The Blood Transcriptome of Experimental Melioidosis Reflects Disease Severity and Shows Considerable Similarity with the Human Disease

Running Title

Common immune responses in mice and humans with melioidosis.

Laura Conejero[*, ¶¶^], Krzysztof Potempa[§^], Christine M. Graham[§^], Natasha Spink[*], Simon Blankley[§], Rungnapa Pankla-Sranujit[¶], Prasong Khaenam[¶], Jacques F. Banchereau[♯], Virginia Pascual[**], Damien Chaussabel[§§], Ganjana Lertmemongkolchai[¶], Anne O'Garra[§, ♯♯ ^], Gregory J. Bancroft[*, ^]

[*]Immunology and Infection Department, Faculty of Infectious and Tropical Diseases, London School of Hygiene & Tropical Medicine, London, UK.

[§]Division of Immunoregulation, MRC National Institute for Medical Research, London, UK.

[¶]The Centre for Research & Development of Medical Diagnostic Laboratories, Faculty of Associated Medical Sciences, Khon Kaen University, Thailand.

[♯]The Jackson Laboratory, Farmington, CT, USA.

[**]Baylor Institute for Immunology Research, Thousand Oaks, Dallas, Texas, USA.

[§§]Benaroya Research Institute, Seattle, WA, USA.

[¶¶]Current address: Vascular Biology & Inflammation Department, Centro Nacional de

Investigaciones Cardiovasculares, Madrid, Spain.

[##]Respiratory Medicine, Imperial College Healthcare NHS Trust, St. Mary's Hospital, London, UK

^Authors contributed equally.

***To whom correspondence should be addressed:**

gregory.bancroft@lshtm.ac.uk

+442079272361

Abstract

Melioidosis, a severe human disease caused by the bacterium *Burkholderia pseudomallei*, has a wide spectrum of clinical manifestations ranging from acute septicaemia to chronic localized illness or latent infection. Mice were intranasally infected with either high or low doses of *B. pseudomallei* to generate either acute, chronic or latent infection and host blood and tissue transcriptional profiles were generated. Acute infection was accompanied by a homogeneous signature associated with induction of multiple innate immune response pathways, such as IL10, TREM1 and IFN-signaling, largely found in both blood and tissue. The transcriptional profile in blood reflected the heterogeneity of chronic infection and quantitatively reflected the severity of disease. Comparison of these mouse blood datasets by pathway and modular analysis with the blood transcriptional signature of patients with melioidosis showed that many genes were similarly perturbed, including IL10, TREM1 and IFN-signaling, revealing the common immune response occurring in both mice and humans.

Introduction

Melioidosis, a severe human disease caused by infection with the environmental Gram-negative bacterium *B. pseudomallei* (1), is an important cause of community-acquired sepsis in Southeast Asia and Northern Australia where it is endemic. However, its known distribution is expanding with cases of melioidosis now being reported in numerous other countries (2). Infection results mainly from inhalation during the rainy season or from percutaneous inoculation (through direct contact with soil and water) (1). The severity of the disease has a wide spectrum of clinical manifestations. Acute fulminant septic illness is characterized by bacteraemia and induction of an extensive cytokine response (3, 4). In contrast, chronic localized infection is generally less severe with internal-organ abscesses and secondary foci in the lung, spleen, liver and skeletal muscle which persist and are difficult to eradicate (5). Sterilizing immunity does not occur and bacteria persist either subclinically in a latent form or as a chronic localized infection. Recurrent melioidosis is common and is usually due to relapse from a persistent focus of infection rather than to reinfection (6). Reactivation in some individuals has also been reported after decades of latency (7). Despite the increased research efforts since *B. pseudomallei* was classified by the USA Centers for Disease Control and Prevention as a Category B bioterrorism agent, the mechanisms of persistence and pathogenesis are still poorly understood and it is still unclear how *B. pseudomallei* can evade the immune response and cause significant human disease.

Acute models of melioidosis have been described in mice, following either intravenous or the more physiologically relevant pulmonary route of *B. pseudomallei* infection, characterized by the presence of bacteremia and uniformly rapid death (5, 8-10). We have also described a model of chronic melioidosis where one of the main features is the heterogeneity of disease progression within the infected animals (11). Together the acute and chronic mouse models share features

resembling the heterogeneity of the clinical manifestations of human melioidosis, however, which of these models more closely resembles the human disease at a molecular as well as clinical level has not been addressed.

Whole genome transcriptional profiling has previously identified a blood signature in patients with septicemic melioidosis, which included genes related to inflammation, antigen processing and presentation, interferons, neutrophils and T cells (12-14). Transcriptional analysis of liver and spleen has been performed in a model of acute melioidosis by intravenous challenge (15) showing altered expression of genes involved in inflammation, TLR signaling and complement pathways. However, the transcriptional signature of acute and chronic melioidosis after infection via the airways has not been reported and the blood transcriptional changes induced by infection with *B. pseudomallei* in experimental models and human disease have not been compared.

In this study we have performed whole genome transcriptional analysis of tissue and blood of mice following pulmonary challenge with *B. pseudomallei*, during acute and chronic infection, and compared this to the blood transcriptional profiles of patients with acute melioidosis in North East Thailand.

Materials and Methods

Bacteria and infection of mice

B. pseudomallei strain 576 was isolated from a patient with a fatal case of human melioidosis in Thailand, and was kindly provided by Dr. Tyrone Pitt (Health Protection Agency, London, UK). Frozen stocks were prepared as previously described (8). All procedures involving live bacteria were performed under ACDP containment level 3 conditions. Female C57BL/6 mice (Harlan Laboratories, Bicester, Oxon, UK) aged 6-10 weeks were used in accordance with the Animals (Scientific Procedures) Act 1986 and approved by the LSHTM Ethical Review Committee (70/6934). Aliquots of bacteria stocks were thawed, diluted in pyrogen-free saline (PFS) and administered intranasally in 50 μL containing 2500 colony-forming units (CFU; acute model) or 100 CFU (chronic model) as previously described(11). Control uninfected mice received 50 μL PFS.

Preparation of mouse organs and determination of tissue and blood bacterial load

At different time points, mice were euthanized to determine the number of CFU in blood, lung, spleen and liver as described previously (11). The limits of detection were 50 CFU/organ or 20 CFU/mL of blood.

Preparation of RNA from mouse whole blood, lung and spleen tissue

Mice were euthanized by terminal anaesthesia with pentobarbital (Euthatal; Merial, Essex, UK) and whole blood collected into Tempus™ tubes (Applied Biosystems, Paisley, UK); 1:2 ratio (blood:Tempus™ reagent). Samples were immediately vortexed to avoid blood clotting and stored at -80°C until RNA extraction. Lung and spleen were harvested in 2 ml cold TRI-Reagent

(RiboPure™ RNA kit, Life Technologies, Austin, TX), homogenized (Polytron PT 1600 E System; Kinematica AG; Luzern, Switzerland) and stored at -80°C until RNA processing. The PerfectPure RNA Blood kit (5 PRIME, Hamburg, Germany) was used for whole blood RNA extraction, following manufacturer's instructions. Globin mRNA was depleted from 2 µg total RNA blood samples using the GLOBIN-clear™ Mouse/rat kit (Ambion, Life Technologies, Austin, TX). Lung and spleen RNA were extracted from tissue homogenates using the RiboPure™ Kit according to manufacturer's instructions. RNA concentration was determined (NanoDrop 1000, Thermo Scientific; Wilmington, USA) and RNA integrity values were assessed (LabChip GX, Caliper, USA). Samples with RNA integrity values >6 were retained for further processing.

Processing mouse RNA for microarray

RNA (300 ng/sample) was amplified (Illumina® TotalPrep RNA Amplification Kit (Applied Biosystems). cRNA (1.5 µg/sample) was hybridized overnight to Illumina Mouse Whole Genome WG6 V2 BeadChips (Illumina). BeadChips were washed, blocked, stained and then scanned on an Illumina iScan, as per manufacturer's instructions. Genome Studio 2.0 software (Illumina) was then used to perform quality control and generate signal intensity values.

Mouse microarray data analysis

Raw background subtracted data from Genome Studio (Illumina) were generated and then analysed using GeneSpring GX 12.1 software (Agilent Technologies). Signal values were set to threshold level 10, log2 transformed, and normalized using a 75th percentile shift algorithm. Next, per-transcript normalization was applied by dividing each messenger RNA transcript to the

median intensity of all the samples. Transcripts were filtered out if at least 10% of the samples failed to pass a 'present' flag cut off set at 0.99 (detection *p*-value > 0.01). The above normalized and quality controlled data were used to identify differentially expressed transcripts having at least two-fold expression value changes from the median of all transcripts in at least 10% of all samples, which passed Mann-Whitney test with Benjamini-Hochberg false discovery rate correction between infected and uninfected mice ($p<0.01$). Heatmap dendrograms of differentially expressed transcripts were generated by clustering with Pearson Uncentered distance metric and the Average Linkage rule on transcripts, unless otherwise stated. All data collected and analysed adhere to the Minimal Information About a Microarray Experiment (MIAME) guidelines. Mouse microarray data were deposited in the Gene Expression Omnibus (GEO: GSE61106). For Figures 1 and 3, significant transcript lists were identified by stringent statistical filtering comparing infected to uninfected mice (Mann-Whitney test plus Benjamini-Hochberg correction $p<0.01$) on data normalized to the median of all transcripts across all samples and passing a detection value ($p<0.01$) and 2-fold expression value across at least 10% of samples within GeneSpring GX12.1 statistical software.

Human microarray data analysis

Whole blood transcriptional microarray data (Illumina Hu6 V2 Beadchips) from our already-described Training and Test cohort of melioidosis patients and healthy controls (Pankla et al., 2009) were re-analysed. The Training group contained 11 patients and 10 healthy controls, while the Test group contained 13 patients and 9 healthy controls. Differentially expressed transcripts were identified from quality control filtered data (transcripts detected in at least 10% of all samples at detection *p*-value < 0.01 and passing a filter of two-fold median intensity value and

200 intensity units across all samples) by Mann-Whitney test and Benjamini-Hochberg false discovery rate correction between patients and healthy controls ($p<0.01$) as above.

Molecular distance to health

Weighted molecular distance to health/uninfected (MDTH) was calculated as previously described (Pankla et al., 2009). Molecular distance to health/uninfected (MDTH) was determined for each dataset by computing the 'molecular distance' between each *B. pseudomallei* infected sample in relation to all uninfected control samples. Qualifying genes differed from the average baseline expression by at least 200 and 2 standard deviations. Nonparametric one-way ANOVA (Kruskal-Wallis) with Dunn's post-hoc test was applied between sample groups $p<0.01$ (**), $p<0.001$ (***) and $p<0.0001$ (****) between uninfected group are indicated. Significant transcripts (811) were identified by comparing *B. pseudomallei* infected mice without bacteraemia (n=7) to uninfected (n=8) (Mann-Whitney test plus Benjamini-Hochberg correction $p<0.01$) on data normalized to the median of all transcripts across all samples and passing a detection value ($p<0.01$) and 2-fold expression value across at least 10% of samples. Lesion scores were applied to generate a supervised heatmap of the identified transcripts. Bacterial counts from spleens were allocated into two groups according to the presence or absence of visible lesions.

Canonical pathway and immune gene analysis

Ingenuity Pathway Analysis (IPA, Qiagen, www.ingenuity.com) knowledge base was used to identify statistically significant pathways (Fisher's exact test, $p<0.05$) associated with differentially expressed transcripts ($p<0.01$ after Mann-Whitney test and Benjamini-Hochberg

correction) following *B. pseudomallei* infection of blood or tissues (lung, spleen). A list of 3,468 genes associated with an immune GO function was generated by search within IPA (19 March 2014) and the Mouse Genome Database (MGD) at the Mouse Genome Informatics website, The Jackson Laboratory, Bar Harbor, Maine. World Wide Web (URL: http://www.informatics.jax.org) (30 March 2014).

Human-mouse orthologs

Of the genes represented in the Illumina whole genome mouse arrays (WG6 V2), 19,945 had annotated Entrez IDs. Of the 20,268 annotated human genes, 15,976 had mouse orthologs based on the IPA knowledge base. For the 1292 genes common between human Training and Test Sets that were significantly changed in human melioidosis, 1131 had mouse orthologs, which were included in subsequent analysis.

Modular analysis

To compare human and mouse data at the modular level, we used the previously described Affymetrix-derived human module system (16) and translated the module lists to accommodate mouse genes based on Entrez IDs. The 2,051 Illumina/AffyIDs used to compute the Affy Modules were entered into IPA, resolving 1,832 genes common to the Human Illumina Hu6 V2 and Mouse WG6 V2 arrays (16). Raw values for the common 1,832 human-mouse genes were used to generate module heatmaps. Percentages of significantly expressed upregulated or downregulated genes (Student's *t*-test, $p<0.05$ comparing infected and uninfected) in each module are represented by colored intensity areas.

Results

IL10 and TREM1-signaling pathways and immune genes are perturbed in blood and lung upon acute infection of mice with B. pseudomallei as shown by transcriptional analysis

Our initial aim was to investigate whether the blood transcriptional signature of mice infected via the airways with *B. pseudomallei* could reflect transcriptional changes in the lung. Genomewide transcriptional profiles were generated from blood and lung samples from acutely infected C57BL/6 mice at days 1 and 3 post intranasal infection with 2500 CFU *B. pseudomallei* 576. Typically, their bacterial load in the lung at day 2 was around 10^6 CFU and increased up to 10^8 CFU by day 3, whereas 10^3 CFU where detected in the blood by day 2 and these mice usually die at day 4 postinfection (data not shown). Unsupervised analysis of blood and lung samples showed that the transcriptional profile of the groups segregated according to infection as shown by Principal Component Analysis (PCA). (fig. S1). Statistical filtering ($p<0.01$ Mann-Whitney, Benjamini-Hochberg correction of infected versus uninfected mice using GeneSpring analysis) resulted in two main dendrogram clusters based on similarity of gene expression, showing 1325 and 2041 transcripts differentially expressed in the blood and lung, respectively, of infected as compared to uninfected mice at day 1 (Fig. 1A). A greater number of differentially expressed genes consisting of 4810 and 4076 transcripts was observed respectively in the blood and lungs of mice at day 3 postinfection compared to uninfected controls (Fig. 1A). The top 5 Ingenuity Pathway Analysis (IPA) Canonical Pathways upregulated at both time points in blood and lung were highly enriched for Granulocyte Adhesion and Diapedesis, IL10 and TREM1-signaling and other pathways involved in innate immunity (Fig. 1B). The IFN-signaling pathway was one of the top upregulated pathways in the blood but not in the lung at the early time point. Other upregulated pathways included Hepatic Fibrosis and Acute Phase Response. The top 5

downregulated pathways in both blood and lung at days 1 and 3 largely consisted of B- and T-cell-associated genes (Fig. 1C). The transcriptional signature showed 561 transcripts were common to the blood and lung at day 1 and represented 404 genes identified by IPA (Fig. 1D). The 1618 common transcripts detected in both blood and lung at day 3 postinfection represented 1236 genes (Fig. 1E).

Further analysis of blood and lung genes associated with acute infection, as defined by both the IPA Knowledge Base and by Gene Ontology (GO) resources available through the Mouse Genome Database, revealed perturbation of a significant number of genes not obviously associated with the immune response (Fig. 2A). In blood at day 1, of 584 upregulated genes, only 267 (46%) were associated with the immune response, and this was increased at day 3 to 473 out of 1522 (31%) total upregulated genes. Similarly, of 829 total upregulated genes in lung at day 1, only 441 (53%) genes were involved in immune function. This increased to 716 immune response genes out of a total of 1795 (40%) genes at day 3 postinfection in the lung. Of 327 genes downregulated in blood at day 1 postinfection, 106 (33%) were associated with immune function and this was increased to 418 (25%) genes associated with immune function out of 1661 at day 3 postinfection. The functional type of these genes are shown in Fig. 2B. A substantial set of genes was common and similarly perturbed in both blood and lung and day 1 and also at day 3 postinfection (Fig. 2C and 2E).

Of the genes that were altered in response to infection at day 1, 507 were only expressed in the blood of which 166 (11%) were immune genes, 1028 were only expressed in the lung, of which 349 (34%) were immune genes, whereas 404 were coexpressed in blood and lung of which 207 (51%) were immune genes (Fig. 2D). These numbers were increased at day 3 postinfection with 1947 genes expressed in blood of which only 407 (21%) were immune genes,

2018 were unique in the lung of which 544 were immune genes and 1236 coexpressed in blood and lung, of which 485 (27%) were immune genes (Fig. 2F). Upregulated genes common to both blood and lung at days 1 and 3 after infection included a large number of chemokines and cytokine genes including IL27, IL1, TNF and IFN-inducible genes, matrix metalloproteases (MMPs), Caspase and ADAM family members. There is altered expression of genes not annotated by IPA/GO as involved in immunity, suggesting these are either physiological changes upon infection, or have an underappreciated role in immunity.

The blood transcriptome of mice chronically infected with B. pseudomallei reflects the disease severity in the tissue

We have established a mouse model of chronic melioidosis where mice are intranasally infected with a low dose of the organisms (11), which together with the acute respiratory models described earlier in this manuscript (1, 2) may represent the spectrum of pathological findings observed in patients with this disease (5). Severity at any one time point, ranges from mice with latent infection with no signs of disease, but from which bacteria cannot be detected in the spleen, through to active infection with the presence of high levels of bacterial burden and severe lesions in the spleen (11). To determine whether the blood transcriptome reflects the heterogeneity of disease progression, blood was harvested from chronically infected mice and RNA was extracted and processed for microarray analysis. Unsupervised PCA analysis of blood transcripts passing detection limits ($p<0.01$) showed that samples from uninfected mice clustered together with those with latent infection and away from mice with either moderate clinical severity (Grade 1-2) or those with clinical severity grade 3 (fig. S1). Upon statistical filtering

(p<0.01 Mann-Whitney, Benjamini-Hochberg correction of infected versus uninfected mice using GeneSpring analysis) segregated mice into groups, either expressing a strong signature, a weak signature or no signature of over and underexpressed transcripts, as compared to uninfected controls (Fig. 3A). In keeping with the PCA analysis, some mice, although infected showed a blood signature which clustered with that of uninfected controls. Upon analysis of gross pathology some mice were found to have no detectable lesions at this time point and were latent. Mice with detectable lesions, ranging in severity from 1 to 3 showed a strong transcriptional blood signature, confirmed quantitatively using the Weighted Molecular Distance to Health (MDTH) algorithm (14) (Fig. 3B). This quantitation using MDTH for measuring the extent of the blood signature correlated with the extent of disease severity based on the spleen lesion score (Fig. 3C). A blood transcriptional signature statistically distinct from that of uninfected mice was detected even in the absence of bacteraemia (Fig. 3D). Mice that were chronically infected with low-dose *B. pseudomallei* could be grouped according their splenic bacterial burden; latent mice (no detectable bacteria); emerging mice (low levels of bacteria with no clinical illness); mice with active disease (high bacterial burdens and clinical disease) (Fig. 3E). As shown in the blood splenic transcriptional signatures again correlated with disease severity (Fig. 3F, G and H).

Blood transcriptional profile of chronically infected mice shows similarity to that of acute infection

Using canonical IPA, many similarities between the blood transcriptome of acute and chronic infection were observed, with the top 5 pathways of upregulated genes comprising Granulocyte Adhesion and Diapedesis, IFN-signaling and TREM1-signaling (Fig. 4A and B). The top 5

downregulated pathways in the blood of both chronically and acutely infected mice consisted of genes related to B Cell Development (Fig. 4C and Fig 1C) .

The splenic signature of chronic infection was also dominated by the Granulocyte Adhesion and Diapedesis pathway, but additionally included Heme Biosynthesis II, LXR/RXR Activation and Hepatic Fibrosis in the top 5 upregulated IPA pathways (Fig. 4B). The downregulated genes in the spleen of mice after chronic infection were dominated by those involved in T-cell responses (Fig. 4C).

Of the upregulated genes in blood, 327 (42%) were immune response genes out of a total of 774, whereas 532 (34%) out of 1571 were upregulated immune response genes in the spleen (Fig. 4D). Of the 496 downregulated genes in the blood, 144 (29%) were immune response genes, whereas 371 (31%) immune response genes out of 1185 were downregulated in the spleen. The functional type of genes altered during chronic infection is shown in Fig. 4E. Many were common and similarly perturbed in both blood and spleen during chronic infection (Fig. 4F).

Comparison of the signatures of chronic blood and spleen showed: 502 unique to the blood of which 141 (28%) were immune genes; 768 genes were common in blood and spleen, of which 331 (43%) were immune genes; and 1988 were unique to the spleen of which 572 (29%) were immune genes (Fig. 4G). A large number of the immune genes common to blood and spleen of chronically infected mice again included chemokines, and cytokine genes such as IL27, IL1, TNF and IFN-inducible genes, MMPs, Caspase and ADAM family members (fig. S2).

Similar immune response pathways and modules are revealed in the blood transcriptome of mice

and humans infected with B. pseudomallei

We then compared the immune signatures of blood and tissue in mice acutely or chronically infected with *B. pseudomallei*, with those of human melioidosis cases (14). We first determined the common sets of genes expressed in our Training and Test set human blood samples (1292 genes; Fig. 5A) from patients with *B. pseudomallei*, who were admitted to hospital with blood culture proven melioidosis (14). IPA was used to select the top 25 ranked pathways across the human and mouse blood datasets based on their average *p*-value (Fig. 5B – where the total 109 pathways significant at $p<0.05$, score 1.31, averaged over the blood transcriptome of human melioidosis and mouse days 1 and 3 acute and chronic melioidosis). The dataset-specific *p*-values are summarized in the first panel in Fig. 5B and include IL10-signaling, TREM1-signaling, Pattern-Recognition Receptors in Recognition of Bacteria and Viruses and TLRs, corroborating our previous IPA analyses. Genes associated with IFN-signaling were now also observed in both human and mouse datasets using this approach (Fig. 5B). The number of genes that are common in each pathway in the human Pankla 1292 gene list and mouse blood day 1, day 3 and chronic and the number of pathway specific genes in each dataset is shown in Fig. 5B. Similar results were obtained by looking at top 10 pathways associated with each dataset (fig. S3).

IFNγ is essential for resistance against infection with *B. pseudomallei* in mice, and is produced by human NK and T cells *in vitro*, but the role of Type I IFNs is not known (17, 18). IFNγ itself was shown to be upregulated in the blood (Fig. 5D), but only in chronically infected mice, whereas Type I IFNs were not. Analysis of expression of genes annotated across all IFN-related IPA-signaling pathways showed clear induction of genes in the blood of acute and chronically infected mice, involving both Type I and Type II signaling (Fig. 5D). Comparison of

these results with the blood of patients with acute melioidosis showed several genes in common to the murine response including SOCS1, STAT1, STAT2, IFITM1 and TAP1 (Fig. 5D).

We then compared gene expression profiles in infected mice and humans using a modular data-mining strategy which analyses clusters of genes that are coordinately regulated (16, 19). The 28 human modules contained from 2 to 192 genes per module, corresponding to 2 to 183 mouse genes with a range of 75-100% mouse to human homology, depending on the module examined. Modular analysis of our original human training and test sets of human melioidosis samples showed clear changes in gene expression in a total of 18 distinct functional modules with increased expression of IFN (M3.1), inflammation (M3.2 and M3.3), neutrophil (M2.2) and myeloid (M2.6) modules and decreased expression of B- and T-cell modules (Fig. 6A and 6B). Of these, 5 of the modules showing upregulation of IFN, inflammation, neutrophil and myeloid genes were similarly altered in the blood of all of the mouse models of acute and chronic melioidosis. Whilst B-cell-associated genes were shown to be downregulated in both acute and chronic melioidosis, the downregulated T-cell-associated genes were only apparently downregulated during acute disease (Fig. 6A) in keeping with the top 10 pathways shown to be downregulated by IPA (fig. S3), although closer analysis of each module demonstrates a number of T-cell-associated genes downregulated at all stages of mouse melioidosis that are common with human disease (Fig. 6A). These data demonstrate highly conserved changes across both mouse and human responses in genes associated with IFNs (e.g. IFITM1, SERPING1, CXCL10), neutrophils (e.g. Arginase 1, MMP9, LCN2), B cells (e.g. FCRLA, VPREB3, BCL11A) and T cells (CCR7, BCL11B, FAIM3) (Fig. 6A, B). Thus, by several different analytical approaches, the host response to *B. pseudomallei* was dominated by upregulation of genes involved in IFN-signaling, phagocyte biology and inflammation and downregulation of

genes involved in B- and T-cell responses, which was conserved between mice and humans. A number of modules showed discordancy between human disease and the experimental models of melioidosis: one myeloid gene module (M1.5) (Fig. 6A) was not globally similar between mouse and human, overrepresented only in the mouse models (Fig. 6A), more detailed inspection of the actual list of genes in that module showed that 20 out of 77 of the genes were upregulated in both mouse and human. Discordancy also included underexpression of genes encoding ribosomal proteins in humans, which were either not perturbed or rather overrepresented in the mouse; and six modules consisting of nonannotated genes underrepresented in human disease but not in the mouse models.

Gene and network level analysis of common mouse and human B. pseudomallei associated transcriptional signatures

Consideration of all genes that were altered in each data set revealed that 26.9% of the genes perturbed in human melioidosis were also similarly regulated in the mouse models (Fig. 7A). However, the highest level of similarity with human was observed at day 3 post acute infection of mice (Fig. 7A). Of the 26.9% genes commonly affected in both human and mouse blood, of those upregulated 118 were immune, 90 were nonimmune functionally (Fig. 7B). Of the 118 upregulated genes, 50 were commonly affected in all mouse models and the human disease, including arginase 1, a number of cytokine genes including IL27, IL18, MMPs, and TLR (fig. S4). The remainder of upregulated immune genes in the blood of melioidosis patients was similarly affected in at least one of the mouse models of *B. pseudomallei* infection, again with the acute day 3 transcriptome showing the greatest similarity (fig. S4). Gene and network level

analysis of the 348 common mouse and human *B. pseudomallei* associated genes showed 179 directly interacting genes, which included members of modules associated with B cells (M1.3) and T cells (M2.8) which were downregulated versus those associated with Inflammation (M3.2 and M3.3), Neutrophils (M2.2), Myeloid (M2.6) and IFN (M3.1) genes, which were upregulated (Fig. 7C).

Discussion

Melioidosis is an increasingly important cause of severe bacterial infection in tropical countries, requiring intense and prolonged antibiotic treatment and for which there is no effective vaccine. Following exposure to *B. pseudomallei* a diverse range of clinical outcomes occurs, ranging from prolonged periods of latency to chronic active disease and acute, lethal sepsis. To further our understanding of immunity and pathogenesis of this infection, we have applied whole genome transcriptional microarray to blood and tissues from mouse models, which reflect each of these stages of infection, and compared these profiles to those of melioidosis patients in N.E. Thailand. We show that the blood transcriptome in mice infected with *B. pseudomallei* accurately reflects disease severity in latent, acute and chronic infection and has substantial similarity to that observed in human meliodosis.

Our initial experiments used an established model of acute pneumonia and lethal sepsis following high dose intranasal challenge of mice with a virulent strain of *B. pseudomallei.* Transcriptional array analysis of blood showed clear and extensive transcriptional responses in infected but not control mice within 24 hrs, which increased by 72 hrs. This included upregulation of numerous proinflammatory gene pathways, consistent with a previous report

following intraperitoneal infection (15). There was also coordinated downregulation, of genes involved in both B and T lymphocyte development and function, possibly due to the egress of these cells from the circulation or apoptosis, as we have previously described in human disease (14, 19, 20). To date, our understanding of events in the lung leading to severe pneumonia after exposure to *B. pseudomallei* via the airways have been limited. Wiersinga and colleagues, using a 33 gene probe array reported induction of several proinflammatory cytokines and chemokines in the lung and bronchoalveolar lavage of infected mice (21). Our whole genome array approach now shows the full scale of the host transcriptional changes in the lung during the development of acute pneumonia. Within 24 hours of exposure to bacteria, some 1400 genes showed altered levels of transcription, and over 3200 genes at the later time point of 72 hrs when clear signs of systemic illness (such as weight loss and bacteraemia) were evident. Most of these genes were not annotated as being directly involved in the immune response using a combination of the genes listed in both GO and IPA, presumably reflecting the massive physiological changes occurring during this time. Functional gene types including GPCR, kinases, phosphatases and transcriptional regulators were annotated in both immune and nonimmune response gene families, with an overabundance of GPCRs and transcriptional regulators associated with the immune response genes after infection. Consistent with the influx of different leucocytes into the lungs (22), expression of multiple genes encoding chemokines, (including CCL2, CCL7 and CXCL2), cytokines (IL6, IL18 and TNF) and the immunoregulatory pathway for IL10 were increased. Increased expression of multiple genes involved in pattern recognition were also observed, consistent with the important role of MyD88/TLR and other innate recognition pathways to resistance against *B. pseudomallei* (23). Several gene expression pathways were found in common between lung and blood of infected mice including, TREM1 and IL10-

signaling and Granulocyte Adhesion and Diapedesis.

Although acute infection of mice is frequently used to study immunity to *B. pseudomallei*, (2) it does not allow investigation of the latent or chronic phases of melioidosis also seen in humans living in endemic areas. To address this, we established a new model using low-dose intranasal challenge into genetically resistant mice which leads to prolonged periods of latency but which eventually progresses in all animals to an active chronic infection, characterized by extensive granulomatous lesions in the spleen containing numerous macrophages and lymphocytes (11). Enumeration of the bacterial burden in the spleen of infected animals throughout this process showed three discrete groups; i) those with undetectable bacteria and no tissue lesions (latent), ii) those with low levels of bacteria but no lesions (emerging), and iii) those with high levels of bacterial burden and obvious lesions (active). Unlike the coordinated and homogenous transcriptional response seen in the acute model, blood from chronically infected mice showed a pronounced heterogeneity in gene expression, which correlated directly with these three stages of infection. A weighted MDTH method showed that during the latent phase of infection, there is no appreciable change in host gene expression above the baseline of uninfected animals, presumably a reflection of the low numbers of bacteria in the tissues not showing an overt immune response. This heterogeneous pattern of gene expression was also observed in the spleen, the main target organ of chronic infection, and again reflected the degree of disease severity as shown by spleen lesion score.

In many bacterial infections including melioidosis, a significant proportion of individuals present with signs of acute disease due to the presence of foci of bacteria in their tissues but are blood culture negative (14). The inability to confirm the identity of the pathogen poses considerable challenges to the management of these patients and new methods to indicate the

presence of infecting organisms are needed. Here we show in mice, that blood transcriptional arrays were sufficiently sensitive to indicate the presence of *B. pseudomallei* lesions in tissues such as liver and spleen, in the absence of bacteraemia. This most likely reflects the migration of activated cells from infected tissues into the circulation and/or the activation of blood leucocytes by subcellular fragments of bacteria, such as LPS or cell wall components, released from the lesions. These data suggest that transcriptional profiling may be a useful tool in blood culture negative patients for confirming the presence of infection, predicting the risks of dissemination and systemic illness and in identifying those likely to relapse following antibiotic treatment by indicating the presence of residual tissue foci of infection.

Using a combination of analytical methods to interrogate gene expression profiles, including GO, IPA, pathways and modular approaches, we evaluated the similarity between mouse and human blood datasets associated with *B. pseudomallei* infection. We first identified a common 1292 gene reference list from two sets of human melioidosis patients (14) that comprised 1131 murine orthologs and then calculated the similarity between human and mouse datasets provided by our different models of acute and chronic *B. pseudomallei* infection. A 26.9% overlap was observed in the combined acute (day 1 and day 3) and chronic *B. pseudomallei* infection mouse models for the common 1292 genes perturbed in human melioidosis. Furthermore, using a focused pathway or modular approach highlighted several individual top pathways that strongly correlated in human and mouse. IPA pathways with the highest averaged overlap of human genes present across mouse pathways included upregulation of Granulocyte Adhesion and Diapedesis, Crosstalk between Dendritic Cells and Natural Killer Cells, IFN-signaling and IL10-signaling, and downregulation of genes associated with B-cell and T-cell pathways. Similarly, infectious disease-based gene modules highlighted that several of

them are similarly activated or repressed in human and mouse models. These included the upregulated modules associated with IFN-inducible, inflammation, neutrophil and myeloid genes, and the downregulated modules associated with B-cell and T-cell genes.

It has been suggested that mouse models poorly mimic human inflammatory diseases (24). This prompted us to evaluate the similarity of mouse and human blood datasets associated with *B. pseudomallei* using the methodology adopted by Seok et al. Using the 1292 human gene reference list comprising 1131 murine orthologs we calculated the similarities between human and mouse datasets by the same approach as Seok et al (Pearson correlation, R^2). We obtained R^2 correlations of 0.95 (human Pankla Test), 0.17 (mouse blood day 1), 0.15 (mouse blood day 3) and 0.15 (mouse blood chronic) when fold-change values of patients to healthy within the Human Pankla Training dataset were taken. The weak correlation values between 0.15 and 0.18 are higher than the human-mouse correlations of 0 to 0.08 reported by Soek et al. In contrast, the murine mouse models of melioidosis strongly correlated with each other (R^2 =0.60-0.62). Using this approach, we obtained strong ($R^2 \geq 0.95$) correlation values for human-human comparisons and weak ($R^2 \geq 0.14$-0.18) correlations for human-mouse comparisons. Seok et al., however, reported stronger ($R^2 \leq 0.49$) human-mouse correlations in individual top pathways, indicating that certain key genes and pathways illustrate similarities between mouse models and human disease, and could help to develop improved animal models. Indeed, our unbiased approach using complementary bioinformatic tools, including GO, pathway and modular analyses, demonstrates that there are clear similarities between mouse and human immune responses, as measured by the blood transcriptome, to *B. pseudomallei* infection. Furthermore of the 348 genes commonly perturbed between mouse and human, 51% were predicted to be directly interacting, the majority of which were involved in immune responses.

Taken together, we show that the level at which one compares human-mouse transcriptional datasets can lead to disparate conclusions regarding the similarity of mouse models to human disease. Perfect correlation between murine and human data should not be expected given the nature of the experimental versus clinical conditions. With the human dataset there is heterogeneity at the individual level, not only in terms of underlying genetic makeup of the individuals but also with regard to the dose received of infectious agent, variance in the virulence of the pathogen as well as variation in the time of sampling during the course infection all of which cannot usually be determined in humans under noncontrolled conditions. In addition, there is additional heterogeneity according to the time of initiation of treatment and type of treatment. In contrast, greater control is possible over these variables (pathogen strain, pathogen dose, mouse strain, time after infection and balanced control of experimental groups) in murine models that allows for the in-depth analysis of perturbed gene expression.

Thus, while comparisons of human and mouse models of disease at the level of a global list of significantly expressed genes from human disease models may poorly correlate with expression profiles from mouse disease models, the identification of key genes or pathways shared in the response in murine and humans may reveal known and potentially novel immune response mechanisms following infection by a pathogen and may help to identify biomarkers relevant to the interplay between host and pathogen. The development of more focused approaches such as modular analysis developed in diverse mouse models is needed for interpreting and further developing animal models of human disease. Our data not only provide new information on the pathogenesis of this infection but also demonstrate that a complementary set of bioinformatics approaches can provide valuable and novel insight into the conserved response pathways in mouse and man, validating mouse models of human disease and providing

approaches to improve them.

Acknowledgements

We thank R. Gilbert and the LSHTM BSF for animal husbandry. We thank A Sesay, H. Jani and the High Throughput Sequencing Facility, NIMR, for microarray. **Author contributions:** GB, AOG, CG, LC conceived the experiments. GB, AOG, JB, VP, DC, GL oversaw the study and data analysis. LC, CG, and NS performed the experiments. LC, KP, CG, SB, RP and PK, analysed the data. GB and AOG wrote the manuscript. All authors discussed results and commented on the manuscript. **Competing Interests:** All authors declare no competing interests.

Data and materials availability: The mouse microarray data is deposited in the Gene Expression Omnibus (GEO: GSE61106).

References

Footnote

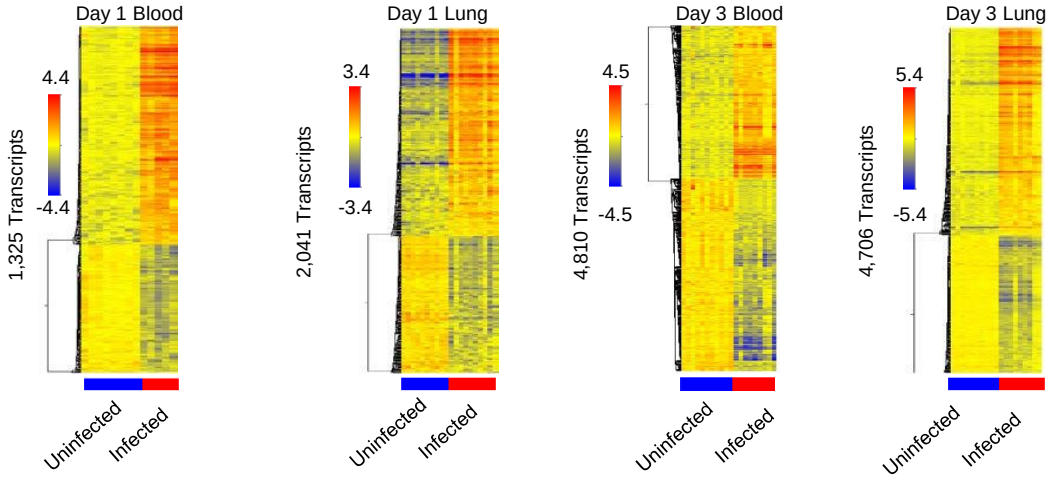
Funding: This project and KP were funded by the ERC 204682, TB-PATH. AOG and CMG funded by the MRC UK.

Address for correspondence:

Dr Gregory J Bancroft

Immunology and Infection Department, Faculty of Infectious and Tropical Diseases, London School of Hygiene & Tropical Medicine, London, UK. WC1E 7HT.

A



B

Top pathways associated with up-regulated genes

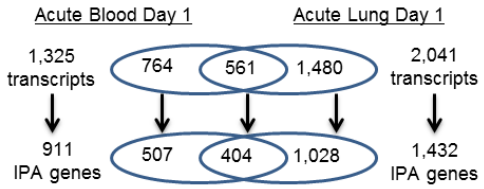
Day 1 Blood	Day 1 Lung	Day 3 Blood	Day 3 Lung
Granulocyte Adhesion and Diapedesis***	Granulocyte Adhesion and Diapedesis***	Granulocyte Adhesion and Diapedesis***	Granulocyte Adhesion and Diapedesis***
IL-10 Signaling	Agranulocyte Adhesion and Diapedesis	IL-10 Signaling**	TREM1 Signaling***
Interferon Signaling	TREM1 Signaling***	LXR/RXR Activation	IL-10 Signaling**
TREM1 Signaling***	Altered T Cell and B Cell Signaling in Rheumatoid Arthritis	TREM1 Signaling***	Agranulocyte Adhesion and Diapedesis
Hepatic Fibrosis / Hepatic Stellate Cell Activation	Acute Phase Response Signaling	Acute Phase Response Signaling	Role of Pattern Recognition Receptors in Recognition of Bacteria and Viruses

C

Top pathways associated with down-regulated genes

Day 1 Blood	Day 1 Lung	Day 3 Blood	Day 3 Lung
B Cell Development*	Regulation of the Epithelial-Mesenchymal Transition Pathway	CD28 Signaling in T Helper Cells	Axonal Guidance Signaling
iCOS-iCOSL Signaling in T Helper Cells	Basal Cell Carcinoma Signaling	iCOS-iCOSL Signaling In T Helper Cells	Cellular Effects of Sildenafil RhoA Signaling
CD28 Signaling in T Helper Cells	Wnt/β-catenin Signaling	Calcium-induced T Lymphocyte Apoptosis	Calcium Signaling
Calcium-induced T Lymphocyte Apoptosis	Axonal Guidance Signaling	Primary Immuno-deficiency Signaling	Phospholipase C Signaling
Role of NFAT in Regulation of the Immune Response	B Cell Development*	T Cell Receptor Signaling	Agranulocyte Adhesion and Diapedesis

D



E

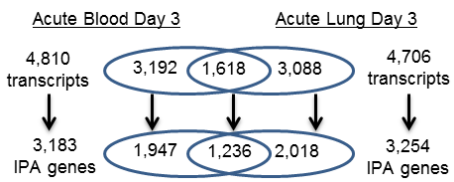


Figure 1. See also Figure S1 that is related to Figure 1A

Fig. 1. Mouse blood and lung transcriptional signatures are associated with similar and distinct canonical pathways after acute *B. pseudomallei* infection.

A. Heatmaps of normalized expression values of blood day 1 (infected n=5, uninfected n=8), lung day1 (infected/uninfected n=10), blood day 3 (infected n=9, uninfected n=11) and lung day 3 (infected n=9, uninfected n=10) differentially expressed transcripts. Scale bar represents log fold change of up (red) or downregulated (blue) transcripts. See also Fig. S1.

B, C. Top 5 significant Ingenuity Pathway Analysis (IPA) pathways ($p<0.01$ after Fisher's Exact Test) associated with up- or downregulated transcripts found within the 1325 blood day 1; 2041 lung day 1; 4810 blood day 3 and 4706 lung day 3 differentially expressed transcripts are shown. Canonical pathways that are common to day 1 blood and lung, to day 3 blood and lung, and to the four datasets are denoted by one to three asterisks, respectively.

D, E. Venn diagram showing the number of overlapping transcripts between blood and lung transcriptional signatures.

Figure 2

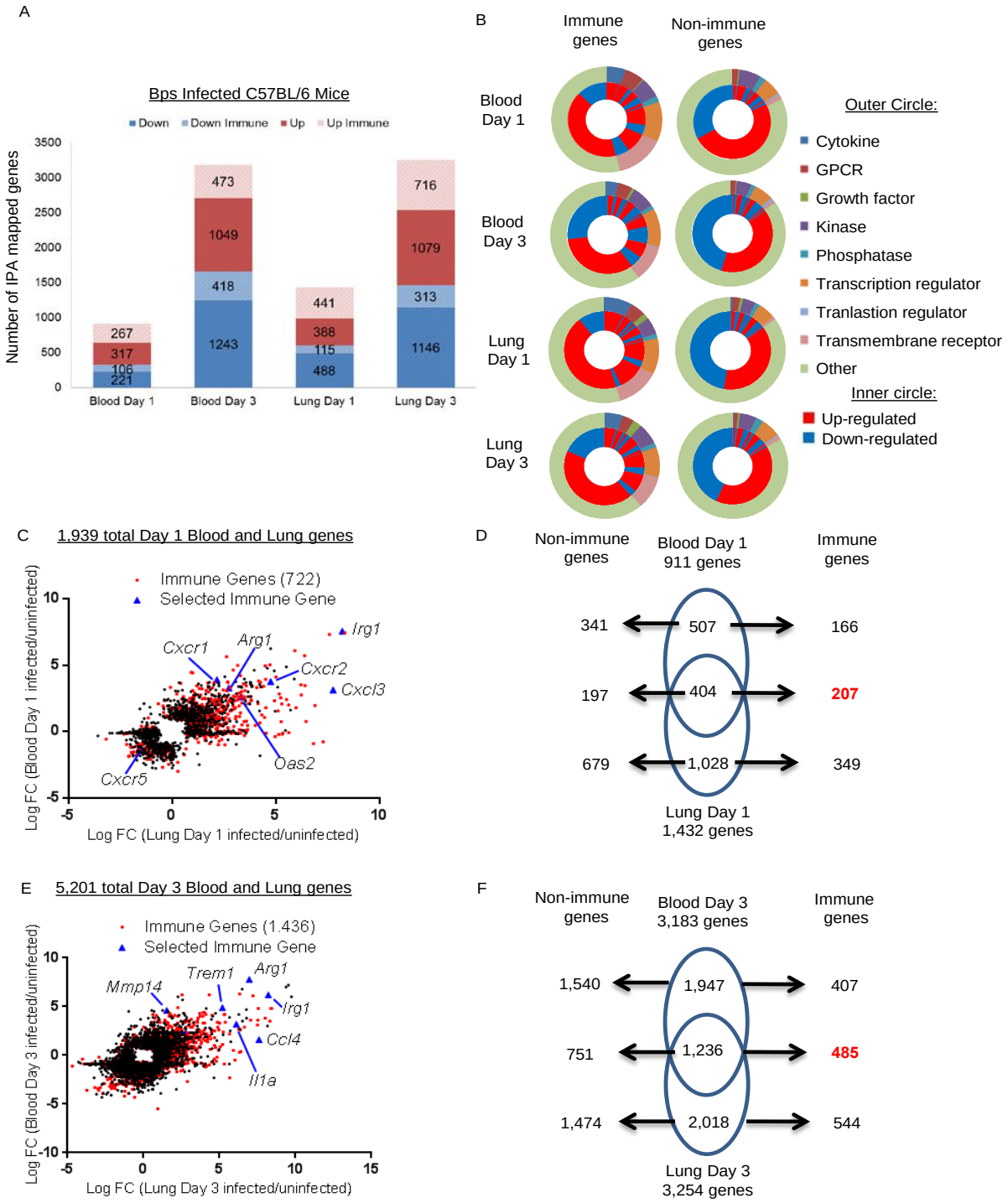

A
Bps Infected C57BL/6 Mice
Down
Down Immune
Up
Up Immune
Number of IPA mapped genes
3500
3000
2500
2000
1500
1000
500
0
267
317
106
221
473
1049
418
1243
441
388
115
488
716
1079
313
1146
Blood Day 1
Blood Day 3
Lung Day 1
Lung Day 3
B
Immune genes
Non-immune genes
Blood Day 1
Blood Day 3
Lung Day 1
Lung Day 3
Outer Circle:
Cytokine
GPCR
Growth factor
Kinase
Phosphatase
Transcription regulator
Tranlastion regulator
Transmembrane receptor
Other
Inner circle:
Up-regulated
Down-regulated
C
1,939 total Day 1 Blood and Lung genes
Immune Genes (722)
Selected Immune Gene
Log FC (Blood Day 1 infected/uninfected)
Log FC (Lung Day 1 infected/uninfected)
Cxcr1
Arg1
Irg1
Cxcr2
Cxcl3
Oas2
Cxcr5
D
Non-immune genes
Blood Day 1
911 genes
Immune genes
341
507
166
197
404
207
679
1,028
349
Lung Day 1
1,432 genes
E
5,201 total Day 3 Blood and Lung genes
Immune Genes (1,436)
Selected Immune Gene
Log FC (Blood Day 3 infected/uninfected)
Log FC (Lung Day 3 infected/uninfected)
Mmp14
Trem1
Arg1
Irg1
Ccl4
Il1a
F
Non-immune genes
Blood Day 3
3,183 genes
Immune genes
1,540
1,947
407
751
1,236
485
1,474
2,018
544
Lung Day 3
3,254 genes

Fig. 2. Immune genes are associated with blood and lung transcriptional signatures following acute *B. pseudomallei* infection.

A. Bar chart of the immune and nonimmune gene distribution within the 911 blood day 1 genes, 3183 blood day 3 genes, 1432 lung day 1 genes and 3254 lung day 3 genes.

B. Twolayer doughnut chart showing the functional annotation of down, down immune, up and up immune genes found within blood day 1, blood day 3, lung day 1 and lung day 3 genes.

C. Scatter plot of total blood and lung genes at day 1 derived by combining 911 blood day 1 and 1432 lung day 1 differentially expressed genes. 722 Immune genes are marked in red.

D. Venn diagram showing the common and dataset-specific genes for blood and lung at day 1 post infection. The number of immune and nonimmune genes is indicated.

E. Scatter plot of total blood and lungs genes at day 3 derived by combining 3183 blood day 3 and 3254 lung day 3 differentially expressed genes. 1436 Immune genes are marked in red.

F. Venn diagram showing the common and dataset-specific immune versus nonimmune genes for blood and lung at day 3 post infection.

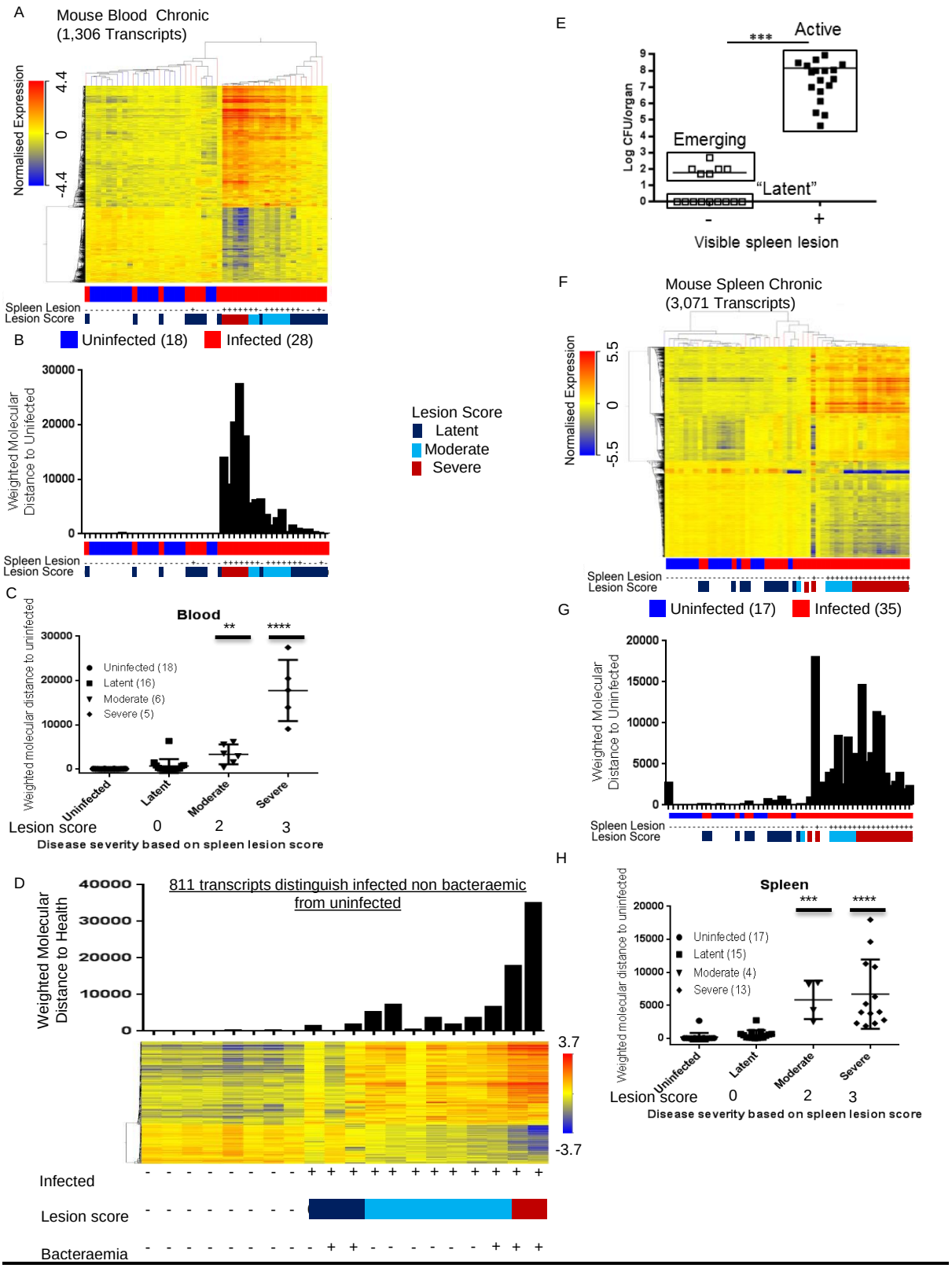
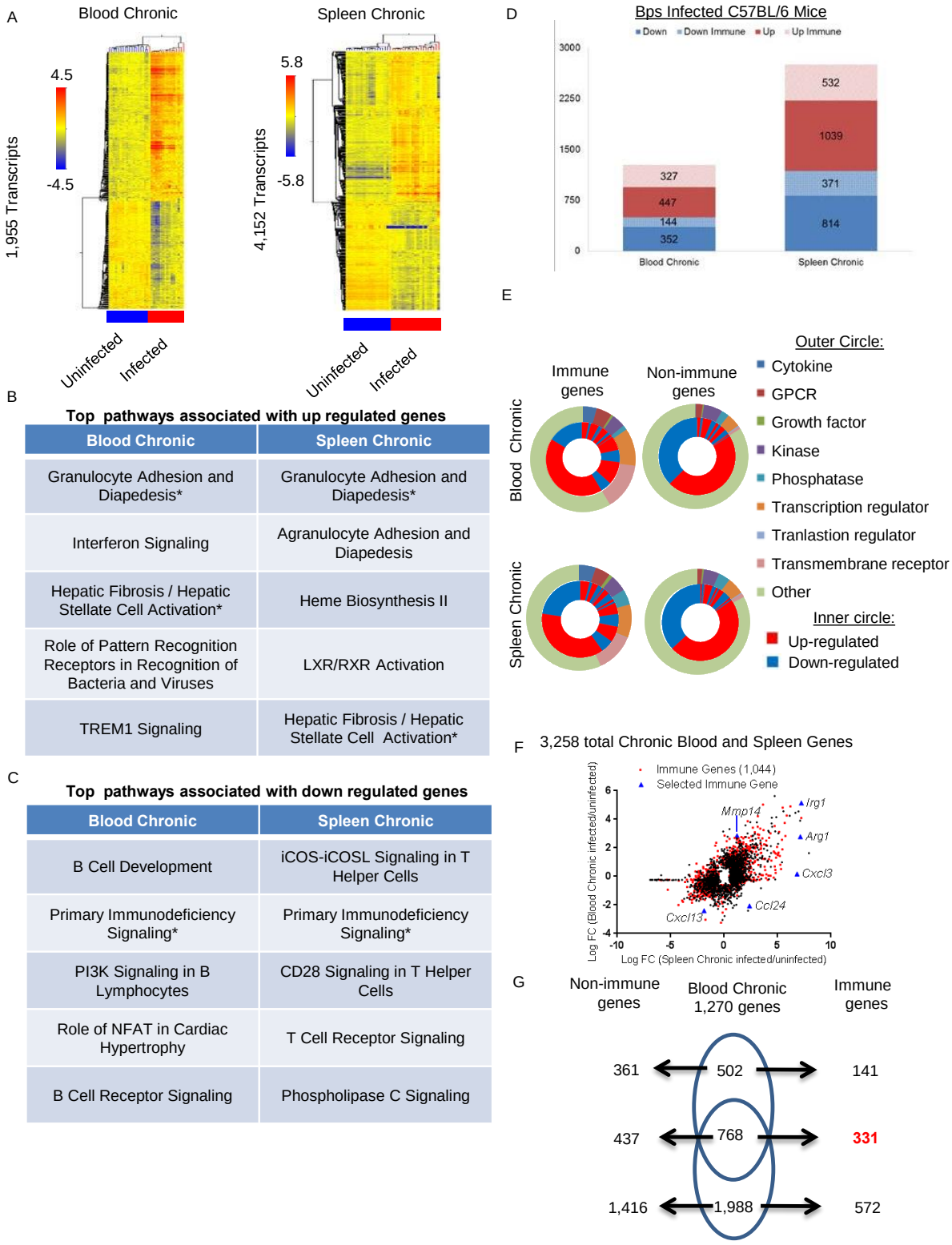

A
Mouse Blood Chronic
(1,306 Transcripts)
Normalised Expression
4.4
0
-4.4
Spleen Lesion
Lesion Score
B
Uninfected (18)
Infected (28)
Weighted Molecular Distance to Unifected
30000
20000
10000
0
Lesion Score
Latent
Moderate
Severe
C
Blood
Weighted molecular distance to uninfected
Uninfected (18)
Latent (16)
Moderate (6)
Severe (5)
Uninfected
Latent
Moderate
Severe
Lesion score
0
2
3
Disease severity based on spleen lesion score
D
811 transcripts distinguish infected non bacteraemic from uninfected
Weighted Molecular Distance to Health
40000
30000
20000
10000
0
3.7
-3.7
Infected
Lesion score
Bacteraemia
E
Active
Emerging
"Latent"
Log CFU/organ
Visible spleen lesion
F
Mouse Spleen Chronic
(3,071 Transcripts)
Normalised Expression
5.5
0
-5.5
Spleen Lesion
Lesion Score
Uninfected (17)
Infected (35)
G
Weighted Molecular Distance to Uninfected
20000
15000
10000
5000
0
H
Spleen
Weighted molecular distance to uninfected
Uninfected (17)
Latent (15)
Moderate (4)
Severe (13)
Uninfected
Latent
Moderate
Severe
Lesion score
0
2
3
Disease severity based on spleen lesion score

Figure 3

Fig. 3. Mouse blood and spleen transcriptional signatures following chronic *B. pseudomallei* infection reveal differences in latent versus active disease state.

(A, F) Unsupervised clustering based heatmaps of normalized expression values of blood chronic (infected n=28, uninfected n=18) and spleen chronic (infected n=35, uninfected n=17) differentially expressed transcripts. Scale bar represents log fold change of up (red) or downregulated (blue) transcripts. See also Fig. S1. (B, G) Molecular distance to health/uninfected (MDTH) is aligned against the hierarchical condition tree generated through unsupervised clustering (top). (C, H) Correlation of individual MDTH to lesion severity score. Horizontal bars represent the median value for each group. (D) Heatmap of 811 significant transcripts that were identified by comparing *B. pseudomallei* infected mice without bacteraemia (n=7) to uninfected (n=8). Lesion scores were applied to generate a supervised heatmap of the identified transcripts. (E) Bacterial counts from spleens were allocated in two groups according to the presence or absence of visible lesions. Each symbol represents an individual organ. The horizontal line represents the mean $p<0.001$ (***) between mice with or without visible lesions. Scale bar represents log fold change of up (red) or downregulated (blue) transcripts (A, D, F).



B

Top pathways associated with up regulated genes

Blood Chronic	Spleen Chronic
Granulocyte Adhesion and Diapedesis*	Granulocyte Adhesion and Diapedesis*
Interferon Signaling	Agranulocyte Adhesion and Diapedesis
Hepatic Fibrosis / Hepatic Stellate Cell Activation*	Heme Biosynthesis II
Role of Pattern Recognition Receptors in Recognition of Bacteria and Viruses	LXR/RXR Activation
TREM1 Signaling	Hepatic Fibrosis / Hepatic Stellate Cell Activation*

C

Top pathways associated with down regulated genes

Blood Chronic	Spleen Chronic
B Cell Development	iCOS-iCOSL Signaling in T Helper Cells
Primary Immunodeficiency Signaling*	Primary Immunodeficiency Signaling*
PI3K Signaling in B Lymphocytes	CD28 Signaling in T Helper Cells
Role of NFAT in Cardiac Hypertrophy	T Cell Receptor Signaling
B Cell Receptor Signaling	Phospholipase C Signaling

Fig. 4. Similar and distinct canonical pathways and immune genes are associated with

blood and spleen transcriptional signatures in active chronic *B. pseudomallei* disease.

A. Heatmaps of infected versus active chronic infection showing unsupervised clustering of 1955 blood chronic (infected n=16, uninfected n=18) and 4152 spleen chronic (infected n=20, uninfected n=17) differentially expressed transcripts. Scale bar represents log fold change of up (red) or down (blue) regulated transcripts.

B, C. Ingenuity Pathway Analysis (IPA) of the up- or downregulated transcripts found within the 1955 blood chronic and 4152 spleen chronic differentially expressed transcripts shows the top 5 significant canonical pathways ($p<0.01$ after Fisher's Exact Test). Pathways that are common to both datasets are marked with a single asterisk.

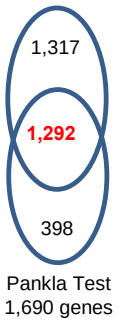
D. Bar chart summarizing the immune and nonimmune gene distribution within the 1955 blood and 4152 spleen transcripts that map to 1270 and 2756 active chronic blood and spleen genes.

E. Twolayer doughnut chart showing the functional annotation of down, down immune, up and up immune genes found within 1270 blood and 2756 spleen genes in active chronic infection.

F. Scatter plot of total blood and spleen genes in active chronic infection derived by combining the 1270 blood and 2756 spleen genes. A list of 1044 Immune genes identified within the 3258 total blood-spleen genes by searching the IPA Knowledge Base and Gene Ontology resources available through MGD are marked in red.

G. Venn diagram showing the common and dataset-specific immune versus nonimmune genes for blood and spleen in active chronic infection. See also Fig. S2 for gene lists.

A

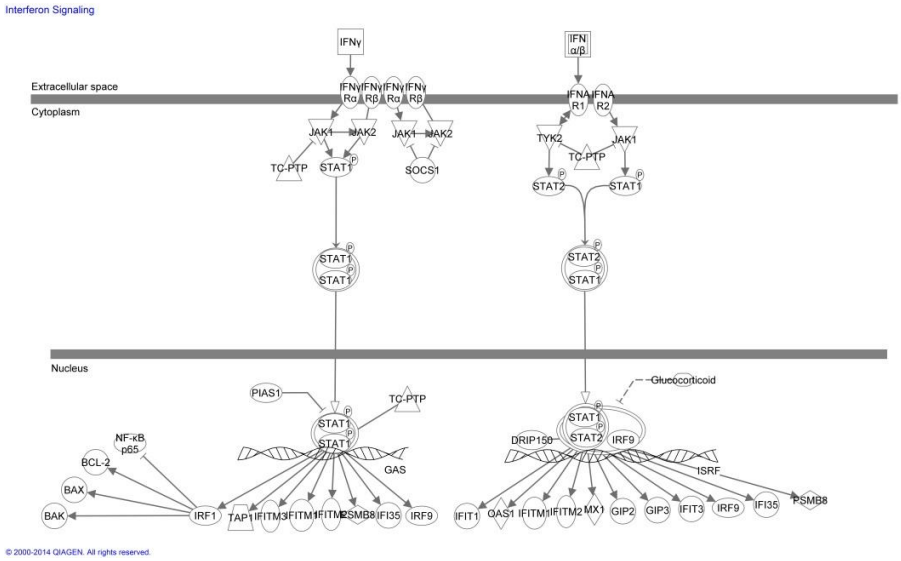


B

Top 25 Mouse-Human Blood Canonical Pathways ranked according to mean negative log p value significance across gene lists	negative log p value significance: Human (1,292 list)	Mouse Day 1	Mouse Day 3	Mouse Chronic	Number of mouse genes that are also in the human 1,292 list: Human (1,292 list)	Mouse Day 1	Mouse Day 3	Mouse Chronic
Hepatic Fibrosis / Hepatic Stellate Cell Activation					21	11	10	9
Altered T Cell and B Cell Signaling in Rheumatoid Arthritis					17	8	10	11
IL-10 Signaling					11	7	7	5
Type I Diabetes Mellitus Signaling					17	10	12	7
Communication between Innate and Adaptive Immune Cells					15	7	9	9
Interferon Signaling					8	7	3	4
Role of Macrophages, Fibroblasts and Endothelial Cells in Rheumatoid Arthritis					36	15	18	14
Granulocyte Adhesion and Diapedesis					14	10	10	11
Dendritic Cell Maturation					20	8	10	10
TREM1 Signaling					9	4	5	6
LXR/RXR Activation					11	5	7	5
Role of Pattern Recognition Receptors in Recognition of Bacteria and Viruses					17	5	8	6
Toll-like Receptor Signaling					12	5	7	5
iCOS-iCOSL Signaling in T Helper Cells					23	10	17	4
Primary Immunodeficiency Signaling					15	5	9	3
T Cell Receptor Signaling					22	8	17	2
Phospholipase C Signaling					31	7	16	5
Crosstalk between Dendritic Cells and Natural Killer Cells					12	7	9	5
Role of NFAT in Regulation of the Immune Response					30	10	18	6
Production of Nitric Oxide and Reactive Oxygen Species in Macrophages					18	5	8	4
CD28 Signaling in T Helper Cells					20	8	16	4
Acute Phase Response Signaling					15	7	10	6
Tec Kinase Signaling					20	10	8	6
p38 MAPK Signaling					19	8	10	7
NF-κB Signaling					24	9	14	9

0 12
-log p-value score

C



D

Interferon Signaling Pathway (IPA)

Bps-infected Blood: Symbol	Mouse Day 1	Mouse Day 3	Mouse Chronic	Human Pankla
BCL2				
IFITM1				
TAP1				
IRF1				
OAS1				
PSMB8				
STAT2				
IFI35				
IFIT1				
IFIT3				
IFNAR2				
PTPN2				
SOCS1				
STAT1				
BAX				
IFNA1/IFNA13				
IFNAR1				
IFNG				
IRF9				
BAK1				
IFNA5				
IFNA21				
IFNB1				
IFNGR1				
IFNGR2				
JAK1				
JAK2				
MED14				
MX1				
PIAS1				
RELA				
TYK2				
Total genes in dataset	13	10	12	8

Log2 of fold-change between infected and uninfected
-3 0 3

640

641 **Fig. 5. Immune pathways including IFN-signaling are significantly associated with mouse**

and human blood transcriptional signatures after *B. pseudomallei* infection

A. Venn diagram shows 1292 overlapping genes between the Pankla Training and the Pankla Test set differentially expressed genes. The significant gene lists were identified by stringent statistical filtering comparing blood from diseased individuals to those from healthy controls (Mann-Whitney test plus Benjamini-Hochberg correction $p<0.01$) on data normalized to the median of all transcripts across all samples and passing a detection value ($p<0.01$) and 2-fold expression value across at least 10% of samples within GeneSpring GX12.1 statistical software.

B. Top 25 ranked pathways across human and mouse blood datasets after *B. pseudomallei* infection based on their average *p*-value. The dataset-specific *p*-values are summarized in the first panel from right. Pathway-by-pathway similarity in relation to number of genes that are common in the human Pankla 1292 gene list and mouse blood day 1, day 3 and chronic are shown in the far right panel. See also Fig. S3 for IPA top10 pathways associated with each dataset.

C. Graphical representation of the IFN-signaling found within IPA.

D. Intensity plot of IFN Pathway-signaling genes across *B. pseudomallei* infected mouse and human blood datasets listed above. Fold changes shown are based on genes found in the blood day 1 (1325 transcripts), blood day 3 (4810), blood chronic (1955) and human blood chronic (1292) with a $p<0.01$ from Mann-Whitney test comparing infected to uninfected followed by Benjamini-Hochberg correction. Scale bar represents log fold change of up- (red) or downregulated (blue) genes.

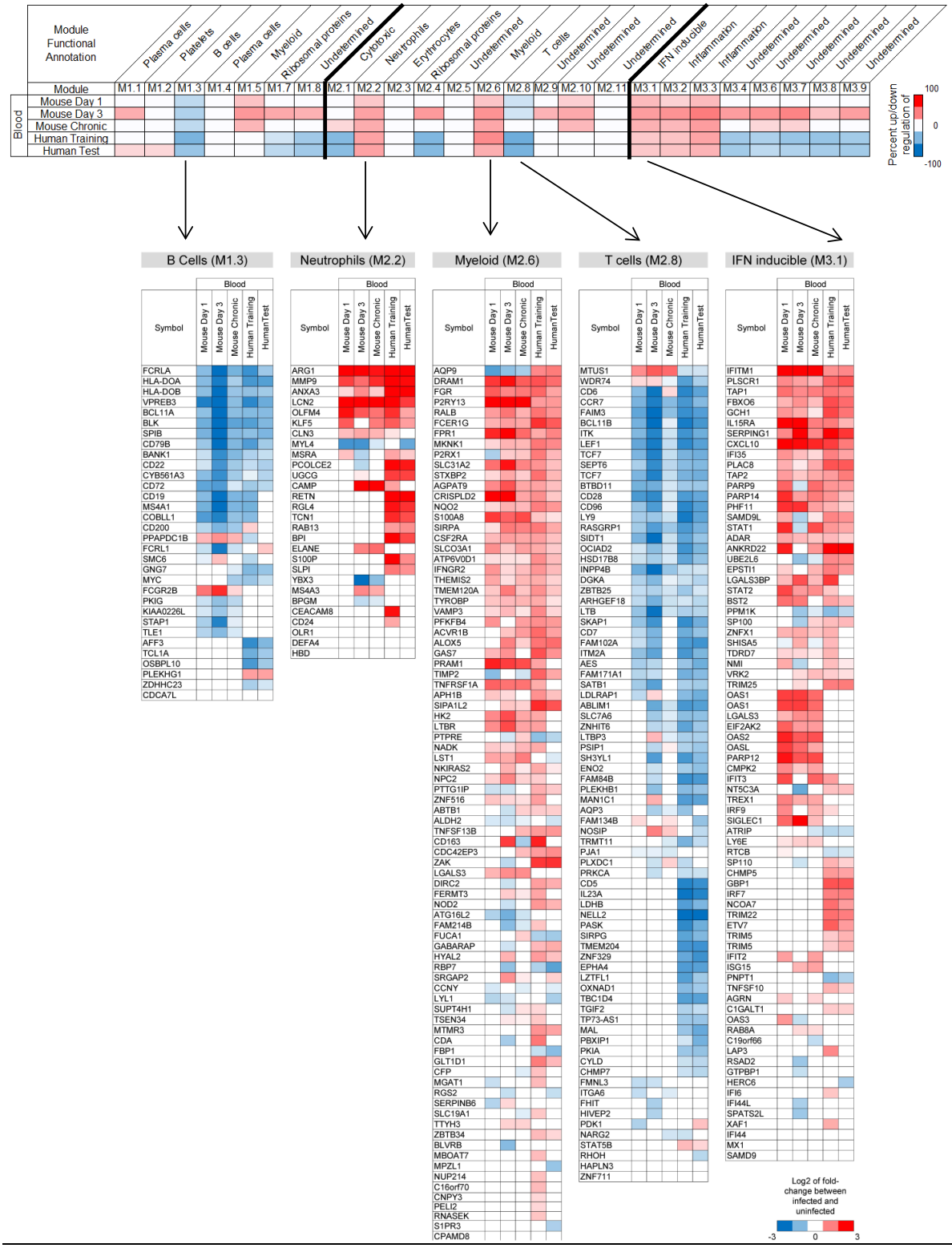

Module Functional Annotation
Plasma cells
Platelets
B cells
Plasma cells
Myeloid
Ribosomal proteins
Undetermined
Cytotoxic
Neutrophils
Erythrocytes
Ribosomal proteins
Undetermined
Myeloid
T cells
Undetermined
Undetermined
Undetermined
IFN inducible
Inflammation
Inflammation
Undetermined
Undetermined
Undetermined
Undetermined
Undetermined
Module
M1.1 M1.2 M1.3 M1.4 M1.5 M1.7 M1.8 M2.1 M2.2 M2.3 M2.4 M2.5 M2.6 M2.8 M2.9 M2.10 M2.11 M3.1 M3.2 M3.3 M3.4 M3.6 M3.7 M3.8 M3.9
Blood
Mouse Day 1
Mouse Day 3
Mouse Chronic
Human Training
Human Test
Percent up/down regulation of
100
0
-100
B Cells (M1.3)
Neutrophils (M2.2)
Myeloid (M2.6)
T cells (M2.8)
IFN inducible (M3.1)
Blood
Symbol
Mouse Day 1
Mouse Day 3
Mouse Chronic
Human Training
HumanTest
Log2 of fold-change between infected and uninfected
-3
0
3

Figure 6A.

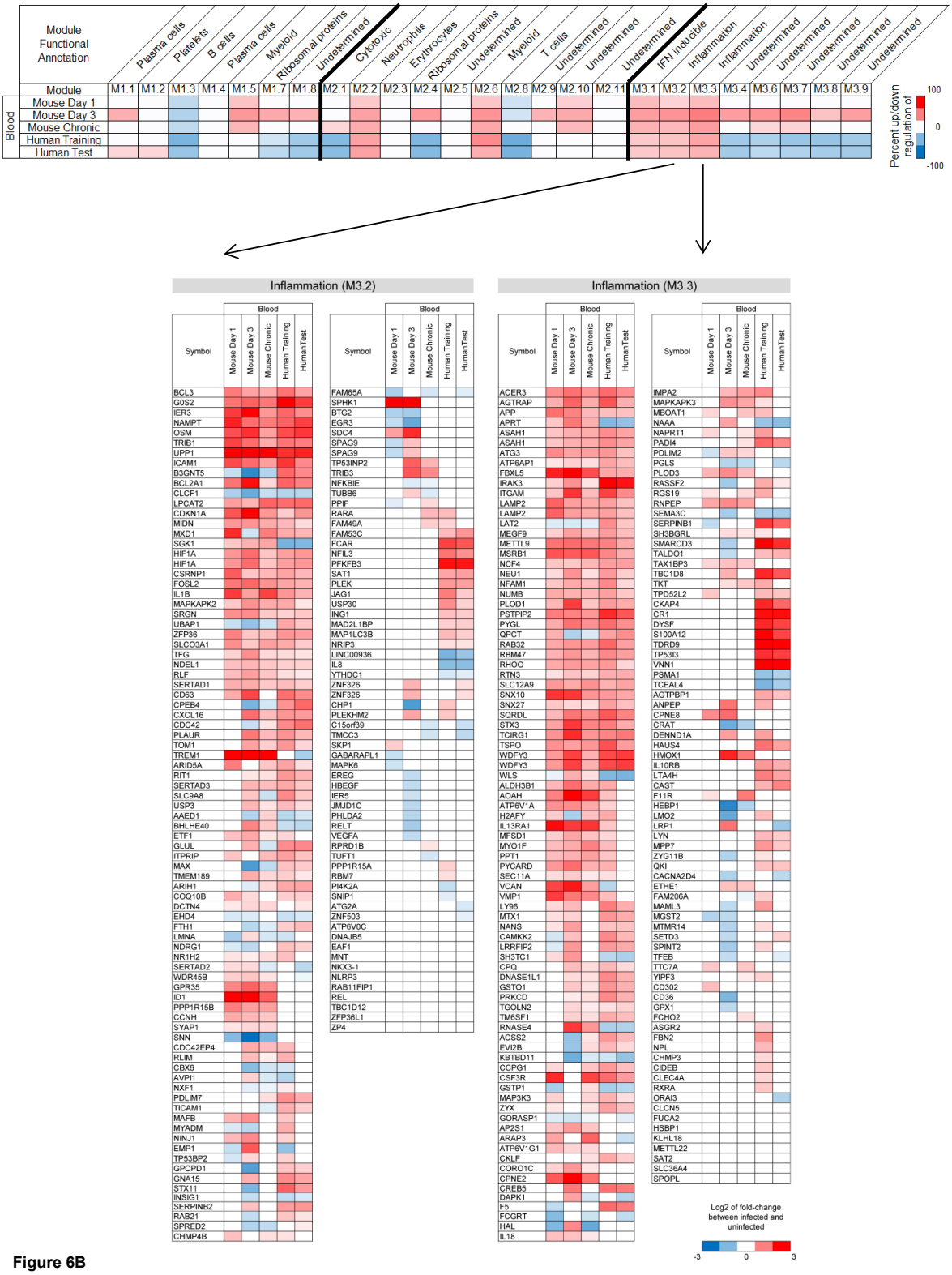
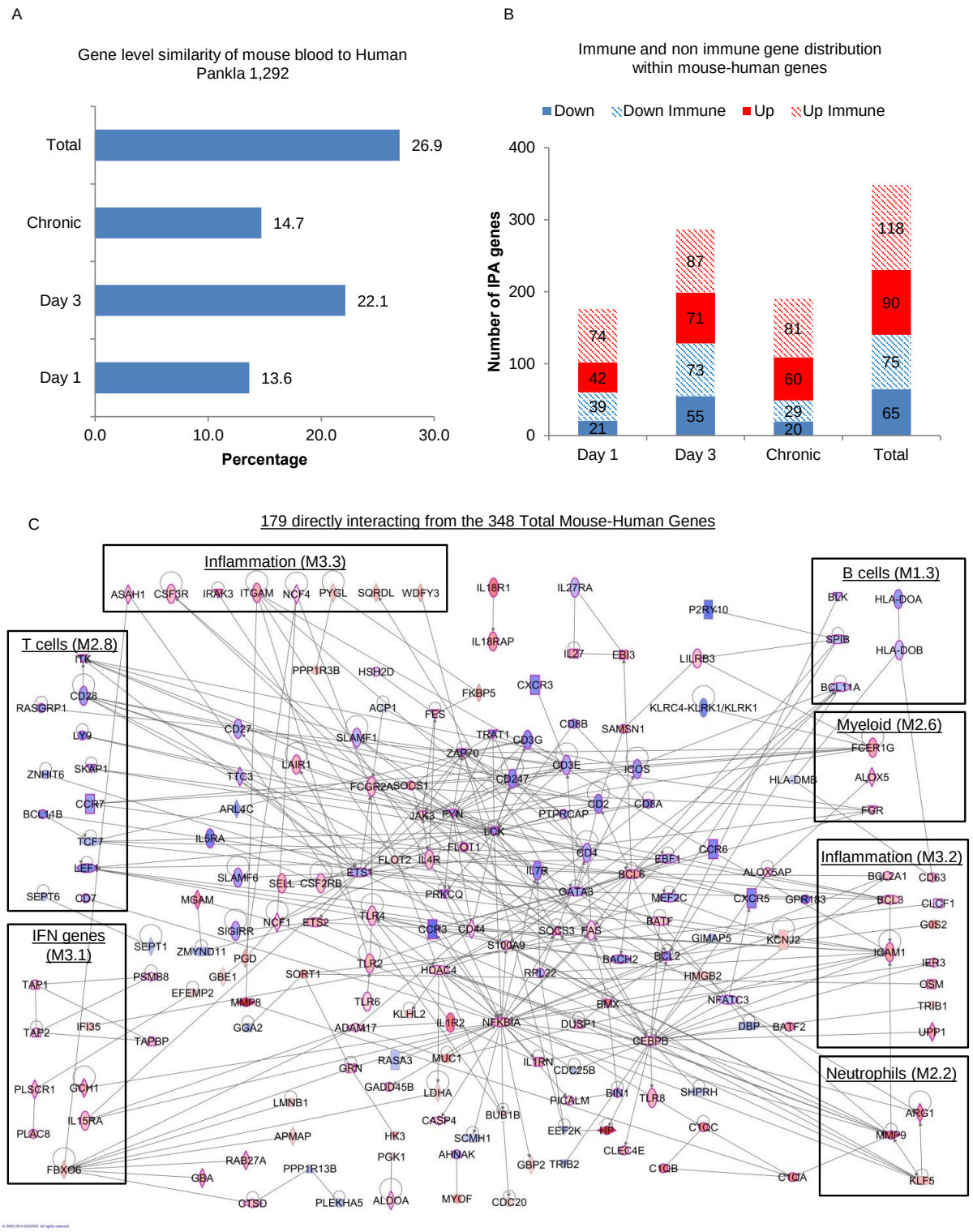

Module Functional Annotation
Plasma cells
Platelets
B cells
Plasma cells
Myeloid
Ribosomal proteins
Undetermined
Cytotoxic
Neutrophils
Erythrocytes
Ribosomal proteins
Undetermined
Myeloid
T cells
Undetermined
Undetermined
Undetermined
IFN inducible
Inflammation
Inflammation
Undetermined
Undetermined
Undetermined
Undetermined
Undetermined
Module
M1.1 M1.2 M1.3 M1.4 M1.5 M1.7 M1.8 M2.1 M2.2 M2.3 M2.4 M2.5 M2.6 M2.8 M2.9 M2.10 M2.11 M3.1 M3.2 M3.3 M3.4 M3.6 M3.7 M3.8 M3.9
Blood
Mouse Day 1
Mouse Day 3
Mouse Chronic
Human Training
Human Test
Percent up/down regulation of
100
0
-100
Inflammation (M3.2)
Inflammation (M3.3)
Blood
Symbol
Mouse Day 1
Mouse Day 3
Mouse Chronic
Human Training
HumanTest
BCL3
G0S2
IER3
NAMPT
OSM
TRIB1
UPP1
ICAM1
B3GNT5
BCL2A1
CLCF1
LPCAT2
CDKN1A
MIDN
MXD1
SGK1
HIF1A
HIF1A
CSRNP1
FOSL2
IL1B
MAPKAPK2
SRGN
UBAP1
ZFP36
SLCO3A1
TFG
NDEL1
RLF
SERTAD1
CD63
CPEB4
CXCL16
CDC42
PLAUR
TOM1
TREM1
ARID5A
RIT1
SERTAD3
SLC9A8
USP3
AAED1
BHLHE40
ETF1
GLUL
ITPRIP
MAX
TMEM189
ARIH1
COQ10B
DCTN4
EHD4
FTH1
LMNA
NDRG1
NR1H2
SERTAD2
WDR45B
GPR35
ID1
PPP1R15B
CCNH
SYAP1
SNN
CDC42EP4
RLIM
CBX6
AVPI1
NXF1
PDLIM7
TICAM1
MAFB
MYADM
NINJ1
EMP1
TP53BP2
GPCPD1
GNA15
STX11
INSIG1
SERPINB2
RAB21
SPRED2
CHMP4B
FAM65A
SPHK1
BTG2
EGR3
SDC4
SPAG9
SPAG9
TP53INP2
TRIB3
NFKBIE
TUBB6
PPIF
RARA
FAM49A
FAM53C
FCAR
NFIL3
PFKFB3
SAT1
PLEK
JAG1
USP30
ING1
MAD2L1BP
MAP1LC3B
NRIP3
LINC00936
IL8
YTHDC1
ZNF326
ZNF326
CHP1
PLEKHM2
C15orf39
TMCC3
SKP1
GABARAPL1
MAPK6
EREG
HBEGF
IER5
JMJD1C
PHLDA2
RELT
VEGFA
RPRD1B
TUFT1
PPP1R15A
RBM7
PI4K2A
SNIP1
ATG2A
ZNF503
ATP6V0C
DNAJB5
EAF1
MNT
NKX3-1
NLRP3
RAB11FIP1
REL
TBC1D12
ZFP36L1
ZP4
ACER3
AGTRAP
APP
APRT
ASAH1
ASAH1
ATG3
ATP6AP1
FBXL5
IRAK3
ITGAM
LAMP2
LAMP2
LAT2
MEGF9
METTL9
MSRB1
NCF4
NEU1
NFAM1
NUMB
PLOD1
PSTPIP2
PYGL
QPCT
RAB32
RBM47
RHOG
RTN3
SLC12A9
SNX10
SNX27
SQRDL
STX3
TCIRG1
TSPO
WDFY3
WDFY3
WLS
ALDH3B1
AOAH
ATP6V1A
H2AFY
IL13RA1
MFSD1
MYO1F
PPT1
PYCARD
SEC11A
VCAN
VMP1
LY96
MTX1
NANS
CAMKK2
LRRFIP2
SH3TC1
CPQ
DNASE1L1
GSTO1
PRKCD
TGOLN2
TM6SF1
RNASE4
ACSS2
EVI2B
KBTBD11
CCPG1
CSF3R
GSTP1
MAP3K3
ZYX
GORASP1
AP2S1
ARAP3
ATP6V1G1
CKLF
CORO1C
CPNE2
CREB5
DAPK1
F5
FCGRT
HAL
IL18
IMPA2
MAPKAPK3
MBOAT1
NAAA
NAPRT1
PADI4
PDLIM2
PGLS
PLOD3
RASSF2
RGS19
RNPEP
SEMA3C
SERPINB1
SH3BGRL
SMARCD3
TALDO1
TAX1BP3
TBC1D8
TKT
TPD52L2
CKAP4
CR1
DYSF
S100A12
TDRD9
TP53I3
VNN1
PSMA1
TCEAL4
AGTPBP1
ANPEP
CPNE8
CRAT
DENND1A
HAUS4
HMOX1
IL10RB
LTA4H
CAST
F11R
HEBP1
LMO2
LRP1
LYN
MPP7
ZYG11B
QKI
CACNA2D4
ETHE1
FAM206A
MAML3
MGST2
MTMR14
SETD3
SPINT2
TFEB
TTC7A
YIPF3
CD302
CD36
GPX1
FCHO2
ASGR2
FBN2
NPL
CHMP3
CIDEB
CLEC4A
RXRA
ORAI3
CLCN5
FUCA2
HSBP1
KLHL18
METTL22
SAT2
SLC36A4
SPOPL
Log2 of fold-change between infected and uninfected
-3
0
3

Figure 6B

Fig. 6. Modular analysis of blood transcriptional profiles reveals a degree of similarity between mouse and human datasets.

Heatmap summary of significant up- or downregulations in modules associated with *B. pseudomallei* mouse and human datasets and passing a *p*-value <0.05 after *t*-test comparing infected to uninfected groups (top). Gene-level analysis of modules that are common across the five datasets are detailed from left to right (A) B cell (M1.3), Neutrophils (M2.2), Myeloid (M2.6), T-cell (M2.8), and IFN Inducible (M3.1); (B) Inflammation (M3.2) and Inflammation (M3.3). Scale bar represents log fold change of up (red) or down (blue) regulated genes.



686

687 **Figure 7. See also Figure S4 that is related to Figure 7B**

Fig. 7. Gene and network level analysis of common mouse-human *B. pseudomallei* associated transcriptional signatures.

A. Percentage similarity of mouse blood *B. pseudomallei* infection-associated datasets to human blood 1292 Pankla gene list. The percentages range from 13.6 to 26.9% with day 1 (13.6%), day 3 (22.1%), Chronic (14.7%) and Total (26.9%) mouse similarity to human Pankla 1292 gene list.

B. Bar chart summarizing the immune and nonimmune gene distribution within 176 day 1, 286 day 3, 190 Chronic and 348 Total mouse-human blood genes associated with *B. pseudomallei* infection.

C. IPA network between 179 of the 348 Total mouse-human blood genes. Genes found within modules detailed in Fig. 6 are boxed, while 130 genes associated with an immune function are outlined in red. Up- and downregulated genes are indicated with red and blue shading.

1. Wiersinga, W. J., T. van der Poll, N. J. White, N. P. Day, and S. J. Peacock. 2006. Melioidosis: insights into the pathogenicity of Burkholderia pseudomallei. *Nature reviews. Microbiology* 4: 272-282.
2. Wiersinga, W. J., B. J. Currie, and S. J. Peacock. 2012. Melioidosis. *The New England journal of medicine* 367: 1035-1044.
3. Friedland, J. S., Y. Suputtamongkol, D. G. Remick, W. Chaowagul, R. M. Strieter, S. L. Kunkel, N. J. White, and G. E. Griffin. 1992. Prolonged elevation of interleukin-8 and interleukin-6 concentrations in plasma and of leukocyte interleukin-8 mRNA levels during septicemic and localized Pseudomonas pseudomallei infection. *Infection and immunity* 60: 2402-2408.
4. Lauw, F. N., A. J. Simpson, J. M. Prins, M. D. Smith, M. Kurimoto, S. J. van Deventer, P. Speelman, W. Chaowagul, N. J. White, and T. van der Poll. 1999. Elevated plasma concentrations of interferon (IFN)-gamma and the IFN-gamma-inducing cytokines interleukin (IL)-18, IL-12, and IL-15 in severe melioidosis. *The Journal of infectious diseases* 180: 1878-1885.
5. White, N. J. 2003. Melioidosis. *Lancet* 361: 1715-1722.

716 6. Limmathurotsakul, D., W. Chaowagul, W. Chierakul, K. Stepniewska, B. Maharjan, V.
717 Wuthiekanun, N. J. White, N. P. Day, and S. J. Peacock. 2006. Risk factors for recurrent
718 melioidosis in northeast Thailand. *Clinical infectious diseases : an official publication of the*
719 *Infectious Diseases Society of America* 43: 979-986.
720 7. Mays, E. E., and E. A. Ricketts. 1975. Melioidosis: recrudescence associated with bronchogenic
721 carcinoma twenty-six years following initial geographic exposure. *Chest* 68: 261-263.
722 8. Haque, A., A. Easton, D. Smith, A. O'Garra, N. Van Rooijen, G. Lertmemongkolchai, R. W. Titball,
723 and G. J. Bancroft. 2006. Role of T cells in innate and adaptive immunity against murine
724 Burkholderia pseudomallei infection. *The Journal of infectious diseases* 193: 370-379.
725 9. Tan, G. Y., Y. Liu, S. P. Sivalingam, S. H. Sim, D. Wang, J. C. Paucod, Y. Gauthier, and E. E. Ooi.
726 2008. Burkholderia pseudomallei aerosol infection results in differential inflammatory responses
727 in BALB/c and C57Bl/6 mice. *Journal of medical microbiology* 57: 508-515.
728 10. Titball, R. W., P. Russell, J. Cuccui, A. Easton, A. Haque, T. Atkins, M. Sarkar-Tyson, V. Harley, B.
729 Wren, and G. J. Bancroft. 2008. Burkholderia pseudomallei: animal models of infection.
730 *Transactions of the Royal Society of Tropical Medicine and Hygiene* 102 Suppl 1: S111-116.
731 11. Conejero, L., N. Patel, M. de Reynal, S. Oberdorf, J. Prior, P. L. Felgner, R. W. Titball, F. J.
732 Salguero, and G. J. Bancroft. 2011. Low-dose exposure of C57BL/6 mice to burkholderia
733 pseudomallei mimics chronic human melioidosis. *The American journal of pathology* 179: 270-
734 280.
735 12. Koh, G. C., R. R. Maude, M. F. Schreiber, D. Limmathurotsakul, W. J. Wiersinga, V. Wuthiekanun,
736 S. J. Lee, W. Mahavanakul, W. Chaowagul, W. Chierakul, N. J. White, T. van der Poll, N. P. Day, G.
737 Dougan, and S. J. Peacock. 2011. Glyburide is anti-inflammatory and associated with reduced
738 mortality in melioidosis. *Clinical infectious diseases : an official publication of the Infectious*
739 *Diseases Society of America* 52: 717-725.
740 13. Koh, G. C., M. F. Schreiber, R. Bautista, R. R. Maude, S. Dunachie, D. Limmathurotsakul, N. P.
741 Day, G. Dougan, and S. J. Peacock. 2013. Host responses to melioidosis and tuberculosis are both
742 dominated by interferon-mediated signaling. *PloS one* 8: e54961.
743 14. Pankla, R., S. Buddhisa, M. Berry, D. M. Blankenship, G. J. Bancroft, J. Banchereau, G.
744 Lertmemongkolchai, and D. Chaussabel. 2009. Genomic transcriptional profiling identifies a
745 candidate blood biomarker signature for the diagnosis of septicemic melioidosis. *Genome*
746 *biology* 10: R127.
747 15. Chin, C. Y., D. M. Monack, and S. Nathan. 2010. Genome wide transcriptome profiling of a
748 murine acute melioidosis model reveals new insights into how Burkholderia pseudomallei
749 overcomes host innate immunity. *BMC genomics* 11: 672.
750 16. Chaussabel, D., C. Quinn, J. Shen, P. Patel, C. Glaser, N. Baldwin, D. Stichweh, D. Blankenship, L.
751 Li, I. Munagala, L. Bennett, F. Allantaz, A. Mejias, M. Ardura, E. Kaizer, L. Monnet, W. Allman, H.
752 Randall, D. Johnson, A. Lanier, M. Punaro, K. M. Wittkowski, P. White, J. Fay, G. Klintmalm, O.
753 Ramilo, A. K. Palucka, J. Banchereau, and V. Pascual. 2008. A modular analysis framework for
754 blood genomics studies: application to systemic lupus erythematosus. *Immunity* 29: 150-164.
755 17. Santanirand, P., V. S. Harley, D. A. Dance, B. S. Drasar, and G. J. Bancroft. 1999. Obligatory role of
756 gamma interferon for host survival in a murine model of infection with Burkholderia
757 pseudomallei. *Infection and immunity* 67: 3593-3600.
758 18. Tippayawat, P., W. Saenwongsa, J. Mahawantung, D. Suwannasaen, P. Chetchotisakd, D.
759 Limmathurotsakul, S. J. Peacock, P. L. Felgner, H. S. Atkins, R. W. Titball, G. J. Bancroft, and G.
760 Lertmemongkolchai. 2009. Phenotypic and functional characterization of human memory T cell
761 responses to Burkholderia pseudomallei. *PLoS neglected tropical diseases* 3: e407.
762 19. Berry, M. P., C. M. Graham, F. W. McNab, Z. Xu, S. A. Bloch, T. Oni, K. A. Wilkinson, R.
763 Banchereau, J. Skinner, R. J. Wilkinson, C. Quinn, D. Blankenship, R. Dhawan, J. J. Cush, A. Mejias,

764 O. Ramilo, O. M. Kon, V. Pascual, J. Banchereau, D. Chaussabel, and A. O'Garra. 2010. An
765 interferon-inducible neutrophil-driven blood transcriptional signature in human tuberculosis.
766 *Nature* 466: 973-977.
767 20. Bloom, C. I., C. M. Graham, M. P. Berry, F. Rozakeas, P. S. Redford, Y. Wang, Z. Xu, K. A.
768 Wilkinson, R. J. Wilkinson, Y. Kendrick, G. Devouassoux, T. Ferry, M. Miyara, D. Bouvry, D.
769 Valeyre, G. Gorochov, D. Blankenship, M. Saadatian, P. Vanhems, H. Beynon, R. Vancheeswaran,
770 M. Wickremasinghe, D. Chaussabel, J. Banchereau, V. Pascual, L. P. Ho, M. Lipman, and A.
771 O'Garra. 2013. Transcriptional blood signatures distinguish pulmonary tuberculosis, pulmonary
772 sarcoidosis, pneumonias and lung cancers. *PloS one* 8: e70630.
773 21. Wiersinga, W. J., M. C. Dessing, and T. van der Poll. 2008. Gene-expression profiles in murine
774 melioidosis. *Microbes and infection / Institut Pasteur* 10: 868-877.
775 22. Easton, A., A. Haque, K. Chu, R. Lukaszewski, and G. J. Bancroft. 2007. A critical role for
776 neutrophils in resistance to experimental infection with Burkholderia pseudomallei. *The Journal*
777 *of infectious diseases* 195: 99-107.
778 23. Wiersinga, W. J., C. W. Wieland, M. C. Dessing, N. Chantratita, A. C. Cheng, D. Limmathurotsakul,
779 W. Chierakul, M. Leendertse, S. Florquin, A. F. de Vos, N. White, A. M. Dondorp, N. P. Day, S. J.
780 Peacock, and T. van der Poll. 2007. Toll-like receptor 2 impairs host defense in gram-negative
781 sepsis caused by Burkholderia pseudomallei (Melioidosis). *PLoS medicine* 4: e248.
782 24. Seok, J., H. S. Warren, A. G. Cuenca, M. N. Mindrinos, H. V. Baker, W. Xu, D. R. Richards, G. P.
783 McDonald-Smith, H. Gao, L. Hennessy, C. C. Finnerty, C. M. Lopez, S. Honari, E. E. Moore, J. P.
784 Minei, J. Cuschieri, P. E. Bankey, J. L. Johnson, J. Sperry, A. B. Nathens, T. R. Billiar, M. A. West,
785 M. G. Jeschke, M. B. Klein, R. L. Gamelli, N. S. Gibran, B. H. Brownstein, C. Miller-Graziano, S. E.
786 Calvano, P. H. Mason, J. P. Cobb, L. G. Rahme, S. F. Lowry, R. V. Maier, L. L. Moldawer, D. N.
787 Herndon, R. W. Davis, W. Xiao, and R. G. Tompkins. 2013. Genomic responses in mouse models
788 poorly mimic human inflammatory diseases. *Proceedings of the National Academy of Sciences of*
789 *the United States of America* 110: 3507-3512.

790