Haplotype Association Mapping Identifies a Candidate Gene Region in Mice Infected With Staphylococcus aureus

Nicole V. Johnson,*1,2 Sun Hee Ahn,†,2 Hitesh Deshmukh,‡ Mikhail K. Levin,§ Charlotte L. Nelson,** William K. Scott,†† Andrew Allen,§ Vance G. Fowler, Jr.,† and Lindsay G. Cowell§†
*Department of Biostatistics and Bioinformatics and †Department of Medicine, Duke University Medical Center, Durham, North Carolina 27710, ‡Department of Pediatrics, University at Buffalo, Buffalo, New York 14222, §Department of Clinical Sciences, University of Texas Southwestern Medical Center, Dallas, Texas 75390, **Duke Clinical Research Institute, Duke University Medical Center, Durham, North Carolina 27705, and ††Dr. John T. Macdonald Foundation Department of Human Genetics and John P. Hussman Institute for Human Genomics, University of Miami Miller School of Medicine, Miami, Florida 33136

ABSTRACT Exposure to Staphylococcus aureus has a variety of outcomes, from asymptomatic colonization to fatal infection. Strong evidence suggests that host genetics play an important role in susceptibility, but the specific host genetic factors involved are not known. The availability of genome-wide single nucleotide polymorphism (SNP) data for inbred Mus musculus strains means that haplotype association mapping can be used to identify candidate susceptibility genes. We applied haplotype association mapping to Perlegen SNP data and kidney bacterial counts from Staphylococcus aureus-infected mice from 13 inbred strains and detected an associated block on chromosome 7. Strong experimental evidence supports the result: a separate study demonstrated the presence of a susceptibility locus on chromosome 7 using consomic mice. The associated block contains no genes, but lies within the gene cluster of the 26-member extended kallikrein gene family, whose members have well-recognized roles in the generation of antimicrobial peptides and the regulation of inflammation. Efficient mixed-model association (EMMA) testing of all SNPs with two alleles and located within the gene cluster boundaries finds two significant associations: one of the three polymorphisms defining the associated block and one in the gene closest to the block, Klk1b11. In addition, we find that 7 of the 26 kallikrein genes are differentially expressed between susceptible and resistant mice, including the Klk1b11 gene. These genes represent a promising set of candidate genes influencing susceptibility to Staphylococcus aureus.

Staphylococcus aureus (S. aureus) causes a diverse array of clinical conditions in humans, ranging from asymptomatic colonization to endocarditis and death. The factors influencing infection severity, however, are not fully understood. While it is clear that bacterial (Ferry et al. 2005) and environmental factors (Laupland et al. 2003) influence the type and severity of S. aureus infections, an increasing body of evidence suggests that host genetics also play an important role. For example, higher rates of S. aureus infection have been observed among genetically distinct populations, including New Zealand Maori (Hill et al. 2001), Canadian Native Americans (Embíl et al. 1994), and Australian Aboriginals (Tong et al. 2009). Familial clusters of recurrent S. aureus infections (Noble et al. 1967; Zimakoff et al. 1988) and rare genetic conditions characterized by susceptibility to S. aureus (e.g. Job Syndrome, Chediak-Higashi Syndrome) (Fagél et al. 1998; Hill et al. 2001) have also been described. In addition, several recent studies have demonstrated the importance of host genetics in susceptibility to S. aureus using inbred strains of mice. For example,
of S. aureus into the peritoneal fluid. Kidney samples were collected from all 76 mice; peritoneal fluid was obtained from 68 of the 76 mice. Kidneys and peritoneal fluid were collected and processed as previously described (Deshmukh et al. 2009). Briefly, kidneys were collected from killed mice and homogenized in phosphate buffered saline (PBS). Kidney homogenate and peritoneal fluid were diluted serially with PBS 10-fold. The serial dilutions (50 µl) were plated on Tryptic Soy Agar plates, incubated (37°C) overnight, and bacterial colonies were counted. For both phenotypes, the average bacterial load in the kidney and peritoneal fluid was counted.

The between-strain variance ($\sigma_b^2$) and within-strain variance ($\sigma_w^2$) were estimated from the trait heritability ($H^2$) was estimated for the kidney count, peritoneal fluid count, and serum IL-6 phenotypes using

$$H^2 = \frac{\sigma_b^2}{\sigma_b^2 + \sigma_w^2}$$

where $\sigma_b^2$ is the between-strain variance and $\sigma_w^2$ is the within-strain variance (Anholt and Mackay 2009). $\sigma_b^2$ and $\sigma_w^2$ were estimated from the between-strain and within-strain mean squared errors obtained.

### MATERIALS AND METHODS

#### Survival data

Mice from 13 inbred strains (N = 142) (A/J, BALB/cByJ, 129S1/SvImJ, AKR/J, NZW/LacJ, PWD/PhJ, KK/HJ, FVB/NJ, BTBR Tfl/J, DBA/2J, C3H/HeJ, NOD/ShiLtJ, and C57BL/6J) were intraperitoneally injected with $10^7$ colony forming units per gram (CFU/g) of the methicillin-susceptible Sanger 476 strain of S. aureus (Table 1). Mice were monitored every eight hours for five days, and survival times were recorded in hours. The median survival time was calculated for each strain and used for HAM. Mice were euthanized using CO₂ asphyxiation if they appeared moribund. Pain and distress were assessed using a numerical scale for the following characteristics: appearance (0 = normal; 1 = lack of grooming; 2 = rough hair coat; 3 = abnormal posture); behavior (0 = normal; 1 = moving slowly; 2 = moving slowly, irregular ambulation; 3 = immobile). A total score (appearance plus behavior) of three indicated significant pain and distress culminating in the early euthanasia of the animal. A log-rank test was used to detect statistically significant differences between the survival curves.

### Table 1 Number of mice injected with S. aureus for each of 13 inbred mouse strains

| Strain          | Survival | Kidney Count | Peritoneal Fluid | Serum IL-6 |
|-----------------|----------|--------------|-----------------|------------|
| A/J             | 15       | 10           | 12              | 3          |
| BALB/cByJ       | 10       | 4            | 4               | 3          |
| 129S1/SvImJ     | 10       | 8            | 7               | 3          |
| AKR/J           | 14       | 5            | 4               | 3          |
| NZW/LacJ        | 10       | 5            | 4               | 4          |
| KK/HJ           | 10       | 5            | 4               | 3          |
| PWD/PhJ         | 10       | 4            | 4               | 3          |
| FVB/NJ          | 8        | 8            | 3               | 3          |
| BTBR Tfl/J      | 10       | 5            | 3               | 3          |
| DBA/2J          | 10       | 5            | 4               | 3          |
| C3H/HeJ         | 15       | 5            | 4               | 3          |
| NOD/ShiLtJ      | 10       | 3            | 3               | 3          |
| C57BL/6J        | 10       | 9            | 12              | 9          |
| **Total**       | 142      | 76           | 68              | 46         |

#### Bacterial load in the kidney and peritoneal fluid

A total of 76 mice from 13 strains were injected with S. aureus as described above (Table 1). Mice were killed 24 hr post infection. Kidneys were collected from all 76 mice; peritoneal fluid was obtained from 68 of the 76 mice. Kidneys and peritoneal fluid were collected and processed as previously described (Deshmukh et al. 2009). Briefly, kidneys were collected from killed mice and homogenized in phosphate buffered saline (PBS). Kidney homogenate and peritoneal fluid were diluted serially with PBS 10-fold. The serial dilutions (50 µl) were plated on Tryptic Soy Agar plates, incubated (37°C) overnight, and bacterial colonies were counted. Both types, the average number of CFU per either gram or milliliter was calculated for each strain and used for HAM. Single-factor ANOVA was used to detect statistically significant differences between the strain mean phenotype values.

#### Serum levels of IL-6

A total of 46 mice from 13 strains were injected with S. aureus as described above (Table 1). Mice were killed at 24 hr post infection. Blood was obtained by intracardiac puncture, serum was separated by centrifugation, and the amount of total protein was measured using the bicinchoninic acid method kit (Pierce). Serum samples were diluted to ensure equal amounts of total protein, and the levels of IL-6 were estimated using an enzyme-linked immunosorbent assay (Due kit, Invitrogen). Single-factor ANOVA was used to detect statistically significant differences between the strain mean phenotype values.

#### Trait heritability

Broad-sense heritability ($H^2$) was estimated for the kidney count, peritoneal fluid count, and serum IL-6 phenotypes using
by single-factor ANOVA. These heritability estimates likely overestimate the true heritability as two of the strains used, C57BL/6J and A/J, were selected for study as representative of extreme S. aureus susceptibility phenotypes (Hegmann and Possidente 1981).

Haplotype association mapping

The overall approach is outlined in Figure 1A. Briefly, genotype data were obtained from the Perlegen SNP database (http://phenome.jax.org/db/q?rtn=projects/projdet&reqprojd=198) for the 13 inbred mouse strains for which phenotype data were collected. SNPs with genotype information available for all 13 strains (3,260,963 SNPs) were used for HAM. Overlapping three-SNP windows were formed using three adjacent SNPs, as described in Fletcher et al. (2004). For each window, mouse strains were assigned to haplotype groups based on the three SNP alleles within the window, thereby forming the strain segregation pattern for the window (Figure 1B). Adjacent three-SNP windows with the same strain segregation pattern were combined to form a single haplotype block (Figure 1C). Haplotype blocks with only a single haplotype were excluded from the analysis. Among the resulting 1,364,341 haplotype blocks, 66,906 unique strain segregation patterns were observed. For each of the four phenotypes, a test of association was conducted for each pattern using the weighted F statistic defined in Fletcher et al. (2004).

To determine which strain segregation patterns demonstrate a significant genotype-phenotype association, we applied a modified version of the minP method described in Ge et al. (2003). Briefly, for each phenotype, a permutation analysis (described below) was conducted to generate a null distribution of F statistics (\(F_p\)) for each strain segregation pattern. The \(F_p\) distribution for a pattern was used to estimate the type I error (\(p_F\)) for the observed F statistic (\(F_p\)) for that pattern by counting the number of \(F_p \geq F_p\) and dividing by the number of permutations (five million). Type I error (\(p_F\)) was estimated for each \(F_p\) in the same way. For each permutation, the minimum \(p_F\) (min \(p_F\)) across all unique strain segregation patterns was captured and used to generate a null distribution of genome-wide minimum \(p_F\) values. Genome-wide type I error was then estimated for each phenotype by dividing the total number of \(p_F\) smaller than the minimum \(p_F\) (min \(p_F\)) for that phenotype by the number of permutations (five million). In addition, for each phenotype, the 5th percentile (\(p_{0.05}\)) of the null min \(p_F\) distribution was used as the significance threshold for identification of significantly associated genomic regions: haplotype blocks for which the observed \(p_F\) was smaller than \(p_{0.05}\) were deemed to be significantly associated with phenotype.

A permutation analysis was conducted for each phenotype as follows. For the bacterial count and serum IL-6 phenotypes, a mean and variance were calculated for each strain from the observed phenotype values of the individual mice and were used to parameterize 13 phenotype distributions. For a single permutation, a strain phenotype value was sampled from each of the 13 distributions, and the set of 13 values was permuted. An \(F_p\) statistic was then calculated for each of the 66,906 unique strain patterns using the set of permuted strain phenotype values. For the survival phenotype, the median survival times for the 13 strains were permuted. For all four phenotypes, five million permutations were run.

The permutations were run on the high-performance compute cluster Lonestar at the Texas Advanced Computing Center (http://www.tacc.utexas.edu/) at the University of Texas at Austin.

Efficient Mixed-Model Association

The population structure and genetic relatedness of inbred mouse strains is known to result in inflated false positives in tests of genotype-phenotype association (Kang et al. 2008). Efficient mixed-model association (EMMA) conducts tests for association on single SNPs with two alleles, correcting for population structure and genetic relatedness in model organism association mapping (Kang et al. 2008). To assess the likelihood that the genomic region identified by the above-described HAM procedure represents population structure artifact, we used EMMA to conduct a test of association for all two-allele SNPs within the region. We used the publicly available R-package implementation of EMMA (available at http://mouse.cs.ucla.edu/emma/). We corrected for multiple testing using a BH-adjusted \(P\) value computed using the R package mt.rawp2adjp. The BH option utilizes the Benjamini and Hochberg step-up FDR-controlling procedure (Benjamini and Hochberg 1995).

Gene expression data

Gene expression studies were conducted on mice from 6 of the 13 inbred mouse strains, the 3 resistant strains (C57BL/6J, NOD/LtJ, and C3H/HeJ) and the 3 highly susceptible strains that had high average
kidney count values (A/J, AKR/J, and BALB/cByJ). Three mice from each strain were infected with *S. aureus* as described above, and blood was taken by intracardiac puncture 2 hr after infection. Blood was also taken from three uninfected mice of each strain. Blood was stored in RNA later at −20 °C. Total RNA was prepared from mouse blood using the Mouse RiboPure Blood RNA isolation kit (Ambion), and globin mRNA was removed using the Globinclean kit (Ambion). All RNA samples passed the quality criteria of the Agilent Bioanalyzer and were used for the analysis. One round of linear amplification was performed for all samples (Ambion MessageAmp Primier). Biotin-labeled cDNA was hybridized to Affymetrix Mouse Genome 430 2.0 Array-Chips for 16 hr at 45 °C following the manufacturer’s instruction. The arrays were then washed and labeled with streptavidin-phycocerythrin (strept-PE), and the signal was amplified using biotinylated anti-streptavidin followed by another round of staining with strep-PE. Washing and staining were performed on the Affymetrix fluidics station according to recommended protocols. Labeled gene chips were scanned using an Affymetrix Genechip Scanner 7G.

**Microarray data analysis**

Preprocessing was conducted using the Robust Multichip Analysis (RMA) (Irizarry *et al.* 2003) implementation in the Bioconductor “affy” package (http://www.bioconductor.org/), with an additional step to account for differences in probe hybridization resulting from SNPs between susceptible and resistant mice. The additional step is referred to as SNP masking (Walter *et al.* 2007) and is applied after background correction and quantile-quantile normalization but prior to the determination of probeset expression values. Genomic locations hybridized by each probe were obtained from the Ensembl database (http://www.ensembl.org/index.html), and these genomic locations were compared to the locations of SNPs for which at least one susceptible and one resistant strain have different alleles. Probes that hybridize to such SNPs within their target transcripts were excluded from the determination of probeset expression values. Probesets with fewer than four remaining probes were excluded from further analysis.

Twenty-five probesets on the Mouse 430 2.0 array were identified as described in Results and analyzed using ANOVA to determine whether there were statistically significant differences in the mean expression levels between susceptible and resistant mice. The following generalized linear model was used:

\[ Y_{ijkl} = B^i_j + B^2_j + T_k + S * T_{k,l} \]

where \( B^1 \) corresponds to IVT batch effects, \( B^2 \) to hybridization batch effects, \( T \) to infection main effects, and \( S + T \) to strain-infection interaction effects. Two factor levels were used for infection state: uninfected and infected. Two factor levels were used for strain: susceptible and resistant. False discovery rate (FDR)-adjusted \( P \) values were calculated using a false discovery rate of 0.1 (Storey and Tibshirani 2003).

**RESULTS**

**Phenotype varies with mouse strain for inbred mice infected with *S. aureus***

A wide range of values was observed for all four phenotypes measured in this study (Figure 2 and File S1). With regard to survival, there were three resistant strains for which none of the mice died (C57BL/6J, NOD/ShiLtJ, and C3H/HeJ), and four highly susceptible strains with median survival times that were ≤ 26 hr (A/J, BALB/cByJ, 129S1/SvImJ, and AKR/J) (Figure 2A). The remaining six strains exhibited intermediate median survival times (Figure 2A). A log-rank test indicated a statistically significant difference between the survival curves (\( P < 10^{-19} \)).

The three resistant strains, along with KK/HJ and FVB/NJ, had low average kidney count values (<10 CFU/g), whereas three of the highly susceptible strains, along with DBA/2J, had high average kidney count values (≥244 CFU/g) (Figure 2B). The remaining four strains, including the highly susceptible strain 129S1/SvImJ, had average kidney count values ranging from 20 CFU/g to 97 CFU/g (Figure 2B).

**HAM identifies a haplotype block on chromosome 7 significantly associated with kidney bacterial counts**

For each phenotype, we identified the smallest HAM \( p \) value genome-wide, \( \min p_{h} \) (Table 2). We found that, for the kidney count phenotype, the \( \min p_{h} \) represents a significant association between phenotype and a single haplotype block 1.9 kb in length located on chromosome 7 at 51,256,409–51,258,299 bp (B7). This block also has the genome-wide \( \max F_{h} \) for the kidney count phenotype. No other block was significantly associated with the kidney count phenotype, and none of the other phenotypes was found to be significantly associated with any block.

The distribution of \( F_{h} \) and \( p_{h} \) throughout the genome (Figure 3) reveals that the \( F_{h} \) corresponding to B7 is much larger than the \( F_{h} \) for any other block. Similarly, the \( p_{h} \) for B7 is much smaller than the \( p_{h} \) for any other block.

The strain-haplotype pattern for the single significantly associated block B7 consists of five haplotypes with the mouse strain assignments shown in Table 3.

Interestingly, 9 of the 13 mouse strains, those assigned to haplotypes GTC, GCC, and AAT, have an average kidney count value < 100 CFU/g.
Of the remaining 4 strains, the A/J, BALB/cByJ, and DBA/2J strains have average kidney counts ranging from 243 CFU/g to 264 CFU/g and share the haplotype GTT. The AKR/J strain does not share its haplotype with any other strain (ACT) and has an average kidney count value an order of magnitude higher than any other strain (10^13 CFU/g).

The block B7 does not contain any genes, but it lies within the cluster of the 26-member extended kallikrein (Klk) gene family of serine peptidases (Table 4). The gene closest to B7 is Klk1b11, located 1.2 kb away.

To assess the likelihood that the genomic region identified by the above-described HAM procedure represents population structure artifact or whether the result may be driven by one or more of the Klk genes, we used EMMA to conduct a test of association between the kidney colonization phenotype and the 486 two-allele SNPs within the Klk gene cluster (Kang et al. 2008). Two of the SNPs are significantly associated with phenotype (P = 0.000003694 and adjusted P = 0.0008977). One of the SNPs lies within the Klk1b11 gene, and the other is one of the three SNPs that define the boundaries of the HAM-detected block B7.

This result indicates that there is association between the HAM-detected genomic region and phenotype even when the association

| Table 2 Genome-wide minimum observed p value and associated type I error estimate for each phenotype |
|-----------------------------------------------------|
| **Phenotype** | **min p_o** | **F_o** | **Type I Error for min p_o** |
|----------------|-------------|--------|-----------------------------|
| Median survival | 0.0000042 | 40.08  | 0.53                        |
| Serum IL6       | 0.000187   | 63.51  | 0.68                        |
| Peritoneal fluid | 0.000055 | 99.80  | 0.24                        |
| Kidney          | 0.000002   | 948.32 | 0.02                        |
that the highly susceptible strains do not effectively clear \textit{S. aureus} from the site of infection, do not mount an effective IL-6–mediated inflammatory response, and are subject to extensive bacterial dissemination.

We applied HAM as described in Pletcher et al. (2004) to the four sets of phenotype data and detected one significant association, which was between bacterial colonization of the kidney and a single haplotype block, B\textsubscript{7}, on chromosome 7. There is strong experimental evidence that B\textsubscript{7} is in fact linked to one or more causal variants: in a separate study conducted by our group, the presence of a causal variant on chromosome 7 was demonstrated using consomic mice created from the highly susceptible \textit{A/J} strain and the highly resistant \textit{C57BL/6j} strain. Mice from the consomic mouse strain created by replacing the \textit{C57BL/6j} chromosome 7 with the \textit{A/J} chromosome 7 (Nadeau et al. 2000) were more susceptible than \textit{C57BL/6j} to \textit{S. aureus} infection, with a median survival time of 2.5 days (Ahn et al. 2010).

B\textsubscript{7} does not contain any genes but lies within the gene cluster of the 26-member extended kallikrein gene family. The kallikrein proteins (KLK) have well-established roles in the degradation of extracellular matrix (ECM), the generation of antimicrobial peptides, and the regulation of immune responses, particularly inflammation [reviewed in Morizane et al. (2010), Sotiropoulou and Pampalakis (2010), and Sotiropoulou et al. (2009)]. Degradation of ECM facilitates the infiltration by immune cells of the skin and other tissues. Thus, reduced ECM degradation by KLK enzymes could inhibit infiltration of the site of infection by host immune cells. Similarly, enhanced ECM degradation by KLK enzymes could facilitate \textit{S. aureus} dissemination. The proteolytic activities of KLKs are important for the generation of antimicrobial peptides, particularly cathelicidins and defensins, which directly kill microbes as well as influence innate immune response processes. In this case also, reduced KLK activity would weaken the host immune response to \textit{S. aureus}.

KLKs play an important role in the regulation of inflammation, particularly through activation of the IL-1\beta precursor and the potent vasoactive peptides bradykinin and kallidin (Moreau et al. 2005). Thus, alterations of KLK activity could result in dysregulation of the inflammatory response. We and von Kockritz-Blickwede et al. (2008) observed increased kidney infiltration by \textit{S. aureus} in susceptible mice relative to resistant mice. In addition, von Kockritz-Blickwede et al. (2008) observed increased lung infiltration by \textit{S. aureus} and erythrocytes, as well as evidence of extensive lung hemorrhage in \textit{A/J} mice. These observations are all consistent with severe, increased microvascular permeability in susceptible mice in response to \textit{S. aureus} infection. Further, von Kockritz-Blickwede et al. (2008) observed increased levels of serum bradykinin in \textit{A/J} mice as well as decreased activated partial thromboplastin time. Both observations provide evidence that susceptible mice experience increased microvascular permeability as a result of increased activation of the kallikrein-kinin or contact system.

Our gene expression microarray results provide further evidence for the role of these genes in mediating host susceptibility to \textit{S. aureus}.

\begin{table}[h]
\centering
\caption{Table 3 Strain-haplotype pattern for the single significantly associated block B\textsubscript{7}}
\begin{tabular}{|c|c|}
\hline
\textbf{Haplotype} & \textbf{Mouse Strains} \\
\hline
ACT & AKR/J \\
GTC & NZW/LacJ \\
GCC & PWD/PhJ \\
GTT & A/J, BALB/cByJ, DBA/2J \\
AAT & 129S1/SvImJ, BTBR t+th/J, C3H/HeJ, C57BL/6J, FVB/NJ, KKHJ, NOD/ShiLtJ \\
\hline
\end{tabular}
\end{table}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{Figure3.png}
\caption{Figure 3 The distribution of \(-\log_{10}(P)\) throughout the genome. For each haplotype block, one black dot is plotted for the block’s \(-\log_{10}(P)\) value. Location is shown on the x axis; the x-axis coordinate of each dot corresponds to the middle of the block. The black circle at the top indicates the \(-\log_{10}(P)\) corresponding to the block B\textsubscript{7}, with a statistically significant association with the kidney count phenotype.}
\end{figure}

\textbf{DISCUSSION}

Exposure to \textit{S. aureus} has a wide variety of outcomes, and there is strong evidence that host genetics play an important role (von Kockritz-Blickwede et al. 2008). We therefore characterized 13 strains of inbred mice for four measures of infection severity and utilized HAM (Pletcher et al. 2004) to identify a single region on chromosome 7 significantly associated with colonization of the kidney by \textit{S. aureus} and containing a family of candidate genes.

Across the 13 mouse strains, we observe a wide variety of phenotype values that are consistent with what is known about \textit{S. aureus} pathogenesis and the host immune response. High bacterial counts in the kidneys are indicative of bacterial dissemination and are correlated with kidney dysfunction in mouse strains susceptible to \textit{S. aureus} infection (Deshmukh et al. 2009; von Kockritz-Blickwede et al. 2008). High bacterial counts in the peritoneal fluid are indicative of deficient host clearing of bacteria (von Kockritz-Blickwede et al. 2008). Similarly, low levels of IL-6 may be indicative of a deficient host immune response, as IL-6 is an important mediator of inflammation and activator of neutrophils and has been shown to be required for successful defense against bacterial pathogens, such as \textit{Streptococcus pneumoniae} (van der Poll et al. 1997) and \textit{Listeria monocytogenes} (Kopf et al. 1994). Taken together, these data indicate
infection. Given the large number of Klk genes, however, the varied and wide-ranging functions of their gene products, and the complicated patterns of coregulation via reciprocal- and auto-proteolysis, a series of gene-specific experiments at the nucleic acid and protein levels are required for each of the 26 genes to disentangle their precise role in host susceptibility to \textit{S. aureus}.

Our study has some limitations. First, we only detected a significantly associated haplotype block for one of four phenotypes. This is likely due to the fact that the HAM approach assumes that phenotypic similarities between mouse strains result from shared underlying genetic variants. Our study phenotypes may result from interactions between many different genetic variants, each of which is shared by only a subset of the strains exhibiting similar phenotypes. In addition, the current study may lack sufficient power to detect the corresponding causal variants. The mouse strain panel used in this study is relatively small. Although some studies have used a similar number of strains and detected significant associations using HAM (Yang et al. 2009), many other studies have used much larger numbers of strains (Bopp et al. 2010; Kirby et al. 2010). It is thus likely that with a larger mouse strain panel, additional associations would be detected.

Finally, survival times were recorded for only five days, resulting in an underestimate of the median survival time, particularly for the resistant strains. This results in an underestimate of the F-statistic sum of squares for the within- and between-haplotype group variability. Although underestimates of the within-group variability may result in false-positive associations, underestimates of the between-group variability may result in false-negative associations.

Despite these limitations, we were able to identify a genomic region significantly associated with susceptibility to \textit{S. aureus} infection, for which there is strong supporting experimental evidence and which implicates a large gene family whose members are promising candidate genes for future biological validation. Future studies in the mouse will identify specific members of the family with a role in \textit{S. aureus} pathogenesis in the murine host and will elucidate the specific mechanisms by which the gene products confer susceptibility. These studies in the mouse will be followed by studies to evaluate the role of the corresponding orthologous genes in human susceptibility to \textit{S. aureus}.

**ACKNOWLEDGMENTS**

This work was supported primarily by National Institute of Allergy and Infectious Diseases (NIAID) R01-AI-068804 to V.G.F. Additional support for N.V.J., M.L., and L.G.C. was provided by NIAID R01-AI0-77706 and a Burroughs Wellcome Fund Career Award to L.G.C. The authors acknowledge the Texas Advanced Computing Center (TACC; http://www.tacc.utexas.edu) at the University of Texas at Austin for providing the high-performance computing resources for this project.

**LITERATURE CITED**

Ahn, S. H., H. S. Deshmukh, N. V. Johnson, L. G. Cowell, T. H. Rude et al., 2010 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. 6: e1001088.

Anholt, R. R. H., and T. F. C. Mackay, 2009 Principles of Behavioral Genetics. Academic Press, London.

Benjamini, Y., and Y. Hochberg, 1995 Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc., B 57: 289–300.

Bopp, S. E., V. Ramachandran, K. Henson, A. Luzader, M. Lindstrom et al., 2010 Genome wide analysis of inbred mouse lines identifies a locus containing Ppar-gamma as contributing to enhanced malaria survival. PLoS ONE 5: e10903.

Burgess-Herbert, S. L., S. W. Tsaih, I. M. Stylianou, K. Walsh, A. J. Cox et al., 2009 An experimental assessment of in silico haplotype association mapping in laboratory mice. BMC Genet. 10: 81.

Cervino, A. C., G. Li, S. Edwards, J. Zhu, C. Laurie et al., 2005 Integrating QTL and high-density SNP analyses in mice to identify Insig2 as a susceptibility gene for plasma cholesterol levels. Genomics 86: 505–517.

---

**Table 4 Genes located in the extended region on chromosome 7 interrogated for candidate genes**

| Genomic Location (bp) | Gene Symbol | Gene Name | Probeset | Differentially Expressed |
|-----------------------|-------------|-----------|----------|--------------------------|
| 50945788              | Klk14       | Kallikrein related-peptidase 14 |          |                          |
| 50967937              | Klk13       | Kallikrein related-peptidase 13 |          |                          |
| 51024267              | Klk12       | Kallikrein related-peptidase 12 |          |                          |
| 51029987              | Klk11       | Kallikrein related-peptidase 11 | Y        |                          |
| 51036424              | Klk10       | Kallikrein related-peptidase 10 | Y        |                          |
| 51047261              | Klk9        | Kallikrein related-peptidase 9  | Y        |                          |
| 51052947              | Klk8        | Kallikrein related-peptidase 8  | Y        |                          |
| 51066814              | Klk7        | Kallikrein related-peptidase 7  | Y        |                          |
| 51079869              | Klk6        | Kallikrein related-peptidase 6  | Y        |                          |
| 51097639              | Klk5        | Kallikrein related-peptidase 5  | Y        |                          |
| 51136542              | Klk4        | Kallikrein related-peptidase 4  | Y        | Upregulated              |
| 51189141              | Klk15       | Kallikrein related-peptidase 15 |          |                          |
| 51206034              | Klk1b8      | Kallikrein 1-related peptidase b8 | Y        | Downregulated            |
| 51222138              | Klk1b1      | Kallikrein 1-related peptidase b1 | Y        | Downregulated            |
| 51231431              | Klk1b9      | Kallikrein 1-related peptidase b9 |          |                          |
| 51251247              | Klk1b11     | Kallikrein 1-related peptidase b11 | Y        | Downregulated            |
| 51268048              | Klk1b26     | Kallikrein 1-related peptidase b26 | Y        |                          |
| 51307660              | Klk1b27     | Kallikrein 1-related peptidase b27 | Y        |                          |
| 51357662              | Klk1b21     | Kallikrein 1-related peptidase b21 | Y        |                          |
| 51368043              | Klk1b22     | Kallikrein 1-related peptidase b22 |          |                          |
| 51392137              | Klk1b16     | Kallikrein 1-related peptidase b16 | Y        | Downregulated            |
| 51443606              | Klk1b24     | Kallikrein 1-related peptidase b24 | Y        | Downregulated            |
| 51453561              | Klk1b3      | Kallikrein 1-related peptidase b3  | Y        |                          |
| 51462805              | Klk1b4      | Kallikrein 1-related peptidase b4  | Y        |                          |
| 51471844              | Klk1b5      | Kallikrein 1-related peptidase b5  | Y        |                          |
| 51480807              | Klk1        | Kallikrein 1               | Y        | Downregulated            |
