

## Blood transcriptomics to characterize key biological pathways and identify biomarkers for predicting mortality in melioidosis

Thatcha Yimthin<sup>a</sup>, Jacqueline Margaret Cliff<sup>db</sup>, Rungnapa Phunpang<sup>c</sup>, Peeraya Ekcharyawat<sup>a,d</sup>, Taniya Kaewarpai<sup>a</sup>, Ji-Sook Lee<sup>b</sup>, Clare Eckold<sup>e</sup>, Megan Andrada<sup>f</sup>, Ekkachai Thiansukhon<sup>g</sup>, Kittisak Tanwisaid<sup>h</sup>, Somchai Chuananont<sup>h</sup>, Chumpol Morakot<sup>i</sup>, Narongchai Sangsa<sup>j</sup>, Wirayut Silakun<sup>k</sup>, Suneey Chayangsu<sup>l</sup>, Noppol Buasi<sup>m</sup>, Nicholas Day<sup>c,n</sup>, Ganjana Lertmemongkolchai<sup>o,p\*</sup>, Wasun Chantratita<sup>q</sup>, T. Eoin West<sup>id r</sup> and Narisara Chantratita<sup>id a,c</sup>

<sup>a</sup>Faculty of Tropical Medicine, Department of Microbiology and Immunology, Mahidol University, Bangkok, Thailand; <sup>b</sup>Faculty of Infectious and Tropical Diseases, Department of Immunology and Infection, London School of Hygiene & Tropical Medicine, London, UK; <sup>c</sup>Faculty of Tropical Medicine, Mahidol-Oxford Tropical Medicine Research Unit, Mahidol University, Bangkok, Thailand; <sup>d</sup>Faculty of Public Health, Department of Microbiology, Mahidol University, Bangkok, Thailand; <sup>e</sup>Faculty of Medicine, Department of Surgery and Cancer, Imperial College London, London, UK; <sup>f</sup>Department of Tropical Medicine, Medical Microbiology, and Pharmacology, John A. Burns School of Medicine, University of Hawaii at Manoa, Honolulu, Hawaii, USA; <sup>g</sup>Department of Medicine, Udon Thani Hospital, Udon Thani, Thailand; <sup>h</sup>Department of Medicine, Nakhon Phanom Hospital, Nakhon Phanom, Thailand; <sup>i</sup>Department of Medicine, Mukdahan Hospital, Mukdahan, Thailand; <sup>j</sup>Department of Medicine, Roi Et Hospital, Roi Et, Thailand; <sup>k</sup>Department of Medicine, Buriram Hospital, Buriram, Thailand; <sup>l</sup>Department of Medicine, Surin Hospital, Surin, Thailand; <sup>m</sup>Department of Medicine, Sisaket Hospital, Sisaket, Thailand; <sup>n</sup>Centre for Tropical Medicine, Nuffield Department of Medicine, University of Oxford, Oxford, UK; <sup>o</sup>Faculty of Associated Medical Science, Department of Clinical Immunology, Khon Kaen University, Khon Kaen, Thailand; <sup>p</sup>The Centre for Research and Development of Medical Diagnostic Laboratories, Khon Kaen University, Khon Kaen, Thailand; <sup>q</sup>Faculty of Medicine Ramathibodi Hospital, Center for Medical Genomics, Mahidol University, Bangkok, Thailand; <sup>r</sup>Division of Pulmonary and Critical Care Medicine, Harborview Medical Center, University of Washington, Seattle, Washington, USA

### ABSTRACT

Melioidosis is an often lethal tropical disease caused by the Gram-negative bacillus, *Burkholderia pseudomallei*. The study objective was to characterize transcriptomes in melioidosis patients and identify genes associated with outcome. Whole blood RNA-seq was performed in a discovery set of 29 melioidosis patients and 3 healthy controls. Transcriptomic profiles of patients who did not survive to 28 days were compared with patients who survived and healthy controls, showing 65 genes were significantly up-regulated and 218 were down-regulated in non-survivors compared to survivors. Up-regulated genes were involved in myeloid leukocyte activation, Toll-like receptor cascades and reactive oxygen species metabolic processes. Down-regulated genes were hematopoietic cell lineage, adaptive immune system and lymphocyte activation pathways. RT-qPCR was performed for 28 genes in a validation set of 60 melioidosis patients and 20 healthy controls, confirming differential expression. *IL1R2*, *GAS7*, *S100A9*, *IRAK3*, and *NFKBIA* were significantly higher in non-survivors compared with survivors ( $P < 0.005$ ) and healthy controls ( $P < 0.0001$ ). The AUROC of these genes for mortality discrimination ranged from 0.80-0.88. In survivors, expression of *IL1R2*, *S100A9* and *IRAK3* genes decreased significantly over 28 days ( $P < 0.05$ ). These findings augment our understanding of this severe infection, showing expression levels of specific genes are potential biomarkers to predict melioidosis outcomes.

**ARTICLE HISTORY** Received 10 September 2020; Revised 10 November 2020; Accepted 26 November 2020

**KEYWORDS** RNA-sequencing; transcriptomics; melioidosis; biomarkers; *Burkholderia pseudomallei*; outcome; immune response


### Introduction

Melioidosis is a severe infectious disease caused by *Burkholderia pseudomallei*, a Gram-negative bacterium and biothreat agent [1]. The disease is highly endemic in the tropics, particularly in Southeast Asia and northern Australia but reported cases are increasing globally. Melioidosis carries a mortality rate of

40% or higher in many endemic regions where resources are limited. This poor outcome from melioidosis has remained unchanged for many years [2,3]. Melioidosis is associated with several host factors, but diabetes is the major risk [4,5]. Pneumonia and bacteremia are the most common manifestations of disease; infections of these systems are frequently

**CONTACT** Narisara Chantratita  [narisara@tropmedres.ac](mailto:narisara@tropmedres.ac)  Faculty of Tropical Medicine, Department of Microbiology and Immunology, Mahidol University, 420/6 Ratchawithi Road, Rachathewi, Bangkok 10400, Thailand;  [thatcha\\_yimt@hotmail.com](mailto:thatcha_yimt@hotmail.com)  Faculty of Tropical Medicine, Mahidol-Oxford Tropical Medicine Research Unit, Mahidol University, 420/6 Ratchawithi Road, Rachathewi, Bangkok 10400, Thailand

\*Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai, Thailand

 Supplemental data for this article can be accessed at <https://doi.org/10.1080/22221751.2020.1858176>

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group, on behalf of Shanghai Shangyixun Cultural Communication Co., Ltd. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

associated with septic shock and contribute to high mortality [2].

A comprehensive understanding of the individual response to infection is necessary to develop effective and targeted therapies. Additionally, biomarkers that predict outcome may be useful to guide patient management. Evaluation of the entire transcriptome of cells offers both the possibilities of characterizing pathways activated in disease and identifying potential biomarkers. In murine melioidosis, blood transcriptomic profiling reveals the regulation of many immune pathways, which reflect severity of disease [6] and can be used to identify a potential marker of acute lung infection [7]. Transcriptomic changes have been reported in human melioidosis during acute infection, highlighting the involvement of host immunity against infection [8]. Recent studies based on microarrays showed that blood transcriptional profiles can distinguish *B. pseudomallei* infection from sepsis caused by other microorganisms [9,10]. These studies suggest that these transcriptomic profiles may be useful in understanding the immune response during infection and serve as informative biomarkers of infection. RNA-sequencing (RNA-seq) is a unbiased approach and powerful tool to define the transcriptome [11]. However, to date, RNA-seq has not been used extensively to characterize human melioidosis. The aims of this study were to use RNA-seq (i) to analyze whole blood transcriptomic profiles of acute melioidosis patients to define biological pathways associated with death, and (ii) to identify host prognostic gene biomarkers that are associated with mortality.

## Methods

### Study design and patients

A prospective study of whole blood transcriptomic analyses in 97 individuals with melioidosis was conducted at seven hospitals in Northeast of Thailand: Udon Thani Hospital, Nakhon Phanom Hospital, Mukdahan Hospital, Roi Et Hospital, Buriram Hospital, Surin Hospital, and Sisaket Hospital. This study was part of a multi-centre study of patients aged  $\geq 15$  years who were culture-positive for *B. pseudomallei* from any type of clinical samples and admitted to the hospitals between January 2015 and December 2019. The inclusion and exclusion criteria were described previously [12]. *B. pseudomallei* were identified by biochemical tests and latex agglutination [13] at the microbiology laboratories of the hospitals and further confirmed by Matrix-Assisted Laser Desorption Ionization Mass Spectrometry (MALDI-TOF MS) as previously described [14]. Whole blood samples were collected at the time of enrolment (within 24 h of culture results, defined

as day 0) and day 5, day 12, and day 28 after enrolment. Clinical information was obtained from the medical records. Mortality of patients was recorded at the hospitals or by phone calls for 28 days of follow-up.

Twenty-three healthy individuals aged  $\geq 18$  years were recruited from Udon Thani Hospital and Mukdahan hospital as baseline controls for discovery and validation data sets. Inclusion and exclusion criteria for these controls were previously described [15].

This study was designed by the process of 3 data sets as follows: discovery set, validation set, and follow-up set as described in Supplementary Figure 1.

### Ethical approval

The study was approved by the ethical committees of Faculty of Tropical Medicine, Mahidol University, Udon Thani Hospital, Nakhon Phanom Hospital, Mukdahan Hospital, Roi Et Hospital, Buriram Hospital, Surin Hospital, and Sisaket Hospital. Written informed consent was obtained from all participants or their representatives.

### Sample collection

Three milliliters of whole blood were collected from melioidosis patients and healthy controls into Tempus™ Blood RNA Tubes (Thermo Fisher Scientific) and stored at  $-20^{\circ}\text{C}$  or  $-80^{\circ}\text{C}$  at the hospitals. The frozen samples were transported on dry ice to the laboratory in Bangkok for RNA extraction.

### RNA extraction

Total RNA was extracted from Tempus-stabilized blood using the MagMAX™ for Stabilized Blood Tubes RNA Isolation Kit (Life technologies). Total RNA concentration and its purity were assessed by determining the A260/280 and A260/230 ratios, respectively on the NanoDrop Spectrophotometer (Thermo Fisher Scientific). RNA integrity number (RIN) was assessed with the Agilent RNA 6000 Pico kit on 2100 Bioanalyzer (Agilent Technologies). Genomic DNA contamination was checked by RT-qPCR using primers for the Peptidylprolyl isomerase A (*PPIA*) gene [16].

### Library preparation for RNA-seq

Libraries were prepared from 50 ng of RNA per sample using Ion AmpliSeq™ Transcriptome Human Gene Expression Kit (Thermo Fisher Scientific). Targets of 20,802 genes were amplified with Ion AmpliSeq™ Transcriptome Human Gene Expression core panel (Life Technologies). The primer

sequences were then digested, and DNA adaptors (Ion PI Adaptor and Ion Xpress Barcode Adaptor, Life Technologies) were ligated to the targets. Adaptor ligated targets were purified using the Agencourt AMPure XP reagent (Beckman Coulter) and eluted into an amplification mix containing Platinum PCR SuperMix High Fidelity and Library Amplification Primer Mix (Life Technologies) for further amplification. Size-selection purification was performed using Agencourt AMPure XP reagent (Beckman Coulter). Amplicons were quantified using a Fragment Analyzer<sup>TM</sup> instrument with a DNF-474 High Sensitivity NGS Fragment Analysis Kit (Advanced Analytical Technologies, INC.). Samples were then pooled together with four samples per pool and performed an emulsion PCR on the Ion Chef System using the Ion PI Hi-Q Chef Kit (Life Technologies). The emulsion PCR samples were loaded on Ion PI v3 chips and sequenced on an Ion Proton System using an Ion PI Hi-Q Sequencing 200 Kit chemistry (Life Technologies) to obtain approximately 200 bp read length.

### Transcriptomic data analysis

Sequencing data were generated using Torrent Suite Software version 5.4.0 with AmpliSeq RNA plugin (Thermo Fisher Scientific) and normalized using reads per million mapped reads (RPM) method. The normalized transcripts were analyzed using GeneSpring GX software version 14.9 (Agilent Technologies) to identify differentially expressed genes (DEGs) within the 10<sup>th</sup> -100<sup>th</sup> percentile. One-way ANOVA was used to compare DEGs among non-survivors, survivors, and healthy controls. Moderated t-test was used to compare DEGs between non-survivors and survivors. An adjusted *P* value < 0.05 was deemed significant (Benjamini-Hochberg correction method). Functional analysis was derived using Metascape tool (<http://metascape.org>). Area under the receiver operating characteristic curves (AUROC) were plotted using GraphPad Prism version 6.0.

DEGs were initially selected for validation based on fold change  $\geq 2$  and adjusted *P* value  $\leq 0.05$  between non-survivors and survivors.

### Quantitative reverse-transcriptase PCR (RT-qPCR)

Two-step RT-qPCR was used to quantitatively validate gene expression. Total RNA from whole blood was converted into cDNA using the iScript<sup>TM</sup> cDNA Synthesis Kit (Bio-Rad). The amplification was performed in duplicate in a total volume of 10  $\mu$ l containing 5  $\mu$ l of iTaq Universal SYBR Green (Bio-Rad), 2  $\mu$ l of 4 ng cDNA, 0.4  $\mu$ l of 10 mM forward primer, 0.4  $\mu$ l of 10 mM reverse primer, 2.2  $\mu$ l of distilled water. The cycle conditions were as follows: 1 cycle of 95°C for

30s followed by 40 cycles of 95°C for 10s and 60°C for 30s. After amplification, melting curve analysis was carried out from 65°C to 95°C. Primers were designed using NCBI PrimerBlast (<https://www.ncbi.nlm.nih.gov/tools/primer-blast/>). All primer pairs are listed in Supplementary Table 1. Peptidylprolyl isomerase A (*PPIA*), Human large ribosomal protein P0 (*RPLP0*), and Tata-box binding protein (*TBP*) were used as reference genes for calculating the relative expression levels of other genes [16] The expression levels were calculated by using the  $2^{-\Delta Ct}$  method, where  $\Delta Ct = \text{mean Ct of target gene} - \text{mean Ct of the three reference genes}$ .

### Statistical analysis

Mann-Whitney or Kruskal-Wallis tests followed by Dunn's multiple comparison tests correction were used to test the difference in gene expressions among subject groups. Mean, median, interquartile range (IQR), standard deviation (SD), area under the receiver operating characteristic curve (AUROC) values and 95% confidence intervals (CI) were assessed using Prism 6 (GraphPad Software). The classification accuracy of the 12 gene signature was determined using the randomForest machine learning R package (v. 4.16) [17] applied to the qRT-PCR data. The AUROC curve was visualised using the pROC package (v. 1.10).

## Results

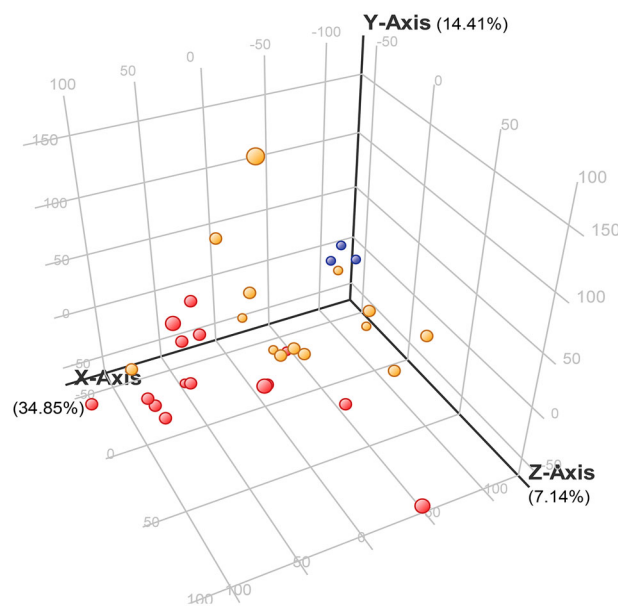
### Whole blood transcriptomic profiles of survivors and non-survivors

To identify genes associated with mortality, we performed whole blood transcriptomic analysis of a discovery set consisting of 29 Thai melioidosis patients, fourteen of whom survived and fifteen of whom died within 28 days, and 3 healthy controls. The clinical characteristics of the patients are shown in Table 1. The quality of 32 RNA samples were analyzed for integrity and read count/mapped read numbers. Overall average RNA integrity numbers (RIN) of 6.0-8.6, average OD ratios 260/280 > 1.8, 260/230 < 1, and average of 22 million reads with mapping rate of >58% were achieved from each cDNA library. Out of 20,802 genes, 18,713 genes with expression values between 10<sup>th</sup> - 100<sup>th</sup> percentiles were further analyzed using one-way ANOVA and 5,189 genes were statistically different among groups as shown in three dimensional principal component analysis (3D-PCA) plots (Figure 1).

Analysis of differentially expressed genes (DEGs) between non-survivors and survivors performed using the moderated t-test method identified 283 DEGs. Hierarchical cluster analysis of these genes

**Table 1.** Characteristics of melioidosis patients and healthy controls.

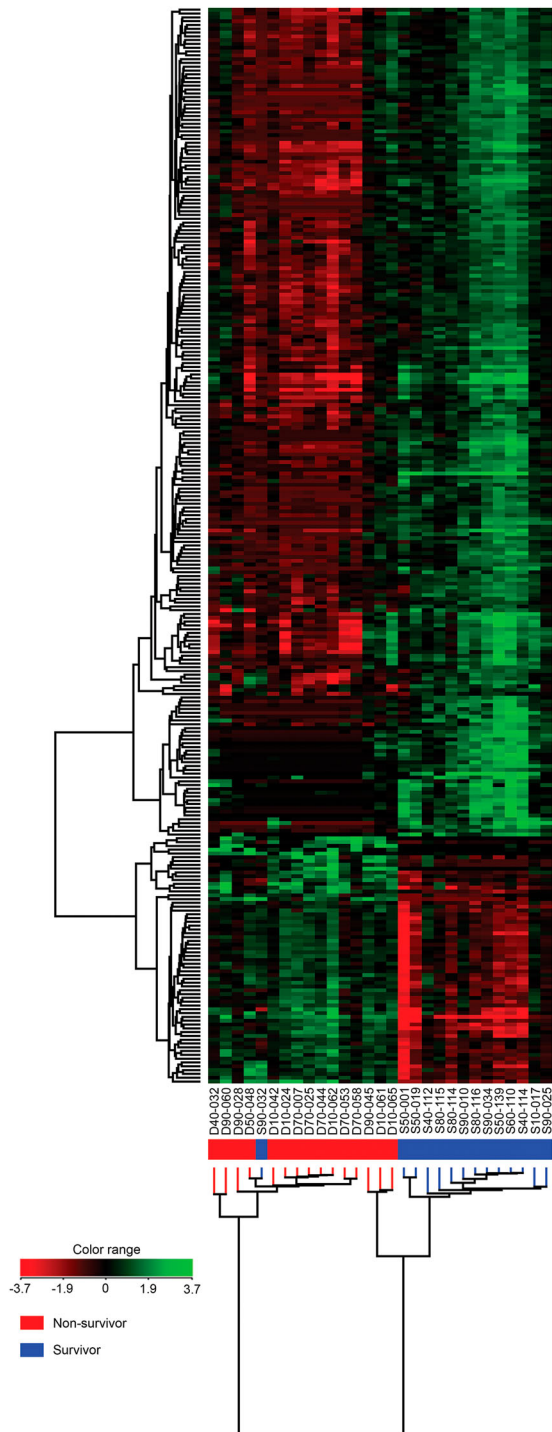
Characteristics	Discovery cohort		Validation cohort		Follow-up cohort (n=8)	Healthy control (n=23)
	Non-survivors (n=15)	Survivors (n=14)	Non-survivors (n=30)	Survivors (n=30)		
<b>Mean age in years (range)</b>	57 (36–81)	51 (28–74)	62 (45–84)	60 (34–80)	50 (32–70)	43 (28–68)
<b>Male (%)</b>	11 (73%)	10 (71%)	26 (87%)	21 (70%)	8 (100%)	14 (61%)
<b>Comorbidity</b>						
Diabetes (%)	9 (60%)	7 (50%)	18 (60%)	17 (57%)	7 (88%)	–
Alcoholism (%)	3 (20%)	6 (43%)	7 (23%)	10 (33%)	3 (38%)	–
Kidney disease (%)	2 (13%)	1 (7%)	5 (17%)	3 (10%)	5 (63%)	–
Hypertension (%)	5 (33%)	2 (14%)	13 (43%)	7 (23%)	3 (38%)	–
Thalassemia (%)	–	2 (14%)	–	1 (3%)	–	–
Cancer (%)	2 (13%)	–	2 (7%)	1 (3%)	–	–
None (%)	3 (20%)	1 (7%)	4 (13%)	4 (13%)	–	23 (100%)
<b>Clinical symptom</b>						
Bacteremia (%)	14 (93%)	8 (57%)	28 (93%)	23 (77%)	7 (88%)	–
<b>Fever</b>						
<15 days (%)	14 (93%)	11 (79%)	28 (93%)	23 (77%)	7 (88%)	–
≥15 days (%)	1 (7%)	3 (21%)	2 (7%)	7 (23%)	1 (13%)	–

**Figure 1.** Three-dimensional principal component analysis (3D-PCA) of differentially expressed genes among non-survivors and survivors and healthy controls. One point per subject in yellow, red, and light blue, represents groups of melioidosis patients who survived ( $n=14$ ) and did not survive ( $n=15$ ), and healthy controls ( $n=3$ ), respectively. Each axis shows percent variation explained by each group.

was generated by GeneSpring (Figure 2). Whole blood of non-survivors presented more down-regulated genes compared to survivors (fold change  $\geq 2$ ). RNA-seq data of 65 up-regulated genes and 218 down-regulated genes with  $P$  value  $\leq 0.05$  and fold change  $\geq 2$  are shown in Supplementary Table 2. In comparison to melioidosis patients who survived, the fold changes of up-regulated genes in non-survivors ranged between 2.00–15.72 and  $P$  value =  $1.70 \times 10^{-3}$  to  $5.47 \times 10^{-9}$ . The fold change of down-regulated genes ranged between 2.00–9.42 and  $P$  value =  $9.50 \times 10^{-5}$  to  $2.54 \times 10^{-9}$ . The volcano plot in Figure 3 shows the distribution and relationship between fold change and  $P$  value of 65 up-regulated genes and 218 down-regulated genes in non-survivors in relation to survivors.

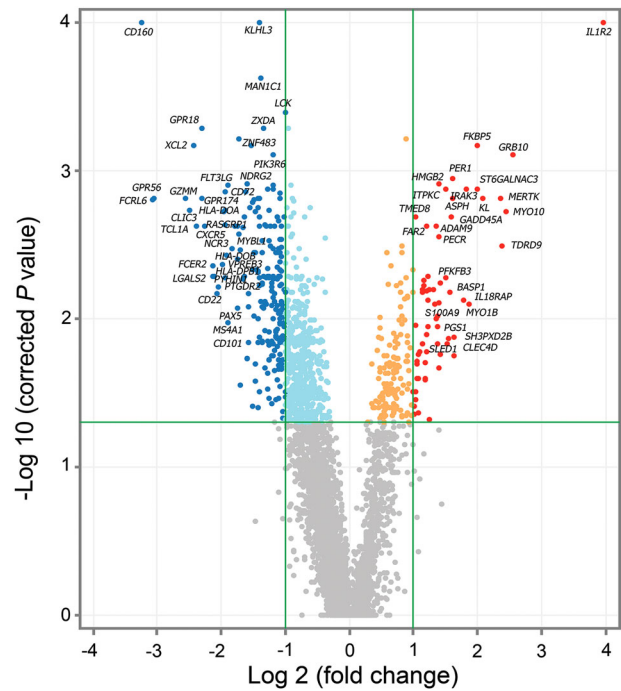
### Functional enrichment analysis of DEGs between survivors and non-survivors

In order to gain insight into the biological function of DEGs, the genes found significantly differential expressed (65 up-regulated and 218 down-regulated) between survivors and non-survivors were analyzed using the Metascape tool. The analysis was based on combined datasets for enrichment analysis, including gene ontology, KEGG pathways, reactome gene sets, canonical pathways, and CORUM complexes. The data in Figure 4 shows that the significant DEGs were involved in functions of host immune response ( $n=7$ ), stress response ( $n=6$ ), cell development ( $n=35$ ), signaling transduction ( $n=23$ ), catabolic process ( $n=16$ ), and metabolic process ( $n=24$ ). The



**Figure 2.** Hierarchical clustering analysis of 283 differentially expressed genes (DEGs) in whole blood of surviving and non-surviving melioidosis patients. High expression of genes is shown in green whereas low expression of genes is shown in red. Each column represents individual subjects and each row in the figure represents one altered gene that significantly expressed at  $P \leq 0.05$  and fold change  $\geq 2$ . Subjects from our study are melioidosis survivors ( $n = 14$ ), melioidosis non-survivors ( $n = 15$ ).

significant 65 up-regulated DEGs in non-survivors were involved in myeloid leukocyte activation ( $n = 14$ ), Toll-like receptor cascades ( $n = 8$ ), and reactive oxygen species metabolic processes ( $n = 8$ ) (Figure 4A) while the majority of 218 down-regulated genes

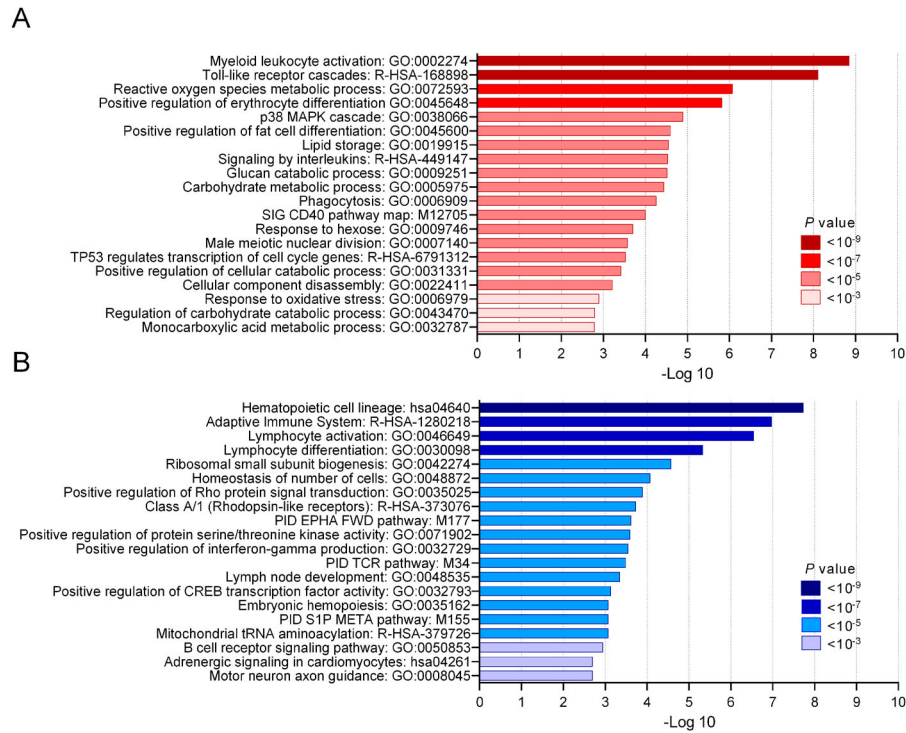


**Figure 3.** Differential expression analysis of survivors compared to non-survivors at the time of diagnosis (day 0). Gene expression profile of patients with melioidosis that survived after 28 days ( $n = 14$ ) compared to patients that did not survive ( $n = 15$ ). Color indicates statistically significant genes (adjusted  $P$  value  $\leq 0.05$ , correction method = Benjamini-Hochberg), dark blue: down-regulated genes  $\geq 2$  fold change, dark red: up-regulated genes  $\geq 2$  fold change with grey corresponding to genes showing no expression change.

set in non-survivors were hematopoietic cell lineage ( $n = 10$ ), adaptive immune system ( $n = 24$ ) and lymphocyte activation ( $n = 23$ ) (Figure 4B). Gene names and details of each functional group are shown in Supplementary Table 3.

### Pathway analysis of DEGs between melioidosis survivors and non-survivors

To gain better understanding of the underlying mechanisms of the 283 altered genes in non-survivors compared to survivors, we performed KEGG pathway analysis. Interestingly, KEGG identified six pathways in immunological response that were associated with 65 up-regulated genes (Supplementary Table 4). These included pathways of Toll-like receptor signaling, Th17 cell differentiation, MAPK, IL-17 signalling, FoxO signalling, HIF-1 signalling. Moreover, KEGG identified seven pathways in immunological response that were associated with 218 down-regulated genes. These included hematopoietic cell lineage, cell adhesion molecules (CAMs), intestinal immune network for IgA production, Th1 and Th2 cell differentiation, Th17 cell differentiation, antigen processing and presentation and B cell receptor signalling pathway.

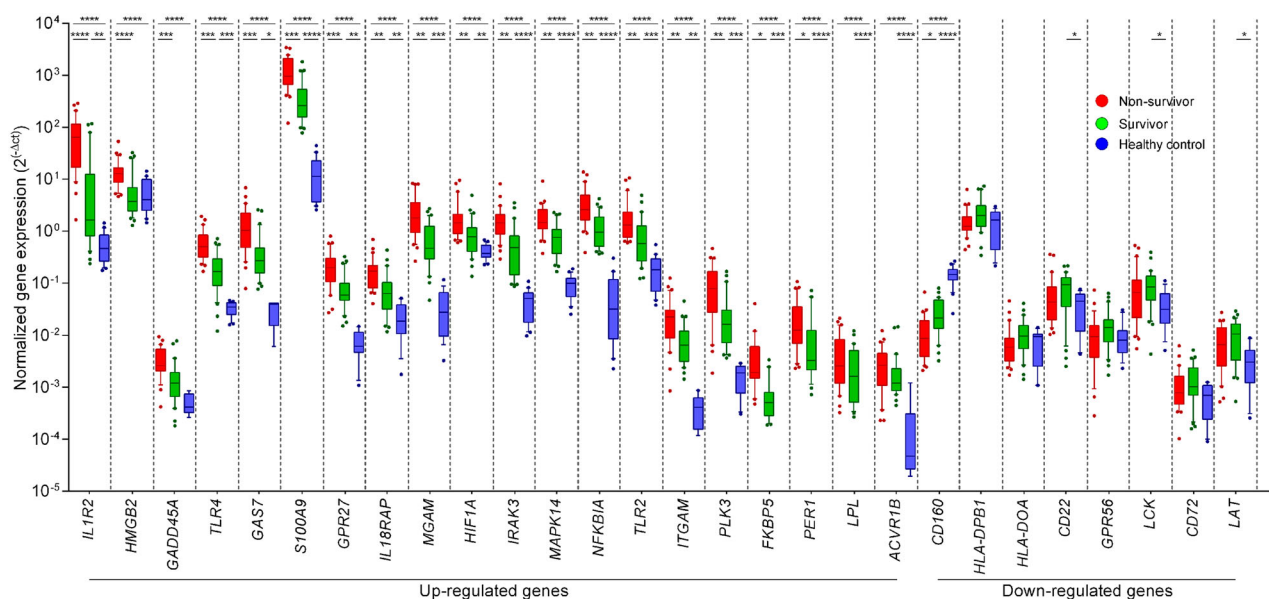


**Figure 4.** Functional enrichment analysis of DEGs in non-surviving melioidosis patients compared with patients that survived. (A) Top 20 enriched terms of 65 up-regulated genes in non-surviving melioidosis patients. (B) Top 20 enriched terms of 218 down-regulated genes in non-surviving melioidosis. Saturation of color corresponds to  $P$  values.

### RT-qPCR validation of DEGs to predict mortality in melioidosis

Twenty-eight DEGs were manually selected to confirm the expression by RT-qPCR in a validation set of 30 non-survivors, 30 survivors and 20 healthy controls. The DEGs were selected according to (i) their degree of alteration (fold changes and  $P$  value) (Supplementary Table 2) and (ii) their functions related with

immunological responses (Supplementary Table 4). These DEGs included 20 up-regulated genes and 8 down-regulated. RT-qPCR results in the validation set confirmed significantly higher expression in non-survivors compared with survivors and healthy controls for 16 of the 20 up-regulated genes and 1 of the 8 down-regulated genes, respectively (Figure 5 and Supplementary Table 5). RT-qPCR in the validation set confirmed significantly lower expression in non-survivors compared



**Figure 5.** Validation of the differential expression analysis of 28 DEGs in whole blood from melioidosis patients. Genes that were found to be differentially expressed in patients with melioidosis that did not survive and survived were validated with real-time qPCR. The Kruskal–Wallis test was performed for comparing three groups. Subjects from our study were melioidosis survivors ( $n = 30$ ), melioidosis non-survivors ( $n = 30$ ), and healthy controls ( $n = 20$ ). \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.005$ , \*\*\*\* $P < 0.0001$ .

with survivors ( $P = 0.016$ ) and healthy controls ( $P < 0.0001$ ) for 1 of 8 down-regulated genes: *CD160*.

### ROC assessment of gene expression as predictive markers for mortality

Receiver operating characteristic (ROC) curves were constructed based on the RT-qPCR results from the validation set of melioidosis patients to examine the classification accuracy of each DEG for distinguishing between non-survivors and survivors (Figure 6A-C). The highest area under the ROC (AUROCC) were obtained from the genes listed in Supplementary Table 6. Among these, *S100A9* showed the highest AUROCC value (0.88) followed by *IL1R2* (0.87) and *TLR4* (0.86). The down-regulated gene with the highest AUROCC was *CD160* (0.77). A combined signature of the expression of the 12 genes with best individual discriminatory ability was able to classify the non-survivors from the survivors in a Random Forest model (AUROCC 0.85, CI = 0.74–0.94), and completely discriminated the melioidosis patients from the healthy controls (Figure 6D).

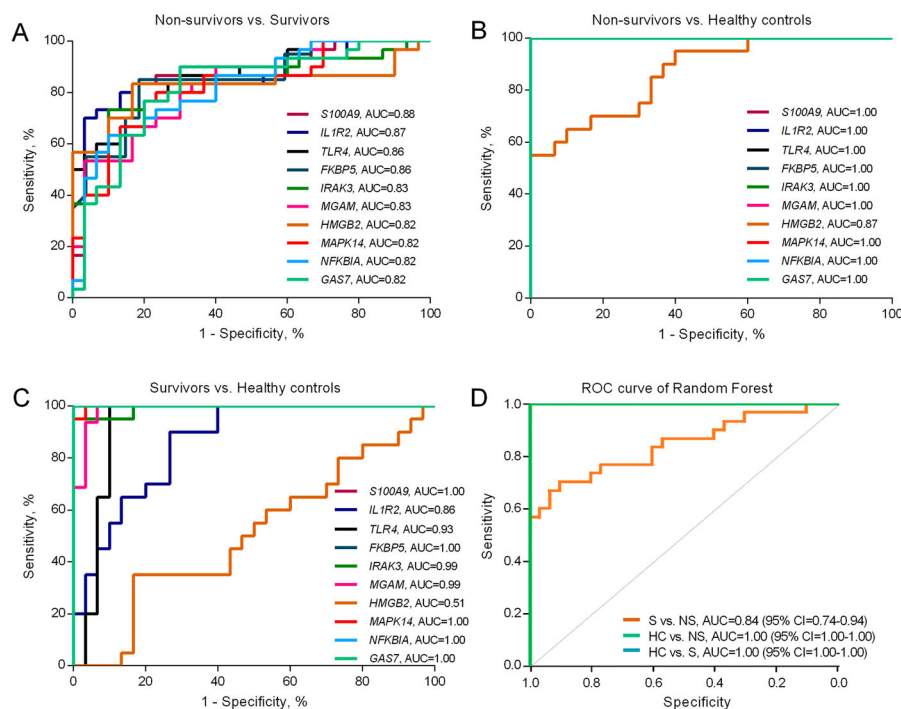
### Trajectory of gene expression profiles in survivors after enrolment

Five up-regulated DEGs (*S100A9*, *IL1R2*, *IRAK3*, *NFKBIA* and *GAS7*) were selected based on

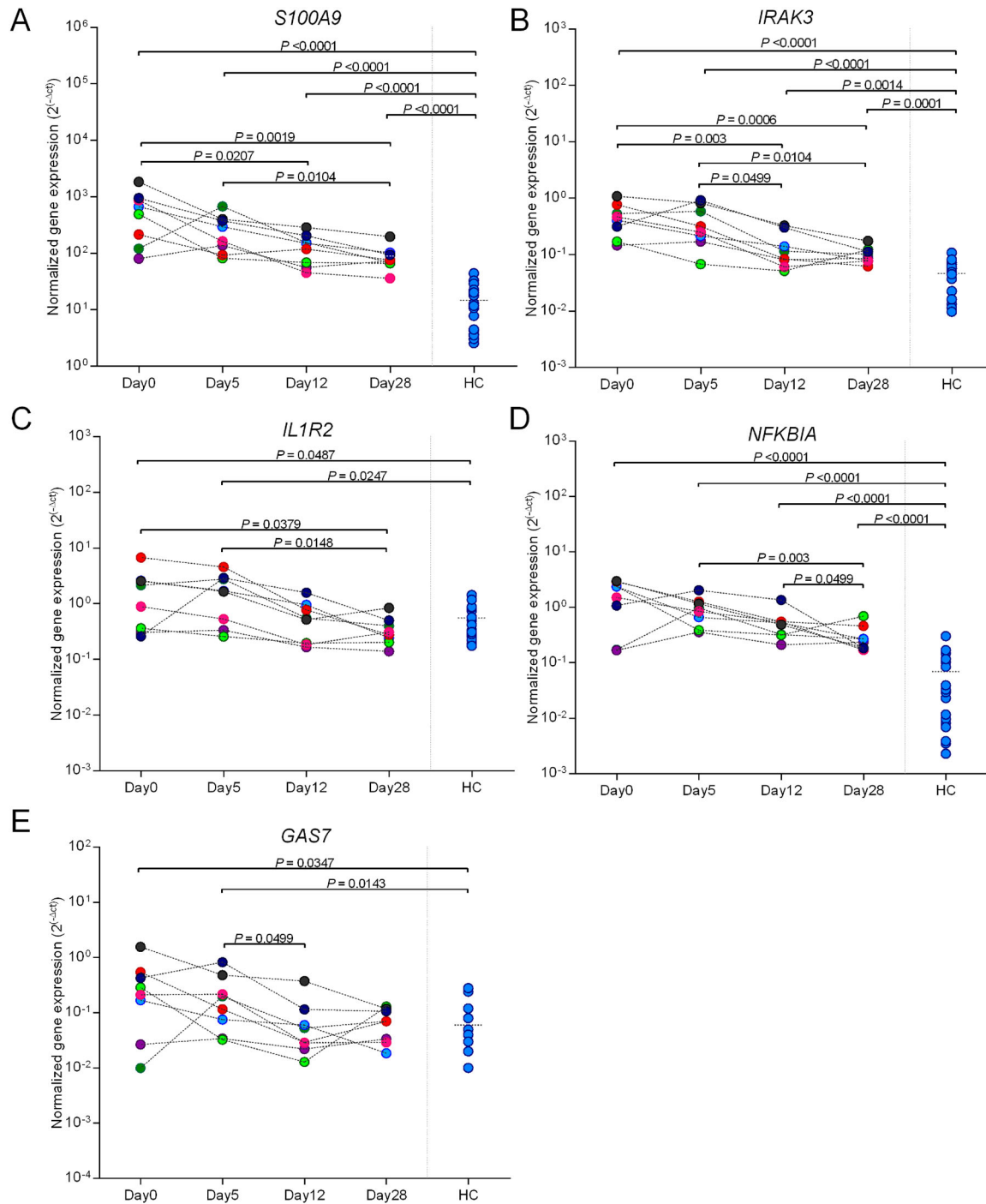
AUROCC  $\geq 0.82$  and whether the genes have secretory functions of proteins as they may be better suited to a point-of-care assay. Gene expression was measured by RT-qPCR in survivors ( $n = 8$ ) at day 0, day 5, day 12, and day 28 to test whether expression decreases as patients recovered. The trend of gene expression at day 0, day 5, day 12, and day 28 were determined by calculating the fold change reduction. None of the five genes had major changes in expression at day 5 but *S100A9*, *IRAK3* and *IL1R2* subsequently had decreased expression over time as patients recovered (Figure 7 and Supplementary Table 7). Expression of *S100A9*, *IRAK3*, *IL1R2* and *NFKBIA* significantly decreased at day 28 relative to day 5. Expression of *S100A9*, *IRAK3*, and *NFKBIA* in patients decreased at day 28 but did not reach to the expression level of healthy controls ( $P < 0.0001$ ). However, expression of *IL1R2* and *GAS7* rapidly decreased to the same level of healthy controls and did not change further after day 12 ( $P < 0.05$ ). The mean fold changes (day 28/day 5) for gene expression of 8 individual patients and 95% CI are shown in Supplementary Table 8.

### Discussion

Our study demonstrated that the whole blood transcriptome of melioidosis patients who survived was distinguishable from non-survivors, with 283 DEGs



**Figure 6.** Area under the receiver operating characteristic curve (AUROCC) of DEGs in discrimination among non-survivors, survivors and healthy controls. (A) AUROCC of 10 DEGs between non-survivors versus survivors. (B) AUROCC of 10 DEGs between non-survivors versus healthy controls. (C) AUROCC of 10 DEGs between survivors versus healthy controls. (D). Random Forest model of a combined gene signature discriminates survivors and non-survivors. The 12 genes which individually discriminated clinical groups with AUROCC  $> 0.80$  in qRT-PCR were combined to create a single model, which was used to classify the separation between survivors (S), non-survivors (NS) and healthy controls (HC) in the qRT-PCR dataset.



**Figure 7.** One month follow-up of *S100A9*, *IRAK3*, *IL1R2*, *GAP7*, and *NFKB1A* in surviving melioidosis patients over the course of illness. Whole blood samples from melioidosis survivors ( $n = 8$ ) were collected at the various times from diagnosis (day 0, day 5, day 12, and day 28). The  $P$ -values were calculated by Mann–Whitney test. Data of healthy individuals were plotted as the controls.

significantly associated with mortality. The majority of these DEGs were related to the immune response, cellular functions and metabolism. Twenty-eight DEGs were selected by functional enrichment and pathway analyses and RT-qPCR of these genes in a validation cohort confirmed 16 up-regulated and 1 down-regulated gene associated with mortality. ROC analyses of the validation set identified the 15 most predictive genes. Subsequent RT-qPCR of four selected genes (*S100A9*, *IRAK3*, *IL1R2*, and *NFKB1A*) in surviving patients followed over time demonstrated a trajectory

expression profile with decreased differential expression by day 12 and day 28 after enrolment.

Genes of melioidosis patients associated with death include *IL1R2*, *IRAK3*, *IL18RAP*, *MGAM*, *LPL*, *HGMB2*, *S100A9*, *GAS7*, *NFKB1A*, *TLR2*, *TLR4*, *MAPK14*, *GPR27*, *HIF1A*, and *ITGAM*. Many of these genes or their proteins have been reported in related studies. Elevation of *IL1R2* expression and soluble *IL1R2* concentrations are correlated with severity of *Escherichia coli* and *Staphylococcus aureus* infections [18]. Increased expression levels of the *IRAK3*



gene are correlated with the development of acute lung injury in patients with severe sepsis [19]. In melioidosis, Wiersinga et al. reported up-regulation of *IRAK3* is related to attenuated capacity of monocytes to respond to *B. pseudomallei* stimulation and this coincided with mortality [20]. In parallel to our study, a recent study reported that extracellular S100A8 and S100A9 (S100A8/A9), a Ca<sup>2+</sup> sensor in cytoskeleton rearrangement and arachidonic acid metabolism, are the key mediators of sepsis secreted from neutrophils and monocytes during inflammation [21]. The S100A9 serve as damage associated molecular patterns and induce pro-inflammatory cytokine expression and secretion via toll-like receptor 4 (TLR4) activation [22,23]. Increasing evidence supports that *NFKBIA*-mediated inflammation is linked to susceptibility to infectious and inflammatory diseases [24–26]. A report demonstrated an up-regulation of *NFKBIA* expression in mouse macrophages in response to *B. pseudomallei* infection [27] and our data confirmed that increased *NFKBIA* expression is associated with fatality in melioidosis patients.

A recent study suggests that *HLA-DPA1* and *-DRB3* are under-expressed in whole blood of sepsis patients caused by *B. pseudomallei*, which distinguished melioidosis from sepsis caused by other organisms [9]. In addition, we found *HLA-DPB1* was down-regulated in non-survivors in our discovery cohort. Our data also revealed that non-survivors had reduced expression of *HLA-DPB1*, *HLA-DOA*, *HLA-DOB*, and *HLA-DRA* representing MHC class II molecules, which are important for antigen presentation. Our results in melioidosis are similar to the results of other studies [28–30] suggesting that non-surviving patients with severe sepsis from melioidosis or other infections exhibit decreased MHC class II expression and that can contribute to persistent failure of T cell activation [31,32]. We did not observe the changes of these MHC class I at transcriptional levels. However, Dunachie et al. showed the presence of MHC class I genes, *HLA-B46*, and *HLA-C\*01* was associated with an increased mortality in an acute melioidosis cohort [8].

Enrichment analysis demonstrated a number of GO terms, including the up-regulation of myeloid leukocyte activation and down-regulation of lymphocyte activation in non-survivors compared with survivors. KEGG pathway analysis revealed many up-regulated genes involved in signal transduction pathways associated with severe melioidosis. Among these, TLRs are known to recognize *B. pseudomallei* LPS and initiate inflammation [33–36] and acute septic melioidosis patients had increased expression of many TLRs in leukocytes [34]. The activation of MAPK signaling and Th17 pathway in melioidosis patients have also been demonstrated in previous studies [37–40]. Multiple signaling pathways were

down-regulated in severe melioidosis suggesting that prolonged bacterial persistence exacerbates inflammatory responses that may lead to immune exhaustion, immune suppression, and poor outcome of the disease.

Expression of several genes, assayed on day 0, had high mortality discrimination, including *S100A9* and *IL1R2*. Notably, expression of these genes decreased significantly in surviving patients by day 12, suggesting that the gene expression tracks with clinical condition. Therefore, these genes and their encoded proteins could be considered as candidate biomarkers for predicting clinical outcomes in patients with melioidosis, and deserve further study in comparison to other clinical and biological prediction tools.

Strengths of our study were the multi-center design, prospective subject enrolment and sample collection, serial sampling over time in a subset of patients, and validation of selected findings. Some limitations are the relatively small number of samples in the discovery cohort, enrolment into our study only after the diagnosis of melioidosis was confirmed (rather than at the time of admission to hospital), and validation of only a subset of genes.

In conclusion, our findings provide new knowledge about transcriptional host responses in circulating leukocytes from hospitalized melioidosis patients and suggest several candidate biomarkers for further study. These data are important to ongoing efforts to reduce the burden of this often severe infection.

## Acknowledgements

We would like to thank staff from Department of Microbiology and Immunology, and Mahidol-Oxford Tropical Medicine Research Unit, Faculty of Tropical Medicine, Mahidol University; Center for Medical Genomics, Faculty of Medicine, Ramathibidi Hospital; Department of Clinical Immunology, Faculty of Associated Medical Science, Khon Kaen University; Udon Thani Hospital; Nakhon Phanom Hospital; Mukdahan Hospital; Roi Et Hospital Buriram Hospital; Surin Hospital and Sisaket Hospital for their assistance.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## Funding

This work was supported by The Royal Society, UK and the Office of the Higher Education Commission of Thailand under Newton Mobility Grant Award (grant number NI170152), and the National Institute of Allergy and Infectious Diseases of the National Institutes of Health, USA (grant number U01AI115520). The content is solely the responsibility of the authors and does not necessarily represent the official views of the funders.

## Author contributions

JMC, TEW, and NC designed the study; TY, JSL, TK, MA, PE, WC, JMC, GL, TEW, and NC conducted the experiments; RP, TY, and TK acquired data; TY, RP, JSL, CE, TK, JMC, TEW, WC, and NC analyzed data; NC provided samples or reagents; TY, TK, MA, JMC, TEW and NC wrote the manuscript.

## ORCID

T. Eoin West  <http://orcid.org/0000-0001-5503-7204>  
Narisara Chantratita  <http://orcid.org/0000-0003-3906-7379>

## References

- Mukhopadhyay C, Shaw T, Varghese GM, et al. Melioidosis in South Asia (India, Nepal, Pakistan, Bhutan and Afghanistan). *Trop Med Infect Dis*. 2018 May 22;3(2):51. doi:10.3390/tropicalmed3020051
- Chin CY, Monack DM, Nathan S. Genome wide transcriptome profiling of a murine acute melioidosis model reveals new insights into how *Burkholderia pseudomallei* overcomes host innate immunity. *BMC Genom*. 2010 Nov 27;11:672. doi:10.1186/1471-2164-11-672
- Wiersinga WJ, Virk HS, Torres AG, et al. Melioidosis [Review]. *Nat Rev Dis Prim*. 2018 Feb 1;4:17107. doi:10.1038/nrdp.2017.107
- Cheng AC, Dance DA, Currie BJ. Bioterrorism, glanders and melioidosis. *Euro Surveill*. 2005 Mar;10(3):11–12. doi:10.2807/esm.10.03.00528-en
- Cheng AC, Currie BJ. Melioidosis: epidemiology, pathophysiology, and management. *Clin Microbiol Rev*. 2005 Apr;18(2):383–416. doi:10.1128/CMR.18.2.383-416.2005
- Conejero L, Potempa K, Graham CM, et al. The blood transcriptome of experimental melioidosis reflects disease severity and shows considerable similarity with the human disease. *J Immunol*. 2015 Oct 1;195(7):3248–3261. doi:10.4049/jimmunol.1500641
- Clark GC, Essex-Lopresti A, Moore KA, et al. Common host responses in murine aerosol models of infection caused by highly virulent gram-negative bacteria from the Genera *Burkholderia*, *Francisella* and *Yersinia*. *Pathogens*. 2019 Sep 21;8(4):159. doi:10.3390/pathogens8040159
- Dunachie SJ, Jenjaroen K, Reynolds CJ, et al. Infection with *Burkholderia pseudomallei* - immune correlates of survival in acute melioidosis. *Sci Rep*. 2017 Sep 22;7(1):12143. doi:10.1038/s41598-017-12331-5
- Pankla R, Buddhisa S, Berry M, et al. Genomic transcriptional profiling identifies a candidate blood biomarker signature for the diagnosis of septicemic melioidosis. *Genome Biol*. 2009;10(11):R127. doi:10.1186/gb-2009-10-11-r127
- Krishnananthasivam S, Jayathilaka N, Sathkumara HD, et al. Host gene expression analysis in Sri Lankan melioidosis patients. *PLoS Negl Trop Dis*. 2017 Jun;11(6):e0005643. doi:10.1371/journal.pntd.0005643
- Chaussabel D, Pascual V, Banchereau J. Assessing the human immune system through blood transcriptomics. *BMC Biol*. 2010 Jul 1;8:84. doi:10.1186/1741-7007-8-84
- Kaewarpai T, Ekchariyawat P, Phunpang R, et al. Longitudinal profiling of plasma cytokines in melioidosis and their association with mortality: a prospective cohort study. *Clin Microbiol Infect*. 2020 Jun;26(6):783.e1–783.e8. doi:10.1016/j.cmi.2019.10.032.
- Duval BD, Elrod MG, Gee JE, et al. Evaluation of a latex agglutination assay for the identification of *Burkholderia pseudomallei* and *Burkholderia mallei*. *Am J Trop Med Hyg*. 2014 Jun;90(6):1043–1046. doi:10.4269/ajtmh.14-0025
- Suttisunhakul V, Pumpuang A, Ekchariyawat P, et al. Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry for the identification of *Burkholderia pseudomallei* from Asia and Australia and differentiation between *Burkholderia* species. *PLoS One*. 2017;12(4):e0175294. doi:10.1371/journal.pone.0175294
- Wright SW, Kaewarpai T, Lovelace-Macon L, et al. A two biomarker model augments clinical prediction of mortality in melioidosis. *Clin Infect Dis*. 2020 Feb 8: cial126. doi:10.1093/cid/ciaa126
- Cliff JM, Cho JE, Lee JS, et al. Excessive cytolytic responses predict tuberculosis relapse after apparently successful treatment. *J Infect Dis*. 2016 Feb 1;213(3):485–495. doi:10.1093/infdis/jiv447
- Breiman L. Random forests. *Mach Learn*. 2001 October;45:5–32. doi:10.1023/A:1010933404324
- Lang Y, Jiang Y, Gao M, et al. Interleukin-1 receptor 2: a new biomarker for sepsis diagnosis and Gram-negative/Gram-positive bacterial differentiation. *Shock*. 2017 Jan;47(1):119–124. doi:10.1097/SHK.0000000000000714
- Pino-Yanes M, Ma SF, Sun X, et al. Interleukin-1 receptor-associated kinase 3 gene associates with susceptibility to acute lung injury. *Am J Respir Cell Mol Biol*. 2011 Oct;45(4):740–745. doi:10.1165/rcmb.2010-0292OC
- Wiersinga WJ, van't Veer C, van den Pangaart PS, et al. Immunosuppression associated with interleukin-1R-associated-kinase-M upregulation predicts mortality in Gram-negative sepsis (melioidosis). *Crit Care Med*. 2009 Feb;37(2):569–576. doi:10.1097/CCM.0b013e318194b1bf
- Wang S, Song R, Wang Z, et al. S100a8/A9 in inflammation. *Front Immunol*. 2018;9:1298. doi:10.3389/fimmu.2018.01298
- Schenten V, Plancon S, Jung N, et al. Secretion of the phosphorylated form of S100A9 from neutrophils is essential for the proinflammatory functions of extracellular S100A8/A9. *Front Immunol*. 2018 Mar 13;9:447. doi:10.3389/fimmu.2018.00447
- Leanderson T, Liberg D, Ivars F. S100a9 as a pharmacological target molecule in inflammation and cancer. *Endocr Metab Immune Disord Drug Targets*. 2015;15(2):97–104. doi:10.2174/1871530315666150316123854
- Zhang GL, Zou YF, Feng XL, et al. Association of the *NFKBIA* gene polymorphisms with susceptibility to autoimmune and inflammatory diseases: a meta-analysis. *Inflamm Res*. 2011 Jan;60(1):11–18. doi:10.1007/s00011-010-0216-2
- Ali S, Hirschfeld AF, Mayer ML, et al. Functional genetic variation in *NFKBIA* and susceptibility to childhood asthma, bronchiolitis, and bronchopulmonary dysplasia. *J Immunol*. 2013 Apr 15;190(8):3949–3958. doi:10.4049/jimmunol.1201015

- [26] von Bernuth H, Puel A, Ku CL, et al. Septicemia without sepsis: inherited disorders of nuclear factor-kappa B-mediated inflammation. *Clin Infect Dis*. 2005 Nov 15;41(Suppl 7):S436–S439. doi:10.1086/431994
- [27] Peng D, Pang F, Cao R, et al. Upregulation of immune process-associated genes in RAW264.7 Macrophage cells in response to *Burkholderia pseudomallei* infection. *Biomed Res Int*. 2018 Jun 4;2018:1235097. doi:10.1155/2018/1235097
- [28] Pachot A, Monneret G, Brion A, et al. Messenger RNA expression of major histocompatibility complex class II genes in whole blood from septic shock patients. *Crit Care Med*. 2005 Jan;33(1):31–38. discussion 236–7. doi:10.1097/01.CCM.0000150958.20209.A3
- [29] Lekkou A, Karakantza M, Mouzaki A, et al. Cytokine production and monocyte HLA-DR expression as predictors of outcome for patients with community-acquired severe infections. *Clin Diagn Lab Immunol*. 2004 Jan;11(1):161–167. doi:10.1128/CDLI.11.1.161-167.2004
- [30] Zhuang Y, Peng H, Chen Y, et al. Dynamic monitoring of monocyte HLA-DR expression for the diagnosis, prognosis, and prediction of sepsis. *Front Biosci (Landmark Ed)*. 2017 Mar 1;22:1344–1354. doi:10.2741/4547
- [31] Holling TM, van der Stoep N, Quinten E, et al. Activated human T cells accomplish MHC class II expression through T cell-specific occupation of class II transactivator promoter III. *J Immunol*. 2002 Jan 15;168(2):763–770. doi:10.4049/jimmunol.168.2.763
- [32] Poluektov YO, Kim A, Sadegh-Nasseri S. HLA-DO and its role in MHC class II antigen presentation. *Front Immunol*. 2013 Aug 29;4:260. doi:10.3389/fimmu.2013.00260
- [33] West TE, Ernst RK, Jansson-Hutson MJ, et al. Activation of toll-like receptors by *Burkholderia pseudomallei*. *BMC Immunol*. 2008 Aug 8;9(46). doi:10.1186/1471-2172-9-46
- [34] Wiersinga WJ, Wieland CW, Dessing MC, et al. Toll-like receptor 2 impairs host defense in gram-negative sepsis caused by *Burkholderia pseudomallei* (melioidosis). *PLoS Med*. 2007 Jul 31;4(7):e248. doi:10.1371/journal.pmed.0040248
- [35] Chantratita N, Tandhavanant S, Myers ND, et al. Screen of whole blood responses to flagellin identifies *TLR5* variation associated with outcome in melioidosis. *Genes Immun*. 2014 Mar;15(2):63–71. doi:10.1038/gene.2013.60
- [36] Sengyee S, Yoon SH, Paksanont S, et al. Comprehensive analysis of clinical *Burkholderia pseudomallei* isolates demonstrates conservation of unique lipid A structure and TLR4-dependent innate immune activation. *PLoS Negl Trop Dis*. 2018 Feb;12(2):e0006287. doi:10.1371/journal.pntd.0006287
- [37] Hii CS, Sun GW, Goh JW, et al. Interleukin-8 induction by *Burkholderia pseudomallei* can occur without Toll-like receptor signaling but requires a functional type III secretion system. *J Infect Dis*. 2008 Jun 1;197(11):1537–1547. doi:10.1086/587905
- [38] Kang WT, Vellasamy KM, Vadivelu J. Eukaryotic pathways targeted by the type III secretion system effector protein, BipC, involved in the intracellular lifecycle of *Burkholderia pseudomallei*. *Sci Rep*. 2016 Sep 16;6:33528. doi:10.1038/srep33528
- [39] Ng MY, Wang M, Casey PJ, et al. Activation of MAPK/ERK signaling by *Burkholderia pseudomallei* cycle inhibiting factor (Cif). *PLoS One*. 2017 Feb 6;12(2):e0171464. doi:10.1371/journal.pone.0171464
- [40] Krishnananthasivam S, Sathkumara HD, Corea E, et al. Gene expression profile of Human Cytokines in response to *Burkholderia pseudomallei* infection. *mSphere*. 2017 Apr 19;2(2):e00121-17. doi:10.1128/mSphere.00121-17