DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD
DASH.HARVARD.EDU

HARVARD LIBRARY
Office for Scholarly Communication

Prediction of Multiple Infections After Severe Burn Trauma

Citation

Yan, Shuangchun, Amy Tsurumi, Yok-Ai Que, Colleen M. Ryan, Arunava Bandyopadhaya, Alexander A. Morgan, Patrick J. Flaherty, Ronald G. Tompkins, and Laurence G. Rahme. 2015. "Prediction of Multiple Infections After Severe Burn Trauma." In Annals of Surgery 261 no. 4: 781–792. doi:10.1097/sla.0000000000000759.

Published Version

doi:10.1097/sla.0000000000000759

Permanent link

http://nrs.harvard.edu/urn-3:HUL.InstRepos:17190533

Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. Submit a story .

Accessibility

Prediction of Multiple Infections After Severe Burn Trauma: a Prospective Cohort Study

Authors

Shuangchun Yan PhD[1,2,3], Amy Tsurumi PhD[1,2,3], Yok-Ai Que MD, PhD[4], Colleen M. Ryan MD[1,3], Arunava Bandyopadhaya PhD[1,2,3], Alexander A. Morgan PhD[7], Patrick J. Flaherty PhD[5,6], Ronald G. Tompkins, MD, ScD[1], Laurence G. Rahme MS, PhD[1,2,3]

Author affiliations

1 Department of Surgery, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts, United States of America;

2 Department of Microbiology and Immunobiology, Harvard Medical School, Boston, Massachusetts, United States of America;

3 Shriners Hospitals for Children Boston, Boston, Massachusetts, United States of America;

4 Department of Fundamental Microbiology, University of Lausanne, Lausanne, 1015, Switzerland;

5 Department of Biomedical Engineering, Worcester Polytechnic Institute, Worchester, Massachusetts, United States of America;

6 Program in Bioinformatics and Computational Biology, Worcester Polytechnic Institute, Worchester, Massachusetts, United States of America;

7 Department of Biochemistry & Stanford Genome Technology Center, Stanford University School of Medicine, Palo Alto, California, United States of America.

Corresponding author:
Laurence G. Rahme, PhD, MS
Harvard Medical School
Massachusetts General Hospital
Email: rahme@molbio.mgh.harvard.edu
Tel. +1 (617) 724-5003; Fax +1 (617) 724-8558

Sources of support: U.S. Army Medical Research Acquisition Act of U.S. Department of Defense; National Institute of General Medical Sciences.

Running head:

Prognostic models of multiple infections

ABSTRACT

Objective To develop predictive models for early triage of burn patients based on hyper-susceptibility to repeated infections.

Background Infection remains a major cause of mortality and morbidity after severe trauma, demanding new strategies to combat infections. Models for infection prediction are lacking.

Methods Secondary analysis of 459 burn patients (≥16 years old) with ≥20% total body surface area burns recruited from six US burn centers. We compared blood transcriptomes with a 180-h cut-off on the injury-to-transcriptome interval of 47 patients (≤1 infection episode) to those of 66 hyper-susceptible patients (multiple [≥2] infection episodes [MIE]). We used LASSO regression to select biomarkers and multivariate logistic regression to built models, accuracy of which were assessed by area under receiver operating characteristic curve (AUROC) and cross-validation.

Results Three predictive models were developed covariates of: (1) clinical characteristics; (2) expression profiles of 14 genomic probes; (3) combining (1) and (2). The genomic and clinical models were highly predictive of MIE status ($AUROC_{Genomic}$ = 0.946 [95% CI, 0.906–0.986]); $AUROC_{Clinical}$ = 0.864 [CI, 0.794–0.933]; $AUROC_{Genomic}$/$AUROC_{Clinical}$ P = 0.044). Combined model has an increased $AUROC_{Combined}$ of 0.967 (CI, 0.940–0.993) compared to the individual models ($AUROC_{Combined}$/$AUROC_{Clinical}$ P = 0.0069). Hyper-susceptible patients show early alterations in immune-related signaling pathways, epigenetic modulation and chromatin remodeling.

Conclusions Early triage of burn patients more susceptible to infections can be made using clinical characteristics and/or genomic signatures. Genomic signature suggests new insights into the pathophysiology of hyper-susceptibility to infection may lead to novel potential therapeutic or prophylactic targets.

Mini-Abstract

Early genomic signature and clinical characteristics of 113 burn patients were used paradigmatically to build three novel predictive models of multiple, repeated infections in burn trauma, which could facilitate early triage of traumatically injured burn patients to prevent or treat sepsis. Genomic signature suggests new mechanistic aspects of hyper-susceptibility to infections.

INTRODUCTION

Although several studies have found association between specific risk factors or clinical characteristics with mortality after trauma,[1-4] studies attempting to apply those clinical characteristics or genomic biomarkers to appreciate susceptibility to infection and build predictive models are currently lacking. Improvements in early care and trauma centers have reduced early mortality considerably.[3,5] However, severe trauma, such as burn trauma, cause immunosuppression which predispose patients to infections. Despite all medical improvements, infections remain a major cause of critical injury-related morbidity and mortality, and recurrent sepsis predisposes patients to multiple organ failure, lengthens hospital stays, and increases costs.[6] Therefore, improvements in prevention and treatment of infections are increasingly important.[7,8] Moreover, the rapid emergence of multi-(MDR) or pan-drug resistant (PDR) pathogens that cause highly problematic acute, persistent or relapsing infections pose a dire threat to healthcare, especially among trauma and surgical patients.[9,10] The increased use of antibiotics has further accelerated their emergence,[11-13] and also increased the challenge of treating polymicrobial wound infections.[14,15] Due to the paucity of novel anti-infectives in development, further improvement in patient care and treatment efficacy may rely heavily on optimizing existing strategies and promoting patients-tailored therapies.[16-18]

Successful personalized approach requires rigorous triaging: early and accurate identification of patients more susceptible to infections could help tailor the anti-infective treatments,[19,20] and especially to elaborate long-term treatment plan. Future successful clinical trials aiming to improve sepsis outcome may also rely on biomarkers to identify the right patients for the right treatment.[21,22] Several studies have reported risk factors associated with increased probability of infection and sepsis in trauma patients,[23-26] but no specific predictive model has been developed. Existing plasma biomarkers such as C-reactive protein (CRP) and procalcitonin (PCT) are mainly used to diagnose sepsis[27,28] rather than reflective of susceptibility or health status. The clinical characteristics measurable rapidly upon admission are the current gold standard for prognosis of general patient's outcome.

As trauma promotes susceptibility to infection and genomic signatures appear to play an increasingly promising role in prognosis,[26,29] we analyzed the blood transcriptome and clinical characteristics data of 113 patients from the 573 thermally injured patients enrolled in the Inflammation and the Host Response to Injury study. Using clinical characteristics available upon admission and early genomic signatures, we developed novel predictive models that would permit early identification of burn patients at high risk of developing repeated infection indicative of an early hyper-susceptible state. The genomic signature suggests new mechanistic aspects for susceptibility to infection after burn trauma.

METHODS

Subject Recruitment and Sample Selection

This study was conducted via secondary use of the clinical and genomic data of the Inflammation and the Host Response to Injury Study ("Glue Grant"). Briefly, 573 burn patients with minimum 20% total burn surface area (TBSA) were enrolled from six institutions between 2003 and 2009 in a prospective, longitudinal study. RNA of leucocytes isolated from whole blood samples were extracted for transcriptome analysis using Affymetrix GeneChip Human Genome U133 Plus 2.0 microarrays at University of Florida–Gainesville, as described previously.[30] The complete inclusion/exclusion criteria are described elsewhere.[31] Permission for this secondary use of the de-identified data was obtained from the Massachusetts General Hospital Institutional Review Board (MGH IRB protocol 2008-P-000629/1).

Our patient inclusion process is summarized in Figure 1. From 573 potential patients in the data pool, we selected for patients that were at least 16 years old with early transcriptome data. We set a 180-h cut-off limit on the injury-to-transcriptome interval to include only samples that were obtained early relative to the recovery process, while still allowing enough samples to remain eligible for biomarker discovery. If multiple blood samples were collected from a patient, only the earliest eligible sample was included. We excluded patients who died within 9 days of blood collection and had fewer than two

infection episodes during this time window (Figure 1; Figure 1A). Our method for collection of data related to clinical characteristics is described elsewhere.[31] To enable direct comparisons, as well as combination of clinical and genomic prediction, we used the same set of patients for both our clinical characteristic and our genomic signature prediction models.

Definition of Outcomes

We defined infections according to the information collected in the Glue Grant database based on previously described standards.[32] Infection episodes were quantified for each patient for up to 60 days after blood sample collection. We developed a decision tree (Figure 1B; Supplemental Digital Content[SDC] Table 1) for evaluating each record based on: (1) time of infection; (2) type of infection; and (3) the pathogen(s) isolated. Since no genotyping data of the isolated pathogen species were available, we were unable to classify whether a later episode was caused by the same strain isolated earlier. However, once a record was counted, the infection type and isolated pathogen combination (e.g. *Pseudomonas aeruginosa* + lung) was put on a "waiting list" for the next 6 days, which likely reduced the likelihood of an infection episode caused by the same isolate from being counted. Subsequent records that were part of the same infection episode were thereby omitted. The patients were separated into two groups based on susceptibility to infection, measured by the number of independent infection episodes recorded. We defined patients with ≤1 infection episodes as the less susceptible control group (N = 47), and patients with ≥2 (multiple) infection episodes (MIE) as the hyper-susceptible case group (N = 66).

Microarray Processing and Filtering

Raw microarray data (.CEL files) were downloaded from the Glue Grant website (http://www.gluegrant.org/trdb/) and filtered using the steps outlined in Figure 1, SDC Table 1 and Figure 1B. We used the gcrma[33] package on the R/Bioconductor platform[34] to normalize 124 blood samples from

124 eligible patients collected within 180 h post-injury. Samples identified as outliers by arrayQualityMetrics[35] were excluded from subsequent analysis. One patient was removed due to incompleteness of clinical data. Two patients' datasets were discarded due to mortality within 9 days after sample collection. After these filtration steps, 113 blood samples were deemed suitable high-quality microarray data sets for subsequent functional analyses, biomarker discovery, and modeling.

We used the EMA package[36] in R software to filter outlying or information-poor probe sets. We eliminated probe sets with a maximum $\log_2$ expression value below 3.5, reducing the number of probe sets from 54,675 to 26,107. Using limma package,[37] we selected 1142 probe sets with an at least 1.5-fold difference between less susceptible patients and hyper-susceptible patients and with an average expression level of at least 3 for functional analyses and biomarker panel selection process.

Statistical Analysis

Clinical data set. Continuous variables are reported as means (standard deviations), or as medians with inter-quartile ranges (IQRs) as indicated. Categorical variables are reported as frequencies and percentages. Demographic variables between less susceptible and hyper-susceptible patients were tested for statistical difference with a Wilcoxon rank sums test, a Chi-square test, or a Fisher's exact test as appropriate. Statistical significance was accepted at $P < 0.05$ (two-tailed when appropriate).

Body mass index (BMI) was calculated as weight/height^2 (kg/m^2). For patients ≥20 years old, BMI categories of underweight, healthy, overweight and obese were define according to BMI numbers: <18.5, 18.5–24.9, 25–29.9, and ≥30, respectively; whereas for patients <20 years old, the same BMI categories were defined using percentile ranking based on Centers for Disease Control and Prevention BMI-for-age growth charts: <5th percentile, 5th to <85th percentile, 85th to <95th percentile, and ≥95th percentile, respectively.

Genomic data set. In our evaluation of significant expression differences between less susceptible

and hyper-susceptible patients, Benjamini-Hochberg multiple-comparison adjustments were applied to control for false discovery rate.

Development of the clinical predictive models. We implemented stepwise logistic regression with an entry level of 0.3 and a stay level of 0.25 to identify significant predictor variables among clinical covariates relevant to the outcome variable of MIE: TBSA, age, BMI, and the presence of inhalation injury. We determined predictive power by calculating area under receiver operating characteristic curve (AUROC), reported with 95% confidence intervals (CIs).

Development of the genomic predictive models. We used the LASSO regularized regression method[38] implemented in the glmnet package[39] in R software to identify probe sets that collectively predicted the likelihood of MIE. We used 10-fold cross-validation (CV) to select the optimal value of LASSO penalty weighting, λ. The value of λ that gave the minimum average binomial deviance plus 1 standard error on the test set, λ_{1se}, was used to select probe sets (Figure 3A). λ_{1se} is a stronger penalty parameter to guard against over-fitting than λ_{min}, which minimizes the average binomial deviance of CV (Figure 3B). This 10-fold CV process was repeated 100 times to generate 100 λ_{1se} values. The median λ_{1se}, 0.0940, yielded selection of a 14-probe-set biomarker panel (Figure 3C; Table 2). Logistic regression was performed to model the MIE outcome with the $\log_2$ expression values of the 14 probe sets as explanatory variables. Furthermore, we conducted multivariate logistic regression with the clinical covariates TBSA, age, and inhalation injury together with the 14 probe sets for the outcome variable of MIE. Leave-one-out cross-validation was used to assess the degree of over-fitting and model performance.

Functional Analysis

Functional and pathway analyses were conducted using Ingenuity IPA (Ingenuity® Systems, www.ingenuity.com) and DAVID.[40]

Software Platform and Package Versions

R (version 2.15.*); EMA package for R (version 1.3.2); pROC package for R (version 1.5.4); limma package for R (version 3.14.4); glmnet package for R (version 1.9-3); arrayQualityMetrics package for R (version 3.14.0); gcrma package for R (version 2.30.0); JMP Pro 10 and SAS 9.3 (SAS Institute Inc., North Carolina, USA).

RESULTS

Clinical Characteristics

From a pool of 573 patients, 124 met our inclusion criteria, of which 11 were unsuitable for modeling, leaving a cohort of 113 patients (Figure 1), including 47 patients less susceptible to infection (control group with ≤1 infection episodes) and 66 hyper-susceptible patients (case group with multiple [≥2] infection episodes [MIE]). The demographics, injury characteristics, and outcomes of these 113 patients are summarized in Table 1.

From 612 microbiological records for the 113 patients in the final cohort, we identified 325 independent infection episodes, 107 (32.9%) of which are polymicrobial at the species level. Twenty-four patients had no infection episodes, 23 had one episode, and 66 had MIE. The less susceptible and hyper-susceptible patients show significantly different clinical characteristics (Table 1). Relative to the control group, hyper-susceptible patients were slightly older (mean, 38.2, SD 16.4 vs 37.0, SD 14.6), had higher TBSA (46%, IQR 35–71 vs 32%, IQR 23–41, $P < 0.0001$), had more inhalation injuries (41/66 [62.1%] vs 8/47 [17.0%], $P < 0.0001$) and were more severely ill (according to their APACHE II score 24, IQR 18–29 vs 13, IQR 9–20, $P < 0.0001$). They also had longer hospital stays (median, 60, IQR 33–71 vs 20, IQR 15–30, $P < 0.0001$), more days on mechanical ventilation (median, 28, IQR13–40 vs 2, IQR 0–5, $P < 0.0001$), and had a higher mortality (18/66 [27.3%] vs 3/47 [6.4%], $P = 0.0029$) (Table 1). The median post-injury interval for the second episode in the case group was 15 days (IQR, 10–20; range, 3–43), a

time window that provides opportunity for prophylactic intervention.

Inhalation injury significantly increased the risk of developing MIE and may be related to pneumonia risk in particular: 78.8% of hyper-susceptible patients had pneumonia vs 10.6% of controls; among cases, 84.7% had both MIE and inhalation injuries, 67.4% had both pneumonia and inhalation injuries. Interestingly, 4/5 of underweight patients had MIE (Table 1), supporting the notion that being overweight and mild obesity may be protective against post-injury infection whereas being underweight increases risk.[32,41]

Burn wound infection and nosocomial pneumonia were the most frequent types of infection observed (Table 1; Figure 2A). *Pseudomonas aeruginosa* and Staphylococci (both *Staphylococcus aureus* and coagulase negative Staphylococci) were the most commonly isolated micro-organisms (Table 1; Figure 2B). *P. aeruginosa* and *Acinetobacter* infections were more common among patients with MIE than controls, suggesting that hyper-susceptible patients were even more susceptible to nosocomial Gram-negative pathogens.

MIE Prediction from Clinical Characteristics

We used stepwise logistic regression to select covariates for modeling from TBSA, age, BMI, and the presence of inhalation injury. The final multivariate logistic regression model included three covariates: TBSA, age, and inhalation injury, which were significant independent predictors of MIE. The AUROC, CV AUROC, sensitivity, and specificity values for the clinical characteristics model are 0.845 (95% CI, 0.773–0.916), 0.838 (95% CI, 0.762–0.914), 0.803 (95% CI, 0.683–0.887), and 0.745 (95% CI, 0.594–0.856), respectively (Figure 3). The model's positive and negative predictive values were 0.815 (95% CI, 0.696–0.843) and 0.729 (95% CI, 0.579–0.843), respectively. Inhalation injury significantly increased MIE incidence (odds ratio [OR], 6.942; 95% CI, 2.482–19.417). Patients who had inhalation injuries were twice as likely to get pneumonia compared to those without them (risk ratio [RR], 2.05; 95% CI, 1.37–

3.07). Among those who had inhalation injuries, 67.4% had pneumonia, and 83.67% had MIE. TBSA (OR, 1.078; 95% CI, 1.040–1.118) and age (OR, 1.040; 95% CI, 1.006–1.075) were also associated with increased infection susceptibility.

MIE Prediction from Genomic Biomarkers in Blood

Ten-fold CV using LASSO regularized regression[38] of the 1142 probe sets that presented a minimum of 1.5-fold change between the two patient groups yielded a minimal set of 14 predictors (probe sets) that together optimized the fit of the model (Figure 4A and 4B). Of these 14 probe sets—which mapped to 12 genes—4 were upregulated and 10 were down-regulated (Table 2, all $P < 0.01$; see Figure 4C for heat map and clustering of patients and biomarkers; see Figure 2 for expression profiles of each probe set). The biological processes associated with each probe set are presented in Table 3 together with the coefficients of the biomarker panel logistic regression model (model intercept = 0.7449; SDC Table 6). The AUROC, CV AUROC, sensitivity, and specificity values for the resulting genomic signature model are 0.946 (95% CI, 0.906–0.986), 0.872 (95% CI, 0.804 - 0.940), 0.924 (95% CI, 0.825–0.972), and 0.830 (95% CI, 0.687–0.919), respectively (Figure 3), confirming the model to be highly sensitive and specific. The positive and negative predictive values of the model were 0.884 (95% CI, 0.779–0.945) and 0.886 (95% CI, 0.746–0.957), respectively. We compared each patient's probability of developing MIE estimated from our clinical or genomic biomarker logistic regression models with each of the observed outcomes, using cut-off points of 30% to 70% as being uncertain. We found that the clinical model correctly predicted outcomes of 73 (65%) patients with certainty. Comparatively, the genomic biomarker model correctly predicted 90 (80%) patients with certainty, showing a 15% improvement over the clinical model. Both models misclassified 9 patients (8%). Collectively, these data suggest that genomic biomarkers may complement triage by clinical characteristics and enhance early prediction of a patient's likelihood to develop MIE.

MIE Prediction from a Combined Model

A multivariate logistic model that included the aforementioned clinical covariates (TBSA, age, presence of inhalation injury) and genomic biomarkers resulted in an AUROC (0.967; 95% CI, 0.940–0.993) that was significantly greater than that for the clinical model ($P = 0.0069$), but not significantly different from that of the genomic biomarker panel model (Figure 3). The positive and negative predictive values of the combined model were 0.881 (95% CI, 0.773–0.943) and 0.848 (95% CI, 0.705–0.932), respectively. The estimates of the above models are listed in SDC Table 6.

Functional and Canonical Pathway Changes in Patients with MIE Revealed by Transcriptome Data Analysis

The 1142 probe sets showing a minimum of 1.5-fold change in hyper-susceptible patients versus less susceptible patients were mapped to 844 annotated genes. We identified functionally related genes among these 884 genes using Gene Ontology (GO). Subsequent analysis of the changes in canonical pathways and functions linked to these 844 genes indicated that hyper-susceptible patients' transcriptomes demonstrated the following early functional changes relative to control transcriptomes: (1) early activation of immune cells, increased chemotaxis and trafficking; (2) decreased expansion of leukocytes, thymocytes, and number of phagocytes, and increased cell death and apoptosis; and (3) suppression of immune cell activation and lymphoid organ development (Table 2). The 1142 probe sets showed enrichment in four main gene ontology biological process categories: (1) immune response; (2) epigenetic modulation of gene expression; (3) transcription; and (4) metabolism (SDC Tables 2). Functional enrichment clustering is also in agreement with the enrichment of the 4 functional groups (SDC Table 3). The top 30 affected pathways were mainly involved in immune cell signaling and cytokine signaling (Figure 5). Canonical pathway analysis using IPA software (Figure 5) largely agrees with KEGG pathway

enrichment analysis using DAVID (SDC Table 5), providing additional confidence. Overall, many of the predicted functional changes (Table 2) are downstream of the affected canonical pathways (Figure 5; SDC Table 5).

Canonical Pathways and T-cell Signaling

Significant changes in IL-8 signaling (17 upregulated and 12 down-regulated genes [17 up/12 down]), Gαq signaling (16 up/9 down), Rho family GTPase signaling (20 up/10 down) and integrin signaling (21 up/9 down) suggest that the adhesion and migration of leukocytes are affected (Table 2; SDC Table 3; and Figure 5). The changes in chemotaxis may be partially caused by the presence of bacteria at wound site, as fMLP signaling pathway (12 up/8 down) suggests. Genes involved in phospholipase C signaling, a regulator of chemotactic response are differentially expressed (20 up/16 down). The increased cell movement, adhesion, and chemotaxis are related to phagocytosis process (e.g. FcγR-mediated phagocytosis, SDC Table 6), clearance of the pathogen from the site of infection, and induced by host damage associated molecular patterns (DAMP).

We found strong evidence that T-cells were also differentially regulated in case patients. Several pathways, including T-cell receptors (TCR) (7 up/16 down), JAK-STAT signaling (9 up/7 down), PKCθ signaling (8 up/15 down), and IL-6 signaling pathway (13 up/6 down) are known to regulate T-cell differentiation, activation, and cytokine production. Changes in iCOS-iCOSL signaling (10 up/14 down), CD28 signaling (11 up/16 down), and IL-2 signaling (7 up/7 down), indicate that T helper cell maturation and proliferation were likely affected. In summary, patient transcriptome data is consistent with compromised cellular immune responses mediated by impaired T-cells signaling.

Functional Enrichment in Histone Modification and Chromatin Remodeling

We found evidence for dramatic epigenetic changes in leukocytes that long precede patient outcome of

MIE. Functions related to epigenetic modulation were commonly enriched in our functional enrichment analyses (SDC Tables 2, 3, and 4). Notably, 42 probe sets (39 genes) have functional annotation associated with chromatin remodeling and histone modifications (SDC Table 4). Two genes from the biomarker panel involved in epigenetic modulation were found to be down-regulated in the case group with MIE: *WHSC1L1*, which encodes a histone lysine methyltransferase; and *SMARCA4*, which encodes an ATP-dependent helicase related to the SWI/SNF chromatin remodeling factor. A multitude of differentially expressed genes encoding histone post-translational modifiers as well as key components of the nucleosome remodeling complex mediating ATP-dependent nucleosome sliding, including *SMARCC1*, *SMARCA4*, *CHD2* and *CHD9*, were down-regulated (SDC Table 4). Other notable histone methyltransferases/demethylases differentially expressed include *KDM4*, *KDM5C*, *KDM6*, *PRDM5*, *SETD2*, *SETDB2*, and *SUZ12*. Genes coding for histone deacetylases/acetyltransferases and associated factors including *HDAC9*, *KAT6A* and *EP400* were down-regulated and histone acetylation recognizing bromodomain containing protein, *BRD2*, was upregulated in the case group. Furthermore, critical non-histone heterochromatin proteins HP1-α and –γ were down-regulated, as well as core histone cluster. Taken together, our data may suggest a global loss of heterochromatin and genome instability, as well as probable gene-specific transcriptional deregulation in hyper-susceptible patients compared to controls.

DISCUSSION

The work presented reports novel predictive models for hyper-susceptibility to infection among traumatically injured patients, using genomic biomarkers and/or clinical characteristics that have not been used to build statistical prognostic models for the purpose of predicting infection outcomes. We provide evidence that our models can identify burn patients at high risk of developing repeated infections indicative of their hyper-susceptible state. To our knowledge, this work is the first to describe such models in trauma patients, and the first to describe functional transcriptome data of burn patients in

relation to infections. The prediction accuracy of hyper-susceptibility to MIE is significantly increased over clinical markers when the genomic signature is used, providing strong evidence of the promising role of genomic biomarkers in prognosis even when used alone. By combining the biomarker panel with clinical characteristics, we demonstrated even better prediction accuracy, supporting the tremendous potential of using genomic signature to increase confidence in data used for treatment decision-making.

Clinical Implications.

We identified two distinct patient groups with different genomic signatures and clinical characteristics, essentially allowing the rapid identification of patients with a high risk of developing MIE following burn trauma. Although burn patients generally suffer from immunosuppression, clinical experience and our data suggest that the severity of immunosuppression and infection outcome vary. These data suggest that patients could potentially receive personalized therapy depending on their susceptibility to infection, triaged by physical exam and a blood test on admission. This information could facilitate the determination of appropriate treatment courses, particularly in regards to antibiotic use, allowing for selective use of prophylactic antibiotics and more objective justification of length of treatment courses. For the patient, this could limit complications related to unneeded antibiotics, reduce the burden of lines needed to deliver the antibiotics, and streamline hospital care. For the population, this could promote antibiotic stewardship, help stem the emergence of resistant organisms, and reduce the cost of care.

Mechanistic Aspects.

Genomic signatures provide insight into the molecular mechanisms of the more susceptible health status, and may aid in the discovery of novel therapeutic targets. Our findings point to novel potential targets for the prevention and/or early treatment of infections. Functional analyses of the 1142 biomarker candidates suggest new aspects into the pathophysiology of susceptibility to MIE after trauma. Susceptibility to MIE

was associated with early alterations in numerous signaling pathways related to innate and adaptive immune responses, and changes in epigenetic modulation and metabolism.

Some of our findings are consistent with previous literature. For instance, upregulation of *THBS1* (thrombospondin 1), to which 3/14 of the biomarker probe sets were mapped, has been associated with complicated recovery in blunt trauma patients,[29] supporting the broad applicability of our approach and findings. The discovery of *THBS1* also supports the potential biological relevance of our biomarkers. Indeed, increased expression of mouse homologue *Thbs1* has been reported to be associated with infection,[42] thrombosis, and increased lipopolysaccharide-induced mortality. Interestingly, *Thbs1* -/- knockout mice show reduced susceptibility to peritoneal sepsis,[43] whereas *Thbs1* over-expressing transgenic mice show impaired wound healing associated with wound angiogenesis inhibition.[44] THBS1 in human wounds could be functioning to provide adhesion target for pathogens through promotion of thrombosis,[45] and/or delayed wound healing, which could lead to increased susceptibility to infection. Thus, building on convergent findings in humans and mice, our data confirm that processes related to coagulation play important roles in sepsis, and suggest that THBS1 could be a novel target for sepsis prevention and treatment.

We showed evidence for increased chemotaxis, cell adhesion, and migration of immune cells, and simultaneously, decreased expansion of immune cells and development of lymphatic system components. This seeming contradiction may well be the consequences of dysfunctional immune system and cytokine signaling, especially in T-cells.

Our data suggest that epigenetic changes occur early on, rather than mainly as a consequence of septic shock. Epigenetic regulation of immune system is a common mechanism for gene expression regulation and it plays a role in long-term immunosuppression after sepsis.[46] Tightly regulated chromatin remodeling is required for transcriptional regulation, which is vital for proper host immune and inflammatory responses.[47] Among the genes associated with epigenetic regulations, several have

confirmed roles in immune responses, such as *KAT6A* and *KDM6B* (SDC Table 4). [46,48-50] Furthermore, our data further supports the notion that genes related to cell-cycle control and DNA repair have roles in both immune responses and tumorigenesis. In summary, the dramatic epigenetic changes could potentially explain why our biomarker panel could predict MIE that occurred weeks later, and the underlying mechanisms that favor infections by Gram-negative opportunistic pathogens.

Implications for Future Research.

With the aforementioned clinical implications and mechanistic aspects, our findings lay the groundwork for a new pathway of investigation potentially applicable to other forms of trauma and possibly even useful in determining patient risk for MIE prior to elective surgical procedures. This study provides a much-needed new direction for future clinical trials. In particular, appropriate biomarkers and additional information regarding patient health status might be essential for successful clinical trials of anti-sepsis drugs.[21,22] Identification of the hyper-susceptible patients could enable more focused study design when expensive/invasive interventions, such as for the testing of cutting-edge technologies or products are involved by directing intervention to those who need it most. Identification of this group early after admission could also allow adjunctive treatments such as immunotherapy, extra-corporeal lipopolysaccharide removal, and other novel treatments to be tested prior to the decline of the patient's clinical status due to MIE.

We envision that the development of a comprehensive diagnostic tool set will depend on the integration of genomic signatures of both host and pathogen. The blood biomarkers reported could be further developed and integrated with other diagnostic tools, such as genomic single nucleotide polymorphisms (SNPs) that predispose certain patients to infection,[51,52] and produce a more comprehensive prognosis of patient susceptibility. Physician decisions rely heavily on blood tests over the course of recovery, and a positive culture is still the most accepted and reliable method for diagnosing infection. Using biomarkers, these blood samples could also allow us to monitor the changes in

susceptibility status and adjust treatments accordingly. Modern molecular based microbiological tests,[53] such as detection of *P. aeruginosa* in wound biopsy using RT-PCR based assays,[54] have been developed but not yet widely utilized. Several molecular early detection kits have become commercially available for diagnosing common bloodstream infections, and have been found to show some promise despite of much room left for improvement.[55,56] Our biomarkers on the host response may work synergistically with these tests to support physician decisions.

The discovery of these biomarkers and the validation of the methods pave the way for identifying biomarkers from other tissues involved in host defense, such as muscle, fat, and skin samples,[57] of which often become available from surgical procedures or wound debridement. Biomarkers from other tissues may further enhance a combined model or perhaps provide even better prognostic value than blood biomarkers and clinical characteristics.

This study is limited by the unavailability of pathogen genotyping information below species level. We could not distinguish whether a reoccurring infection was caused by persistent or MDR pathogen, and could not identify biomarkers that can potentially differentiate susceptibility to different pathogens, such as Gram positive/negative bacteria, and even to species level. Nonetheless, our 6-day window (SDC Figure 1B) was designed to minimize infection episodes caused by the same strain(s). Our definition of hyper-susceptibility is based on natural definition of having repeated infections. Changing this definition, for example, to having at least three infection episodes, did not significantly change the biomarkers identified (data not shown). However, the *P* values for differential gene expression and clinical characteristics became less significant, suggesting either the criterion is not the best cut off point to separate two different groups, or that the statistical power is reduced due to smaller number of patients in the hyper-susceptible group.

Although this work and our model focused on thermally injured trauma patients, our approach is potentially applicable to other types of trauma and surgical patients. In this study, to ensure portability of

our models, we carried out rigorous internal CV to ensure robustness of our regression models. However, due to the novelty of this clinical and transcriptome dataset, independent cohort data was unavailable for CV. Although our dataset is the largest of its kind to date, the sample size is still too small to build a larger panel without risking over-fitting the model. Our genomics data warrant future trials with a larger randomized cohort study, as well as mechanistic interrogations using animal models. Our findings open new avenues for the prevention and treatment of repeated infections in critical care, and provide novel components for the development of integrated prognosis and diagnosis using biomarkers, SNPs and pathogen detection. Future studies should investigate the potential broad applicability, and assess whether early triage based on predictive models can improve outcomes of trauma patients.

Acknowledgements

This work was supported by the U.S. Army Medical Research Acquisition Act of U.S. Department of Defense, Congressionally Directed Medical Research Programs (CDMRP), Defense Medical Research and Development Program (DMRDP) Basic Research Award, W81XWH-10-DMRDP-BRA to LGR. The investigators acknowledge the contribution of the Inflammation and the Host Response to Injury Large-Scale Collaborative Project Award #5U54GM062119 from the National Institute of General Medical Sciences. We thank W. Xu, W. Xiao, and A. A. Tzika for suggestions on the data analysis.

Disclaimer

The Inflammation and the Host Response to Injury "Glue Grant" program is supported by the National Institute of General Medical Sciences. This manuscript was prepared using a dataset obtained from the Glue Grant program and does not necessarily reflect the opinions or views of the Inflammation and the Host Response to Injury Investigators or the NIGMS.

Trial Registration clinicaltrials.gov Identifier NCT00257244

References

1. Morris JA, MacKenzie EJ, Damiano AM, et al. Mortality in trauma patients: the interaction between host factors and severity. *J Trauma.* 1990;30:1476-1482.

2. Kraft R, Herndon DN, Al-Mousawi AM, et al. Burn size and survival probability in paediatric patients in modern burn care: a prospective observational cohort study. *Lancet.* 2012;379:1013-1021.

3. Ryan CM, Schoenfeld DA, Thorpe WP, et al. Objective estimates of the probability of death from burn injuries. *N Engl J Med.* 1998;338:362-366.

4. Osler T, Glance L, Buzas JS, et al. A trauma mortality prediction model based on the anatomic injury scale. *Ann Surg.* 2008;247:1041-8.

5. MacKenzie EJ, Rivara FP, Jurkovich GJ, et al. A National Evaluation of the Effect of Trauma-Center Care on Mortality. *N Engl J Med.* 2006;354:366-378.

6. Church D, Elsayed S, Reid O, et al. Burn wound infections. *Clin Microbiol Rev.* 2006;19:403-434.

7. Bloemsma GC, Dokter J, Boxma H, et al. Mortality and causes of death in a burn centre. *Burns.* 2008;34:1103-1107.

8. Ingraham AM, Xiong W, Hemmila MR, et al. The attributable mortality and length of stay of trauma-related complications: a matched cohort study. *Ann Surg.* 2010;252:358-62.

9. Kesarwani M, Hazan R, He J, et al. A quorum sensing regulated small volatile molecule reduces acute virulence and promotes chronic infection phenotypes. *PLoS Pathog.* 2011;7:e1002192.

10. Bandyopadhaya A, Kesarwani M, Que Y-A, et al. The quorum sensing volatile molecule 2-amino acetophenon modulates host immune responses in a manner that promotes life with unwanted guests. *PLoS Pathog.* 2012;8:e1003024.

11. Boucher HW, Talbot GH, Bradley JS, et al. Bad bugs, no drugs: no ESKAPE! An update from the Infectious Diseases Society of America. *Clin Infect Dis.* 2009;48:1-12.

12. Avni T, Levcovich A, Ad-El DD, et al. Prophylactic antibiotics for burns patients: systematic review and meta-analysis. *BMJ (Clinical research ed.).* 2010;340:c241.

13. Cohen NR, Lobritz MA, and Collins JJ. Microbial Persistence and the Road to Drug Resistance. *Cell Host Microbe.* 2013;13:632-642.

14. Pirnay J-P, De Vos D, Cochez C, et al. Molecular Epidemiology of Pseudomonas aeruginosa Colonization in a Burn Unit: Persistence of a Multidrug-Resistant Clone and a Silver Sulfadiazine-Resistant Clone. *J Clin Microbiol.* 2003;41:1192-1202.

15. De Vos D, Lim AJ, Pirnay P, et al. Analysis of epidemic Pseudomonas aeruginosa isolates by isoelectric focusing of pyoverdine and RAPD-PCR: modern tools for an integrated anti-nosocomial infection strategy in burn wound centres. *Burns.* 1997;23:379 - 386.

16. Brunkhorst FM, Oppert M, Marx G, et al. Effect of empirical treatment with moxifloxacin and meropenem vs meropenem on sepsis-related organ dysfunction in patients with severe sepsis: a randomized trial. *JAMA.* 2012;307:2390-2399.

17. Schuetz P, Litke A, Albrich WC, et al. Blood biomarkers for personalized treatment and patient management

485 decisions in community-acquired pneumonia. *Curr Opin Infect Dis.* 2013;26:159-67.

486 18. Härtel C, Deuster M, Lehrnbecher T, et al. Current approaches for risk stratification of infectious complications
487 in pediatric oncology. *Pediatr Blood Cancer.* 2007;49:767-73.

488 19. Angus DC. The search for effective therapy for sepsis: back to the drawing board? *JAMA.* 2011;306:2614-2615.

489 20. Kuehn BM. Guideline Promotes Early, Aggressive Sepsis Treatment to Boost Survival. *JAMA.* 2013;309:969-
490 970.

491 21. Schuetz P, Haubitz S, and Mueller B. Do sepsis biomarkers in the emergency room allow transition from
492 bundled sepsis care to personalized patient care? *Curr Opin Crit Care.* 2012;18:341-9.

493 22. Angus DC, and van der Poll T. Severe sepsis and septic shock. *N Engl J Med.* 2013;369:840-851.

494 23. Nichols RL, Smith JW, Klein DB, et al. Risk of infection after penetrating abdominal trauma. *N Engl J Med.*
495 1984;311:1065-1070.

496 24. Kisat M, Villegas CV, Onguti S, et al. Predictors of sepsis in moderately severely injured patients: an analysis of
497 the national trauma data bank. *Surg Infect (Larchmt).* 2013;14:62-68.

498 25. Wibbenmeyer L, Danks R, Faucher L, et al. Prospective analysis of nosocomial infection rates, antibiotic use,
499 and patterns of resistance in a burn population. *J Burn Care Res.* 2006;27:152-160.

500 26. Boomer JS, To K, Chang KC, et al. Immunosuppression in patients who die of sepsis and multiple organ failure.
501 *JAMA.* 2011;306:2594-2605.

502 27. Lavrentieva A, Papadopoulou S, Kioumis J, et al. PCT as a diagnostic and prognostic tool in burn patients.
503 Whether time course has a role in monitoring sepsis treatment. *Burns.* 2012;38:356-363.

504 28. Schultz L, Walker SAN, Elligsen M, et al. Identification of predictors of early infection in acute burn patients.
505 *Burns.* 2013;39:1355-1366.

506 29. Cuenca AG, Gentile LF, Lopez MC, et al. Development of a Genomic Metric That Can Be Rapidly Used to
507 Predict Clinical Outcome in Severely Injured Trauma Patients. *Crit Care Med.* 2013;41:1175-85.

508 30. Laudanski K, Miller-Graziano C, Xiao W, et al. Cell-specific expression and pathway analyses reveal alterations
509 in trauma-related human T cell and monocyte pathways. *Proc Natl Acad Sci U S A.* 2006;103:15564-15569.

510 31. Xiao W, Mindrinos MN, Seok J, et al. A genomic storm in critically injured humans. *J Exp Med.*
511 2011;208:2581-2590.

512 32. Jeschke MG, Finnerty CC, Emdad F, et al. Mild Obesity Is Protective After Severe Burn Injury. *Ann Surg.*
513 2013;Publish Ahead of Print:1.

514 33. Wu Z, Irizarry RA, Gentleman R, et al. A model based background adjustment for oligonucleotide expression
515 arrays. *Johns Hopkins University, Dept. of Biostatistics Working Papers.* 2004;

516 34. Gentleman RC, Carey VJ, Bates DM, et al. Bioconductor: open software development for computational biology
517 and bioinformatics. *Genome Biol.* 2004;5:R80.

518 35. Kauffmann A, Gentleman R, and Huber W. arrayQualityMetrics--a bioconductor package for quality assessment
519 of microarray data. *Bioinformatics.* 2009;25:415-416.

520 36. Servant N, Gravier E, Gestraud P, et al. EMA - A R package for Easy Microarray data analysis. *BMC Res Notes.*
521 2010;3:277.

522 37. Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray
523 experiments. *Stat Appl Genet Mol Biol.* 2004;3:Article3.

524 38. Tibshirani R. Regression Shrinkage and Selection via the Lasso. *J R Stat Soc Series B Stat Methodol.*
525 1996;58:267-288.

526 39. Friedman J, Hastie T, and Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate
527 Descent. *J Stat Softw.* 2010;33:1-22.

528 40. Huang DW, Sherman BT, Stephens R, et al. DAVID gene ID conversion tool. *Bioinformation.* 2008;2:428-430.

529 41. Wacharasint P, Boyd JH, Russell JA, et al. One size does not fit all in severe infection: obesity alters outcome,
530 susceptibility, treatment, and inflammatory response. *Crit Care.* 2013;17:R122.

531 42. Johnson CA, Kleshchenko YY, Ikejiani AO, et al. Thrombospondin-1 Interacts with Trypanosoma cruzi Surface
532 Calreticulin to Enhance Cellular Infection. *PLoS One.* 2012;7:e40614.

533 43. McMaken S, Exline MC, Mehta P, et al. Thrombospondin-1 Contributes to Mortality in Murine Sepsis through
534 Effects on Innate Immunity. *PLoS One.* 2011;6:e19654.

535 44. Streit M, Velasco P, Riccardi L, et al. Thrombospondin-1 suppresses wound healing and granulation tissue
536 formation in the skin of transgenic mice. *EMBO J.* 2000;19:3272-3282.

537 45. Shannon O. Platelets interact with bacterial pathogens. *Thromb Haemost.* 2009;102:613-4.

538 46. Carson WF, Cavassani KA, Dou Y, et al. Epigenetic regulation of immune cell functions during post-septic
539 immunosuppression. *Epigenetics.* 2011;6:273-283.

540 47. Smale ST. Selective transcription in response to an inflammatory stimulus. *Cell.* 2010;140:833-844.

541 48. Perez-Campo FM, Costa G, Lie-a-Ling M, et al. The MYSTerious MOZ, a histone acetyltransferase with a key
542 role in haematopoiesis. *Immunology.* 2013;139:161-165.

543 49. De Santa F, Narang V, Yap ZH, et al. Jmjd3 contributes to the control of gene expression in LPS-activated
544 macrophages. *EMBO J.* 2009;28:3341-3352.

545 50. Kruidenier L, Chung C-W, Cheng Z, et al. A selective jumonji H3K27 demethylase inhibitor modulates the
546 proinflammatory macrophage response. *Nature.* 2012;488:404-408.

547 51. Netea MG, Wijmenga C, and O'Neill LAJ. Genetic variation in Toll-like receptors and disease susceptibility.
548 *Nat Immunol.* 2012;13:535-542.

549 52. Bronkhorst MWGA, Lomax MAZ, Vossen RHAM, et al. Risk of infection and sepsis in severely injured
550 patients related to single nucleotide polymorphisms in the lectin pathway. *Br J Surg.* 2013;100:1818-1826.

551 53. Jannes G, and De Vos D. A review of current and future molecular diagnostic tests for use in the microbiology
552 laboratory. *Methods Mol Biol.* 2006;345:1-21.

553 54. Pirnay J-P, De Vos D, Duinslaeger L, et al. Quantitation of Pseudomonas aeruginosa in wound biopsy samples:
554 from bacterial culture to rapid 'real-time' polymerase chain reaction. *Crit Care.* 2000;4:255.

555 55. Chang S-S, Hsieh W-H, Liu T-S, et al. Multiplex PCR system for rapid detection of pathogens in patients with
556 presumed sepsis - a systemic review and meta-analysis. *PLoS One.* 2013;8:e62323.

557 56. Skvarc M, Stubljar D, Rogina P, et al. Non-culture-based methods to diagnose bloodstream infection: Does it
558 work? *Eur J Microbiol Immunol (Bp).* 2013;3:97-104.

559 57. Apidianakis Y, Que YA, Xu W, et al. Down-regulation of glutatione S-transferase 4 (hGSTA4) in the muscle of
560 thermally injured patients is indicative of susceptibility to bacterial infection. *FASEB J.* 2012;26:730-737.

561

Figure Legends

Figure 1. Sample selection process.

[a]Development of predictive models and discovery of biomarkers.

Figure 2. Type of infections and isolated pathogens. A. Types of infection. One case of pseudomembranous colitis represents 0.2%. **B.** The percentage of isolated pathogens among all infection records.

Figure 3. Clinical and genomic prediction models. ROC curves of the clinical model, genomic model, and combined model, and their respective AUROC, cross-validated (CV) AUROC, sensitivities, and specificities; 95% CIs are reported in parentheses. The blue, orange, and black lines are the ROC curves for the biomarker panel model, clinical model, and combined model, respectively.

Figure 4. Biomarker selection by LASSO regularized regression. A. A representative repetition of 10-fold CV LASSO that chose 14 probe sets at λ1se. The first vertical dotted line corresponds to the λ_{min} that minimized binomial deviance during CV. The second dotted line corresponds to λ_{1se}, used for the selection of 14 probe sets as shown in B. **B.** LASSO coefficient profile plot of the coefficient paths. At $\lambda_{1se,}$ as shown with the dotted line, 14 probe sets have their coefficients significantly different from zero and thus were chosen as part of the biomarker panel. **C.** Heat map showing the expression levels of the 14 probe sets selected by LASSO as covariates for the genomic model. Each column corresponds to one of the 113 patient samples. Each row corresponds to one of the 14 probe sets. Whenever available, gene names were provided (see Table 2 for Affymetrix probe identification). The heat map color-coding is based on probe-set-specific, re-normalized expression values, with red signifying upregulation, blue

signifying down-regulation, and white indicating no difference in the hyper-susceptible patients compared to the controls. Patients that developed MIE are labeled red and those that had <2 infection episodes are labeled green at the bottom of the heat map.

Figure 5. Pathways significantly altered. Top 30 pathways significantly altered in case group with MIE. X-axis is the negative log *P* value calculated from Fisher's exact test right-tailed. Red/Green inside bars are the number of upregulated/down-regulated genes. The total number of genes in a pathway is indicated in the parenthesis after pathway name. *P* value is calculated by Fisher's exact test by IPA software.

List of Supplemental Digital Content

SDC Figure 1. **A.** Timeline of the study. **B.** Decision tree used to define independent infection episodes using available clinical and microbiological records. Overriding rules of the decision tree are as included below the table and also described in the methods section.

SDC Figure 2. Expression profile of 12 genes in the biomarker panel. A total of 14 probe sets mapped to 12 genes are shown as scatter plot overlaid with notched box plots. *P* values were calculated using limma package in R software using moderated t-statistics and then adjusted for multiple comparisons using B-H method. Each data point in the scatter plot corresponds to a sample from a patient, and color-coded based on the total infection episodes the patient had from 2 days to 60 days after blood collection.

SDC Table 1. Infection episode decision table. Alternative presentation of the decision tree, complementary to SDC Figure 1B.

SDC Table 2. Term centric singular enrichment in gene ontology biological process and molecular function of the 1142 probe sets. Abbreviations: BP, biological process; MF, molecular function. Adjusted *P* value is based on Benjamini method. Color shading indicates whether this term is associated with one of the four functional categories: immune responses, epigenetic modulation, transcription and metabolism. Light green represents "associated". Dark green represents "highly associated". The color-coding is manually curated.

SDC Table 3. Term centric functional annotation clustering that shows annotation groups that are enriched for the 1142 probe sets. Top 50 clusters were included. The rest of the 50 clusters are decreasing in statistical significance and not shown. Abbreviations: BP, biological process. MF: molecular function. Adjusted *P* value is based on Benjamini method.

SDC Table 4. Genes involved in epigenetic modulation and chromatin remodeling from the 1142 probe sets. Adjusted *P* value is based on B-H method. Gene symbols in bold are the genes that are part of the biomarker panel.

SDC Table 5. KEGG pathway enrichment analysis using DAVID. The results are consistent with IPA pathway enrichment analysis.

SDC Table 6. Estimates of multivariate logistic regression models.

Mini-Abstract

Early genomic signature and clinical characteristics of 113 burn patients were used paradigmatically to build three novel predictive models of multiple, repeated infections in burn trauma, which could facilitate early triage of traumatically injured burn patients to prevent or treat sepsis. Genomic signature suggests new mechanistic aspects of hyper-susceptibility to infections.

ABSTRACT

Objective To develop predictive models for early triage of burn patients based on hyper-susceptibility to repeated infections.

Background Infection remains a major cause of mortality and morbidity after severe trauma, demanding new strategies to combat infections. Models for infection prediction are lacking.

Methods Secondary analysis of 459 burn patients (≥16 years old) with ≥20% total body surface area burns recruited from six US burn centers. We compared blood transcriptomes with a 180-h cut-off on the injury-to-transcriptome interval of 47 patients (≤1 infection episode) to those of 66 hyper-susceptible patients (multiple [≥2] infection episodes [MIE]). We used LASSO regression to select biomarkers and multivariate logistic regression to built models, accuracy of which were assessed by area under receiver operating characteristic curve (AUROC) and cross-validation.

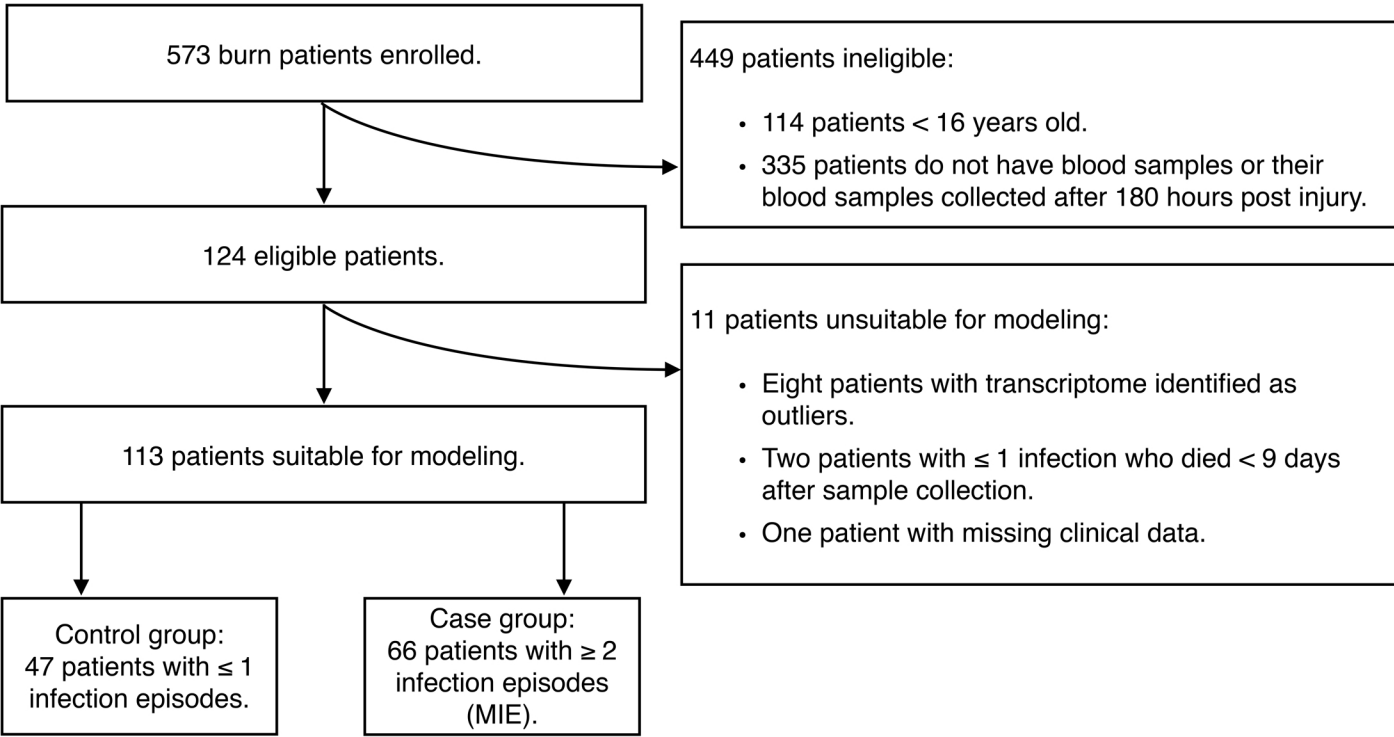
Results Three predictive models were developed covariates of: (1) clinical characteristics; (2) expression profiles of 14 genomic probes; (3) combining (1) and (2). The genomic and clinical models were highly predictive of MIE status ($AUROC_{Genomic}$ = 0.946 [95% CI, 0.906–0.986]); $AUROC_{Clinical}$ = 0.864 [CI, 0.794–0.933]; $AUROC_{Genomic}/AUROC_{Clinical}$ P = 0.044). Combined model has an increased $AUROC_{Combined}$ of 0.967 (CI, 0.940–0.993) compared to the individual models ($AUROC_{Combined}/AUROC_{Clinical}$ P = 0.0069). Hyper-susceptible patients show early alterations in immune-related signaling pathways, epigenetic modulation and chromatin remodeling.

Conclusions Early triage of burn patients more susceptible to infections can be made using clinical characteristics and/or genomic signatures. Genomic signature suggests new insights into

the pathophysiology of hyper-susceptibility to infection may lead to novel potential therapeutic or prophylactic targets.

Type of file: figure
Label: Figure 1
Filename: Fig1.pdf

Figure 1.


573 burn patients enrolled.
449 patients ineligible:
• 114 patients < 16 years old.
• 335 patients do not have blood samples or their blood samples collected after 180 hours post injury.
124 eligible patients.
11 patients unsuitable for modeling:
• Eight patients with transcriptome identified as outliers.
• Two patients with ≤ 1 infection who died < 9 days after sample collection.
• One patient with missing clinical data.
113 patients suitable for modeling.
Control group:
47 patients with ≤ 1 infection episodes.
Case group:
66 patients with ≥ 2 infection episodes (MIE).

Type of file: figure
Label: Figure 2
Filename: Fig2.pdf

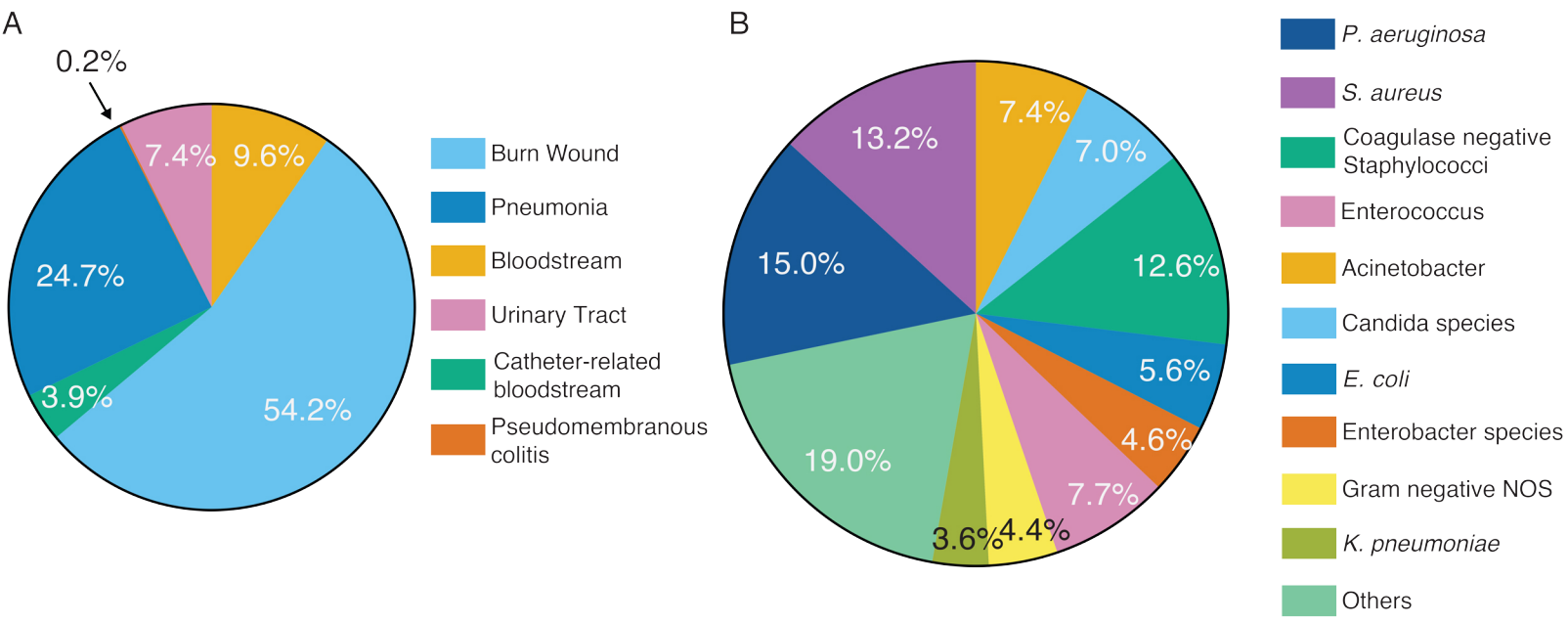
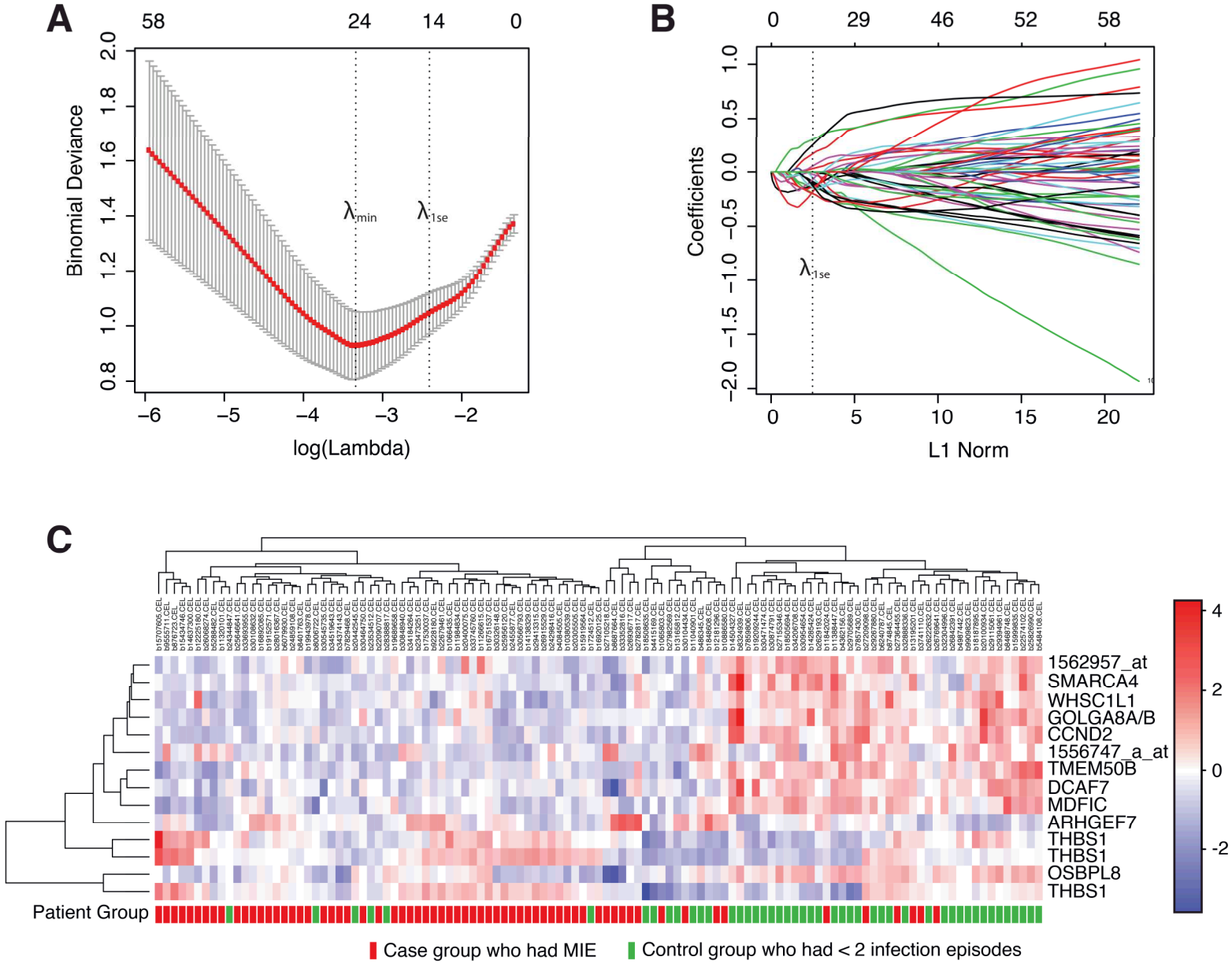


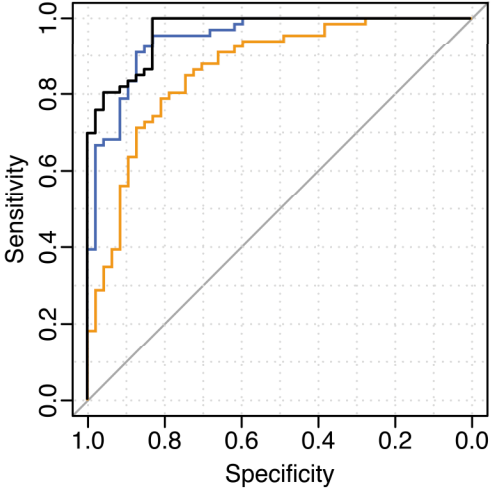
Figure 2. Type of infections and isolated pathogens.

Type of file: figure
Label: Figure 3
Filename: Fig3.tif


A
58
24
14
0
Binomial Deviance
2.0
1.8
1.6
1.4
1.2
1.0
0.8
λmin
λ1se
−6
−5
−4
−3
−2
log(Lambda)
B
0
29
46
52
58
Coefficients
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
λ1se
0
5
10
15
20
L1 Norm
C
1562957_at
SMARCA4
WHSC1L1
GOLGA8A/B
CCND2
1556747_a_at
TMEM50B
DCAF7
MDFIC
ARHGEF7
THBS1
THBS1
OSBPL8
THBS1
Patient Group
4
2
0
-2
Case group who had MIE
Control group who had < 2 infection episodes

Type of file: figure
Label: Figure 4
Filename: Fig4.tif

Figure 4.

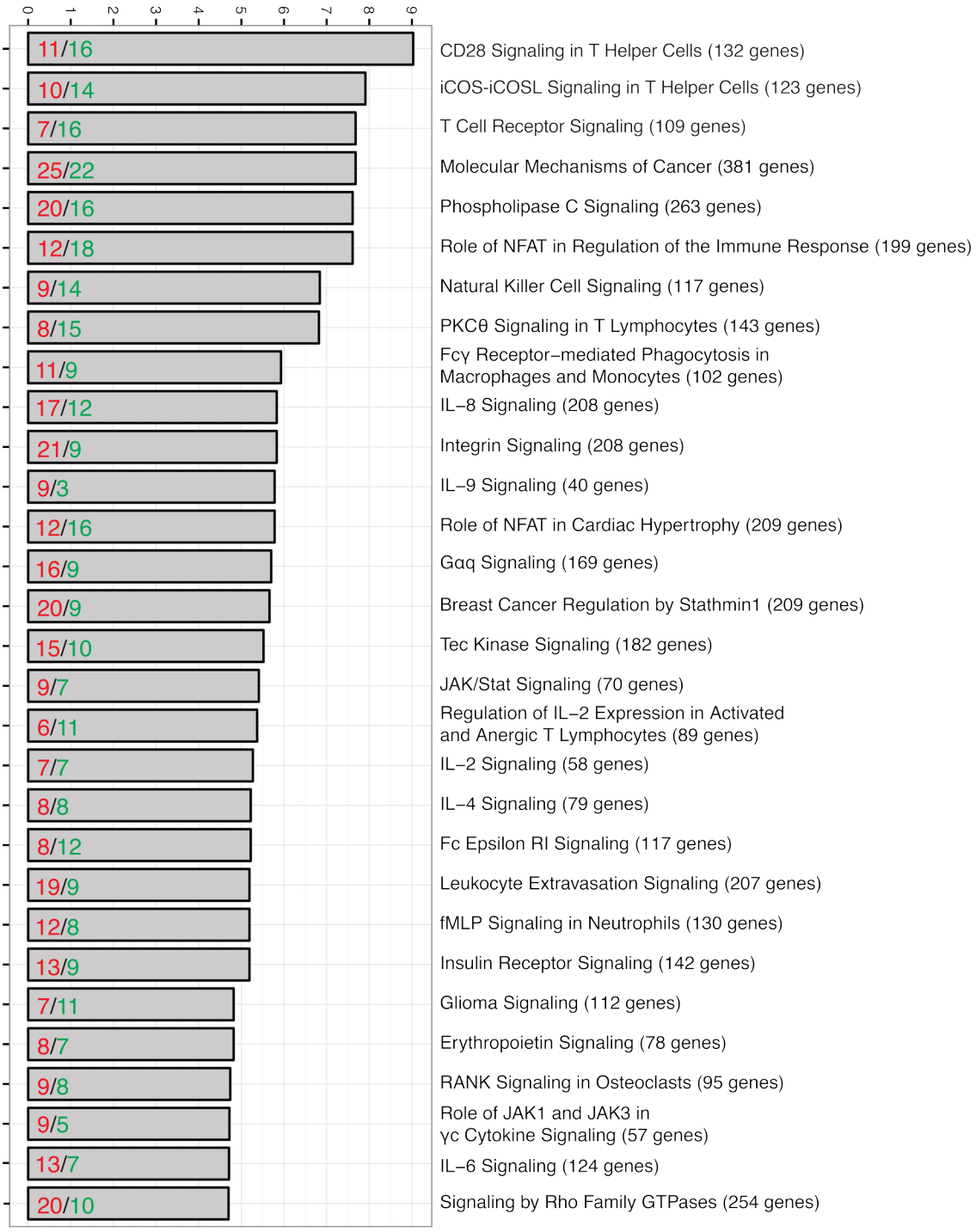


	AUROC (95% CI)	CV AUROC (95% CI):
Clinical:	0.845 (0.773 - 0.916)	0.838 (0.762 - 0.914)
Genomic:	0.946 (0.906 - 0.986)	0.872 (0.804 - 0.940)
Combined:	0.967 (0.940 - 0.993)	0.888 (0.826 - 0.949)

	Sensitivity (95% CI)	Specificity (95% CI)
Clinical:	0.803 (0.683 - 0.887)	0.745 (0.594 - 0.856)
Genomic:	0.924 (0.825 - 0.972)	0.830 (0.687 - 0.919)
Combined:	0.894 (0.788 - 0.953)	0.830 (0.687 - 0.919)

Type of file: figure
Label: Figure 5
Filename: Fig5.pdf

Figure 5.



Red/Green
Red: the number of up-regulated genes
Green: the number of down-regulated genes

Type of file: table
Label: Tables 1 - 3
Filename: Tables.docx

Table 1. Demographics and clinical characteristics of participants.

	All (n=113)	Controls (≤1 Infectious Episodes) (n=47)	Cases (≥2 Infectious Episodes [MIE]) (n=66)	*P* value
Age when injured, mean (SD), y	37.7 (15.6)	37.0 (14.6)	38.2 (16.4)	0.681
Sex, n (%) males	90 (79.6%)	40 (85.1%)	50 (75.8%)	0.218
BMI Category, n (%)				0.888
Underweight	5 (4.4%)	1 (2.1%)	4 (6.1%)	
Healthy	44 (38.9%)	19 (40.4%)	25 (37.9%)	
Overweight	35 (31.0%)	15 (31.9%)	20 (30.3%)	
Obese	29 (25.7%)	12 (25.6%)	17 (25.8%)	
Severity of Injury				
APACHE II Score, median (IQR)	20 (12-26)	13 (8-20)	24 (18-28)	<0.001*
Burns size of TBSA, % (IQR)	40 (28-56)	32 (23-40)	46 (35-70)	<0.001*
Presence of Inhalation Injury, n (%)	49 (43.4%)	8 (17.0%)	41 (62.1%)	<0.001*
Outcome				
Hospital Stay, d (IQR)	35 (19-62)	20 (15-27)	60 (33-71)	<0.001*
Hospital Stay of Survived, d (IQR)	36 (19-62)	20.5 (15-27)	61 (44-72)	<0.001*
Days on Ventilation, d (IQR)	13 (2-33)	2 (0-5)	28 (13-40)	<0.001*
Day of Death Since Injury, d (IQR)	34 (18-63)	21 (18-21)	35.5 (18-65)	0.3753
Mortality, no. (%)	21 (18.6%)	3 (6.38%)	18 (27.3%)	0.0029*
Number of Records by Type of Infection, n (%)				
Burn wound	332 (54.2%)	24 (60%)	308 (53.8%)	
Pneumonia	151 (24.7%)	8 (20%)	143 (25.0%)	
Bloodstream	59 (9.6%)	1 (2.5%)	58 (10.1%)	
Urinary tract	45 (7.4%)	7 (17.5%)	38 (6.6%)	
Catheter-related bloodstream	24 (3.9%)	0 (0%)	24 (4.2%)	
Pseudomembranous colitis	1 (0.2%)	0 (0%)	1 (0.2%)	
Number of Records by Isolated Pathogens, n (%)				
P. aeruginosa	92 (15.0%)	4 (10%)	88 (15.4%)	
S. aureus	81 (13.2%)	7 (17.5%)	74 (13.0%)	
Coagulase negative Staphylococci	77 (12.6%)	6 (15.0%)	71 (12.4%)	
Enterococcus	47 (7.7%)	4 (10.0%)	43 (7.5%)	
Acinetobacter	45 (7.4%)	1 (2.5%)	44 (7.7%)	
Candida species	43 (7.0%)	0 (0%)	43 (7.5%)	
E. coli	34 (5.6%)	1 (2.5%)	33 (5.8%)	
Enterobacter species	28 (4.6%)	1 (2.5%)	27 (4.7%)	
Gram negative NOS	27 (4.4%)	0 (0%)	27 (4.7%)	
K. pneumoniae	22 (3.6%)	0 (0%)	22 (3.8%)	

Others	116 (18.9%)	16 (40%)	100 (17.5%)

*$P < 0.05$.

Abbreviations: BMI, body mass index; IQR, inter-quartile range; TBSA, total body surface area.

Table 2. The 14 probe sets in the biomarker panel.

Probe set	Gene Symbol	Gene Name	Gene Ontology Biological Process Annotation	Fold Change	Coeffi cients	*P* value
Upregulated						
201109_s_at	THBS1	thrombospondin 1	Angiogenesis, regulation of cytokine production, regulation of endothelial cell proliferation, regulation of antigen processing and presentation, regulation of immune system process	3.37	0.560	<0.001
201110_s_at	THBS1	thrombospondin 1	Same as above	2.31	0.100	0.001
201108_s_at	THBS1	thrombospondin 1	Same as above	2.02	0.824	0.001
235412_at	ARHGEF7	Rho guanine nucleotide exchange factor (GEF) 7	Apoptotic process, signal transduction, epidermal growth factor receptor signaling pathway, small GTPase mediated signal transduction, apoptotic signaling pathway, lamellipodium assembly	1.86	0.747	0.017
Down-regulated						
217599_s_at	MDFIC	MyoD family inhibitor domain containing	Transcription, activation of JUN kinase activity, virus-host interaction, regulation of Wnt receptor signaling pathway, negative regulation of protein import into nucleus, positive regulation of viral transcription	-2.34	-0.289	<0.001
200951_s_at	CCND2	cyclin D2	Positive regulation of cyclin-dependent protein kinase activity, cell cycle, cell division	-2.21	0.292	<0.001
228986_at	OSBPL8	oxysterol binding protein-like 8	Lipid transport, negative regulation of sequestering of triglyceride, fat cell differentiation	-1.98	0.111	<0.001
224730_at	DCAF7	DDB1 and CUL4 associated factor 7	Multicellular organismal development, protein ubiquitination	-1.87	-0.908	<0.001
222907_x_at	TMEM50B	transmembrane protein 50B	NA	-1.80	-0.335	<0.001
208797_s_at	GOLGA8A/ GOLGA8B	golgin A8 family, member B	NA	-1.78	-1.068	<0.001
217656_at	SMARCA4	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4	Negative regulation of transcription from RNA polymerase II promoter, chromatin remodeling, negative regulation of cell growth, negative regulation of androgen receptor signaling pathway, etc.	-1.59	0.252	<0.001
221248_s_at	WHSC1L1	Wolf-Hirschhorn syndrome candidate 1-like 1	Transcription, regulation of transcription, cell growth, histone methylation, cell differentiation, histone lysine methylation	-1.51	-0.676	<0.001
1556747_a_ at	NA	NA	NA	-1.66	-0.786	0.005
1562957_at	NA	NA	NA	-1.64	-0.409	<0.001

P values were adjusted for multiple comparisons based on Benjamini-Hochberg method during the fold-change calculation of 26,107 probes after initial filtering (see Methods).

Table 3. Predicted early functional changes in case group that had MIE.

Functions annotation	*P* value	Activation z-score	# of genes
Increased			
Chemotaxis	<0.001	3.924	55
Chemotaxis of cells	<0.001	3.924	54
Homing of cells	<0.001	3.815	59
Chemotaxis of leukocytes	<0.001	3.795	37
Chemotaxis of phagocytes	<0.001	3.546	30
Chemotaxis of myeloid cells	<0.001	3.501	29
Homing of leukocytes	<0.001	3.484	41
Replication of Influenza A virus	<0.001	3.413	38
Replication of virus	<0.001	3.314	64
Leukocyte migration	<0.001	3.088	100
Inflammatory response	<0.001	3.085	72
Viral infection	<0.001	3.046	166
Cytostasis	<0.001	2.913	30
Replication of RNA virus	<0.001	2.782	56
Cell movement	<0.001	2.766	173
Migration of cells	<0.001	2.619	161
Tyrosine phosphorylation of protein	<0.001	2.456	29
Recruitment of cells	<0.001	2.451	34
Recruitment of granulocytes	<0.001	2.405	26
Polarization of leukocytes	<0.001	2.337	13
Recruitment of leukocytes	<0.001	2.333	33
Adhesion of immune cells	<0.001	2.271	40
Recruitment of myeloid cells	<0.001	2.263	27
Adhesion of blood cells	<0.001	2.250	41
Cell viability	<0.001	2.240	112
Orientation of macrophages	<0.001	2.200	6
Attachment of cells	<0.001	2.166	18
Disassembly of focal adhesions	<0.001	2.164	7
Formation of membrane ruffles	<0.001	2.137	12
Cell survival	<0.001	2.101	121
Cell movement of neutrophils	<0.001	2.067	37
Invasion of breast cancer cell lines	<0.001	2.064	25
Orientation of cells	<0.001	2.028	19
Decreased	<0.001		
Development of lymphoid organ	<0.001	-3.241	30
Development of lymphatic system component	<0.001	-2.970	41
Bacterial infection	<0.001	-2.890	47
Expansion of leukocytes	<0.001	-2.753	25
Expansion of lymphocytes	<0.001	-2.635	21
Development of lymph node	<0.001	-2.608	14
Morphology of germinal center	<0.001	-2.415	11

Morphology of lymph follicle	<0.001	-2.415	15
Expansion of blood cells	<0.001	-2.384	26
Encephalitis	<0.001	-2.374	27
Inflammation of organ	<0.001	-2.362	97
Quantity of neutrophils	0.0011	-2.208	23
Development of thymocytes	<0.001	-2.189	13
Quantity of granulocytes	<0.001	-2.133	36
Organismal death	<0.001	-2.074	196

An absolute z-score of ≥2 was designated as significant by the IPA software. The numbers of genes used to predict functional changes are indicated in the column with the heading "# of genes".