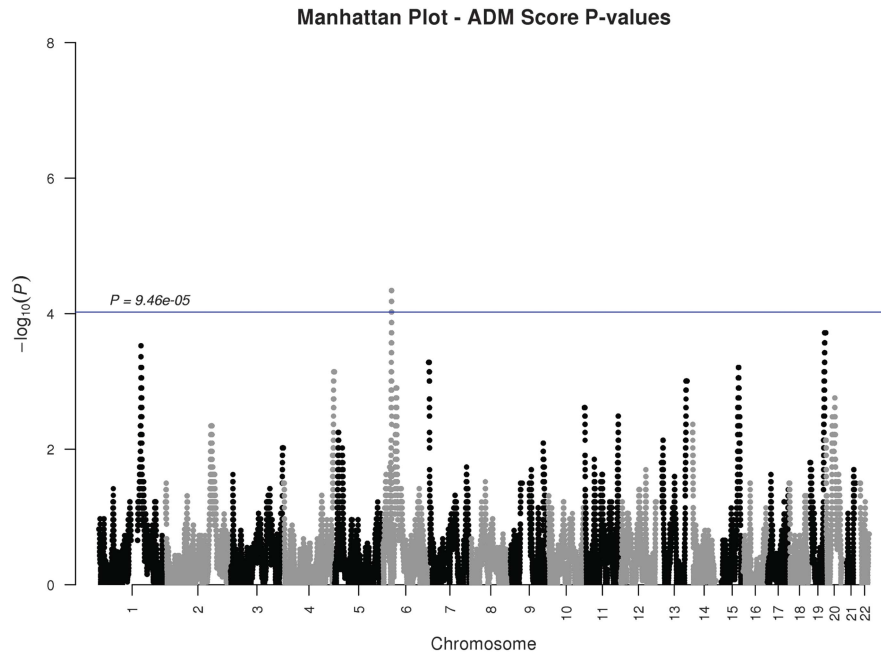
The genetic basis of human susceptibility to *S. aureus* infection is poorly understood.<sup>19</sup> The current investigation is the first to evaluate genetic susceptibility to *S. aureus* in AA and is one of the first genetic studies involving infectious disease risk in AAs. The study was possible by using AM as an alternate analytical approach to consider the genetic basis of SAB in AA populations. Using this approach, one region on chromosome 6 met genome-wide significance thresholds.

This region overlaps the HLA class II region previously implicated in a GWAS of susceptibility to *S. aureus* infection.<sup>7</sup> HLA has been identified as a potentially important determinant of the innate and adaptive immune response to infections caused by *S. aureus* and other Gram-positive bacteria.<sup>7,20–25</sup> The current study adds to the body of evidence identifying HLA as a potential determinant of host susceptibility to *S. aureus* infection.

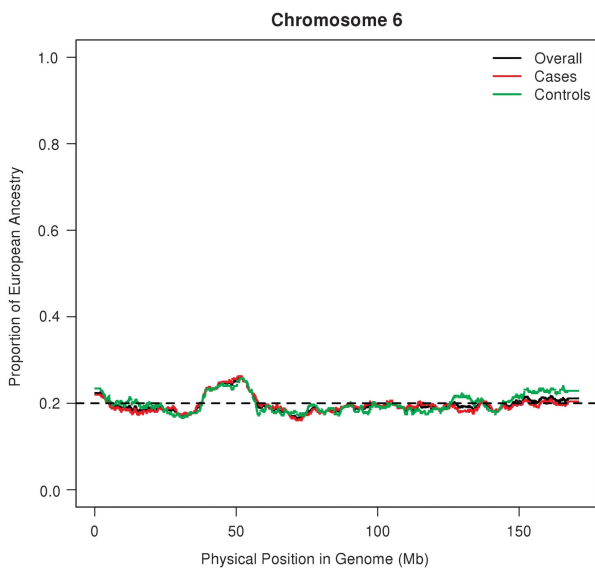
In our recent GWAS of 4701 unique white subjects with culture-confirmed *S. aureus* infection and 45 344 matched controls, two imputed SNPs near the genes encoding HLA-DRA and HLA-DRB1 achieved genome-wide significance (for example, rs115231074: odds ratio (OR), 1.22;  $P = 1.3 \times 10^{-10}$ ; rs35079132: OR, 1.24;  $P = 3.8 \times 10^{-8}$ ) and one adjacent genotyped SNP was nearly genome-wide significant (rs4321864: OR, 1.13;  $P = 8.8 \times 10^{-8}$ ).<sup>7</sup> These SNPs were located near HLA-DRA and HLA-DRB1 genes in the HLA class II region of human chromosome 6. In the current study, there is significant evidence of increased European ancestry in AA SAB cases in the same region (whereby there is a higher degree of European ancestry in this region than in the genome as a whole in both cases and controls), although the strongest signal is located 5' of the HLA-DRA gene. The current study identifies the same region previously identified in our prior GWAS and extends the association of HLA class II to SAB in multiple ethnic groups. Collectively, these results suggest that efforts to follow-up the association of SAB with HLA class II should consider European alleles and haplotypes in both European and African-descended populations.

The findings of the current study are consistent with a growing body of evidence that associates genetic variation in HLA with susceptibility to Gram-positive bacteria, including *S. aureus*. First, variations within the HLA class II region are the only ones to date to have been implicated on a genome-wide significance level as being associated with susceptibility to *S. aureus*.<sup>7</sup> Second, haplotypes across the HLA class II region have been associated with invasive *Streptococcus pyogenes* infection<sup>20</sup> and determine severity of response to bacterial superantigens from both *Streptococcus pyogenes*<sup>21</sup> and *S. aureus*.<sup>22</sup> Third, toxic shock syndrome toxin (TSST-1), a *S. aureus* superantigen, binds HLA-DR1<sup>23,24</sup> and is important in the pathogenesis of SAB and endocarditis.<sup>26</sup> Fourth, nasal colonization with *S. aureus* is associated with the HLA-DR3 and HLA-DR7 class II serotypes.<sup>25</sup> Taken together, these results suggest that further study is needed to better understand the role of variation within the HLA class II region in the pathogenesis of *S. aureus* infection.

The impact of host genetic variation on susceptibility to *S. aureus* infection is unresolved. Rates of *S. aureus* infections are significantly higher among distinct ethnic populations, including Australian<sup>27</sup> and Canadian<sup>28</sup> aboriginal populations, New Zealand Maori<sup>29</sup> and AAs.<sup>12,13</sup> Although a twin study failed to demonstrate



**Figure 1.** ADM score *P*-values. The blue line indicates a significance threshold of 9.46e-05, corresponding to empirical multiplicity adjusted *P*-values.



**Figure 2.** Average local ancestry of 565 AA samples in chromosome 6. Visually, there is no notable difference in the estimated proportions of European ancestry between cases and controls; however, a region of 52 SNPs (from 32.37728 to 32.66094 Mb) was found to exhibit statistically significant increased European ancestry.

an association between nasal carriage of monozygotic twin status,<sup>30</sup> a recent 20-year nationwide cohort study drawing from the entire Danish population of >8 million individuals showed evidence of familial clustering of SAB in first-degree relatives.<sup>6</sup> Among the 34 774 individuals with a first-degree relative (index case patient) previously hospitalized with SAB who were followed for a mean of 7.8 years, a higher rate of SAB was observed among first-degree relatives than among the background Danish population (standardized incidence ratio (SIR): 2.49; (95% confidence

interval (CI): 1.95–3.19)). This estimate was significantly higher if the index case patient was a sibling (SIR: 5.01; (95% CI: 3.30–7.62)) than a parent and highest in siblings of individuals who developed non-hospital acquired SAB (SIR: 5.66; (95% CI: 3.47–9.24)). These findings persisted after adjustment for a number of comorbid conditions.

The current study has limitations. Though we were able to identify a single genome-wide significant admixture signal, it is possible that additional sites are not identified owing to sample size. Although we estimated *a priori* that we had 90% power to detect single loci explaining differences between observed AA and European rates,<sup>12,13</sup> the difference in admixture proportion detected on chromosome 6 was more modest, suggesting that single locus explanation is unlikely and that the current sample size is underpowered to identify multiple loci with lesser admixture signals. The use of admixture analysis in a control set is a useful way to filter results that are not associated with the disease trait; however, as noted earlier, we cannot be certain that the controls were sufficiently exposed to *S. aureus* to develop bacteremia if susceptible. This limited the utility of using admixture proportions in controls because we do not know their exposure status with certainty. Finally, both cases and controls exhibited evidence for increased European ancestry in this region. However, this does not necessarily argue against the region having a role in susceptibility to infection, as many of the control subjects may simply have had insufficient exposure to *S. aureus* to develop bloodstream infections if susceptible.

Despite these limitations, the current manuscript makes a key observation. Using an innovative statistical approach that allows analysis of an underrepresented study population, we have again identified HLA as a promising and biologically plausible determinant of genetic susceptibility to *S. aureus* at a genome-wide level of statistical significance. These results provide further support for association of SAB susceptibility with the HLA class II region. When placed in context, these results suggest that future studies of HLA-mediated susceptibility to SAB should consider the role of European class II haplotypes in determining risk in multiple ethnic groups with European admixture.

## MATERIALS AND METHODS

### Study sample

Our study used a case–control design. Data for cases were obtained from the *S. aureus* Bacteremia Group repository,<sup>11,31</sup> which has prospectively cataloged clinical data, bloodstream *S. aureus* isolates and/or human DNA from all consenting patients with SAB at Duke University Medical Center since 1994. Cases were unique AA adult inpatients with monomicrobial SAB. Controls are age-matched adult AA inpatients with no current or past *S. aureus* infection. The study was approved by the Duke University Institutional Review Board (IRB). All participants provided informed consent according to IRB policies. Patients dying of SAB prior to consent were included using IRB-approved policies for decedent research.

### Genotyping and quality control (QC)

The Illumina Multi-Ethnic Genotyping Array (MEGA, comprising 1 779 819 markers; San Diego, CA, USA), developed to capture common and rare variation in a multi-ethnic sample and particularly enriched for variations initially detected in African and AA samples, was used to genotype 672 samples (450 AA SAB, 194 AA controls, plus 28 QC samples) at the University of Miami Hussman Institute for Human Genomics (HIHG) Center for Genome Technology. One CEPH standard sample was included on each 96-well plate to ensure reproducibility of results and check for plate rotation errors. Samples were processed 8 per MEGA ex-array and randomized with respect to case/control status prior to laboratory analysis. MEGA array images were analyzed using Illumina GenomeStudio (San Diego, CA, USA), using the manufacturer's protocols and cluster files. Called genotypes were exported from GenomeStudio for QC analysis.

First, replicate CEPH samples were checked for concordance; plates with discordant CEPH replicates were re-examined and rerun if necessary. Following this step, formatted data analysis files were extracted for subsequent statistical analysis.

### Power and sample size

Prior to initiating the study, we conducted analytic power calculations using standard power formulas.<sup>32</sup> These formulas involve four model parameters:  $q_A$ , the average amount of African ancestry;  $p_A$ , the African risk allele frequency;  $p_E$ , the European risk allele frequency; and  $\gamma$ , the risk conferred by each copy of the risk alleles compared with the non-risk homozygote. A significance threshold of  $3e-5$  was used to account for the number of hypotheses planned to be tested. All parameters were calibrated using the published rates of invasive multidrug-resistant *S. aureus* in AAs and European-descended populations.<sup>12,13</sup> Thus it was shown that, if a single genetic locus explained the observed differences in AA and European rates,<sup>12,13</sup> then we have >95% power to detect loci conferring increased risk for SAB in a sample of 450 AA cases and 194 AA controls.

### QC filtering

Following genotyping, 430 SAB cases and 200 AA controls remained. Several QC filters were applied to the genotyped data: sample and SNP genotyping success rates ( $\geq 95\%$ ) (removal of 8 cases, 0 controls and 250 956 markers), Hardy–Weinberg equilibrium ( $P < 0.001$ ) (removal of 7900 markers), gender discrepancies (removal of 11 cases and 13 controls) and identity-by-descent analysis (relatedness estimate  $\geq 0.10$ ) (removal of 17 cases and 8 controls). Duplicate SNPs were also excluded (24 026 removed) as well as SNPs on chromosomes coded 0 (8422 removed), SNPs on chromosome 26 (265 removed) and SNPs on non-zero chromosomes with base-pair location 0 (17 removed). A total of 390 AA SAB cases, 175 AA controls and 229 860 SNPs remained for statistical analysis after QC filtering.

### Descriptive statistics

Selected baseline characteristics of the SAB cases and AA controls are summarized with frequencies (percentages). Comparisons between SAB cases and AA controls were made using Pearson's chi-square test when cell frequencies were sufficient; otherwise Fisher's exact test was used.

### Inference of local ancestry using LAMP-LD

Local ancestry is defined as the genetic ancestry of an individual at a particular chromosomal location, where an individual can have 0, 1 or 2

copies of an allele derived from each ancestral population.<sup>33</sup> Several approaches have been proposed for local ancestry estimation in two-way admixtures, with methods that explicitly model the linkage disequilibrium (LD) structure within the ancestral populations showing the highest accuracy in AAs.<sup>34–38</sup> In this analysis, we use LAMP-LD,<sup>39</sup> a window-based algorithm combined with a hierarchical Hidden Markov Model to represent haplotypes in the ancestral population. Phased CEU and YRI haplotypes from the HapMap Phase 3 project<sup>40</sup> were used as reference panels. A total of 229 860 SNPs were analyzed.

### Association statistics

The MIXSCORE software<sup>41</sup> was used to evaluate two association statistics on the filtered admixed data. First, a case-only ADM<sup>41</sup> was computed to test the hypothesis that the proportion of European ancestry at the candidate locus differs from the genome-wide proportion (local admixture enrichment). Second, because the causal SNP may have different allele frequencies in ancestral populations, a mixed  $\chi^2(1df)$  score (MIX)<sup>35</sup> was also calculated to jointly evaluate both differences in SNP allele frequencies in cases vs controls and differential admixture; the  $P$ -value cutoff for this statistic was 5.00e-08.

### Adjustment for multiple testing

As regions of the genome can span multiple loci, resulting in highly correlated tests, a standard Bonferroni correction is conservative. Therefore, we calculated multiplicity adjusted  $P$ -values using the permutation-based step-down maxT algorithm of Westfall and Young.<sup>42</sup> Complete details of the method and its implementation relative to this study are given in Supplementary Materials.

## CONFLICT OF INTEREST

VGF served as Chair of V710 Scientific Advisory Committee (Merck) and has received grant support from Cerexa/Actavis, Pfizer, Advanced Liquid Logics, NIH, MedImmune, Cubist/Merck, Karius, Contrafect and Genentech. He has NIH STTR/SBIR grants pending from Affinergy and Locust, Medical Surface, Inc.; has been a paid consultant for Achaogen, Astellas, Arsanis, Affinergy, Basilea, Bayer, Cerexa, Contrafect, Cubist, Debiopharm, Durata, Grifols, Genentech, MedImmune, Merck, Medicines Co., Pfizer, Novartis, Novadigm, Theravance and xBiotech; has received honoraria from Theravance and Green Cross; and has a patent pending in sepsis diagnostics. No other authors have any commercial or other association that might pose a conflict of interest.

## ACKNOWLEDGEMENTS

This work was supported by R01-AI-068804 (to VGF). The sponsor had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

## AUTHOR CONTRIBUTIONS

VGF participated in conceiving the study idea and the study design, secured the grant funding, oversaw execution of the study and drafted the manuscript. DDC and G-JD conducted the primary analysis and drafted the manuscript. ASA participated in conceiving the study idea and overseeing the analytical aspects of the study. FR was responsible for enrollment of subjects and data collection. CA performed descriptive analysis on cases and controls. JTT, SAM and MS provided critical input on the manuscript. DMD oversaw genotyping of the study samples. SG performed quality control of the genotyping data. WKS oversaw the analytical aspects of the study and the genotyping of study samples. All authors read, edited and approved the final manuscript.

## REFERENCES

- 1 Messina JA, Thaden JT, Sharma-Kuinkel BK, Fowler VG Jr. Impact of bacterial and human genetic variation on *Staphylococcus aureus* infections. *PLoS Pathog* 2016; **12**: e1005330.
- 2 Ahn SH, Deshmukh H, Johnson N, Cowell LG, Rude TH, Scott WK *et al*. Two genes on A/J chromosome 18 are associated with susceptibility to *Staphylococcus aureus* infection by combined microarray and quantitative trait locus (QTL) analyses. *PLoS Pathog* 2010; **6**: e1001088.

- 3 Johnson NV, Ahn SH, Deshmukh H, Levin MK, Nelson CL, Scott WK *et al*. Haplotype association mapping identifies a candidate gene region in mice infected with *Staphylococcus aureus*. *G3 (Bethesda)* 2012; **2**: 693–700.
- 4 Bonnefont CM, Rainard P, Cunha P, Gilbert FB, Toufeer M, Aurel MR *et al*. Genetic susceptibility to *S. aureus* mastitis in sheep: differential expression of mammary epithelial cells in response to live bacteria or supernatant. *Physiol Genomics* 2012; **44**: 403–416.
- 5 Griesbeck-Zilch B, Osman M, Ch Kühn, Schwerin M, Bruckmaier RH, Pfaffl MW *et al*. Analysis of key molecules of the innate immune system in mammary epithelial cells isolated from marker-assisted and conventionally-selected cattle. *J Dairy Sci* 2009; **92**: 4621–4633.
- 6 Oestergaard LB, Christiansen MN, Schmiegelow MD, Skov RL, Andersen PS, Petersen A *et al*. Familial clustering of *Staphylococcus aureus* bacteremia in first-degree relatives: a Danish nationwide cohort study. *Ann Intern Med* 2016; **165**: 390–398.
- 7 DeLorenze GN, Nelson CL, Scott WK, Allen AS, Ray GT, Tsai AL *et al*. Polymorphisms in HLA class II genes are associated with susceptibility to *Staphylococcus aureus* infection in a white population. *J Infect Dis* 2016; **213**: 816–823.
- 8 Smith MW, Patterson N, Lautenberger JA, Truelove AL, McDonald GJ, Waliszewska A *et al*. A high-density admixture map for disease gene discovery in African Americans. *Am J Hum Genet* 2004; **74**: 1001–1013.
- 9 Darvasi A, Shifman S. The beauty of admixture. *Nat Genet* 2005; **37**: 118–119.
- 10 Ye Z, Vasco DA, Carter TC, Brilliant MH, Schrodi SJ, Shukla SK. Genome wide association study of SNP-, gene-, and pathway-based approaches to identify genes influencing susceptibility to *Staphylococcus aureus* infections. *Front Genet* 2014; **5**: 125.
- 11 Nelson CL, Pelak K, Podgoreanu MV, Ahn SH, Scott WK, Allen AS *et al*. A genome-wide association study of variants associated with acquisition of *Staphylococcus aureus* bacteremia in a healthcare setting. *BMC Infect Dis* 2014; **14**: 83.
- 12 Kallen AJ, Mu Y, Bulens S, Reingold A, Petit S, Gershman K *et al*. Active Bacterial Core Surveillance (ABCs) MRSA Investigators of the Emerging Infections Program. Health care-associated invasive MRSA infections. *JAMA* 2010; **304**: 641–648.
- 13 Klevens RM, Morrison MA, Nadle J, Petit S, Gershman K, Ray S *et al*. Active Bacterial Core Surveillance (ABCs) MRSA Investigators. Invasive methicillin-resistant *Staphylococcus aureus* infections in the United States. *JAMA* 2007; **298**: 1763–1771.
- 14 Zhu X, Luke A, Cooper RS, Quertermous T, Hanis C, Mosley T *et al*. Admixture mapping for hypertension loci with genome-scan markers. *Nat Genet* 2005; **37**: 177–181.
- 15 Bock CH, Schwartz AG, Ruterbusch JJ, Levin AM, Neslund-Dudas C, Land SJ *et al*. Results from a prostate cancer admixture mapping study in African-American men. *Hum Genet* 2009; **126**: 637–642.
- 16 Freedman ML, Haiman CA, Patterson N, McDonald GJ, Tandon A, Waliszewska A *et al*. Admixture mapping identifies 8q24 as a prostate cancer risk locus in African-American men. *Proc Natl Acad Sci USA* 2006; **103**: 14068–14073.
- 17 Scherer ML, Nalls MA, Pawlikowska L, Ziv E, Mitchell K, Huntsman S *et al*. Admixture mapping of ankle-arm index: identification of a candidate locus associated with peripheral arterial disease. *J Med Genet* 2010; **47**: 1–7.
- 18 Elbein SC, Das SK, Hallman DM, Hanis CL, Hasstedt SJ. Genome-wide linkage and admixture mapping of type 2 diabetes in African American families from the American Diabetes Association GENNID (Genetics of NIDDM) Study Cohort. *Diabetes* 2009; **58**: 268–274.
- 19 Shukla SK, Rose W, Schrodi SJ. Complex host genetic susceptibility to *Staphylococcus aureus* infections. *Trends Microbiol* 2015; **23**: 529–536.
- 20 Kotb M, Norrby-Teglund A, McGeer A, El-Sherbini H, Dorak MT, Khurshid A *et al*. An immunogenetic and molecular basis for differences in outcomes of invasive group A streptococcal infections. *Nat Med* 2002; **8**: 1398–1404.
- 21 Nooh MM, El-Gengehi N, Kansal R, David CS, Kotb M. HLA transgenic mice provide evidence for a direct and dominant role of HLA class II variation in modulating the severity of streptococcal sepsis. *J Immunol* 2007; **178**: 3076–3083.
- 22 Llewelyn M, Sriskandan S, Peakman M, Ambrozak DR, Douek DC, Kwok WW *et al*. HLA class II polymorphisms determine responses to bacterial superantigens. *J Immunol* 2004; **172**: 1719–1726.
- 23 Kim J, Urban RG, Strominger JL, Wiley DC. Toxic shock syndrome toxin-1 complexed with a class II major histocompatibility molecule HLA-DR1. *Science* 1994; **266**: 1870–1874.
- 24 Lavoie PM, Thibodeau J, Cloutier I, Busch R, Sékaly RP. Selective binding of bacterial toxins to major histocompatibility complex class II-expressing cells is controlled by invariant chain and HLA-DM. *Proc Natl Acad Sci USA* 1997; **94**: 6892–6897.
- 25 Kinsman OS, McKenna R, Noble WC. Association between histocompatibility antigens (HLA) and nasal carriage of *Staphylococcus aureus*. *J Med Microbiol* 1983; **16**: 215–220.
- 26 Salgado-Pabón W, Breshears L, Spaulding AR, Merriman JA, Stach CS, Horswill AR *et al*. Superantigens are critical for *Staphylococcus aureus* Infective endocarditis, sepsis, and acute kidney injury. *MBio* 2013; **4**: e00494–13.
- 27 Maguire GP, Arthur AD, Boustead PJ, Dwyer B, Currie BJ. Clinical experience and outcomes of community-acquired and nosocomial methicillin-resistant *Staphylococcus aureus* in a northern Australian hospital. *J Hosp Infect* 1998; **38**: 273–281.
- 28 Embil J, Ramotar K, Romance L, Alfa M, Conly J, Cronk S *et al*. Methicillin-resistant *Staphylococcus aureus* in tertiary care institutions of the Canadian prairies 1990–1992. *Infect Control Hosp Epidemiol* 1994; **15**: 646–651.
- 29 Hill PC, Wong CG, Voss LM, Taylor SL, Pottumarthy S, Drinkovic D *et al*. Prospective study of 125 cases of *Staphylococcus aureus* bacteremia in children in New Zealand. *Pediatr Infect Dis J* 2001; **20**: 868–873.
- 30 Andersen PS, Pedersen JK, Fode P, Skov RL, Fowler Jr VG, Stegger M *et al*. Influence of host genetics and environment on nasal carriage of *Staphylococcus aureus* in Danish middle-aged and elderly twins. *J Infect Dis* 2012; **206**: 1178–1184.
- 31 Fowler Jr VG, Olsen MK, Corey GR, Woods CW, Cabell CH, Reller LB *et al*. Clinical identifiers of complicated *Staphylococcus aureus* bacteremia. *Arch Intern Med* 2003; **163**: 2066–2072.
- 32 Montana G, Pritchard JK. Statistical tests for admixture mapping with case-control and cases-only data. *Am J Hum Genet* 2004; **75**: 771–789.
- 33 Thornton TA, Bermejo JL. Local and global ancestry inference and applications to genetic association analysis for admixed populations. *Genet Epidemiol* 2014; **38** (Suppl 1): S5–S12.
- 34 Pasaniuc B, Kennedy J, Mandouli I. Imputation-Based Local Ancestry Inference in Admixed Populations. *Proceedings of the 5th International Symposium on Bioinformatics Research and Applications*. Fort Lauderdale, FL, USA, 13–16 May 2009, pp 221–233.
- 35 Pasaniuc B, Sankararaman S, Kimmel G, Halperin E. Inference of locus-specific ancestry in closely related populations. *Bioinformatics* 2009; **25**: i213–i221.
- 36 Price AL, Tandon A, Patterson N, Barnes KC, Rafaels N, Ruczinski I *et al*. Sensitive detection of chromosomal segments of distinct ancestry in admixed populations. *PLoS Genet* 2009; **5**: e1000519.
- 37 Sundquist A, Fratkin E, Do CB, Batzoglou S. Effect of genetic divergence in identifying ancestral origin using HAPAA. *Genome Res* 2008; **18**: 676–682.
- 38 Seldin M, Pasaniuc B, Price AL. New approaches to disease mapping in admixed populations. *Nat Rev Genet* 2011; **12**: 523–528.
- 39 Baran Y, Pasaniuc B, Sankararaman S, Torgerson DG, Gignoux C, Eng C *et al*. Fast and accurate inference of local ancestry in Latino populations. *Bioinformatics* 2012; **28**: 1359–1367.
- 40 The International Hapmap Consortium. A second generation human haplotype map of over 3.1 million SNPs. *Nature* 2007; **449**: 851–861.
- 41 Pasaniuc B, Zaitlen N, Lettre G, Chen GK, Tandon A, Linda Kao WH *et al*. Enhanced statistical tests for GWAS in admixed populations: assessment using African Americans from CARE and a breast cancer consortium. *PLoS Genet* 2011; **7**: e1001371.
- 42 Westfall PH, Young SS. *Resampling-Based Multiple Testing: Examples and Methods for P-Value Adjustment*. John Wiley & Sons: New York, USA, 1993.



This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/4.0/>

© The Author(s) 2017

Supplementary Information accompanies this paper on Genes and Immunity website (<http://www.nature.com/gene>)