Journal: *BioRxiv*

Manuscript title: **“Data driven analysis reveals shared transcriptome response and immune cell composition across aetiologies of critical illness.”**

Running title: **“Shared transcriptome and immune responses in the critically ill.”**

Authors:

Zsolt Zador¹, Alexander Landry¹, Michael Balas¹, John C .Marshall², Michael D. Cusimano¹

- 1) Department of Surgery, Li Ka Shing Knowledge Institute, St. Michael’s Hospital,
Toronto, ON, Canada
- 2) Critical Illness and Injury Research Centre, Li Ka Shing Knowledge Institute, St.
Michael’s Hospital, Toronto, ON, Canada

Corresponding Author:

Zsolt Zador, MD, PhD

Department of Surgery, St. Michael’s Hospital

30 Bond Street, Toronto, ON, Canada M5B 1W8

Email: zadzso@gmail.com

Phone: 00-416-864-5312

Disclosures, Conflict of interest: none

Author Contributions

ZZ: study design, data analysis, manuscript preparation

APL: data analysis, manuscript revisions

MB: data analysis, manuscript revisions

JCM: study design

MDC: manuscript revisions

Competing Interests

The authors declare no competing interests.

Availability of Materials and Data

All data referenced in this study is publicly available through the open repository Gene

Expression Omnibus (<https://www.ncbi.nlm.nih.gov/geo/>).

Data driven analysis reveals shared transcriptome response and immune cell fractions across different aetiologies of critical illness.

Abstract

Sepsis and trauma are frequent and challenging health problems in critical care. The diversity of patient response to these conditions complicates both disease management and outcome prediction. Whole blood transcriptomics allows the analysis of response in the critically ill at a molecular level. Prior results in this field demonstrate robust and diverse genomic response in the acute phase and others have shown shared biological mechanisms across wide disease aetiologies. We hypothesize that specific biological mechanisms, particularly those related to immune processes, are shared between sepsis and trauma cohorts. These may serve as universal markers for patients vulnerable to a complicated clinical course and/or mortality. We present a systems level analysis of gene expression for a total of 317 patients with abdominal sepsis (51), pulmonary sepsis (108) or trauma (158) and compare them to healthy controls (68). Our results confirm that immune processes are shared across disease aetiologies in critical illnesses. We identify two consistent and distinct subgroups of critical illness: 1) increased dendritic cell and CD4 T helper fractions but suppressed neutrophils and 2) high neutrophils and otherwise suppressed leukocyte fractions. These subgroups validate in an independent cohort of 181 paediatric patients suffering from septic shock of diverse aetiologies. Furthermore, we found immune and inflammatory processes derived from gene co-expression networks were downregulated in subgroup 1. This paralleled prior results which find similar leukocyte configuration by deconvolving whole blood transcriptomics, and this leukocyte configuration associates with greater susceptibility to multi organ failure. We are the first to identify a patient subgroup

with a preserved leukocyte configuration across aetiologies of critical illness, which may serve as a universal predictor of complicated clinical course/poor outcome.

Introduction

Host responses to critical illness such as infection or trauma are highly variable between patients (1–3). Specific examples of this complexity include the reprioritization patterns in gene expression (2,4,5) and the coexistence of proinflammatory and immunosuppressed states (1–3,6,7) in critically ill patients. Understanding this diversity in systemic response is highly important to treatment planning and recovery prediction. Analysis of whole blood transcriptomics (WBT) may facilitate this by identifying individual patient response to a variety of stresses leading to marker discovery and improved outcome prediction. It can also give valuable insight into biological processes and mechanisms of disease. WBT involves rapid isolation of total blood leukocytes followed by purification of total RNA and high throughput quantification of gene expression levels (1,2). WBT analysis has been applied to predict outcome in trauma (5), identify clinically relevant subgroups in sepsis (6) and explain diverse response to vaccination (8). Physiological processes such as aging have also been associated with a transcriptomics pattern to establish the “molecular age” of patients (9).

“Systems level analysis” of transcriptomics data can associate gene expression patterns with molecular processes and ultimately biological/clinical phenotypes. Two well-established and widely used methods in systems biology are gene expression networks (10,11) and the deconvolution of cellular fractions in bulk tissue (12). Gene expression networks are able to identify mechanisms relevant to physiological and pathological states by detecting clusters of highly co-expressed transcripts (13–15). An additional advantage of this approach is the ability to capture low and additive gene expression signals which may otherwise be overlooked by conventional differential gene expression analysis. This network-based technique has been successfully applied in exploring complex biological phenotypes like Huntington’s disease (16), peripheral nerve regeneration (13), malignant

primary brain tumors (17) or weight gain (11). Deconvolution of bulk transcriptomics data uses cell-specific markers to estimate the fraction of corresponding cellular types present in a tissue sample (12). This technique has provided insight into the immune landscape of cancer (18), the tumor microenvironment in glioblastoma multiforme (19) and the immune cell composition of sera in a variety of diseases (20).

Studies have shown shared transcriptomics responses across different aetiologies of sepsis (6) and a large scale comparison of different diseases also suggested substantial overlap in WBT signal (4). However, the characteristics of these shared mechanisms and their relevance remains less explored. We firstly hypothesize that transcriptomics response to a variety of critical illnesses organizes into biological processes and an immune profile which are shared across aetiologies. We further hypothesize responses that are more predictive of good outcome than others. In the first part of our study we applied gene co-expression networks to open source transcriptomics data (2,21) and explore biological mechanisms in critical illness: abdominal sepsis, pulmonary sepsis and trauma. Our research is novel in the following: 1) contemporary studies largely rely on differential gene expression analysis to explore underlying mechanisms in critical illness, while our study uses co-expression network to identify relevant genes. This approach is rooted in biological reality, and may capture additive biological signal which could otherwise be overlooked by conventional techniques; 2) Our study is the first to explore the heterogeneity of patient response across critical illness with different aetiologies (sepsis due to community acquired pneumonia/abdominal source and trauma). We confirm that immune processes are shared across these conditions and we further explore the immune cell composition to identify consistently homogenous patient subgroups in critical illness.

Methods

Database

We used open source datasets from the open source genetic repository Gene Expression Omnibus (GEO) (22): GSE11375 (5) for trauma and GSE65682 (MARS Consortium) for abdominal (23) and pulmonary (24) sepsis. Demographics for the cohort are summarized in Table 1. For verification of our results we use a paediatric cohort of 181 patients (GSE66099) suffering from septic shock (25).

Table 1: Study demographics. A: control, B: trauma, C: CAP, D: abdominal sepsis, E: Septic shock (unknown aetiology) *Of available data. #Validation cohort

GEO entry	N	Mean age (SD)	Sex (M:F)	Patient distribution					Mortality (%)
				A	B	C	D	E	
GSE11375	184	33.1 (11.1)	115:69	26	158	0	0	0	7/184 (3.8)
GSE65682	201	58.1 (17.3)	114:87	42	0	108	51	0	31/201 (15.4)
GSE66099 [#]	228	N/A	N/A	47	0	0	0	181	N/A
Overall*	613	46.2 (19.2)	229:156	68	158	108	51	181	38/385 (9.9)

Systems level analysis of whole blood gene expression data

The trauma cohort consisted of 158 adult patients under the age of 55 who suffered severe blunt trauma. Blood samples were taken within 12 hours of the injury and processed for gene expression analysis as described by Xiao et al (2). The cohort of patients suffering from sepsis (GSE65682) was divided into two groups based on the admitting: sepsis due to community acquired pneumonia (CAP) and abdominal sepsis (AS). Blood samples were taken within 24 hours of admission to critical care (23,24). To demonstrate the generalizability of our findings, we chose a large dataset of 276 paediatric patients, the Genomics of Pediatric SIRS and Septic Shock Investigators (GPSSSI)(25). In our verification subgroup, we had 181 unique patients who have been admitted with a diagnosis of septic

shock from a variety of infectious sources and blood sample drawn on the first day of admission.

Gene co-expression patterns were identified using the well-established “Weighted Gene Correlation Network Analysis” (WGCNA) to detect “modules” (clusters) of strongly co-expressed genes (10). Per these previously described techniques, we first computed an “adjacency matrix”, a symmetrical matrix carrying soft-thresholded Pearson correlations between each gene pair. Computation of the similarity measures is detailed by Langfelder et al (10). This was then converted into a biologically-inspired topological overlap map (TOM), wherein pairwise gene similarities were derived from comparing their connectivity profiles (26). Hierarchical clustering with subsequent “dynamic” tree-cut (27) ultimately yielded gene modules, whose biological functions were annotated with the *Database for Annotation, Visualization and Integrated Discovery* (DAVID) (28), an open-source bioinformatics resource. Additionally, representative module “meta-genes” for each sample were computed as the first principal component of their constituent genes’ expression values, a well-established approach. We demonstrated associations between gene modules (thereby biological mechanisms) and external traits by correlating module meta genes with trait labels (for example as healthy vs septic) as described previously (10). Shared mechanisms as annotated by DAVID were identified based on ontological similarities using the GoSemSim package (29). This method determines similarities between Gene Ontology (GO) terms based on graph-based measures, wherein mechanisms are associated based on evolutionary and biological relationships and similarity is computed based on graph topology structure. Methods are described further in ref (29).

Deconvolution of leukocyte cell factions

We used the well-established technique of CIBERSORT (12) to estimate the immune cell composition of whole blood. This method relies upon a previously derived reference gene expression matrix generated from a collection of pure lymphocyte fractions. The signature matrix was truncated to the genes detectable in all 3 cohorts (514 genes out of 547 genes in the original signature matrix). Within this signature matrix, each cell type is assigned a series of markers, which are then incorporated into nu-support vector regression to derive cell fractions. The method is further described in ref (12). The CIBERSORT technique is implemented from R package *epiDISH*.

Patient subgroup discovery and validation

We used k-means clustering to identify patient subclasses based on immune cell landscape. This is a widely used technique in computational biology which initiates with a specified number of clusters (k) and assigns data points to each cluster based on a distance measure (30). During an iterative process, the position of the cluster centres in the data space are adjusted based on the cluster memberships as points are being added. Data points are reassigned as the cluster centre is being adjusted. The process continues until there is further adjustment required to the cluster centres (clusters are deemed stable). The optimal class numbers were estimated using the “elbow” method, where the number of clusters are weighed against the percent of variance explained by the clusters (31,32).

To validate the distinct composition of each cluster we first optimized a logistic regression model using backward elimination then trained the optimized model on the trauma cohort. We then validated on the two datasets with sepsis (abdominal and respiratory source). We also tested internal validation on the trauma dataset using k-fold cross-validation. Model performance was assessed using area under the receiver operator curve (33,34).

Results

Overlapping biological processes in critical care patients.

Gene co-expression analysis, performed separately on each disease etiology with respective controls included, identified multiple expression modules which predominantly enrich in processes relating to inflammation and cell division (Figure 1).

Figure 1: Shared transcriptome response across aetiologies of critical illness. Dendrogram summaries of gene co-expression networks derived from WBT in Trauma (A), Abdominal sepsis (B) and Community-acquired pneumonia (C). Color-coded bar graph under each dendrogram depicts co-expression modules, with grey representing unclassified genes. Only modules significantly different between disease and healthy controls are labeled (Mann Whitney $p < 0.05$). Panel D shows heatmap summaries of ontological similarities between modules in sepsis vs trauma. Note the overlap for inflammation and immune processes

across different aetiologies of critical illness. Panel D *bottom right*: GO terms enriching in each co-expression module. CAP: community acquired pneumonia, AS: Abdominal sepsis

Processes with the greatest degree of similarity between all disease types in GOSemSim analysis mapped largely to immune/inflammatory processes. This approach yielded a similarity metric greater than 0.75 between all modules which significantly differed between disease and controls.

Distinct patterns of leukocyte compositions preserved across different aetiologies of critically illness

Motivated by the overlap in immunological and inflammatory processes we analyzed the immune cell composition from WBT of all 317 patients with critical illness. We applied the deconvolution technique CIBERSORT to compute relative fractions of 22 leukocyte subtypes from whole blood transcriptomics. This technique uses cell-specific marker transcripts to “deconvolve” lymphocyte cell fractions from bulk gene expression (12). Such deconvolution techniques are well established and have reshaped our understanding of the cell composition of “bulk tissue”. They have been applied in a number of studies to estimate not only leukocyte composition but also the immune landscape (18), microenvironment (19), and purity (35) of tumours. Deconvolution of leukocyte fractions showed complex but consistent patterns across trauma and sepsis caused by community acquired pneumonia or abdominal source (Figure 2A-C).

Figure 2: Shared immune cell configurations across aetiologies of critical illness. A-C: Heat map summaries of leukocyte composition in trauma (A), sepsis due to community acquired pneumonia (B) and sepsis due to an abdominal source (C). Sidebar indicates cluster assignment: black-subgroup 1, red-subgroup 2. Patients are presented in rows and cell types in columns. D-F: Bargraph summaries of differential leukocyte fractions showing suppressed neutrophils, elevated dendritic cells and CD4 memory T cells dominating subgroup 1 (top) and elevated neutrophils with otherwise suppressed leukocyte levels seen in subgroup 2 (bottom). Note the pattern remains consistent across the three independent aetiologies. Red bars indicate significant differences (t-test, $p < 0.05$).

The broad characteristics of these distinct groups were as follows: subgroup 1: a higher fraction of neutrophils with relative suppression of leukocyte fractions, subgroup 2: a higher proportion of activated CD4 memory cells associated with resting/activated dendritic cells. We will be referring to these as subgroup 1 and subgroup 2 from here onwards. Comparative analysis of cell fraction for the two consistent subgroups is summarized in Figure 2D-F. Comparative analysis of patient subgroups did not reveal any differences in the clinical/demographics labels (age, gender or mortality) available in the data repositories (Table 1). We have further validated the distinct leukocyte composition of these two subgroups by training a logistic regression model on the trauma cohort, with subgroup as a

binary response variable and cell fractions as regressors, to independently validate the model on the septic cohorts. Area under the ROC curve was over 0.98 for all groups, demonstrating that the two subgroups were highly distinct from one another and consistent in all 3 cohorts (Supplementary figure 1). For further verification, we demonstrated the generalizability of our results in a robust pediatrics dataset consisting of 181 patients with admitting diagnosis of sepsis from a variety of sources. Cluster detection using the same method as the discovery cohorts gave subgroup with similar characteristics: one group with CD4 helper/dendritic cell dominant (corresponding to subgroup 1) and a second neutrophil dominant corresponding to subgroup 2 (Supplementary figure 2).

Suppressed immune and inflammatory mechanisms in distinct subgroup of critically ill

We finally analysed correlation between the patient subgroups and biological mechanisms in critical illness (Figure 3). This was done by correlating module meta genes (Figure 1) with group labels of individual patients (healthy, subgroup 1, subgroup 2). We found that modules that enrich in inflammatory/immune mechanisms were of highest significance. The expression of these modules (meta-genes) were higher in the subgroup 2 (neutrophil-dominated) and lower in subgroup 1 (CD4 T helpers/dendritic cell dominated). Analysis of our validation cohort paralleled these results: module meta genes from the pediatric cohort suffering from septic shock of diverse aetiologies correlated well with subgroup labels: subgroup 2 showed significantly higher correlation with innate immune mechanisms compared to subgroup 1 (Supplementary figure 3).

Figure 3: Boxplot summary of module meta-gene correlation with patient subgroup phenotypes: H: healthy, subgroup 1: CD4 T helper, dendritic cell predominant and neutrophil depleted, subgroup 2: Neutrophil predominant. Note the lower relative expression level of module eigengenes for subgroup 1 versus subgroup 2. All modules plotted as significantly different between subgroups 1 and 2 (MW $p < 0.05$).

Discussion

Our results are the first to show two groups of patients characterized by different patterns of immune response to severe community acquired pneumonia, abdominal sepsis and trauma. One consistent patient subgroup consists depleted neutrophils, high memory CD4 T and dendritic cells counts (subgroup 1) and a second is characterized by high neutrophil counts (subgroup 2). Our results are the first to show that patterns of leukocyte counts (subgroup 2). Our results are the first to show that patterns of leukocyte composition are preserved across these different critical illnesses and may give basis for a “pan-disease” marker for patients with a complicated clinical course. In addition, our study confirms the consistent but diverse activation of host immune mechanisms in both sepsis and trauma (1, 2–4). (1).

Analysis of serum transcriptomics in critical illness has demonstrated a large scale reprioritization of gene expression, also referred to as a “genomic storm”, following burns

and trauma (2). Part of this effect is due to the up-regulation of innate immune mechanisms, microbial recognition and inflammation, and down-regulation of antigen presentation and adaptive immunity. Complicated recovery (ie: prolonged hospital stay, no recovery or death) was associated with upregulated cytokines IL-6 and IL-10, transcription factors p38, and MAPK signalling cascades, while antigen presentation and T-cell regulation is suppressed (2). The current models of the immune/inflammatory paradigm suggest two consecutive mechanisms for the initial surge of inflammatory processes (Systemic Inflammatory Response Syndrome) followed by a compensatory anti-inflammatory mechanism. These recent genome wide results have suggested a new model consisting of the parallel upregulation of innate immunity with the suppression of adaptive immunity (2). In a longitudinal analysis of gene expression levels and cell fractions in the hyperacute phase of trauma showed a relatively limited upregulation of transcriptomics (3). Only 4.2%, 21.4% and 21.0% of total genes were differentially expressed in samples taken within 2 hours of injury, 24 and 72 hours of admission, respectively. Biological processes associated with Multi Organ Dysfunction Syndrome (MODS) were cell death and survival pathways rather than a pro-inflammatory state suggested by prior concepts. We used co-expression networks to detect gene modules in sepsis and trauma patients, yielding processes enriching in inflammation, innate immunity and cell division. These results are in line with previous findings wherein comparison of transcriptomic responses to fecal peritonitis and community acquired pneumonia showed 64% overlap in upregulated genes with trauma (6).

Multiple configurations in leukocyte fractions have been suggested to associate with increased morbidity/mortality in trauma and sepsis (36,37). These configurations were proposed to contribute to the dysregulated immune response and immune suppressed state seen in critical illness. Studies using flow cytometry have shown higher risk of MODS

with leukopenia, increased NK dim cells and reduced $\delta\gamma$ low T lymphocytes in septic patients (36). Analysis of activation markers CD11b and HLE for neutrophils, and ICAM-1 for monocytes showed reduced signals for patients with higher mortality in sepsis (37). Neutrophils have a central role in triggering and amplifying innate immunity. One of the subgroups identified in our study is dominated by a high fraction of neutrophils while other leukocyte fractions are suppressed. We were, however, limited to assessing cell fractions and lack information on the biological activity of these fractions. The second subgroup is characterized by elevated dendritic cells and memory CD4 T fractions. Memory CD4 T cells are derived from activated CD4 T cells (38); while their role is not well defined, they are proposed as early effectors of cytokine response, helpers of B/T cell response and effector responses as reviewed in ref 36 (39). However, it has been suggested that the apoptotic response in sepsis causes the recovery of an antigen specific cell pool with narrower repertoire, translating to a dysregulated immune response (40). Dendritic cells are regarded as the link between innate and adaptive immunity as they stimulate the adaptive immune system through their antigen presenting and secretory functions (41). In response to microbial stimuli, dendritic cells express co-stimulants such as CD40 and CD86 through a process termed maturation (42). The depletion of dendritic cells was shown to associate with increased frequency of ICU acquired infection for patients with septic shock (7) and in burns (43). Beyond just down-regulation, dendritic cells were proposed to contribute to immune-suppressed states in sepsis through alterations in their differentiation (44), phenotypical changes, and immunological functions (45). Furthermore, an important dichotomy in dendritic cell function is the secretion of IL-12 versus IL-10. IL-12 polarizes naïve Th to Th 1 promoting immune response by linking innate to adaptive immunity (46). Conversely, IL-10 promotes regulatory T cells and suppresses Th 1 priming, NKs and

macrophages, hence promoting an immune-suppressed state (47). Given that our analysis is limited to cell fractions and lacks data on the functional state of these cells it is difficult to draw conclusions on the immune state of these subgroups. We were nevertheless able to correlate gene expression patterns that annotated to immunity and inflammation with the two subgroups in leukocyte configuration. We found downregulation of both immunity and inflammatory mechanisms in the first subgroup of depleted neutrophils, increased dendritic and CD4 helper cell fractions implying an immunosuppressed state. Supporting this interpretation are prior studies that correlate pure fractions of leukocytes to clinical course and outcome. Deconvolution of leukocyte fractions in trauma showed increased risk of Multi Organ Dysfunction Syndrome (MODS) with elevated dendritic cell counts but depleted neutrophil levels (3) which corresponds to the configuration of subgroup 1. Given the limitation of our dataset further studies are required to confirm the association of subgroup 2 with MODS in critical illness.

Conclusion

Our study identifies two patient consistent subgroups in diverse forms of critical illness, distinguished based on their leukocyte composition. This preserved configuration may serve as a universal marker of multiple organ dysfunction for the critically ill, and allow for more targeted approaches to outcome prediction and treatment.

References

1. Davenport EE, Burnham KL, Radhakrishnan J, Humburg P, Hutton P, Mills TC, Rautanen A, Gordon AC, Garrard C, Hill AVS, et al. Genomic landscape of the individual host response and outcomes in sepsis: A prospective cohort study. *Lancet Respir Med* (2016) **4**:259–271. doi:10.1016/S2213-2600(16)00046-1
2. Xiao W, Mindrinos MN, Seok J, Cuschieri J, Cuenca AG, Gao H, Hayden DL, Hennessy L, Moore EE, Minei JP, et al. A genomic storm in critically injured humans. (2011) **208**:2581–2590. doi:10.1084/jem.20111354
3. Cabrera CP, Manson J, Shepherd JM, Torrance HD, Watson D, Longhi MP, Hoti M, Patel MB, Dwyer MO, Nourshargh S, et al. Signatures of inflammation and impending multiple organ dysfunction in the hyperacute phase of trauma: A prospective cohort study. (2017)1–21. doi:10.1371/journal.pmed.1002352
4. Wang L, Oh WK, Zhu J. Disease-specific classification using deconvoluted whole blood gene expression. *Sci Rep* (2016) **6**:1–13. doi:10.1038/srep32976
5. Warren HS, Elson CM, Hayden DL, Schoenfeld DA, Cobb JP, Maier R V, Moldawer LL, Moore EE, Harbrecht BG, Pelak K, et al. A Genomic Score Prognostic of Outcome in Trauma Patients. (2009) **5**:3–10. doi:10.2119/molmed.2009.00027
6. Burnham KL, Davenport EE, Radhakrishnan J, Humburg P, Gordon AC, Hutton P, Svoren-jabalera E, Garrard C, Hill AVS, Hinds CJ, et al. Shared and Distinct Aspects of the Sepsis Transcriptomic Response to Fecal Peritonitis and Pneumonia. (2017) **196**:328–339. doi:10.1164/rccm.201608-1685OC
7. Grimaldi D, Louis S, Pène F, Sirgo G, Rousseau C, Claessens YE, Vimeux L, Cariou A, Mira JP, Hosmalin A, et al. Profound and persistent decrease of circulating dendritic cells is associated with ICU-acquired infection in patients with septic shock. *Intensive*

- Care Med* (2011) **37**:1438–1446. doi:10.1007/s00134-011-2306-1
8. Obermoser G, Presnell S, Domico K, Xu H, Wang Y, Anguiano E, Thompson-Snipes LA, Ranganathan R, Zeitner B, Bjork A, et al. Systems scale interactive exploration reveals quantitative and qualitative differences in response to influenza and pneumococcal vaccines. *Immunity* (2013) doi:10.1016/j.immuni.2012.12.008
9. Peters MJ, Joeheanes R, Pilling LC, Schurmann C, Conneely KN, Powell J, Reinmaa E, Sutphin GL, Zhernakova A, Schramm K, et al. The transcriptional landscape of age in human peripheral blood. *Nat Commun* (2015) **6**: doi:10.1038/ncomms9570
10. Langfelder P, Horvath S. WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics* (2008) **9**: doi:10.1186/1471-2105-9-559
11. Ghazalpour A, Doss S, Zhang B, Wang S, Plaisier C, Castellanos R, Brozell A, Schadt EE, Drake TA, Lusis AJ, et al. Integrating genetic and network analysis to characterize genes related to mouse weight. *PLoS Genet* (2006) **2**:1182–1192. doi:10.1371/journal.pgen.0020130
12. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M, Alizadeh AA. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods* (2015) **12**:453–457. doi:10.1038/nmeth.3337
13. Chandran V, Coppola G, Nawabi H, Omura T, Versano R, Huebner EA, Zhang A, Costigan M, Yekkirala A, Barrett L, et al. A Systems-Level Analysis of the Peripheral Nerve Intrinsic Axonal Growth Program. *Neuron* (2016) **89**:956–970. doi:10.1016/j.neuron.2016.01.034
14. Parikshak NN, Swarup V, Belgard TG, Irimia M, Ramaswami G, Gandal MJ, Hartl C, Leppa V, Ubieta L, Huang J, et al. Genome-wide changes in lncRNA, splicing, and regional gene expression patterns in autism. *Nature* (2016) **540**:423–427.

doi:10.1038/nature20612

15. Horvath S, Dong J. Geometric interpretation of gene coexpression network analysis. *PLoS Comput Biol* (2008) **4**:24–26. doi:10.1371/journal.pcbi.1000117
16. Langfelder P, Cattle JP, Chatzopoulou D, Wang N, Gao F, Al-Ramahi I, Lu XH, Ramos EM, El-Zein K, Zhao Y, et al. Integrated genomics and proteomics define huntingtin CAG length-dependent networks in mice. *Nat Neurosci* (2016) **19**:623–633.
doi:10.1038/nn.4256
17. Horvath S, Zhang B, Carlson M, Lu K V., Zhu S, Felciano RM, Laurance MF, Zhao W, Qi S, Chen Z, et al. Analysis of oncogenic signaling networks in glioblastoma identifies ASPM as a molecular target. *Proc Natl Acad Sci* (2006) **103**:17402–17407.
doi:10.1073/pnas.0608396103
18. Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Ou Yang TH, Porta-Pardo E, Gao GF, Plaisier CL, Eddy JA, et al. The Immune Landscape of Cancer. *Immunity* (2018) **48**:812–830.e14. doi:10.1016/j.immuni.2018.03.023
19. Q. W, B. H, X. H, H. K, M. S, L. S, A.C. deCarvalho, S. L, P. L, Y. L, et al. Tumor Evolution of Glioma-Intrinsic Gene Expression Subtypes Associates with Immunological Changes in the Microenvironment. *Cancer Cell* (2017) **32**:42–56.e6.
doi:10.1016/j.ccell.2017.06.003
20. Wang L, Oh WK, Zhu J. Disease-specific classification using deconvoluted whole blood gene expression. *Sci Rep* (2016) **6**:1–13. doi:10.1038/srep32976
21. Scicluna BP, van Vught LA, Zwinderman AH, Wiewel MA, Davenport EE, Burnham KL, Nürnberg P, Schultz MJ, Horn J, Cremer OL, et al. Classification of patients with sepsis according to blood genomic endotype: a prospective cohort study. *Lancet Respir Med* (2017) **5**:816–826. doi:10.1016/S2213-2600(17)30294-1

22. Edgar R. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res* (2002) **30**:207–210. doi:10.1093/nar/30.1.207
23. Scicluna BP, Wiewel MA, van Vught LA, Hoogendijk AJ, Klarenbeek AM, Franitza M, Toliat MR, Nürnberg P, Horn J, Bonten MJ, et al. Molecular Biomarker to Assist in Diagnosing Abdominal Sepsis upon ICU Admission. *Am J Respir Crit Care Med* (2018) **197**:1070–1073. doi:10.1164/rccm.201707-1339LE
24. Scicluna BP, Klein Klouwenberg PMC, Van Vught LA, Wiewel MA, Ong DSY, Zwinderman AH, Franitza M, Toliat MR, Nürnberg P, Hoogendijk AJ, et al. A molecular biomarker to diagnose community-acquired pneumonia on intensive care unit admission. *Am J Respir Crit Care Med* (2015) doi:10.1164/rccm.201502-0355OC
25. Sweeney TE, Shidham A, Wong HR, Khatri P. A comprehensive time-course-based multicohort analysis of sepsis and sterile inflammation reveals a robust diagnostic gene set. *Sci Transl Med* (2015) **7**:1–16. doi:10.1126/scitranslmed.aaa5993
26. Zhang B, Horvath S. A General Framework for Weighted Gene Co-expression Network Analysis. (2005)
27. Methods T, Clustering H, Author D, Langfelder P, Zhang B, Horvath S, Langfelder MP, Gpl L, Contains D, Date RC. Package ‘dynamicTreeCut.’ (2016)1–14.
28. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* (2009) **4**:44–57. doi:10.1038/nprot.2008.211
29. Yu G, Li F, Qin Y, Bo X, Wu Y, Wang S. GOSemSim: An R package for measuring semantic similarity among GO terms and gene products. *Bioinformatics* (2010) **26**:976–978. doi:10.1093/bioinformatics/btq064
30. Tarca AL, Carey VJ, Chen X wen, Romero R, Drăghici S. Machine learning and its

- applications to biology. *PLoS Comput Biol* (2007) doi:10.1371/journal.pcbi.0030116
31. Thorndike RL. Who belongs in the family? *Psychometrika* (1953)
doi:10.1007/BF02289263
32. Ketchen DJ, Shook CL. The Application of Cluster Analysis in Management Strategic Research and Analysis Critique. *Strateg Manag J* (1996)
33. Zador Z, Sperrin M, King AT. Predictors of outcome in traumatic brain injury: New insight using receiver operating curve indices and Bayesian network analysis. *PLoS One* (2016) **11**:1–18. doi:10.1371/journal.pone.0158762
34. Zador Z, Huang W, Sperrin M, Lawton MT. Multivariable and Bayesian Network Analysis of Outcome Predictors in Acute Aneurysmal Subarachnoid Hemorrhage: Review of a Pure Surgical Series in the Postinternational Subarachnoid Aneurysm Trial Era. *Oper Neurosurg* (2017) **0**:1–8. doi:10.1093/ons/opx163
35. Aran D, Sirota M, Butte AJ. Systematic pan-cancer analysis of tumour purity. *Nat Commun* (2015) **6**:1–11. doi:10.1038/ncomms9971
36. Manson J, Cole E, Ath HD De, Vulliamy P, Meier U, Pennington D, Brohi K. Early changes within the lymphocyte population are associated with the development of multiple organ dysfunction syndrome in trauma patients. *Crit Care* (2016)1–10.
doi:10.1186/s13054-016-1341-2
37. Muller Kobold AC, Tulleken JE, Zijlstra JG, Sluiter W, Hermans J, Kallenberg CGM, Cohen Tervaert JW. Leukocyte activation in sepsis: Correlations with disease state and mortality. *Intensive Care Med* (2000) doi:10.1007/s001340051277
38. MacLeod MKL, Kappler JW, Marrack P. Memory CD4 T cells: Generation, reactivation and re-assignment. *Immunology* (2010) **130**:10–15. doi:10.1111/j.1365-2567.2010.03260.x

39. MacLeod MKL, Clambey ET, Kappler JW, Marrack P. CD4 memory T cells: What are they and what can they do? *Semin Immunol* (2009) doi:10.1016/j.smim.2009.02.006
40. Cabrera-Perez J, Condotta SA, Badovinac VP, Griffith TS. Impact of sepsis on CD4 T cell immunity. *J Leukoc Biol* (2014) doi:10.1189/jlb.5MR0114-067R
41. Worbs T, Hammerschmidt SI, Förster R. Dendritic cell migration in health and disease. *Nat Rev Immunol* (2017) **17**:30–48. doi:10.1038/nri.2016.116
42. BANCHEREAU Jacques, M. SR. Dendritic cells and the control of immunity. *Nature* (1998) **392**:245–252.
43. Arpa DN, Amato G, Amelio DL, Napoli B, Pileri D, Cataldo V, Mogavero R, Lombardo C, Conte F. DECREASE OF CIRCULATING DENDRITIC CELLS IN BURN PATIENTS. (2007) **XX**:199–202.
44. Pastille E, Didovic S, Brauckmann D, Rani M, Agrawal H, Schade FU, Zhang Y, Flohe SB. Modulation of Dendritic Cell Differentiation in the Bone Marrow Mediates Sustained Immunosuppression after Polymicrobial Sepsis. *J Immunol* (2011) **186**:977–986. doi:10.4049/jimmunol.1001147
45. Schefold JC, Zuckermann-Becker H, Volk HD, Meisel C. Phenotype changes and impaired function of dendritic cell subsets in patients with sepsis: A prospective observational analysis. *Crit Care* (2009) **13**:1–12. doi:10.1186/cc7969
46. Moser M, Murphy KM. Dendritic cell regulation of TH1-TH2 development. *Nat Immunol* (2000) doi:10.1038/79734
47. De AK, Laudanski K, Miller-Graziano CL. Failure of monocytes of trauma patients to convert to immature dendritic cells is related to preferential macrophage-colony-stimulating factor-driven macrophage differentiation. *J Immunol (Baltimore, Md 1950)* (2003) **170**:6355–6362. doi:10.4049/jimmunol.170.12.6355

Supplementary material

	Endotype	Mean age (SD)	Sex (M:F)	Mortality (%)
TRAUMA	Subgroup 1	33.4 (10.9)	75:37	6/112 (5.4)
	Subgroup 2	35.5 (11.9)	26:20	1/46 (2.2)
ABDOMEN	Subgroup 1	62.2 (13.1)	13:13	4/26 (15.4)
	Subgroup 2	61.5 (10.9)	14:11	4/25 (16.0)
CAP	Subgroup 1	60.6 (16.2)	43:29	15/72 (20.8)
	Subgroup 2	61.8 (17.2)	20:16	8/36 (22.2)

*No parameters have $p < 0.05$ when comparing endotypes of the same cohort (t-test for age and chi square for sex and mortality).

Supplementary table 1: Comparative analysis of patient endotypes distinguished based on immune cell compositions.

Supplementary figure 1: Validation of patient subgroups based on leukocyte composition. Area under (AUC) the receiver operator curve (ROC) show highly accurate classification of patient subgroups (AUC ≥ 0.98). AUC values are show as AUC (95% CI).

A

B

Supplementary figure 2: Verification of patient subgroups; one with increased neutrophils and second group with relatively suppressed neutrophils and dominant dendritic cells, activated CD4 memory cells. A: Heatmap summary of clustering. Black: Subgroup 1, Red: Subgroup 2. B: Barplot representing differential fractions between subgroups 1 and 2. Red bars indicate significant difference between subgroups (MW $p < 0.05$)

Supplementary figure 3: Verification of immune gene modules and meta gene correlations with patient subgroups in paediatric septic shock. Dendrograms with merged dynamic tree cut (A) gave gene co-expression modules with similar annotations as adult cohorts of trauma and sepsis. Panel B: Correlation of meta genes with subgroup memberships (Supplementary figure 2) paralleled results from adult cohort, further suggesting subgroup 2 with higher immune activity versus subgroup 1 corresponding to an immune suppressed state. 0: healthy, 1: group 1 and 2: group 2