INTRODUCTION

Staphylococcus aureus is a complex human pathogen due to its ability to survive both as a carriage organism, and behave as an opportunistic pathogen in a susceptible host. It is a leading cause of invasive bacterial infection that contributes to substantial morbidity and mortality worldwide. This bacterium can cause a variety of diseases ranging from mild to severe skin and soft tissue infections, keratitis, and osteomyelitis to life-threatening tissue infections, bacteremia, pneumonia, endocarditis, and sepsis (Lowy, 1998; Rehm, 2008). Even though S. aureus colonizes human anterior nares and other sites on skin in 30–50% of the general population (Graham et al., 2006; Kuehnert et al., 2006; Gordon and Lowy, 2008), not everyone who is colonized gets infected. One reason could be differences in genetic susceptibility to colonization and infections. The established role of host susceptibility in other infectious diseases lends support for a role of host genetics in S. aureus infection (de Bakker and Telenti, 2010). Genetic susceptibility to S. aureus infections is expected to be complex because this pathogen uses a wide variety of virulence factors that interact with several host pathways to cause disease in humans. Numerous alleles segregating at a large number of loci contribute to complex disease susceptibility (Yang et al., 2011) with contributions from both common and rare alleles. Because genetic risk factors for S. aureus infections have not been previously studied on a genome-wide scale, we utilized the Personalized Medicine Research Project (PMRP) of Marshfield Clinic—a large, population-based biobank of DNA samples—to perform a preliminary genome-wide association study (GWAS) of laboratory-confirmed S. aureus infections to discover the underlying host-pathogen interactions.

MATERIALS AND METHODS

SUBJECTS

This study utilized Marshfield Clinic’s PMRP biobank, a cohort of ~20,000 individuals from 14 Zip Codes surrounding Marshfield, Wisconsin, USA. DNA samples from 3234 PMRP participants were genotyped at > 500,000 SNPs as part of the NHGRI/NIGMS-funded eMERGE network (McCarty et al., 2011). These genotypic data are linked to longitudinal...
electronic medical records (EMR) at the Marshfield Clinic and have served as a powerful resource in previous studies (McCarty et al., 2011; Cross et al., 2012; Hebbring et al., 2013). The population of PMRP is stable and highly homogeneous with a predominant Northern European genetic background and so carries lower risk of confounding by population stratification. All individuals in the study provided informed consent and the study was approved by the Marshfield Clinic IRB.

**STUDY DESIGN**
This study examined host susceptibility genes for *S. aureus* infection regardless of *S. aureus* colonization status. We performed a case/control GWAS on a subgroup of the PMRP population: 3234 subjects (309 cases/2925 controls). All study subjects were over 49 years of age. Cases were defined as individuals who had at least one laboratory confirmed test in their medical record of a disease caused by *S. aureus*. Controls were subjects who did not have any evidence of infection due to *S. aureus* in their EMR. We reasoned that because cases had to have evidence of laboratory confirmed *S. aureus* infection in the EMR, patients with no evidence of infection in the EMR would form an ideal control group. The results using our control group are likely very similar to those that would be obtained using population-based controls, as is often employed in GWAS studies (Burton et al., 2007). Only subjects with self-identified Northern European ancestry were included; therefore, both cases and controls had the same genetic ancestry. Additionally, a principal components analysis of their genome-wide genotypes in all study subjects (without knowledge of case/control status) was performed, revealing no evidence of population substructure (Supplemental Figure 1). With this filtered set of case/control subjects, we performed three levels of statistical analysis: SNP-based, Gene-based, and Pathway-based, to identify novel polymorphisms, genes, and pathways involved in susceptibility to *S. aureus* diseases. This hierarchical investigation of genetic effects allows for the incorporation of a gradation of biological information into the statistical tests—the SNP-based scan is the most comprehensive and agnostic to prior biological knowledge, the gene-based analysis uses positional information and collapses effects from the same protein-coding region, and the pathway-based analysis incorporates information obtained from various molecular biology studies.

**PHENOTYPES**
We extracted demographic and medical information on all case/control subjects from the Marshfield Clinic EMR. All cases had an active infection that had yielded *S. aureus* as the major or the only bacterium on a culture plate from a clinical sample such as blood, sputum, etc. Hospital surveillance subjects who were positive for *S. aureus* colonization by PCR were excluded from the study. Age, sex, body mass index (BMI), and Type 2 Diabetes (T2D) status as determined from the EMR were tested for association with case/control status. The binary variables were analyzed using a Fisher’s exact test, and the continuous variables were analyzed through a two-tailed T-test to test mean differences between cases and controls.

**GENOTYPIC DATA**
The Illumina 660W-QUAD Beadchip array (Illumina, San Diego, California, USA) was used to generate genotype data on over 500 K SNPs from cases and controls. To ease automation of analysis, only data obtained from autosomes were analyzed. After filtering out SNPs with low minor allele frequency (<0.01), missing genotype data (≥ 0.05 of the study population with missing genotypes) or a significant departure from Hardy-Weinberg equilibrium (HWE) (p < 1E-5), there were 508,921 SNPs that passed the quality control screening and were used in subsequent analysis. For a sample to be included in the analysis, we ensured that it had at least 99% of the non-missing SNPs and for each SNP to be included in the analysis, we required that the SNP should have at least 95% of the non-missing subjects.

**POPULATION STRATIFICATION ESTIMATION USING EIGENSOFT**
We used the EIGENSOFT program (Price et al., 2006) to examine population stratification in our dataset. The program combines population genetics methods and the use of PCA to explicitly capture ancestry differences between cases and controls along continuous axes of variation.

**SNP-BASED ANALYSIS USING PLINK AND R-PACKAGE**
PLINK (Purcell et al., 2007), a whole genome association analysis tool set, was used to perform the filtering and QC procedures (as described under “Genotypic data” above) on the raw dataset to generate the data set employed in this study. R is a free software programming language and software environment for statistical computing and graphics (R. Core Team, 2013). The glm() function within R was used to establish the logistic model assuming an additive mode of inheritance with adjustment for the following risk factors: age, gender, BMI, diabetes and the top three principal components estimated from EIGENSOFT for population stratification. The p-values calculated from the logistic model were used to test the associations between SNPs and case/control phenotype.

**GENE-BASED ANALYSIS USING VEGAS**
VEGAS is a software program that tests association between a gene and phenotype trait. VEGAS uses SNP-level data to incorporate information from a full set of markers annotated to each gene and accounts for linkage disequilibrium (LD) between markers (Neale and Sham, 2004; Liu et al., 2010). All SNPs annotated to a gene were used in the calculation, and the method adjusted by the linkage disequilibrium structure by using HapMap data. Monte Carlo simulations from multivariate Gaussian random variables and Cholesky decomposition matrices were employed by VEGAS to produce disease association p-values per gene, correcting for the correlation structure between nearby SNPs. This type of analysis has several advantages including the collapsing of effects for all genotyped SNPs within each gene, and reducing the multiple testing burdens. The LD structure of each gene region is factored into the analysis through applying decomposition matrices in the analysis, effectively factoring-out the correlational structure between tightly linked SNPs. Additionally, gene-based results enable the subsequent use of many pathway analysis packages designed to use a single measurement from each gene.
PATHWAY-BASED ANALYSIS BY DAVID, PANTHER, AND INGENUITY

PATHWAY ANALYSIS (IPA) PROGRAMS

All genes with a p-value = 0.01 from the VEGAS analysis were used as input to perform separate pathway analysis by DAVID, PANTHER, and IPA. The DAVID analyses consisted of three steps: measurement of the functional relationship of gene pairs, a DAVID agglomeration procedure to partition genes into functional gene groups, and visualization of the results (Huang et al., 2007, 2009a,b). PANTHER is a publicly-available database having gene ontology, functional annotation, and evolutionary conservation information. PANTHER Pathways are available within the database and these pathways were used in our analysis. We calculated p-values for enrichment of S. aureus-associated genes within specific PANTHER pathways using a standard hypergeometric statistical approach. The statistical over-representation test was implemented in the PANTHER program. A binomial test was used to compare our gene list to a reference list (all human genes) to determine over- or under-representation of genes from our list in PANTHER gene function categories using an experiment-wise approach (experiment-wise α = 0.05) (Mi and Thomas, 2009). Gene-based analyses using IPA (Ingenuity® Systems, www.ingenuity.com, Redwood City, California, USA, www.ingenuity.com) were used to generate protein-protein interaction networks.

RESULTS

POPULATION STRATIFICATION

Using EIGENSOFT to perform PCA of the 508,921 genotypes from all study subjects (without knowledge of case/control status), we found no evidence of strong population stratification (Supplemental Figure 1). Therefore none of the 3234 study subjects were excluded because of population stratification.

STUDY SUBJECT CHARACTERISTICS

It has been previously noted that male sex, age, high BMI and T2D are risk factors for invasive S. aureus infection (Graffunder and Venezia, 2002). As expected, the percent of males was significantly greater in the cases than the control group (51 vs. 39%; p = 5.50E-05) (Table 1). There was no significant difference in mean age of case and controls. The prevalence of T2D was significantly higher among cases than controls (22.3 vs. 12.7%; p = 1.03E-05) and so was the mean BMI (cases: 32.1 vs. control: 29.5 kg/m^2; p = 3.50E-8). Our findings are consistent with those of previous reports.

Q-Q PLOT

We performed Q-Q analysis on the p-values obtained using logistic model assuming additive model of inheritance (Figure 1). The plot showed no evidence of population stratification, confounding effects, or systematic bias in the results from the statistical routines employed.

HWE

Eight hundred forty three SNPs were excluded from the analysis due to departure from HWE exceeding α = 1E-05.

SNP-BASED ANALYSES

SNP associations were tested using logistic regression analysis after adjusting for risk factors such as age, gender, BMI, diabetes and three principal components. No single SNP in the GWAS reached the level of genome-wide significance (p < 5 × 10^{-8}) (Figure 2). However, four SNPs exceeded the p < 10^{-5} threshold, including two SNPs (rs2455012 and rs7152530) with a p < 10^{-7} (Table 2). Out of these four SNPs, two were intronic (PDE4B and TXNRD2), one was intergenic with respect to VRK1 and BCL11B, and one was in the 3’UTR of PNPLA5. Of the four SNPs on chromosome 14 (Table 2), rs1899234 was in weak linkage disequilibrium (LD) with the other three SNPs (r^2 < 0.42) which were in strong LD with each other (r^2 = 0.80). The two SNPs in XRN1 were in strong LD (r^2 = 0.92), two of the SNPs on chromosome 22 (rs470093 and rs9614174) were in moderate LD (r^2 = 0.49), and the two SNPs on chromosome 19 exhibited low LD (r^2 = 0.09).

VEGAS-BASED GENE ANALYSES

Table 3 shows the top 15 genes ranked by their p-values from the VEGAS analysis and four additional interesting genes that could potentially have a role in S. aureus-caused diseases based on their known involvement in immune and inflammatory processes. The topmost hit was NMRK2 (or ITGB1BP3; p = 1.20E-05) which encodes nicotinamide riboside kinase 2. Two of the 15 genes also featured in the list from the SNP-based analysis: DAPK3 (p = 5.10E-05) and XRN1 (p = 1.85E-04).

PATHWAY ANALYSES

The top-ranked 196 genes (p-value = 0.01) resulting from the VEGAS analysis were selected for the DAVID analyses but none of the gene groups were statistically significant. One of the top gene groups (gene group 1; Supplemental Table 1) included CST8 (cystatin 8), SERPINA6 (serine peptidase inhibitor, clade A, member 6), SERPINA10 (serine peptidase inhibitor, clade A, member 10), and SPINK1 (serine peptidase inhibitor, Kazak type 1). The enriched genes in group 2 (Supplemental Table 1) included four keratin genes: KRT24 (form intermediate filament), KRT82 (type II hair keratin), KRT12 (type I intermediate filament keratin 12), and KRT75 (form intermediate filament in the cytoplasm of epithelial cells). As expected, the PANTHER analysis also suggested enrichment for intermediate filament cytoskeleton pathway (Supplemental Table 2). Using the same gene set input as in DAVID and PANTHER, the IPA was also explored with the intent...
Ye et al. Genetic susceptibility to \textit{S. aureus} infection

**FIGURE 1 |** The Q-Q plot of the \( p \)-values from all 508,921 SNPs. The x-axis shows the expected \(-\log_{10}(p\text{-value})\). The y-axis shows the observed \(-\log_{10}(p\text{-value})\).

**FIGURE 2 |** The manhattan plot of the \( p \)-values from all 508,921 SNPs. The x-axis shows the chromosome numbers. The y-axis is the \(-\log_{10}(p\text{-value})\). The blue line is \( p \)-value of \(10^{-5}\) whereas the red line shows the genome wide significance \( p \)-value of \(5 \times 10^{-8}\).

of hypothesis-generating protein-protein interaction networks associated our gene data set. The IPA yielded 12 protein networks (data not shown) of which network 7, involving cell death, and survival, appeared interesting (Supplemental Figure 2).

**DISCUSSION**

Multiple diseases (endocarditis, skin infections, etc.) caused by \textit{S. aureus} are mediated by two main classes of virulence factors, adhesion and secretory proteins, which interact with host receptors and inflammatory/anti-inflammatory molecules to produce the disease phenotype (Gordon and Lowy, 2008). Adhesion proteins help the pathogen to attach to the skin and survive on the epidermis and in the sub-epidermal layer through a repertoire of molecules collectively known as MSCRAMMS (microbial surface components recognizing adhesive matrix molecules). The MSCRAMMS can bind to fibronectin, fibrinogen, and platelets among others. Subsequent to attachment, \textit{S. aureus} can secrete tissue and organ-specific virulence proteins (e.g., coagulase, proteases, toxins, superantigens) with a wide range of virulence functions that enable the pathogen to infect its host.

So far, most of the genetic susceptibility data related to \textit{S. aureus} has been limited to \textit{S. aureus} colonization. For example, the glucocorticoid receptor gene polymorphisms are associated with carriage risk (van den Akker et al., 2006), whereas \textit{DEFB1} has
been shown to promote persistent colonization by modulating beta-defensin expression in keratinocytes (Nurjadi et al., 2013). In contrast, the Danish middle-aged/elderly twin study showed that host genetics had a modest influence only on *S. aureus* carrier state (Andersen et al., 2012). While our manuscript was under review, a study by Nelson et al did not find any SNP that has genome wide significance for association with *S. aureus* bacteremia (SAB) although an intronic SNP in CDON was speculated to be associated with complicated SAB (Nelson et al., 2014). In a report by Stappers et al, four SNPs from three toll-like receptor (TLR) genes, TLR1, TLR2, and TLR6, increase the susceptibility to complicated skin and soft tissue infections caused by staphylococci, streptococci, and enterococci (Stappers et al., 2014).

In our multi-tiered GWAS-based investigation, we have identified a number of potentially interesting genes that need further investigation. Although the individual SNP results did not pass genome-wide significance, the top-tier SNPs, Gene-based, and Pathway-based results were enriched for genes that have plausible functions in bacterial infections. This includes genes that have roles in intracellular signaling, inflammation, zinc transport, and integrin binding. Thus, these genes have relatively high prior probabilities for involvement in *S. aureus* infection susceptibility—an aspect of the results that we believe supports considerable interest in subsequent studies to follow-up on these findings.

Notably, two genes (DAPK3 and XRN1) were identified in both the SNP-based and gene-based analyses. DAPK3 is a protein kinase that modulates apoptosis-related signaling pathways (Wu et al., 2010) and, in interaction with Rhod, modulates actin filament assembly and focal adhesion reorganization (Nehru et al., 2013). *S. aureus* is known to induce apoptosis of host cells during host invasion, leading to a compromised host immune response (Haslinger-Löffler et al., 2005). XRN1 encodes a 5′–3′ exonuclease family member involved in cellular mRNA turnover (Nagarajan et al., 2013). The gene is shown to complete host mRNA degradation initiated by viral pathogens (Gaglia et al., 2012). These functions suggest possible roles for DAPK3 and XRN1 in susceptibility to diseases caused by *S. aureus*.

Some of the other genes we identified have been implicated in infectious disease processes. *PDE4B* is involved in modulating bacteria-induced inflammation (Komatsu et al., 2013). *PNPLA5* appears to be critical for autophagosome functions, including microbial clearance (Dupont et al., 2014). Mammalian hosts are known to reduce the level of free zinc to thwart pathogen growth (Kehl-Fie and Skaar, 2010) and it is plausible that SLC39A8, a zinc transporter, could be involved. *BCL11B* encodes a transcriptional repressor involved in T-cell development (Wakabayashi et al., 2003). *NMRK2*, an integrin beta1 binding protein, could function in host responses to bacterial function based on the finding that host fibronectin forms a bridge between *S. aureus* fibronectin-binding proteins and host cell beta1 integrins during *S. aureus* cellular invasion (Fowler et al., 2000). Keratin intermediate filaments are shown to have a protective role during infection with *Bartonella henselae* in cat scratch disease (Zhu et al., 2013). Interleukin 1 cytokine family members (*IL1A, IL1B, IL1R1, IL1B, IL1R1*).

---

Table 2 | Top 20 SNPs, their chromosomal locations, associated genes, major and minor alleles, minor allele frequency of cases and controls, p-value and odds ratio and 95% confidence interval.

| db SNP ID | Chr | Position (hg19) | Nearby genes | Variant location | Major/minor alleles | Case MAF* | control MAF | P-value | Odds ratio (95% CI) |
|-----------|-----|----------------|--------------|-----------------|---------------------|-----------|-------------|--------|------------------|
| rs2455012 | 1   | 66520998       | *PDE4B*      | Intron          | C/T                 | 0.096015  | 0.045190    | 6.68E-07 | 2.17 (1.59, 2.93) |
| rs7152530 | 14  | 99641215       | *VRK1, BCL11B* | Intergenic      | G/A                 | 0.270227  | 0.366746    | 8.47E-07 | 0.62 (0.51, 0.75) |
| rs3804047 | 22  | 19879637       | *TXNRD2*     | Intron          | T/C                 | 0.368932  | 0.281143    | 3.32E-06 | 1.52 (1.27, 1.81) |
| rs470093  | 22  | 44276171       | *PNPLA5*     | 3’ UTR          | G/T                 | 0.211974  | 0.146838    | 7.19E-06 | 1.63 (1.31, 2.02) |
| rs1381281 | 14  | 98615771       | *VRK1, BCL11B* | Intergenic      | A/G                 | 0.322006  | 0.410024    | 1.07E-05 | 0.67 (0.56, 0.80) |
| rs13107325 | 4  | 103188709      | *SLC39A8*    | Exon B          | G/A                 | 0.113269  | 0.065299    | 1.09E-05 | 1.88 (1.41, 2.47) |
| rs1892234 | 14  | 99720904       | *VRK1, BCL11B* | Intergenic      | T/C                 | 0.364078  | 0.455897    | 1.38E-05 | 0.68 (0.57, 0.81) |
| rs6948646 | 7   | 152945284      | *ACTR3B, DPP6*| Intergenic      | G/A                 | 0.444984  | 0.362393    | 1.41E-05 | 1.46 (1.23, 1.73) |
| rs1012203 | 9   | 104626708      | *GRIN3A, CYLC2* | Intergenic      | A/G                 | 0.048544  | 0.104103    | 1.59E-05 | 0.44 (0.29, 0.62) |
| rs987514  | 14  | 98628943       | *VRK1, BCL11B* | Intergenic      | C/T                 | 0.330097  | 0.415356    | 1.95E-05 | 0.68 (0.56, 0.81) |
| rs11950651 | 5  | 103009768      | *NUDT12, EFNA5* | Intergenic      | C/T                 | 0.032382  | 0.079829    | 2.09E-05 | 0.37 (0.22, 0.57) |
| rs12160908 | 22 | 25928620       | *LRF6L, ADRK2* | Intergenic      | C/T                 | 0.067961  | 0.127401    | 2.13E-05 | 0.49 (0.35, 0.67) |
| rs12696090 | 3  | 158669688      | *MFSD1, IQCJ-SCHIP1* | Intergenic | G/T                 | 0.150458  | 0.226635    | 2.72E-05 | 0.61 (0.48, 0.76) |
| rs9614174 | 22  | 44082454       | *EFCA6B*     | Intron          | A/C                 | 0.215210  | 0.151795    | 2.78E-05 | 1.56 (1.26, 1.92) |
| rs7255123 | 19  | 3958397       | *DAPK3*      | 3’ near gene    | G/A                 | 0.263754  | 0.198632    | 4.11E-05 | 1.51 (1.24, 1.84) |
| rs6135407 | 20  | 15397676       | *MACROD2*    | Intron          | C/T                 | 0.220065  | 0.158858    | 4.91E-05 | 1.54 (1.25, 1.90) |
| rs4807532 | 19  | 3928369       | *ATCAY*      | 3’ near gene    | G/A                 | 0.461165  | 0.375897    | 5.32E-05 | 1.42 (1.20, 1.69) |
| rs7643377 | 3   | 14211469        | *XRN1*       | Intron          | C/T                 | 0.483819  | 0.400342    | 5.70E-05 | 1.41 (1.19, 1.67) |
| rs2535368 | 7   | 83157207       | *SEMA3E*    | Intron          | G/A                 | 0.202265  | 0.144494    | 5.90E-05 | 1.56 (1.25, 1.93) |
| rs9867210 | 3   | 142132749      | *XRN1*       | Intron          | T/C                 | 0.495146  | 0.411453    | 6.17E-05 | 1.41 (1.19, 1.67) |

*Minor allele frequency.
and IL1RL2) are known mediators of immune and inflammatory responses (Garlanda et al., 2013).

It is known that many Mendelian and oligogenic immunodeficiency disorders confer risk to staphylococcal infection including lymphocyte deficiencies such as severe combined immunodeficiency, chronic granulomatous disease, and hyper-IgE syndrome (Stephan et al., 1993; Grimbacher et al., 1999; Van de Vosse et al., 2009). These disorders are typified by highly disruptive mutations occurring in genes central to lymphoid cell competency including STAT3, JAK3, DOCK8, and CD18, among others (Hogg et al., 1999; Kalman et al., 2004; Jiao et al., 2008; Zhang et al., 2009). However, most cases of severe staphylococcal infection are not attributable to these more rare conditions and have unknown genetic etiology.

In summary, this preliminary GWAS applied a SNP-to-gene-to-disease-pathway approach to identify susceptibility genes against a broad umbrella of laboratory confirmed S. aureus infections. While no one SNP and gene was found to be highly significant in this study, we suspect that for a versatile pathogen like S. aureus, that needs to overcome barriers presented by a variety of tissues and defense systems to infect various sites in the body, there are bound to be several genes involved in host susceptibility. Not everyone exposed to a virulent or a colonizing strain of S. aureus has similar severity of infection. It is reasonable to speculate that effects of variants segregating at multiple genes contribute to the severity of S. aureus infection. Similarly, there could be protective alleles that may lower the risk of clinically-attended infection as well. Additional studies will be needed to confirm these findings but eventually functional studies will be needed to illuminate the detailed mechanisms of how these variants confer predisposition to infection. Once consensus disease loci and pathways are identified, they can serve as targets for future pharmaceutical development and further elucidation of how aberrant cellular processes/signaling give rise to Staphylococcus-induced pathologies.

ACKNOWLEDGMENTS
We would like to thank Crystal Jacobson, Amy Aswani, and Rob Strenn for very helpful informatics expertise and Dr. Matt Hall for valuable discussions and advice regarding the clinical aspects of S. aureus infection. We would also like to sincerely thank the PMRP participants. The project described was supported by the Clinical and Translational Science Award (CTSA) program, previously through the National Center for Research Resources (NCRR) grant 1UL1RR025011 and the National Center for Advancing Translational Sciences (NCATS) grant 9U54TR00021, and now by the NCATS grant UL1TR000427. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. The eMERGE Network is funded by the NHGRI, with additional funding from the National Institute of General Medical Sciences through the following grants: U01HG004438 to the Center for Inherited Disease Research; U01HG004608 to Essentia Institute for Rural Health/Marshfield Clinic Research Foundation. This study was also funded through support from the Marshfield Clinic, the Wisconsin Genomics Initiative, and philanthropic gifts in support of medical research at the Marshfield Clinic.

SUPPLEMENTARY MATERIAL
The Supplementary Material for this article can be found online at: http://www.frontiernet.org/journal/10.3389/2014.00125/abstract

Supplemental Figure 1 | A principal components analysis of genome-wide genotypes in all study subjects.

Supplemental Figure 2 | One of the IPA outputs of the protein-protein interaction network using VEGAS based 196 genes (p \( \leq 0.01 \)) as input.

Three genes, DAPK3, KRT12, and TXNOR2 have been shown to be in direct network UBC (ubiquitin C). Boxes in pink color are input genes from the VEGAS analysis. Network shapes Symbols: \( \phi = \text{enzyme}; \circ = \text{Other}; \bigcirc = \text{Kinase}; \bigotimes = \text{direct interaction}; \bigotimes = \text{indirect interaction}. Additional details about the symbols can be found at www.ingenuity.com.

REFERENCES
Andersen, P. S., Pedersen, J. K., Fode, P., Skov, R. L., Fowler, V. G. Jr., Stegger, M., et al. (2012). Influence of host genetics and environment on nasal carriage of Staphylococcus aureus in Danish middle-aged and elderly twins. J. Infect. Dis. 206, 1178–1184. doi: 10.1093/infdis/jis491

Burton, P. R., Clayton, D. G., Carden, L. R., Craddock, N., Deloukas, P., Duncanson, A., et al. (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447, 661–678. doi: 10.1038/nature05911

Cross, D. S., McCarty, C. A., Hytopoulos, E., Beggs, M., Nolan, N., Harrington, D., et al. (2012). Coronary risk assessment among intermediate risk patients using

Ye et al. Genetic susceptibility to S. aureus infection.
Kalman, L., Lindegren, M. L., Kobrynski, L., Vogt, R., Hannon, H., Howard, J. T., et al. (2004). Mutations in genes required for T-cell development: a review. *Hum. Mol. Genet.* 13, 2401–2408. doi: 10.1093/hmg/ddh268

Kuhnert, M. J., Kriszun-Moran, D., Hill, H. A., McQuillan, G., McAllister, S. K., Fosheim, G., et al. (2006). Prevalence of *Staphylococcus aureus* nasal colonization in the United States, 2001-2002. *J. Infect. Dis.* 193, 169–171. doi: 10.1086/499632

Liu, J. Z., McRae, A. F., Nyholt, D. R., Medland, S. E., Wan, N. R., Brown, K. M., et al. (2010). A versatile genome-based test for genome-wide association studies. *Am. J. Hum. Genet.* 87, 139–145. doi: 10.1016/j.ajhg.2010.06.009

Lowy, F. D. (1998). *Staphylococcus aureus* infections. *N. Engl. J. Med.* 339, 520–532. doi: 10.1056/NEJM199808203390906

McCarty, C. A., Chisholm, R. L., Chute, C. G., Kulo, I. J., Jarvi, G. P., and Larson, E. B. (2011). The EMERGE Network: a consortium of biorepositories linked to electronic medical records data for conducting genomic studies. *BMC Med. Genomics* 4:13. doi: 10.1186/1755-8794-4-13

Nurjadi, D., Herrmann, E., Hinderberger, I., and Zanger, P. (2013). Impaired β-defensin expression in human skin links DEFB1 promoter polymorphisms with persistent *Staphylococcus aureus* nasal carriage. *J. Infect. Dis.* 207, 666–674. doi: 10.1093/infdis/jit735

Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., Reich, D., et al. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* 38, 904–909. doi: 10.1038/ng1847

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., et al. (2007). PLINK: a toolset for whole-genome association and population-based linkage analysis. *Am. J. Hum. Genet.* 81, 559–575. doi: 10.1086/519795

R Core Team. (2013). "R: a language and environment for statistical computing," in *R Foundation for Statistical Computing*, Vienna, Austria. Available online at: http://www.R-project.org

Rehm, S. J. (2008). *Staphylococcus aureus*: the new adventures of a legendary pathogen. *Clever. Clin. Med.* 7, 177–180, 183–186, 190–192. doi: 10.3949/ccmj.7.3.177

Stappers, M. H., Thys, Y., Oosting, M., Plantinga, T. S., Ioana, M., Reimnitz, P., et al. (2014). TLR1, TLR2, and TLR6 gene polymorphisms are associated with increased susceptibility to complicated skin and skin structure infections. *J. Infect. Dis.* 209, 1099–1105. doi: 10.1093/infdis/jiu908. [Epub ahead of print]

Stephan, J. L., Vlekova, V., Le Deist, F., Blanche, S., Donadieu, J., De Saint-Basile, G., et al. (1993). Severe combined immunodeficiency: a retrospective single-center study of clinical presentation and outcome in 117 patients. *J. Pediatr.* 123, 564–572. doi: 10.1016/0022-3476(93)90390-T

Van de Vosse, E., van Wengen, A., van Geelen, J. A., de Boer, M., Roos, D., and van Dissel, J. T. (2009). A novel mechanism in NCF1 in an adult CDG patient with a liver abscess as first presentation. *J. Hum. Genet.* 54, 313–316. doi: 10.1038/jhg.2009.24

van den Akker, E. L., Nouwen, J. L., Melles, D. C., van Rossum, E. F., and Melles, D. C., van Rossum, E. F., Koper, J. W., van den Akker, E. L., Nouwen, J. L., Melles, D. C., van Rossum, E. F., and Koper, J. W., Uitterlinden, A. G., et al. (2006). *Staphylococcus aureus* nasal carriage is associated with glucocorticoid receptor gene polymorphisms. *J. Infect. Dis.* 194, 814–818. doi: 10.1086/506367

Wakabayashi, Y., Watanabe, H., Inoue, J., Takeda, N., Sakata, J., Mishima, Y., et al. (2003). Bcl11b is required for differentiation and survival of alphabeta T lymphocytes. *Nat. Immunol.* 4, 533–539. doi: 10.1038/ni927

Wu, Y., Yan, Q., Zuo, J., Saiyin, H., Jiang, W., Qiao, S., et al. (2010). Link of Dlk/ZIP kinase to cell apoptosis and tumor suppression. *Biochem. Biophys. Res. Commun.* 392, 510–515. doi: 10.1016/j.bbrc.2010.01.054
Ye et al. (2011). Genome partitioning of genetic variation for complex traits using common SNPs. *Nat. Genet.* 43, 519–525. doi: 10.1038/ng.823

Zhang, Q., Davis, J. C., Lamborn, I. T., Freeman, A. F., Jing, H., Favreau, A. J., et al. (2009). Combined immunodeficiency associated with DOCK8 mutations. *N. Engl. J. Med.* 361, 2046–2055. doi: 10.1056/NEJMoa0905506

Zhu, C., Bai, Y., Liu, Q., Li, D., Hong, J., Yang, Z., et al. (2013). Depolymerization of cytokeratin intermediate filaments facilitates intracellular infection of HeLa cells by Bartonella henselae. *J. Infect. Dis.* 207, 1397–1405. doi: 10.1093/infdis/jit040

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.