Diagnostic host gene signature to accurately distinguish enteric fever from other febrile diseases

Blohmke CJ[1]*, Muller J[2], Gibani MM[1], Dobinson H[1], Shrestha S[1], Perinparajah S[1], Jin C[1], Hughes H[1], Blackwell L[1], Dongol S[3], Karkey A[3], Schreiber F[4], Pickard D[4], Basnyat B[3], Dougan G[4], Baker S[5], §Pollard AJ[1], §Darton TC[1,5,6]

§Joint senior authorship

Affiliations:

1 Oxford Vaccine Group, Centre for Clinical Vaccinology and Tropical Medicine, Department of Paediatrics, and the Oxford National Institute of Health Research Biomedical Centre, University of Oxford, UK.

2 The Jenner Institute, University of Oxford, Old Road Campus Research Building, Oxford, UK.

3 Oxford University Clinical Research Unit, Patan Academy of Healthy Sciences, Kathmandu, Nepal.

4 Infection Genomics Program, The Wellcome Trust Sanger Institute, Hinxton, UK.

5 The Hospital for Tropical Diseases, Wellcome Trust Major Overseas Programme, Oxford University Clinical Research Unit, Ho Chi Minh City, Vietnam.

6 Department of Infection, Immunity and Cardiovascular Disease, University of Sheffield, Sheffield, United Kingdom.

Correspondence:
*Christoph J. Blohmke,
NIHR Oxford Biomedical Research Centre,
The Churchill Hospital, Oxford OX3 7LE, UK.
eMail: christoph.blohmke@paediatrics.ox.ac.uk

ABSTRACT

Misdiagnosis of enteric fever is a major global health problem resulting in patient mismanagement, antimicrobial misuse and inaccurate disease burden estimates. Applying a machine-learning algorithm to host gene expression profiles, we identified a diagnostic signature which could accurately distinguish culture-confirmed enteric fever cases from other febrile illnesses (AUROC>95%). Applying this signature to a culture-negative suspected enteric fever cohort in Nepal identified a further 12.6% as likely true cases. Our analysis highlights the power of data-driven approaches to identify host-response patterns for the diagnosis of febrile illnesses. Expression signatures were validated using qPCR highlighting their utility as PCR-based diagnostic for use in endemic settings.

Enteric fever, a disease caused by systemic infection with *S. enterica* serovars Typhi or Paratyphi A, accounts for 13.5 to 26.9 million illness episodes worldwide each year.[1,2] In resource-limited tropical settings these infections are endemic and the accurate diagnosis of patients presenting with undifferentiated fever is challenging.

Diagnostic tests for enteric fever rely on microbiological culture or detection of a serological response to infection, and are often unavailable or insufficiently sensitive and specific.[3] Blood culture remains the reference standard against which new diagnostic tests are evaluated, and the sensitivity for this test can reach 80% under optimal conditions[4] but low blood volumes and uncontrolled antibiotic use often result in decreased sensitive in the field. New diagnostic approaches are urgently needed to enable the accurate detection of enteric fever cases in endemic settings, to guide management of febrile patients, appropriate use of antimicrobials, and to identify populations likely to benefit from vaccine implementation.

Most common tests used for acute infectious disease diagnosis employ methods to directly detect the disease-causing pathogen, either by culture, antigen detection or amplification of genetic material by PCR. An alternative approach is to identify a set of human host immune responses, which together may generate a specific pattern associated with individual infections or pathogens. With an increasing quantity of molecular host response data being generated by high-throughput methods – including whole blood gene expression profiling – differences in the activation status of the immune response network during infection may be a tractable diagnostic approach. Recently small sets containing 2-3 genes have been described, the expression of which can accurately differentiate between viral or bacterial infection, and active or latent tuberculosis.[5,6] Merging available well-characterised datasets derived from human clinical samples representative of a variety of fever-causing infections common in tropical settings presents an invaluable resource to identify host immune response patterns specific for enteric fever.

As a human restricted infection, the development of enteric fever diagnostics has been hindered by the lack of reliable *in vivo* models. Using data from a series of controlled human infection models (CHIM)[4,7] or *S.* Typhi or *S.* Paratyphi A infection, whole blood gene transcriptional responses were identified and then further characterised using samples collected from febrile patients in an endemic setting (Kathmandu, Nepal). Integrating these data with publically available human gene transcription datasets, we employed a machine learning algorithm to identify an expression signature that could accurately distinguish blood culture-confirmed EF cases in both the controlled environment (CHIM) and endemic setting from other febrile disease aetiologies and non-infected individuals (healthy controls).[8-12]

Results

Transcriptional profiles in response to enteric fever are similar in challenge study and endemic cohorts

We recently described the molecular response profile of acute enteric fever in individuals participating in the typhoid CHIM, which was characterized by innate immunity, inflammatory and interferon signalling patterns.[13]

To compare responses to enteric fever occurring during natural infection in an endemic area, we generated transcriptional profiles in samples collected from culture-confirmed enteric fever patients (*S.* Typhi: '03NP-ST'; *S.* Paratyphi: '03NP-SPT'), healthy community controls ('03NP-CTRL') and febrile, culture-negative suspected enteric fever cases ('03NP-sEF') recruited in Nepal (Kathmandu; Study: '03NP') (**Figure 1a**). We detected significant differential expression (DE; FDR<0.05, FC±1.25) of 4,308 and 4,501 genes in enteric fever patients with confirmed *S.* Typhi (*n*=19) and *S.* Paratyphi (*n*=12) bacteraemia, respectively, when compared with healthy community controls (*n*=47; **Figure 1b**). Similar numbers of genes were differentially expressed in samples collected at the time of enteric fever diagnosis in healthy adult volunteers challenged with either *S.* Typhi ('T1-ST') or *S.* Paratyphi ('P1-SPT') in a CHIM (**Figure 1b**).[7,13]

As comparison of host responses at the gene level can be difficult to interpret, we performed Gene Set Enrichment Analysis (GSEA)[14] of blood transcriptional modules (BTMs) as a conceptual framework to interpret the host responses in the context of biological pathways and themes.[15] Overall, between 54 and 74 BTMs were significantly enriched (BH adjusted p<0.01) in blood culture-confirmed enteric fever cases in the CHIM and natural infection and CHIM participants who did not develop enteric fever (measured at day 7 post-challenge - 'nD7) (**Supplementary Table 1**). The majority of BTMs enriched in cases from the enteric fever CHIM were also enriched in naturally infected cases from Nepal (56%-69, **Supplementary Table 1**–red squares). Positively enriched modules represented cell cycle (CCY), type I/II interferon and innate antiviral responses (IFN), dendritic cell (DC), innate immunity, inflammation and monocyte (Infl./Mono) signatures. In contrast, T cell (TC) signatures were down-regulated in patients with confirmed enteric fever, as we have previously described (**Figure 1c-e**).[13] In addition a number of modules including inflammasome receptors (M53), monocyte enrichment (M118.0, M118.1, M81, M4.15, M23, M73, M64, S4) and inflammatory responses (M33) were significantly enriched in the CHIM but not in cases from Nepal. Single sample GSEA (ssGSEA) demonstrated the similar enrichment pattern for a selection of IFN and DC signatures between individuals with confirmed typhoid and paratyphoid fever in the CHIM and naturally infected cases (**Figure 1f**). Overall, we observed marked similarity in the gene transcription responses between acute enteric fever cases from the CHIM and an endemic setting in Nepal.

Responses of febrile, culture-negative samples in Nepal

In culture-negative, suspected enteric fever patients ('sEF') from Nepal, we detected differential expression of 3,517 genes when compared with healthy community controls (**Supplementary Figure 1b**). While we observed 2,843 genes as commonly expressed in all three Nepali patient cohorts (03NP-ST, 03NP-SPT and 03NP-sEF), an additional 582, 756 and 183 genes were uniquely expressed by subjects with confirmed *S.* Typhi, *S.* Paratyphi or suspected enteric fever, respectively (**Supplementary Figure 1a&b**). Unsupervised hierarchical clustering of these patients based on their expression of the 500 most variable genes in the Nepal cohort demonstrated clustering into three groups (**Figure 1g**): Group 1 contained mostly healthy control participants; Group 2 contained mostly patients with suspected enteric fever; and Group 3 contained a mixture of patients with suspected enteric fever, and blood culture-confirmed *S.* Typhi or *S.* Paratyphi infection.

Using ssGSEA we observed a heterogeneous BTM enrichment pattern with broad variability in normalized enrichment scores across suspected enteric fever patients (depicted by the interdecile range; **Supplementary Figure 1c**). The most consistent positively or negatively enriched modules represented cell cycle, IFN, inflammatory responses, DC and some NK cell signatures (green cluster) and TC and BC related signatures (red cluster), respectively. In contrast, heterogeneous enrichment in which approximately half of participant samples demonstrated up or down regulation was observed in BTMs representing TC activation patterns, protein folding and metabolism (brown cluster), or in innate response and monocyte signatures (purple cluster) (**Supplementary Figure 1c**). These febrile patients were considered clinically to have enteric fever, and were therefore treated as such, however their heterogeneous gene transcription profiles suggest that any one of several different aetiologies may have precipitated hospital presentation. Further evidence to this is that in a recent RCT a higher proportion of culture-negative cases responded to fluoroquinolones rather than a 3rd generation cephalosporin, possibly due to the frequency of murine and scrub typhus in this population, however distinguishing between these infections is currently difficult.

Multi-cohort data quality assessment

In order to address the potential over-diagnosis of enteric fever and associated inappropriate antimicrobial use, we next aimed to identify a set of genes whose expression is able to differentiate enteric fever from other common febrile conditions found in tropical settings. We repurposed publically available datasets describing host transcriptional response in two malaria,[10,16] four tuberculosis,[8,17] and four dengue cohorts (**Supplementary Table 2**).[9,18,19] We designed a discovery cohort consisting of control samples from each respective study (n=220 community controls or convalescent samples, 'CTRL'), 74 enteric fever ('EF'), 94 blood stage *P. falciparum* ('bsPf'), 67 dengue ('DENV') and 54 active pulmonary tuberculosis ('PTB') cases. An independent validation cohort consisted of 109 CTRLs, 50 EF, 19 bsPf, 49 DENV, and 97 PTB samples (**Figure 2**). Finally,

a cohort of 'unknown' samples was created consisting of febrile culture-negative, febrile suspected EF cases from Nepal ('sEF'), and samples collected from CHIM study participants who did not elop enteric fever after challenge at day 7 ('nD7') and their respective pre-challenge baseline samples ('D0') (**Figure 2**). Using Principle Component Analysis (PCA) to assess the variability at the level of gene expression between the cohorts indicated some distinct clustering between cases (**Supplementary Figure 2a**), for each infection whereas no such differences were observed with the comparator CTRL samples (**Supplementary Figure 2b**).

Five genes sufficiently distinguish EF from other febrile infections

With these data, we aimed to build a classifier containing a minimum set of genes that could discriminate culture-confirmed enteric fever cases from individuals with other causes of fever (class: 'Rest', consisting of CTRLs, DENV, PTB and bsPf) (2-class classification, **Figure 2**) using a Guided Regularized Random Forest (GRRF) algorithm.[20] Genes were ranked by frequency of selection in each of 100 iterations, and applying a selection threshold of ≥25%, we identified a putative diagnostic signature containing *STAT1* (98% of iterations), *SLAMF8* (76%), *PSME2* (39%), *WARS* (37%), and *ALDH1A1* (36%) (**Figure 3a**). With this 5-gene signature we were able to predict which individuals in the validation cohort had enteric fever with a sensitivity and specificity of 97.1% and 88.0%, respectively (AUROC: 96.7%) (**Figure 3b**, **Supplementary Table 3a**). Of blood culture-confirmed enteric fever cases in the validation cohort, 6/51 were misclassified as 'Rest' (i.e. classification probability>0.5, **Figure 3c**-top), and 8/274 samples belonging to class 'Rest' were classified as enteric fever. These included six tuberculosis and one dengue case, and a pre-challenge baseline sample from a CHIM participant (**Figure 3c**-bottom).

To allow comparison between the different disease conditions, we quantified expression of the 5 genes identified in each sample using the z-score of the geometric mean of the expression values (expression score). Significant differences in expression scores were observed between the enteric fever samples and all other conditions in both the discovery (top) and the validation (bottom) cohort (**Figure 3d**). Of note, there were no significant differences between the scores calculated for the control samples derived from endemic areas or naïve, healthy controls from the CHIM, indicating the homogeneity of expression to these genes in healthy controls from different study and geographical locations.

The design of discovery and validation cohorts is likely to have an impact on the diagnostic signature selected, and we therefore exchanged the validation and discovery cohort and re-ran the analysis. Although in this experiment 4 instead of 5 genes were selected (using a threshold ≥25%), most genes included were also part of the initial signature (*STAT1*, *SLAMF8*, *WARS*) and the high predictive accuracy was maintained (AUROC: 97.2%) (**Supplementary Figure 3a&b**). These results

demonstrate the ability of a small number of genes to accurately predict true EF cases from other febrile illnesses caused by another bacterial pathogen (TB), and of parasitic or viral origin.

Multiclass prediction accurately classifies three of five conditions simultaneously

Given the apparent success of small gene expression signatures in classifying two distinct groups, we sought to leverage the overall dataset and the GRRF algorithm to identify a signature that could accurately classify more than two classes simultaneously. We re-analysed the data preserving the original class labels (i.e. CTRL, bsPf, DENV, PTB and EF) and performed the iterative feature selection step using the GRRF algorithm (**Figure 2**–"multiclass classification"). Applying a ≥25% selection threshold to ranked features identified 7 genes (*RFX7*, *C1QB*, *ANKRD22*, *WARS*, *BATF2*, *STAT1*, and *C1QC*) able to discriminate the classes (**Figure 3e**). Prediction of the validation cohort using this 7-gene signature indicated good sensitivity and specificity for accurately classifying CTRL, bsPf and EF cases, however the identification of DENV and PTB was less accurate (**Figure 3f**, **Supplementary Table 3b**). Analysis of individual gene expression levels in each group indicated that *RFX7* was only upregulated in bsPf samples, while *STAT1*, *WARS* as well as *ANKRD22* and *BATF2* were all strongly upregulated in EF. Expression of these genes in PTB and DENV samples was variable accounting for the lower performance of the signature in these conditions (**Supplementary Figure 4a&b**).

Prediction of unknown samples

Given the superior performance of the 2-class diagnostic signature, our subsequent analyses focused on using the initial 5-genes identified to ascertain whether enteric fever was the likely true underlying aetiology of suspected febrile, blood culture-negative cases in Nepal (sEF; *n*=71), part of the unknown cohort (**Figure 2**). Included in this cohort were 144 samples originating from the challenge study with known class membership confirming the correct classification of 94.4% of the samples by the GRRF algorithm (**Supplementary Table 4**).

Classification of these sEF cases predicted 9/71 (12.6%) febrile, culture-negative patients to be true enteric fever cases and the remaining samples to belong to class 'Rest' (**Figure 4a**). Relating the gene expression scores to the predicted class probabilities indicated no clear separation of scores according to the predicted class (**Figure 4b**). Furthermore, comparing the expression score of febrile, culture-negative samples with culture confirmed enteric fever in Nepal showed a marked overlap, indicating that these scores alone are insufficient for 2-class discrimination (**Figure 4c**).

Diagnostic validation by qPCR

Finally, to validate the induction of the diagnostic gene signature in blood culture-confirmed enteric fever cases, we performed high-throughput qPCR in samples collected during an independent typhoid CHIM (**Supplementary Table 2**)[21] and in the Nepali cohort. Transcription of the 5-gene signature was increased at the time of diagnosis in most participants with culture-confirmed enteric fever in both sample sets (**Figure 4d&e**). Two CHIM participants diagnosed with typhoid infection and one patient infected with *S.* Paratyphi in Nepal showed low expression of all genes and a resulting low expression score (**Figure 4e**–black arrows). In contrast, one day 7 sample from a participant not diagnosed with enteric fever demonstrated high expression of the putative diagnostic gene signature (**Figure 4e**–black arrows).

As surrogate disease severity markers, temperature showed poor correlation with the expression score in both CHIM and endemic setting culture-confirmed enteric fever cases (**Figure 4f&g**–left). In contrast, C-reactive protein levels (only available for CHIM participants) were significantly associated with the expression score of the 5-gene signature (**Figure 4f&g**–right) thus underlining the relevance of this signature in reflecting the clinical presentation of enteric fever. In the Nepal cohort, gene expression also strongly correlated between the array and qPCR data (**Supplementary Figure 6**). Overall these results verify the strong expression of the putative diagnostic signatures in samples from patients with acute enteric fever and underline the clinical plausibility through association with disease severity parameters.

Discussion

New approaches to diagnose patients with enteric fever are urgently needed, as currently available methods are antiquated and unreliable. New diagnostic modalities are required, to both improve the immediate management of patients, and to increase the accuracy of disease burden measurements to support targeted vaccine implementation. Here we demonstrate a reproducible host expression signature of 5 genes (*STAT1*, *SLAMF8*, *PSME2*, *WARS*, and *ALDH1A1*) able to discriminate EF cases from other common causes of fever in the tropics with an accuracy of >96%. To our knowledge, this exceeds the performance of all previously described enteric fever diagnostic methods, which often perform less well when assessed using samples collected directly from patients or participants. Moreover, application of high-throughput methods such as functional genomics, to this major health concern,[22] underscores the importance and tangible benefits of applying 'omics-technologies' to combatting infectious diseases in the most needy populations.[23] While further optimisation work is required, validating the expression of our signature using conventional methods such as qPCR demonstrates feasibility of further development into an affordable diagnostic test for use in endemic settings.[24]

The degree of perturbation of molecular responses occurring during enteric fever can be confounded by the duration of clinical illness (ranging in 12hrs to ≥3 days in the CHIM and patients from Nepal. , respectively) or the specific pathogen (*S.* Typhi or *S.* Paratyphi). This may hinder identification of a reproducible gene expression signature reliably expressed in various settings. The responses to *S.* Typhi and *S.* Paratyphi cases in Nepal were remarkably similar, with the majority of DE genes overlapping between the two groups, which is unsurprising given the close genetic relatedness of both pathogens.[25] Enrichment of BTMs resembled responses described previously by us[13,26] and underlined the concordance between culture-confirmed enteric fever cases from Oxford and Nepal despite the possible differences between challenge and currently circulating strains.

Despite the multiple redundancies incorporated into human immune pathways driven by successful evolution,[27] our data suggest that the pattern of immune response activation is sufficiently specific to allow identification of the causative pathogen. For example, while immune responses during enteric fever and TB are broadly characterized by IFN-signalling, we and others have reported that this response during acute *S.* Typhi infection appears to be skewed towards a type-II pattern likely associated with neutrophils and NK cells rather than the type-I dominated profile found in TB.[7,8,13,28-31] Application of computational methods to large datasets including host gene expression has been shown to be an effective approach to capture such differential activation of immune pathways.[5,6] Two of the genes identified in our 5-gene diagnostic signature are important entities in the IFN-γ signalling cascade (*STAT1*, *WARS*), which has been broadly implicated in the responses to enteric fever, TB,[8] dengue,[32] and *P. falciparum*[33] infection. The discriminatory impact of increased expression of these genes identified in our analysis, however, suggests that there are distinct differences during the

responses to these very different pathogens sufficient to discriminate underlying disease aetiology[34,35] possibly based on subtle metabolic differences.[13,36] While STAT1 and WARS are markers of an IFN-γ response, SLAMF8 is surface-expressed protein[37] found in macrophages, DCs and neutrophils and induced by IFN-γ or Gram-negative bacteria.[38] SLAMF8 negatively regulates ROS production through inhibition of NADPH oxidase 2 (NOX2) in the bacterial phagosome and reduces ROS-induced inflammatory cell migration.[39] While oxidative stress is a common response to infection, *Salmonella* survival is reduced in SLAMF1-deficient mice and can interfere with localization of functional NOX2 in *Salmonella*-containing vacuoles (SCVs), linking SLAM proteins and oxidative stress.[40] PSME2 is one of two interferon-inducible subunits of the 20S immunoproteasome (IP) regulator 11S and is involved in immune responses and antigen processing.[41] The 20S IP can be induced by oxidative stress and preferentially hydrolyses non-ubiquitinated proteins.[42,43] Thus, genes involved in these processes may be exploited to distinguish between pathogens inducing oxidative stress from those also triggering ubiquitination.[44,45] While ALDH1A1 has not specifically been linked with responses to invasive bacterial infections, it is involved in gut-homing of TCs through expression of retinoic acid,[46,47] a phenotype we have observed following infection with *S.* Typhi.[26] C1QB and C1QC are well-known subunits of the complement subcomponent C1q and, together with ANKRD22 (involved in cell cycle control[48]), have previously been described as part of a signature able to distinguish active from latent TB.[17] The function of the transcription factor *RFX7* is largely unknown, but has been found to be strongly up-regulated during blood stage malaria and its selection in our 7-gene signature is therefore likely to be driving the classification of malaria cases.

Of note, while multiclass classification is difficult to perform and here merely serves as demonstration that data driven approaches may be capable of performing this task, it is interesting to observe increased misclassification rates specifically in the DENV and TB groups. In the validation cohort, the majority of misclassified DENV cases were identified as enteric fever (5/49) or TB (9/49), and misclassified TB samples as enteric fever (13/97) or DENV (23/97), possibly reflecting the overlapping immune response seen due to the intracellular nature of all three pathogens. In the TB group, 15/97 samples were misclassified as controls, compared with one DENV sample being misclassified as such for example, potentially owing to the broad clinical phenotype or lack of inflammatory/immune responses seen in the peripheral blood during tissue specific pulmonary TB infection.

Overall, the genes identified in both signatures through our unbiased selection approach are supported by previous studies including those aiming to develop predictive diagnostic signatures.[8,17,49] In the era of biological 'big data', several studies have explored the utility of gene transcription signatures capable of discriminating viral aetiologies, viral or bacterial infections as well as acute or latent tuberculosis.[5,6,17,50-53] Only in the tuberculosis studies have such signatures been identified from samples collected in high-incidence, disease endemic settings and been further validated against other

disease processes including (but not limited to) pneumonia, sepsis, and streptococcal and staphylococcal infections.[6,8,50] Herberg *et al.* demonstrated that distinction between viral and bacterial infections could be achieved based on two genes only.[5] In contrast most efforts undertaken to diagnose active TB employ biomarker signatures ranging in size from 3-86 genes, possibly due to broad and heterologous molecular responses seen in response to differing clinical phenotypes of infection. In our analysis we specifically focused on pathogens with the potential to cause undifferentiated febrile illnesses in tropical settings. While the clinical presentation and epidemiology of the infections chosen may be sufficient to distinguish the aetiologies clinically, enteric fever has a broad differential diagnosis and is frequently over-diagnosed in the absence of confirmatory laboratory results. Notably, despite the high prediction accuracy of the signatures identified in our analysis, this type of data modelling is highly dependent on the quality and availability of suitable input datasets. Although an increasing amount of data is accumulating in the public domain, few well-defined datasets of samples representing a larger repertoire of febrile illnesses are available. For example, rickettsial infection is likely to underlie a large burden of the culture-negative cases in Nepal, however no gene expression datasets exist and the lack of adequate confirmatory diagnostic tests further hinders the inclusion of such data in our analysis.

Although the 5-gene signature achieved high accuracy in identifying enteric fever cases, several culture-confirmed cases were misclassified. Metadata from samples collected in the Oxford CHIM indicate that the majority of these misclassified samples had a temperature below 37°C (5/6) and were diagnosed beyond 7 days after challenge (4/6), which, in our CHIM experience, is likely to indicate a less severe disease phenotype. In contrast, six nD7 samples from the Oxford CHIM (part of the unknown cohort) classified as enteric fever showed some sign of response either based on increased cytokines, temperature or a positive stool culture (data not shown). Because our analysis was purely data driven and not motivated by clinical suspicion, we believe that these observations and the significant association of the gene expression scores with CRP provide sufficient evidence that these study participants had infection despite not meeting our study endpoint definitions for enteric fever.

In summary, our work demonstrates how a large gene expression dataset derived from challenge study cohorts and settings endemic for febrile infectious diseases can be exploited for diagnostic biomarker discovery. Verification of the putative diagnostic signature using qPCR in independent validation sets indicates that a diagnostic test derived from these gene expression data could be developed for deployment in resource-limited settings. The application of purely data-driven analyses to large and well-defined host-pathogen datasets derived from disease relevant populations may enable us to develop a single, highly accurate diagnostic signature which would allow rapid identification of the main fever-causing aetiologies from readily available biological specimens.

Online Methods

Typhoid challenge model

Samples included in the discovery cohort were collected during a typhoid dose-escalation study in which 41 healthy adult volunteers ingested a single dose of *S.* Typhi Quailes strain following pre-treatment with 120 mL sodium bicarbonate solution (Study: T1). In this study, one of two doses were administered: 1-5x10^3 (*n*=21) and 1-5x10^4 (*n*=20).[4] Samples used in the validation cohort were collected from a second typhoid challenge model performed as part of a vaccine efficacy study (Study: T2), in which healthy adult volunteers ingested a single dose of *S.* Typhi Quailes strain (1-5x10^4, *n*=99) 4 weeks after oral vaccination with Ty21a, M01ZH09 or placebo.[54] Lastly, samples collected from the control arm of a further vaccine efficacy challenge study, in which participants received meningococcal ACWY-CRM conjugate vaccine (MENVEO®, GlaxoSmithKline) prior to challenge, were used for the independent qPCR validation experiment.[21] The clinical and molecular results of these studies have been described previously.[4,7,21,54] In all typhoid challenge studies participants were treated with a 2-week course of antibiotics at the time of diagnosis (fever ≥38°C sustained for ≥12hrs and/or positive blood culture), or at day 14 post-challenge if diagnostic criteria were not reached.

Paratyphoid challenge model

Clinical samples for paratyphoid infection were collected during a dose-escalation study, as previously described (P1).[7] Briefly, 40 healthy adult volunteers were challenged with a single oral dose of virulent *S.* Parayphi A (strain NVGH308) bacteria, which as before, was suspended in 30mL sodium bicarbonate solution [17.5mg/mL], and after pre-treatment with 120mL sodium bicarbonate solution. Oral challenge inocula was given at one of two dose levels, low (*n*=20; median [range]=0.9×10^3 CFU [0.7×10^3–1.3×10^3]) or high dose (*n*=20; median [range]=2.4×10^3 CFU [2.2×10^3–2.8×10^3]). Criterion for diagnosis were either microbiological (≥1 positive blood culture collected after day 3) and/or clinical (fever ≥38°C sustained for ≥12hrs). Participants were ambulatory and followed up as outpatients at least daily after challenge when safety, clinical, and laboratory measurements were performed.[7]

Endemic Cohort

To validate the gene transcriptional signatures in a relevant patient cohort, blood samples were collected from three cohorts at Patan Hospital or the Civil Hospital both located in the Lalitpur Sub-Metropolitan City area of Kathmandu Valley in Nepal. Firstly, blood samples were collected as part of a diagnostics study[55] from febrile patients presenting to hospital and diagnosed with blood culture-confirmed *S.* Typhi (*n*=19) or *S.* Paratyphi A (*n*=12) infection and febrile patients who were blood

culture negative (*n*=71). Samples from a cohort of healthy control volunteers (*n*=44) were also collected as part of this study.

Gene expression arrays sample processing

In the human challenge studies (T1, T2, and P1), peripheral venous blood (3mL) was collected in Tempus™ Blood RNA tubes (Applied Biosystems) before challenge (baseline, pre-challenge controls, 'D0', *n*=166) and at paratyphoid diagnosis ('SPT', *n*=18) or typhoid diagnosis ('ST', n=75). In those challenged but who did not develop enteric fever within 14 days of challenge, gene expression was measured at the median day of diagnosis of the diagnosed group in the appropriate studies and this day was termed 'nD7' (*n*=73). In Nepal, blood was collected when patients presented to hospital (n=102) and from healthy controls (*n*=44) (**Figure 1A**, **Supplementary Table S1**). Total RNA was extracted from all samples using the Tempus™ Spin RNA Isolation kit (Life Technologies). Where applicable, 50ng of RNA was used for hybridization into Illumina HT-12v4 bead-arrays (Illumina Inc.) at the Wellcome Trust Sanger Institute (Hinxton, UK) or The Wellcome Trust Centre for Human Genetics (Oxford, UK) and fluorescent probe intensities captured with the GenomeStudio software (Illumina Inc.). For the paratyphoid CHIM (P1) RNA gene expression was determined using RNA sequencing. Briefly, libraries were prepared using a poly-A selection step to exclude ribosomal RNA species (read length: 75bp paired-end) and samples were subsequently multiplexed in 95 samples/lane over 10 lanes plus one 5-plex pool run on 1 lane and sequenced using a Illumina HiSeq 200 V4.

Data pre-processing

Paired-end reads were adapter removed and trimmed from 75 to 65bp using trimmomatic v0.35[56] and only reads exceeding a mean base quality 5 within all sliding windows of 5bp were mapped to the Gencode v25/hg38 transcriptome using STAR aligner v2.5.2b keeping only multi mapped reads mapping to at most 20 locations. featureCounts from the subread set of tools v1.5.1 was used to quantify reads in Gencode v25 basic gene locations with parameters -C -B -M -s 2 -p -S fr. Between-sample normalization was performed using TMM (Trimmed Mean of M-values) normalization as implemented in the edgeR[57] package and we used principle component analysis (PCA) as quality control step and excluded 2 samples, which were clear outliers due to also failing QC during the library preparation. Counts were converted into $\log_2$ counts per million (cpm) values with 0.5 prior counts to avoid taking the logarithm of zero and were then taken forward to the multi-cohort quality control. Illumina HT-12v4 bead array data were pre-processed by background subtraction, quantile normalization and $\log_2$-transformation using the limma package in R.[58] Probes were collapsed to HUGO gene identifiers keeping only the highest expressed probe.

Data download

Previously published whole blood transcriptional array data was downloaded from the Gene Expression Omnibus (GEO) data repository. In this study we specifically focused on studies

investigating blood stage *Plasmodium falciparum* (bsPf; two cohorts of blood-stage, HIV-negative malaria cohorts; children and adults),[10,16] acute uncomplicated dengue (DENV; four adult South-East Asian cohorts of uncomplicated dengue fever patients),[9,18,19] and active pulmonary tuberculosis (PTb; four cohorts of active, pulmonary TB HIV-negative adults from Africa and the UK),[8,17] all infections which present with undifferentiated fever and are relevant to areas where enteric fever is endemic (**Supplementary Table S2**). Raw data were downloaded from GEO using the getGEO-function[59] and quantile normalization with detection p-values and control probes where available. Probes were collapsed to HUGO gene identifiers keeping only the highest expressed probe.

Data processing and cohort Quality Control

Probe sequences on microarrays may not correspond to the most recent release of the human reference genome that was used for the RNAseq alignment. In order to mitigate this potential discrepancy we re-annotated the probes to the Gencode v25/hg38. The new annotations were used as gene names for each probe. To avoid uninformative genes and gender bias only probes common to all datasets, not located on sex chromosomes and with an expression above the lowest tertile of the average expression (12,821 probes) were used and a 'superset' was created by merging the expression data from all studies into one large data matrix. In order to avoid platform or study related artefacts between the data we applied surrogate variable analysis (sva)[60] to remove batch effects based on study ID while preserving the disease condition (i.e. control or individual infection).

Diagnostic signature identification

For classification analyses, we separated the superset into a discovery cohort and a validation cohort. To ensure heterogeneity and optimal feature identification we restricted the discovery cohort to samples solely generated on Illumina platforms and ensured inclusion of EF samples from Oxford and Nepal. In order to establish a validation cohort we casted a wider net and permitted studies generated on other platforms including Affymetrix due to the limited amount of suitable datasets available in the public domain. In addition, to predict unknown samples by applying the signatures identified in this study, we separated the febrile, culture-negative suspected enteric fever cases, samples at day 7 after challenge of those who stayed well and their respective pre-challenge control samples from the superset into a cohort of samples of unknown aetiology (Unknown Cohort) (**Figure 2**).

Only the discovery cohort was used for feature selection using Guided Regularized Random Forest (GRRF)[20] as implemented in the R package RRF v1.7[61] with gamma = 0.5 and parameter mtry tuning was performed using the tuneRRF command. Feature selection was repeated on 100 iterations of bootstrapped subsets of about 70% of the data in the discovery cohort. To assess model performance, predictions on the held out 30% of the discovery cohort were performed and balanced accuracies[62] were recorded to account for class imbalances. Genes were then ranked by the frequency of positive gene selection by GRRF (based on mean Gini) during the 100 iterations and only genes included in at

least 25% of the selection rounds were included in the diagnostic signature and used for prediction of the independent validation cohort as well as the samples belonging to the unknown cohort (**Figure 2**).

High-throughput qPCR validation

We performed TaqMan gene expression assays to validate gene expression levels in samples from Nepal and a subset of individuals from the Oxford challenge studies. A panel of 24 probes were measured in triplicates on a 192.24 Fluidigm chip using the Biomark at the Weatherall Institute for Molecular Medicine (WIMM) single cell facility. Four samples and one probe failed in the quality control and were removed from the analysis. Raw Ct values were normalized to the housekeeping gene cyclophilin A (PPIA) ($^{\Delta}$Ct values) and subsequently to control samples (healthy controls) to achieve ΔΔCt values.

Statistical analysis

All data were processed in R version 3.2.4. Comparison of groups in Figure 3d were performed using Student's t-Test and correlations between clinical parameters and expression scores were performed using Pearson correlation and correlation between array and qPCR expression as performed using Spearman correlations (alternative: two-sided).

Data deposition:

The datasets generated in these studies were deposited at GEO: GSE113867.

Acknowledgements

We gratefully acknowledge the assistance of the participants who have taken part in the study both in Oxford and Nepal. We are also grateful for the support from Laura B. Martin and GSK Vaccines for Global Health for setting up the paratyphoid challenge model and providing the *S.* Paratyphi O:2 antigen and the *S.* Paratyphi challenge strain (NVGH308). We are grateful for the support from Myron M. Levine for providing the *S.* Typhi Quailes strain used in the typhoid CHIM.

This work was supported by the Bill and Melinda Gates Foundation (OPP1089317 and OPP1084259); funding for the challenge studies was provided by a Wellcome Trust Strategic Translational Award (grant number 092661 to AJP); the European Vaccine Initiative (ref: PIM); European Commission FP7 grant "Advanced Immunization Technologies" (ADITEC); and the NIHR Oxford Biomedical Research Centre (Clinical Research Fellowship to TCD; oxfordbrc.nihr.ac.uk). For the support of the fluidigm experiment we acknowledge the Weatherall Institute of Molecular Medicine (WIMM) single cell facility, which was supported by the MRC funded Oxford Consortium for Single-cell Biology (MR/M00919X/1), and the Oxford-Wellcome Trust Institutional Strategic Support Fund.

TCD, CJ, CJB, CSW, AJP are supported by the NIHR Oxford Biomedical Research Centre (Oxford University Hospitals NHS Trust, Oxford), TCD is an NIHR funded Academic Clinical Lecturer; SB is a Sir Henry Dale Fellow, jointly funded by the Wellcome Trust and the Royal Society (100087/Z/12/Z); AJP is a Jenner Investigator, James Martin Senior Fellow and an NIHR Senior Investigator. The views expressed in this article are those of the author(s) and not necessarily those of the NHS, the NIHR, or the Department of Health.

Author Contributions: The study in Nepal was designed by TCD, AJP and CJB and run by BB, AK, SD, SB and TCD. Challenge studies at Oxford were designed and performed by AJP, TCD, MMG, HD, CJ, CJB. Laboratory work and data generation was performed by CJB, SS, SP, HH, LB, FS, DP and GD. The computational analysis was conceptualized and executed by JM and CJB. The manuscript was conceptualized and written by CJB and TCD. All authors critically reviewed the manuscript.

Competing Interest Statement: All authors declare not competing interests.

References

1 Mogasale, V. *et al.* Burden of typhoid fever in low-income and middle-income countries: a systematic, literature-based update with risk-factor adjustment. *Lancet Glob Health* **2**, e570-580, doi:10.1016/S2214-109X(14)70301-8 (2014).

2 Buckle, G. C., Walker, C. L. & Black, R. E. Typhoid fever and paratyphoid fever: Systematic review to estimate global morbidity and mortality for 2010. *J Glob Health* **2**, 010401, doi:10.7189/jogh.02.010401 (2012).

3 Parry, C. M., Wijedoru, L., Arjyal, A. & Baker, S. The utility of diagnostic tests for enteric fever in endemic locations. *Expert Rev Anti Infect Ther* **9**, 711-725, doi:10.1586/eri.11.47 (2011).

4 Waddington, C. S. *et al.* An outpatient, ambulant-design, controlled human infection model using escalating doses of Salmonella Typhi challenge delivered in sodium bicarbonate solution. *Clin Infect Dis* **58**, 1230-1240, doi:10.1093/cid/ciu078 (2014).

5 Herberg, J. A. *et al.* Diagnostic Test Accuracy of a 2-Transcript Host RNA Signature for Discriminating Bacterial vs Viral Infection in Febrile Children. *JAMA* **316**, 835-845, doi:10.1001/jama.2016.11236 (2016).

6 Sweeney, T. E., Braviak, L., Tato, C. M. & Khatri, P. Genome-wide expression for diagnosis of pulmonary tuberculosis: a multicohort analysis. *Lancet Respir Med* **4**, 213-224, doi:10.1016/S2213-2600(16)00048-5 (2016).

7 Dobinson, H. C. *et al.* Evaluation of the clinical and microbiological response to Salmonella Paratyphi A infection in the first paratyphoid human challenge model. *Clin Infect Dis*, doi:10.1093/cid/cix042 (2017).

8 Berry, M. P. *et al.* An interferon-inducible neutrophil-driven blood transcriptional signature in human tuberculosis. *Nature* **466**, 973-977, doi:10.1038/nature09247 (2010).

9 Hoang, L. T. *et al.* The early whole-blood transcriptional signature of dengue virus and features associated with progression to dengue shock syndrome in Vietnamese children and young adults. *J Virol* **84**, 12982-12994, doi:10.1128/JVI.01224-10 (2010).

10 Idaghdour, Y. *et al.* Evidence for additive and interaction effects of host genotype and infection in malaria. *Proc Natl Acad Sci U S A* **109**, 16786-16793, doi:10.1073/pnas.1204945109 (2012).

11 Naim, A. N., Tolfvenstam, T., Fink, K. & Hibberd, M. L. Genome-wide gene expression analysis of human whole-blood samples during acute dengue disease and early convalescence. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE28991 (2016).

12 Obermeyer, Z. & Emanuel, E. J. Predicting the Future - Big Data, Machine Learning, and Clinical Medicine. *N Engl J Med* **375**, 1216-1219, doi:10.1056/NEJMp1606181 (2016).

13 Blohmke, C. J. *et al.* Interferon-driven alterations of the host's amino acid metabolism in the pathogenesis of typhoid fever. *J Exp Med* **213**, 1061-1077, doi:10.1084/jem.20151025 (2016).

14 Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**, 15545-15550, doi:10.1073/pnas.0506580102 (2005).

15 Li, S. *et al.* Molecular signatures of antibody responses derived from a systems biology study of five human vaccines. *Nat Immunol* **15**, 195-204, doi:10.1038/ni.2789 (2014).

16 Subramaniam, K. S. *et al.* The T-Cell Inhibitory Molecule Butyrophilin-Like 2 Is Up-regulated in Mild Plasmodium falciparum Infection and Is Protective During Experimental Cerebral Malaria. *J Infect Dis* **212**, 1322-1331, doi:10.1093/infdis/jiv217 (2015).

17 Kaforou, M. *et al.* Detection of tuberculosis in HIV-infected and -uninfected African adults using whole blood RNA expression signatures: a case-control study. *PLoS Med* **10**, e1001538, doi:10.1371/journal.pmed.1001538 (2013).

18 Kwissa, M. *et al.* Dengue virus infection induces expansion of a CD14(+)CD16(+) monocyte population that stimulates plasmablast differentiation. *Cell Host Microbe* **16**, 115-127, doi:10.1016/j.chom.2014.06.001 (2014).

19 Tolfvenstam, T. *et al.* Characterization of early host responses in adults with dengue disease. *BMC Infect Dis* **11**, 209, doi:10.1186/1471-2334-11-209 (2011).
20 Deng, H. & Runger, G. Gene selection with guided regularized random forest *Pattern Recognition* **46**, 3483-3489 (2013).
21 Jin, C. *et al.* Efficacy and immunogenicity of a Vi-tetanus toxoid conjugate vaccine in the prevention of typhoid fever using a controlled human infection model of Salmonella Typhi: a randomised controlled, phase 2b trial. *Lancet*, doi:10.1016/S0140-6736(17)32149-9 (2017).
22 Escadafal, C. *et al.* New Biomarkers and Diagnostic Tools for the Management of Fever in Low- and Middle-Income Countries: An Overview of the Challenges. *Diagnostics (Basel)* **7**, doi:10.3390/diagnostics7030044 (2017).
23 Baker, S. Genomic medicine has failed the poor. *Nature* **478**, 287, doi:10.1038/478287a (2011).
24 Jiang, L. *et al.* Solar thermal polymerase chain reaction for smartphone-assisted molecular diagnostics. *Sci Rep* **4**, 4137, doi:10.1038/srep04137 (2014).
25 McClelland, M. *et al.* Comparison of genome degradation in Paratyphi A and Typhi, human-restricted serovars of Salmonella enterica that cause typhoid. *Nat Genet* **36**, 1268-1274, doi:10.1038/ng1470 (2004).
26 Salerno-Goncalves, R. *et al.* Challenge of Humans with Wild-type Salmonella enterica Serovar Typhi Elicits Changes in the Activation and Homing Characteristics of Mucosal-Associated Invariant T Cells. *Front Immunol* **8**, 398, doi:10.3389/fimmu.2017.00398 (2017).
27 Nish, S. & Medzhitov, R. Host defense pathways: role of redundancy and compensation in infectious disease phenotypes. *Immunity* **34**, 629-636, doi:10.1016/j.immuni.2011.05.009 (2011).
28 Spees, A. M. *et al.* Neutrophils are a source of gamma interferon during acute Salmonella enterica serovar Typhimurium colitis. *Infect Immun* **82**, 1692-1697, doi:10.1128/IAI.01508-13 (2014).
29 Thompson, L. J. *et al.* Transcriptional response in the peripheral blood of patients infected with Salmonella enterica serovar Typhi. *Proc Natl Acad Sci U S A* **106**, 22433-22438, doi:10.1073/pnas.0912386106 (2009).
30 Blohmke, C. J. *et al.* Induction of Cell Cycle and NK Cell Responses by Live-Attenuated Oral Vaccines against Typhoid Fever. *Front Immunol* **8**, 1276, doi:10.3389/fimmu.2017.01276 (2017).
31 Manca, C. *et al.* Hypervirulent M. tuberculosis W/Beijing strains upregulate type I IFNs and increase expression of negative regulators of the Jak-Stat pathway. *J Interferon Cytokine Res* **25**, 694-701, doi:10.1089/jir.2005.25.694 (2005).
32 De La Cruz Hernandez, S. I. *et al.* A strong interferon response correlates with a milder dengue clinical condition. *J Clin Virol* **60**, 196-199, doi:10.1016/j.jcv.2014.04.002 (2014).
33 Miller, J. L., Sack, B. K., Baldwin, M., Vaughan, A. M. & Kappe, S. H. Interferon-mediated innate immune responses against malaria parasite liver stages. *Cell Rep* **7**, 436-447, doi:10.1016/j.celrep.2014.03.018 (2014).
34 Munoz-Jordan, J. L., Sanchez-Burgos, G. G., Laurent-Rolle, M. & Garcia-Sastre, A. Inhibition of interferon signaling by dengue virus. *Proc Natl Acad Sci U S A* **100**, 14333-14338, doi:10.1073/pnas.2335168100 (2003).
35 Obiero, J. M. *et al.* Impact of malaria preexposure on antiparasite cellular and humoral immune responses after controlled human malaria infection. *Infect Immun* **83**, 2185-2196, doi:10.1128/IAI.03069-14 (2015).
36 Zhang, Y. J. *et al.* Tryptophan biosynthesis protects mycobacteria from CD4 T-cell-mediated killing. *Cell* **155**, 1296-1308, doi:10.1016/j.cell.2013.10.045 (2013).
37 van Driel, B. J., Liao, G., Engel, P. & Terhorst, C. Responses to Microbial Challenges by SLAMF Receptors. *Front Immunol* **7**, 4, doi:10.3389/fimmu.2016.00004 (2016).

38 Wang, G. *et al.* Cutting edge: Slamf8 is a negative regulator of Nox2 activity in macrophages. *J Immunol* **188**, 5829-5832, doi:10.4049/jimmunol.1102620 (2012).

39 Wang, G. *et al.* Migration of myeloid cells during inflammation is differentially regulated by the cell surface receptors Slamf1 and Slamf8. *PLoS One* **10**, e0121968, doi:10.1371/journal.pone.0121968 (2015).

40 Fang, F. C. Antimicrobial actions of reactive oxygen species. *MBio* **2**, doi:10.1128/mBio.00141-11 (2011).

41 Nandi, D., Jiang, H. & Monaco, J. J. Identification of MECL-1 (LMP-10) as the third IFN-gamma-inducible proteasome subunit. *J Immunol* **156**, 2361-2364 (1996).

42 Seifert, U. *et al.* Immunoproteasomes preserve protein homeostasis upon interferon-induced oxidative stress. *Cell* **142**, 613-624, doi:10.1016/j.cell.2010.07.036 (2010).

43 Dubiel, W., Pratt, G., Ferrell, K. & Rechsteiner, M. Purification of an 11 S regulator of the multicatalytic protease. *J Biol Chem* **267**, 22369-22377 (1992).

44 Cirillo, S. L. *et al.* Protection of Mycobacterium tuberculosis from reactive oxygen species conferred by the mel2 locus impacts persistence and dissemination. *Infect Immun* **77**, 2557-2567, doi:10.1128/IAI.01481-08 (2009).

45 Spooner, R. & Yilmaz, O. The role of reactive-oxygen-species in microbial persistence and inflammation. *Int J Mol Sci* **12**, 334-352, doi:10.3390/ijms12010334 (2011).

46 Iwata, M. *et al.* Retinoic acid imprints gut-homing specificity on T cells. *Immunity* **21**, 527-538, doi:10.1016/j.immuni.2004.08.011 (2004).

47 Molotkov, A. & Duester, G. Genetic evidence that retinaldehyde dehydrogenase Raldh1 (Aldh1a1) functions downstream of alcohol dehydrogenase Adh1 in metabolism of retinol to retinoic acid. *J Biol Chem* **278**, 36085-36090, doi:10.1074/jbc.M303709200 (2003).

48 Yin, J. *et al.* ANKRD22 promotes progression of non-small cell lung cancer through transcriptional up-regulation of E2F1. *Sci Rep* **7**, 4430, doi:10.1038/s41598-017-04818-y (2017).

49 Zak, D. E. *et al.* A blood RNA signature for tuberculosis disease risk: a prospective cohort study. *Lancet* **387**, 2312-2322, doi:10.1016/S0140-6736(15)01316-1 (2016).

50 Anderson, S. T. *et al.* Diagnosis of childhood tuberculosis and host RNA expression in Africa. *N Engl J Med* **370**, 1712-1723, doi:10.1056/NEJMoa1303657 (2014).

51 Andres-Terre, M. *et al.* Integrated, Multi-cohort Analysis Identifies Conserved Transcriptional Signatures across Multiple Respiratory Viruses. *Immunity* **43**, 1199-1211, doi:10.1016/j.immuni.2015.11.003 (2015).

52 Mahajan, P. *et al.* Association of RNA Biosignatures With Bacterial Infections in Febrile Infants Aged 60 Days or Younger. *JAMA* **316**, 846-857, doi:10.1001/jama.2016.9207 (2016).

53 Zaas, A. K. *et al.* Gene expression signatures diagnose influenza and other symptomatic respiratory viral infections in humans. *Cell Host Microbe* **6**, 207-217, doi:10.1016/j.chom.2009.07.006 (2009).

54 Darton, T. C. *et al.* Using a Human Challenge Model of Infection to Measure Vaccine Efficacy: A Randomised, Controlled Trial Comparing the Typhoid Vaccines M01ZH09 with Placebo and Ty21a. *PLoS Negl Trop Dis* **10**, e0004926, doi:10.1371/journal.pntd.0004926 (2016).

55 Darton, T. C. *et al.* Identification of Novel Serodiagnostic Signatures of Typhoid Fever Using a Salmonella Proteome Array. *Front Microbiol* **8**, 1794, doi:10.3389/fmicb.2017.01794 (2017).

56 Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120, doi:10.1093/bioinformatics/btu170 (2014).

57 Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140, doi:10.1093/bioinformatics/btp616 (2010).

58 Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* **43**, e47, doi:10.1093/nar/gkv007 (2015).

59 Davis, S. & Meltzer, P. S. GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. *Bioinformatics* **23**, 1846-1847, doi:10.1093/bioinformatics/btm254 (2007).
60 Leek, J. T. *et al.* sva: Surrogate Variable Analysis. *R Package Version* **3.24.4** (2017).
61 Deng, H. Guided Random Forest in the RRF Package. *arXiv:1306.0237* (2013).
62 Brodersen, K. H., Ong, C. S., Stephan, K. E. & Buhmann, J. M. The balanced accuracy and its posterior distribution. . *Pattern Recognition (ICPR), 2010 20th International Conference on, 3121-3124*, doi:10.1109/ICPR.2010.764 (2010).

Figure Legends and Figures:

Figure 1: Overview of Oxford and Nepal comparison. (**a**) Overview of enteric fever cohorts used in this study (T1: Typhoid CHIM study 1; T2: Typhoid CHIM study 2; P1: Paratyphoid CHIM; 03NP: Nepali cohort. ST: *S.* Typhi; SPT: *S.* Paratyphi; sEF: suspected Enteric Fever; D0: day of challenge which represents the control samples in the Oxford CHIM; CTRL: endemic community controls; nD7: day 7 after challenge in participants who stayed well in the CHIM; BC+: Blood-culture positive; BC-: Blood-culture negative; Dx: Diagnosis). (**b**) Volcano plots of up (red) and down (blue) regulated genes in *S.* Typhi and *S.* Paratyphi positive individuals (Nepal and Oxford). Black numbers indicate the up- and down-regulated genes. (**c**) Circular plot depicting the overlap of BTMs between enteric fever and nD7 samples from Oxford and Nepal. Tracks (from outer to inner): cohort and samples; BTM labels; direction of enrichment (blue: down; red: up). Cords represent overlap of enrichment in given cohorts (red: overlap between P1-SPT and T1-ST; green: overlap between T1-nD7 and P1-nD7; blue: overlap of 03NP-ST with P1-SPT and T1-ST; purple: overlap of 03NP-SPT with P1-SPT and T1-ST; yellow: overlap between 03NP-SPT and 03NP-ST). (**d-e**) Scatter plots of BTMs enriched ($p>0.05$) in blood-culture positive samples in Nepal (y-axis) versus Oxford (x-axis) for typhoid fever (d) and paratyphoid fever (e). (**f**) Single-sample GSEA Normalised Enrichment Scores (NES) of IFN and DC BTMs of individuals with blood-culture confirmed enteric fever in Nepal and Oxford. (**g**) Heatmap of the 500 most variably expressed genes in samples of the Nepali cohort. Bar graph on top of the heatmap shows temperature of each individual at the time of sampling.

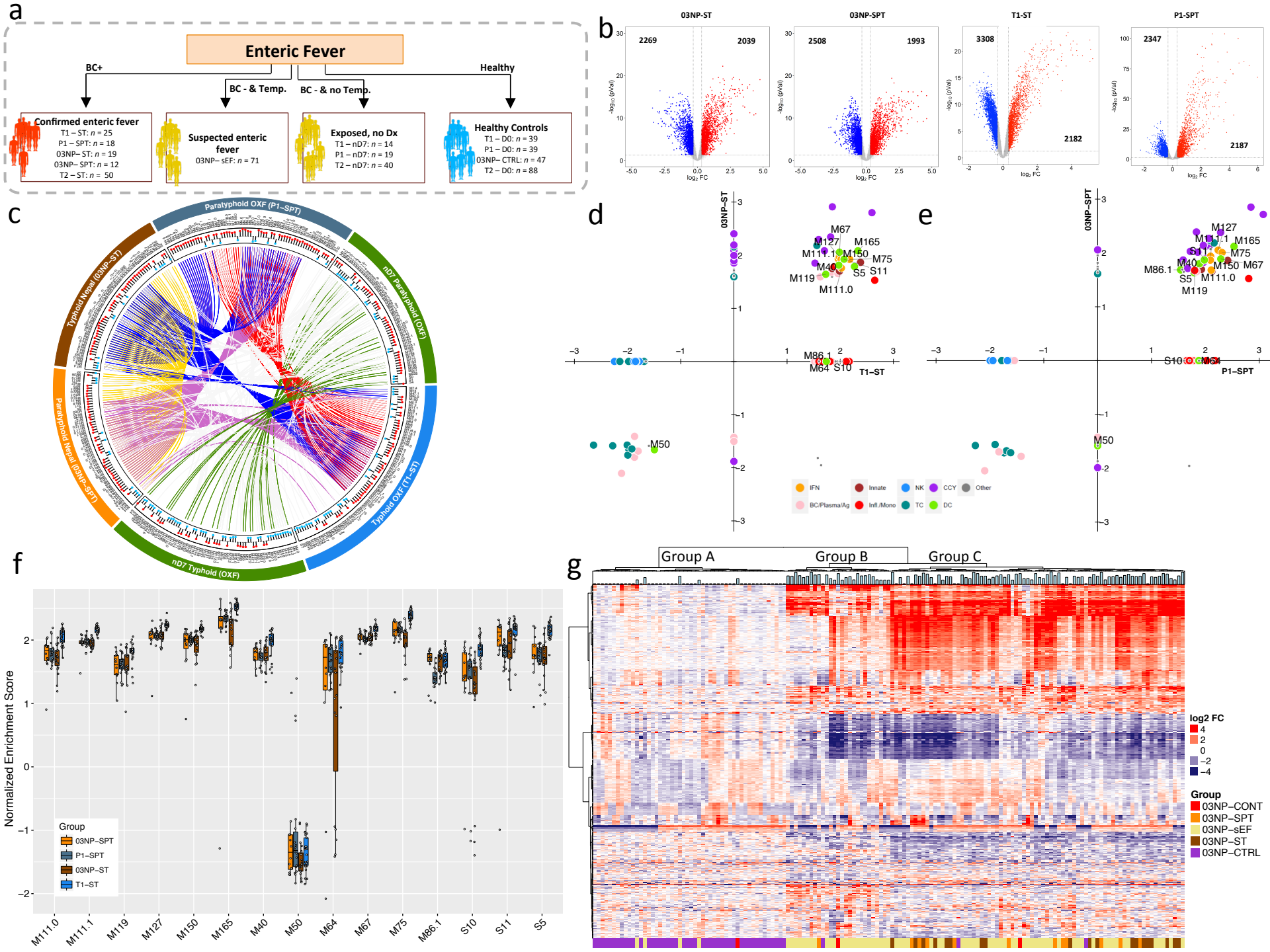
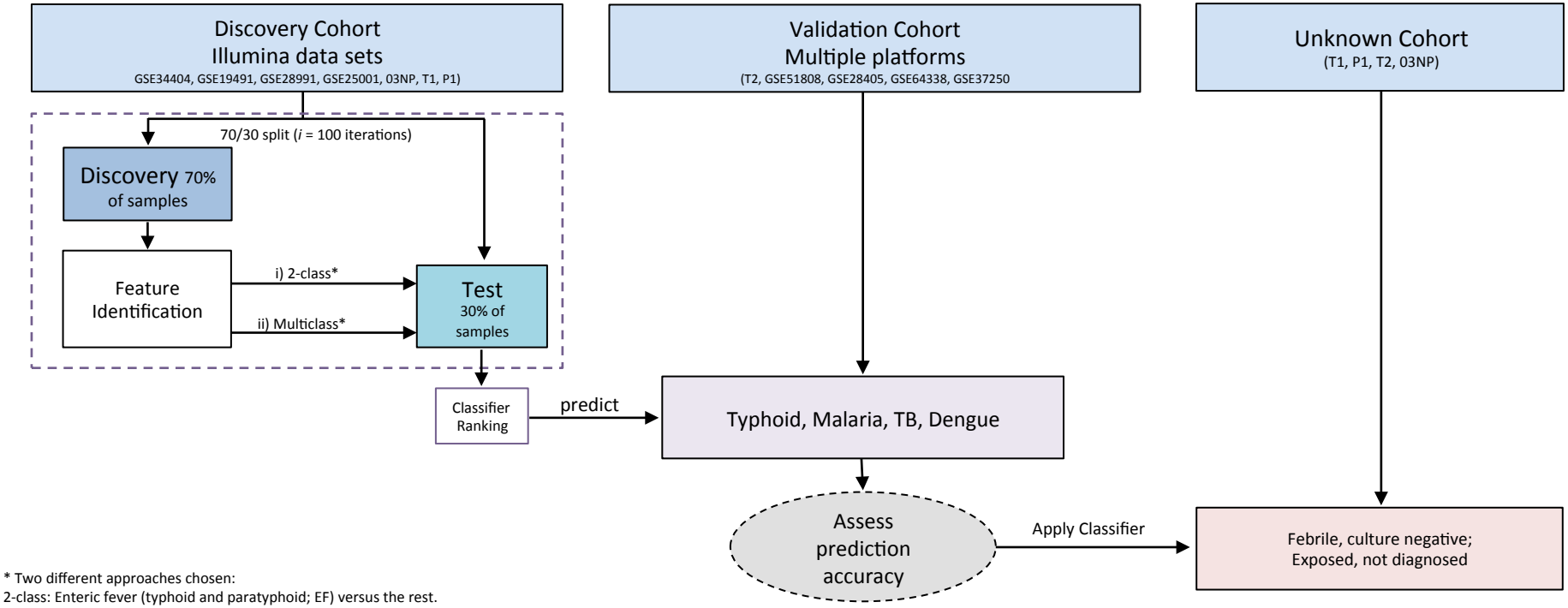

a
Enteric Fever
BC+
BC - & Temp.
BC - & no Temp.
Healthy
Confirmed enteric fever
T1 – ST: n = 25
P1 – SPT: n = 18
03NP– ST: n = 19
03NP– SPT: n = 12
T2 – ST: n = 50
Suspected enteric fever
03NP– sEF: n = 71
Exposed, no Dx
T1 – nD7: n = 14
P1 – nD7: n = 19
T2 – nD7: n = 40
Healthy Controls
T1 – D0: n = 39
P1 – D0: n = 39
03NP– CTRL: n = 47
T2 – D0: n = 88
b
03NP-ST
03NP-SPT
T1-ST
P1-SPT
2269
2039
2508
1993
3308
2182
2347
2187
log2 FC
c
Paratyphoid OXF (P1–SPT)
nD7 Paratyphoid (OXF)
Typhoid OXF (T1–ST)
nD7 Typhoid (OXF)
Paratyphoid Nepal (03NP–SPT)
Typhoid Nepal (03NP–ST)
d
03NP–ST
T1–ST
M67
M165
M127
M111.1
M150
M75
M40
S5
S11
M119
M111.0
M86.1
M64
S10
M50
IFN
Innate
NK
CCY
Other
BC/Plasma/Ag
Infl./Mono
TC
DC
e
03NP–SPT
P1–SPT
f
Normalized Enrichment Score
Group
03NP–SPT
P1–SPT
03NP–ST
T1–ST
M111.0
M111.1
M119
M127
M150
M165
M40
M50
M64
M67
M75
M86.1
S10
S11
S5
g
Group A
Group B
Group C
log2 FC
Group
03NP–CONT
03NP–SPT
03NP–sEF
03NP–ST
03NP–CTRL

Figure 2: Flow diagram of machine learning analysis. The discover cohort consisted of only Illumina datasets and was used for feature selection using the GRRF algorithm. For the validation cohort Affymetrix datasets were also included. A cohort of unknown samples consisted of pre-challenge baseline samples of participants who stayed well following challenge, their respective nD7 samples (7 days after challenge), and febrile, culture-negative suspected enteric fever (sEF) cases from Nepal. Refer to Supplementary Table S2 for study identifiers. 03NP: Nepali cohort. T1: Oxford typhoid CHIM study 1. T2: Oxford typhoid CHIM study 2; P1: Oxford paratyphoid CHIM.



* Two different approaches chosen:
2-class: Enteric fever (typhoid and paratyphoid; EF) versus the rest.
Multiclass: Simultaneously differentiation between all classes (EF, bsPf, DENV, PTB, CTRL).

Figure 3: Identification of diagnostic signatures. (**a**) Ranking of genes by their selection frequency into the diagnostic signature out of 100 iterations during the 2-class classification. A cut-off of 25% was selected to detect 5-gene putative diagnostic signature (orange bar). (**b**) Performance of the 5-gene classifier when predicting the class membership of the validation cohort. (**c**) Top: Probability of an EF sample to be classified as non-EF (>0.5). Bottom: Probability of sample belonging to 'rest' to be classified as EF (>0.5). (**d**) Combined expression score for samples based on the 5-gene signature for samples in the discovery cohort (top) and validation cohort (bottom). Ox.CTRL: Oxford controls (D0); CTRL: Nepali control samples. PTB: pulmonary TB; DENV: Dengue samples; bsPf: blood-stage *P. falciparum*; SPT: *S.* Paratyphi; ST: *S.* Typhi. (**e**) Ranking of genes by their selection frequency into the diagnostic signature out of 100 iterations during the multiclass classification. A cut-off of 25% was selected to detect a 7-gene putative diagnostic signature (orange bar). (**f**) Classification probabilities for each sample of the validation cohort based on the 7-gene signature. Significance levels in panel d were determined using the Student's *t*-Test: *$p<0.05$; **$p<0.01$; ****$p<0.0001$.

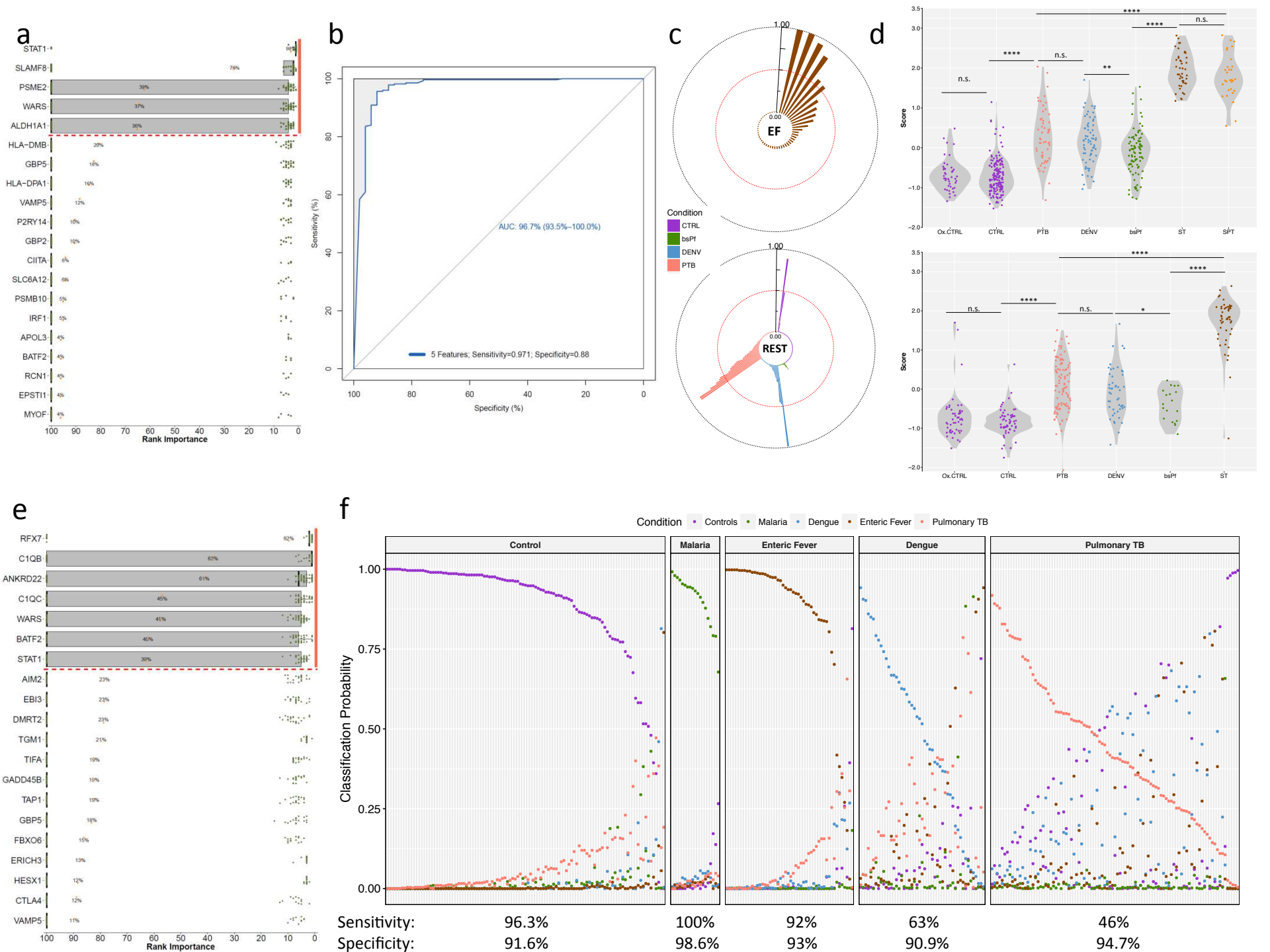
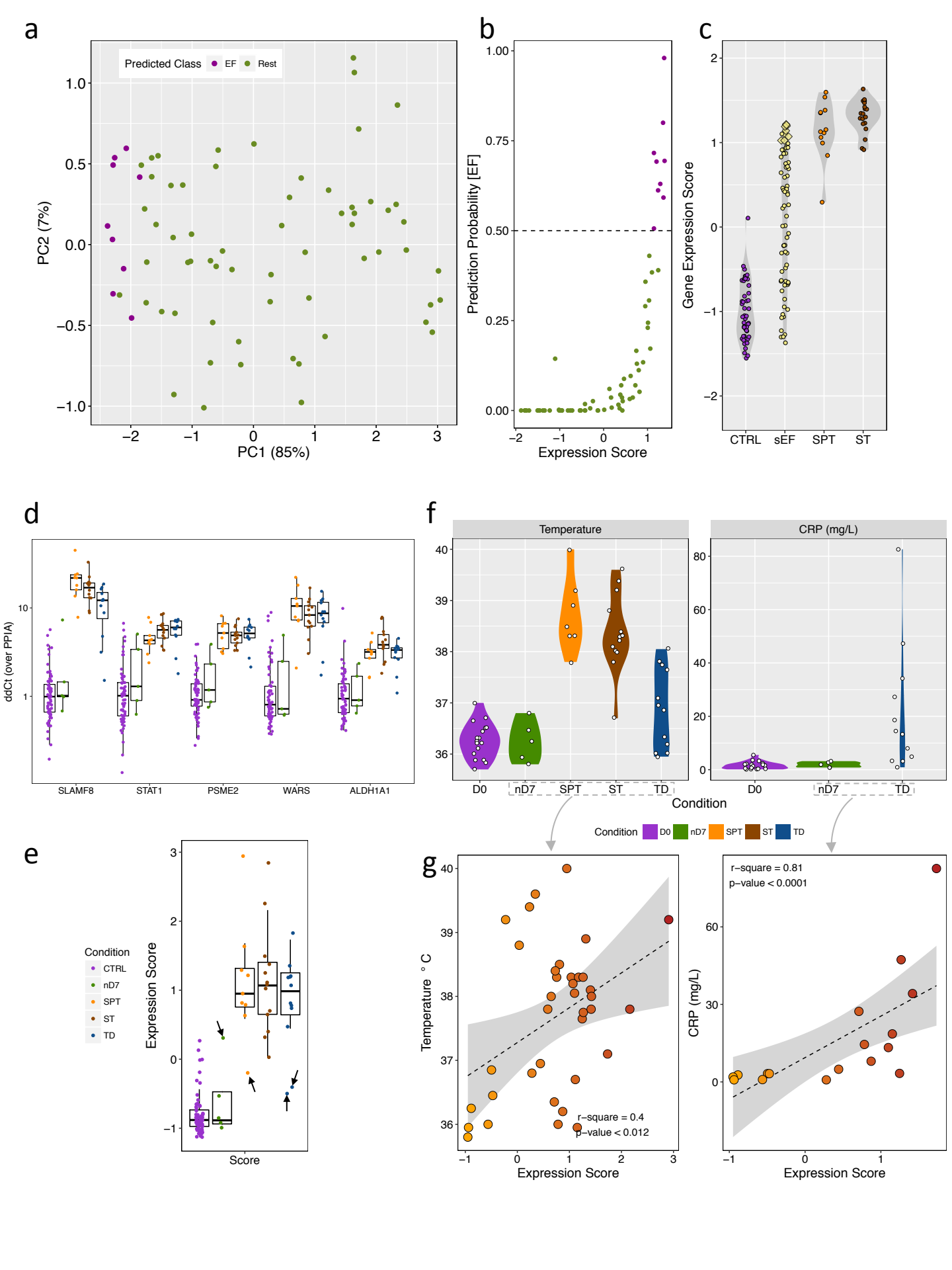

a
STAT1
SLAMF8
PSME2
WARS
ALDH1A1
HLA-DMB
GBP5
HLA-DPA1
VAMP5
P2RY14
GBP2
CIITA
SLC6A12
PSMB10
IRF1
APOL3
BATF2
RCN1
EPSTI1
MYOF
Rank Importance
b
Sensitivity (%)
Specificity (%)
AUC: 96.7% (93.5%–100.0%)
5 Features; Sensitivity=0.971; Specificity=0.88
c
EF
REST
Condition
CTRL
bsPf
DENV
PTB
d
Score
Ox.CTRL
CTRL
PTB
DENV
bsPf
ST
SPT
n.s.
****
**
*
e
RFX7
C1QB
ANKRD22
C1QC
WARS
BATF2
STAT1
AIM2
EBI3
DMRT2
TGM1
TIFA
GADD45B
TAP1
GBP5
FBXO6
ERICH3
HESX1
CTLA4
VAMP5
Rank Importance
f
Condition
Controls
Malaria
Dengue
Enteric Fever
Pulmonary TB
Control
Malaria
Enteric Fever
Dengue
Pulmonary TB
Classification Probability
Sensitivity: 96.3% 100% 92% 63% 46%
Specificity: 91.6% 98.6% 93% 90.9% 94.7%

Figure 4: Prediction of Nepali unknown samples using the 2-class and qPCR validation. (**a**) PCA of sEF samples based on the 5-gene signature coloured by predicted class membership (EF: purple; green: rest). (**b**) Dot plot of prediction probability of being class EF versus the expression score calculated on the bases of the 5-gene signature. (**c**) qPCR gene expression scores of the 5-gene signature (ΔΔCT over PPIA) for CTRLs, sEF, SPT and ST samples from Nepal. Yellow diamonds in the sEF category represent the 9 patients classified as EF based on the RF algorithm. (**d**) qPCR expression values (ΔΔCt over PPIA) of the 5-gene signature in control samples (Oxford and Nepal), samples at day 7 after challenge of participants who stayed well following challenge with *S.* Typhi (nD7), *S.* Paratyphi (SPT) or *S.* Typhi (ST) in Nepal, or typhoid diagnosis after challenge (TD). Colour legend in panel (e). (**e**) Combined qPCR expression score of the 5-gene signature. Black arrows indicate outlier samples. (**f**) Temperature and CRP for samples of which data was available (CRP was only measured in the Oxford CHIM). (**g**) Spearman's rank correlation of the 5-gene combined expression score and temperature (left; only nD7 and TD samples from the Oxford CHIM and SPT and ST cases from Nepal were included) and CRP (right; CRP was only available for Oxford CHIM samples and we excluded D0 baseline measures) at presentation to hospital (Nepal), diagnosis (Oxford CHIM) or day 7 after challenge in those who stayed well (Oxford CHIM).


a
Predicted Class
EF
Rest
PC2 (7%)
PC1 (85%)
b
Prediction Probability [EF]
Expression Score
c
Gene Expression Score
CTRL
sEF
SPT
ST
d
ddCt (over PPIA)
SLAMF8
STAT1
PSME2
WARS
ALDH1A1
e
Condition
CTRL
nD7
SPT
ST
TD
Expression Score
Score
f
Temperature
CRP (mg/L)
D0
nD7
SPT
ST
TD
Condition
g
Temperature ° C
r−square = 0.4
p−value < 0.012
CRP (mg/L)
r−square = 0.81
p−value < 0.0001
Expression Score

Supplementary Figure Legends

Supplementary Figure 1: Differentially expressed genes and BTMs of sEF cases from Nepal. (**a**) Volcano plot of differentially expressed genes in the Nepali sEF cohort. (**b**) Venn diagram representing the overlap of DE genes in the Nepali ST, SPT and sEF cases. (**c**) ssGSEA heatmap of BTMs significantly expressed ($p<0.05$) in at least 60% of sEF samples. Bar plot panel represents the interdecile range (IDR) for each BTM across all sEF cases. NES: Normalized Enrichment Score.

Supplementary Figure 2: Superset quality control. (**a**) PCA plot based on the 500 most variable genes (IQR) of enteric fever cases (EF), malaria cases (bsPf), dengue cases (DENV) and TB cases (PTB) after batch correction. (**b**) PCA of all control samples for each disease cohort after batch correction.

Supplementary Figure 3: Signature identification using a re-designed discovery and validation cohort. (**a**) Ranking of genes by their selection frequency into the diagnostic signature out of 100 iterations during the 2-class classification. A cut-off of 25% was chosen to detect a putative diagnostic signature consisting of 4 genes (orange bar). (**b**) Prediction of the validation cohort using the 4 genes identified in (a).

Supplementary Figure 4: Expression of the 7 target genes identified during the multiclass classification analysis in each sample of the discovery (**a**) and validation cohort (**b**).

Supplementary Figure 5: Prediction of Oxford CHIM samples part of the unknown cohort. (**a**) PCA of Oxford pre-challenge baseline samples and nD7 samples based on the expression values for the 5-gene diagnostic signature (2-class classification) coloured by predicted class membership (green: REST, purple: EF). (**b**) Dot plot of prediction probabilities against a combined expression score for each sample coloured by predicted class membership.

Supplementary Figure 6: Spearman correlation of expression values of the 5-gene diagnostic signature derived from microarrays or qPCR.

Supplementary Tables:

Supplementary Table 1: Overlap of BTMs between different study groups (in percent).

Supplementary Table 2: Datasets included in this study.

Supplementary Table 3. (**a**) Contingency table of class membership following the 2-class classification. (**b**) Contingency table of class membership following the multiclass classification.

Supplementary Table 4: Prediction accuracy and overview of misclassified samples following prediction using the 5-gene 2-class signature of the Oxford samples included in the unknown cohort.

Supplementary Table 5: Class memberships of Oxford CHIM and Nepali samples included in the unknown cohort following the prediction using the 7-gene multiclass signature.