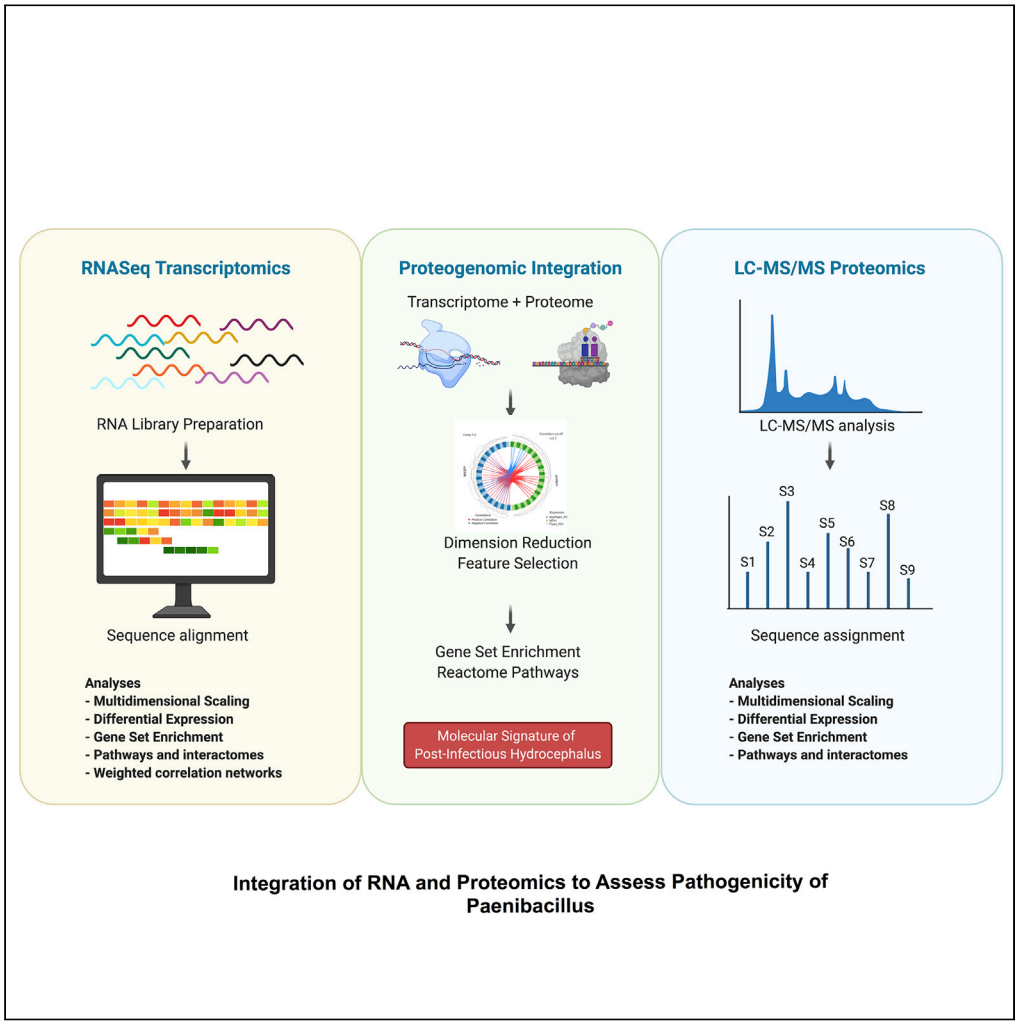
iScience

CellPress
OPEN ACCESS

Article

Immune activation during *Paenibacillus* brain infection in African infants with frequent cytomegalovirus co-infection



Albert M. Isaacs, Sarah U. Morton, Mercedeh Movassagh, ..., David D. Limbrick, Jr., Joseph N. Paulson, Steven J. Schiff

steven.j.schiff@gmail.com

Highlights

There is a characteristic immune response to *Paenibacillus* brain infection

There is a characteristic immune response to CMV brain infection

The matching immune response validates pathogen genomic presence

The combined results support molecular infection causality

Isaacs et al., iScience 24, 102351
April 23, 2021
https://doi.org/10.1016/j.isci.2021.102351

iScience

CellPress
OPEN ACCESS

Article

Immune activation during *Paenibacillus* brain infection in African infants with frequent cytomegalovirus co-infection

Albert M. Isaacs,[1,2,18] Sarah U. Morton,[3,4,18] Mercedeh Movassagh,[5] Qiang Zhang,[6] Christine Hehnly,[7,8] Lijun Zhang,[7] Diego M. Morales,[9] Shamim A. Sinnar,[10,11] Jessica E. Ericson,[12] Edith Mbabazi-Kabachelor,[13] Peter Ssenyonga,[13] Justin Onen,[13] Ronnie Mulondo,[13] Mady Hornig,[14] Benjamin C. Warf,[15] James R. Broach,[7,8,19] R. Reid Townsend,[6,19] David D. Limbrick, Jr.,[9,19] Joseph N. Paulson,[16,19] and Steven J. Schiff[10,17,19,20,*]

SUMMARY

Inflammation during neonatal brain infections leads to significant secondary sequelae such as hydrocephalus, which often follows neonatal sepsis in the developing world. In 100 African hydrocephalic infants we identified the biological pathways that account for this response. The dominant bacterial pathogen was a *Paenibacillus* species, with frequent cytomegalovirus co-infection. A proteogenomic strategy was employed to confirm host immune response to *Paenibacillus* and to define the interplay within the host immune response network. Immune activation emphasized neuroinflammation, oxidative stress reaction, and extracellular matrix organization. The innate immune system response included neutrophil activity, signaling via IL-4, IL-12, IL-13, interferon, and Jak/STAT pathways. Platelet-activating factors and factors involved with microbe recognition such as Class I MHC antigen-presenting complex were also increased. Evidence suggests that dysregulated neuroinflammation propagates inflammatory hydrocephalus, and these pathways are potential targets for adjunctive treatments to reduce the hazards of neuroinflammation and risk of hydrocephalus following neonatal sepsis.

INTRODUCTION

Hydrocephalus is a severe brain disorder in children (Isaacs et al., 2018; Dewan et al., 2018) and the most common indication for pediatric neurosurgery worldwide (Simon et al., 2008). Globally, the most frequent antecedent of hydrocephalus is infection such as neonatal or infant sepsis (Warf, 2005). Postinfectious hydrocephalus (PIH) accounts for up to 60% (Warf and East African Neurosurgical Research, 2010) of the nearly 400,000 infants who develop hydrocephalus each year, principally in low- and middle-income countries (Dewan et al., 2018). Strategies to prevent PIH have been thwarted for two principal reasons. First, the key pathogens responsible for the underlying infections that precede PIH are rarely identified and thus treatment of the underlying infections cannot be optimized (Sinnar and Schiff, 2020). Second, we have not identified the specific underlying inflammatory response that causes the hydrocephalus (McAllister et al., 2015; Karimy et al., 2020). Despite recent clinical efforts to optimize surgical treatment (Kulkarni et al., 2017), the outcomes of treatment for PIH early in life remains disappointing (Sinnar and Schiff, 2020; McAllister et al., 2015; Karimy et al., 2020). Characterizing the inflammatory responses that create PIH is necessary to develop adjunctive strategies that can reduce the likelihood of subsequently developing hydrocephalus.

The underlying pathogenic basis for PIH can be complex. A recent pan-microbial approach uncovered *Paenibacillus spp.* as a PIH pathogen associated with severe ventriculitis in a cohort of Ugandan children (Paulson et al., 2020). However, there were also frequent co-infections with cytomegalovirus (CMV). It is unclear whether the bacteria or virus dominates the inflammatory response or whether they act synergistically. To unravel the immune contributions from multiple pathogens, we performed whole transcriptome and protein mass spectrometry of cerebrospinal fluid (CSF) from 64 infants with PIH and 36 control non-postinfectious hydrocephalus (NPIH) infants without acute inflammation to examine the immunopathogenesis of PIH

[1]Department of Neuroscience, Washington University School of Medicine, St. Louis, MO 63110, USA

[2]Department of Clinical Neurosciences, University of Calgary, Calgary, AB T2N 1N4, Canada

[3]Division of Newborn Medicine, Boston Children's Hospital, Boston, MA 02115, USA

[4]Department of Pediatrics, Harvard Medical School, Boston, MA 02115, USA

[5]Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA 02115, USA

[6]Department of Medicine, Washington University School of Medicine, St. Louis, MO 63110, USA

[7]Institute for Personalized Medicine, Pennsylvania State University, Hershey, PA 17033, USA

[8]Department of Biochemistry and Molecular Biology, Pennsylvania State University, State College, PA 16801, USA

[9]Department of Neurosurgery, Washington University School of Medicine, St. Louis, MO 63110, USA

[10]Center for Neural Engineering, Pennsylvania State University, State College, PA 16801, USA

[11]Department of Medicine, Pennsylvania State University College of Medicine, Hershey, PA 17033, USA

[12]Department of Pediatrics, Pennsylvania State College of Medicine, Hershey, PA 17033, USA

Continued

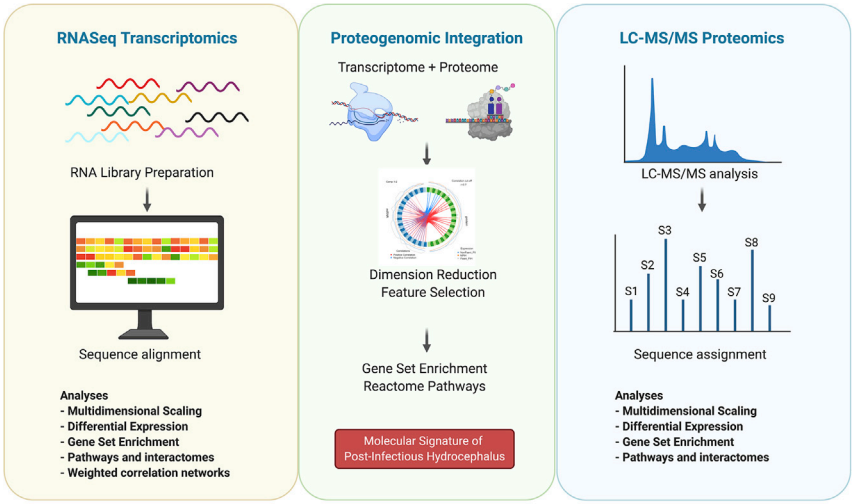


Figure 1. Schematic of proteogenomics experimental and data analysis workflow

Output data from concurrent proteomics and RNA-seq on the same samples were preprocessed, normalized, batch-controlled, and explored independently. Differentially expressed transcriptomic and proteomic data were integrated following dimension reduction and feature selection. Gene ontology enrichment was evaluated and common reactome pathways were visualized to identify prominent molecular pathways implicated in the pathophysiology of postinfectious hydrocephalus. Adapted from "PTMScan Workflow," by BioRender.com (2021). Retrieved from https://app.biorender.com/biorender-templates.

(Paulson et al., 2020; Rohlwink et al., 2019; Morales et al., 2012). Matching deep-scale proteomics (i.e., high gene coverage) with transcriptomics (Nesvizhskii, 2014; Zhang et al., 2014) enabled optimal characterization of the host response (Figure 1).

RESULTS

Patient characteristics

A total of 100 patients (64 PIH, 36 NPIH), all 3 months of age or less and with weight greater than 2.5 kg, were recruited. Infants with PIH had either a history of febrile illness and/or seizures preceding the onset of hydrocephalus or alternative findings such as imaging and endoscopic results indicative of prior ventriculitis (septations, loculations, or deposits of debris within the ventricular system). None of the infants with PIH had a history of hydrocephalus at birth. Patients with NPIH had findings of hydrocephalus that was of non-infectious origin on computed tomography (CT) or at endoscopy, including a structural cause (obstruction of the aqueduct of Sylvius, Dandy-Walker cyst, or other congenital malformation) or evidence of hemorrhage (bloody CSF). Before CSF acquisition, none of the 100 patients had undergone surgery on the nervous system (reservoirs, shunt, third ventriculostomy, or myelomeningocele closure), or had evidence of communication between the nervous system and skin such as meningocele, encephalocele, dermal sinus tract, or fistula.

CSF samples acquired directly from the cerebral ventricles of all patients during surgical treatment of hydrocephalus (shunt insertion or endoscopic treatment) in Uganda were placed into cryotubes with DNA/RNA preservative (DNA/RNA Shield, Zymo Corp). Samples were then frozen in either liquid nitrogen or at −80°C and shipped cryogenically to the United States for further processing. Previous 16S rRNA sequencing had recovered *Paenibacillus spp.* from 59% of the infants with PIH in this study, of which 8 and 27 patients also had human herpesvirus 5 (CMV) in the CSF and blood (Paulson et al., 2020), respectively. These previous findings directed stratification of our patients into "Paeni-positive" or "Paeni-negative," "CMV-CSF-positive" or "CMV-CSF-negative," and "CMV-blood-positive" or "CMV-blood-negative." Comparison of demographic and clinical attributes between Paeni-positive and Paeni-negative patients showed that Paeni-positive patients had higher white blood cell (WBC)

[13]CURE Children's Hospital of Uganda, Mbale, Uganda

[14]Department of Epidemiology, Columbia University Mailman School of Public Health, New York, NY 10032, USA

[15]Department of Neurosurgery, Harvard Medical School, Boston, MA 02115, USA

[16]Department of Biostatistics, Product Development, Genentech Inc., South San Francisco, CA 94080, USA

[17]Center for Infectious Disease Dynamics, Departments of Neurosurgery, Engineering Science and Mechanics, and Physics, The Pennsylvania State University, University Park, PA 16802, USA

[18]These authors contributed equally

[19]Senior authors

[20]Lead contact

*Correspondence Steven J. Schiff, W311 Millennium Science Complex, The Pennsylvania State University, University Park, PA 16802, USA: steven.j.schiff@gmail.com

https://doi.org/10.1016/j.isci.2021.102351

Table 1. Demographics and clinical characteristics of patients based on *Paenibacillus* spp. status

		Paeni-negative n = 62		Paeni-positive n = 38
	All patients n = 100	NPIH n = 36	PIH n = 26	PIH n = 38
Age in days, mean (SD)	57 (24)	43 (27)	75 (11)	59 (17)
Sex				
Male (%)	51 (51)	16 (44)	14 (54)	21 (55)
Female (%)	49 (49)	20 (56)	12 (46)	17 (45)
Positive CMV status				
Blood	27	9	8	10
CSF	8	0	2	6
CSF WBC [1.0 × 10^3]/μL, mean (SD)	30 (62)	5 (0.5)	5 (0.0)	72 (86.0)
Peripheral blood WBC [1.0 × 10^3]/μL, mean (SD)	10.3 (3.6)	8.8 (2.6)	9.8 (3.0)	11.9 (4.1)
Hemoglobin g/dL, mean (SD)	11.5 (2.2)	13.0 (2.8)	10.5 (1.3)	10.8 (1.3)
Hematocrit %, mean (SD)	36.8 (7.4)	41.6 (9.3)	33.6 (4.2)	34.4 (4.1)

CSF, cerebrospinal fluid; CMV, cytomegalovirus; PIH, postinfectious hydrocephalus; NPIH, non-postinfectious hydrocephalus; Paeni positive, CSF *Paenibacillus* spp.-positive; Paeni-negative, CSF *Paenibacillus* spp. negative.

counts in peripheral blood (11.9 versus 9.3 × 10^3/μL, p = 0.002) and CSF (71.6 versus 5.0 × 10^3/μL, p < .001), and lower blood hemoglobin levels (10.8 g/dL versus 12.0 g/dL, p = 0.039) than those of Paeni-negative patients. There were no significant differences in age or gender between the groups (Table 1).

Protein expression

Tandem mass spectrometry (MS/MS) spectra obtained from a standard 10-channel peptide mass tagging (TMT-10) system and compared with the current UniRef database yielded quantitative proteomic data (Mertins et al., 2018; Chen et al., 2012; McAlister et al., 2012; Morales et al., 2012). Median values of peptide intensities were assigned to unique proteins and were used to infer relative protein abundances. Peptide identifications that are non-unique by the principle of parsimony were excluded from protein quantification (Koskinen et al., 2011). Batch effects were corrected, and intensity values were quantile normalized. Multi-dimensional scaling (MDS) analysis demonstrated the clustering of infants with PIH and NPIH and distinguished Paeni-positive from Paeni-negative patients (Figure 2A). Of 616 proteins identified, 292 were differentially expressed based on *Paenibacillus spp.* status: 144 and 148 proteins were up- or downregulated in the Paeni-positive group, respectively (Figure 2B). Gene set enrichment analysis of the differentially expressed proteins identified that the predominant enriched functions were involved in neuroinflammation, particularly neutrophil-mediated inflammation, negative regulation of proteolysis and peptidase activity, and modulation of extracellular matrix and structure (Figure 2C).

RNA expression

Gene read counts (25 million reads per sample) were aggregated from expression levels estimated from paired-end RNA sequencing (RNA-seq) data mapped to the human reference genome hg38 with STAR and quantified with RESM (Table S1) (Dobin et al., 2013; Li and Dewey, 2011; Paulson et al., 2017). Expression of at least 1 count per million in a minimum of 18 samples was required for inclusion in the analysis. MDS analysis demonstrated the clustering of infants with PIH and NPIH (Figure 3A). Of 11,114 genes, 2,161 were differentially expressed, and the hierarchical clustering of differentially expressed genes distinguished Paeni-positive from Paeni-negative patients (Figure 3B). Gene ontology analysis of the differentially expressed genes based on *Paenibacillus spp.* status demonstrated enrichment for genes predominantly related to response to bacteria, host immune regulation, cell motility, migration, and adhesion (Figure 3C).

Differential expression analysis of long non-coding genes

When aligning RNA, we detected 116 long non-coding RNAs (lncRNAs), out of 1,220 lncRNAs annotated in Ensembl. Of the 116, 70 had at least one count (read) in at least 20% of the samples. Ninety samples had at least one lncRNA present. The five most highly expressed lncRNAs were OIP5-AS1, FAM211A-AS1,

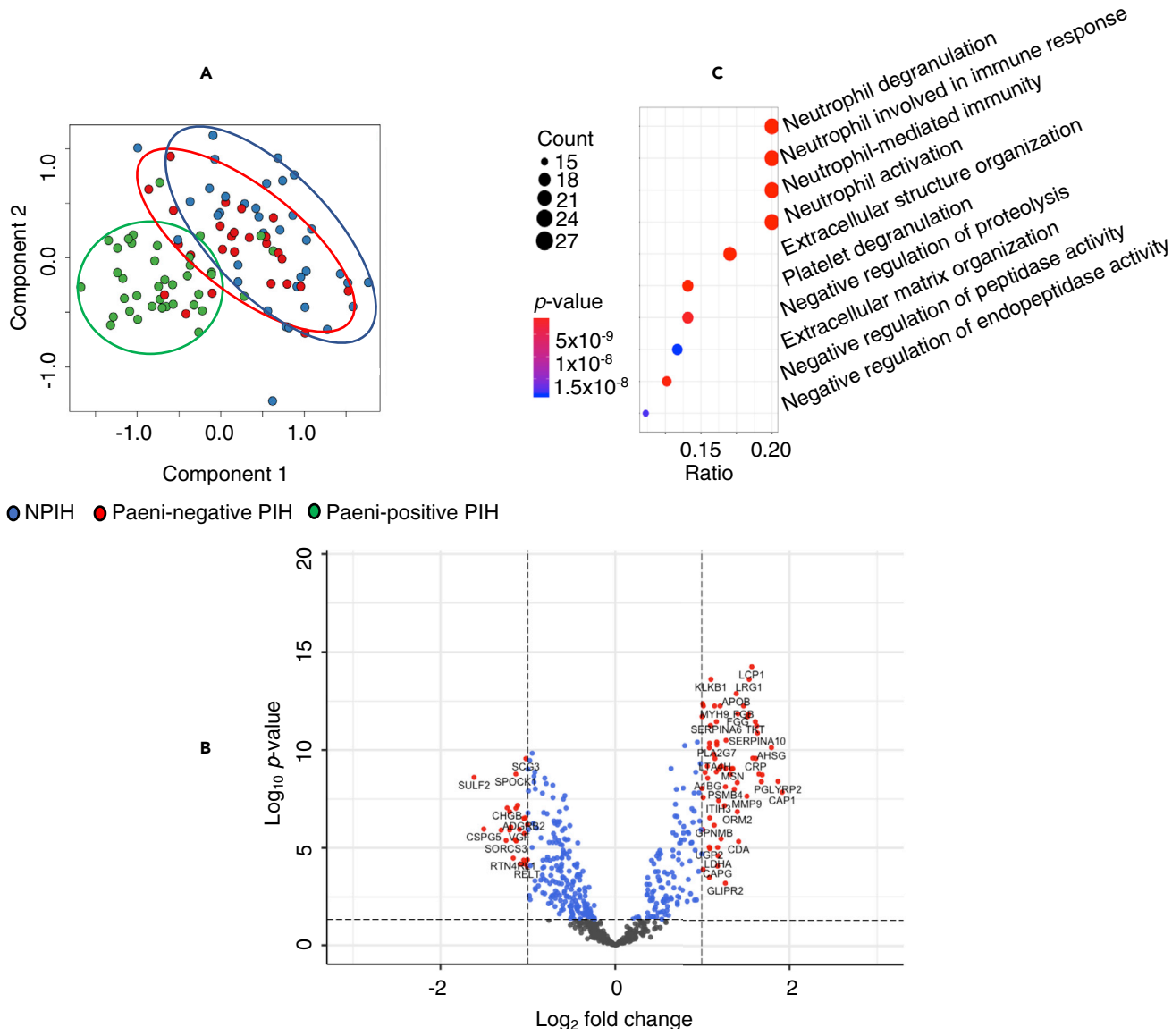


Figure 2. Proteomic profile of infants with postinfectious hydrocephalus (PIH) and those without (NPIH), based on 16s rRNA-determined *Paenibacillus spp.* status

(A) Multidimensional scaling of normalized protein abundances demonstrating clustering of *Paenibacillus* spp.-positive (Paeni-positive) infants from the remaining groups. The abscissa and ordinate of the scatterplot of individual participants represent the first and second components, respectively. Each oval (not drawn to scale) encircles majority of patients belonging to the color-matched group, with blue as infants in the NPIH group, red as Paeni-negative PIH infants, and green as Paeni-positive infants.

(B) Volcano plot demonstrating differential expression of genes between Paeni-positive and Paeni-negative infants. Adjustment for cytomegalovirus status did not change the differential expressions between groups. The vertical lines crossing the positive and negative abscissae demarcate fold changes of 1 and -1, respectively, and the horizontal dashed line crosses the ordinate at the alpha significance level of 0.05. Each point represents a differentially expressed protein, and those with an absolute fold change greater than 1 that met the significance level (red points) were selected for gene set enrichment analyses. Genes with enrichment below the statistically significant threshold are displayed in gray.

(C) Gene ontology analyses of differentially expressed proteins of infants based on 16s rRNA-determined *Paenibacillus* spp. status. There was enrichment for functions associated with neuroinflammation, extracellular matrix structure, and cell-cell adhesion among Paeni-positive infants compared with Paeni-negative infants. The abscissae (ratio) of the dot plots correspond to the number of proteins per total number of proteins, and the size of each circle reflects the relative number of proteins expressed that are enriched for the corresponding function (ordinates).

C1orf132, LINC00963, and ZNF518A. The most variable lncRNAs included OIP5-AS1, FAM211A-AS1, C1orf132, LINC00963, as well as FAM157C. In performing differential expression, we identified 35 lncRNAs significantly associated with Paeni-positive samples with a threshold of false discovery rate <0.05. In

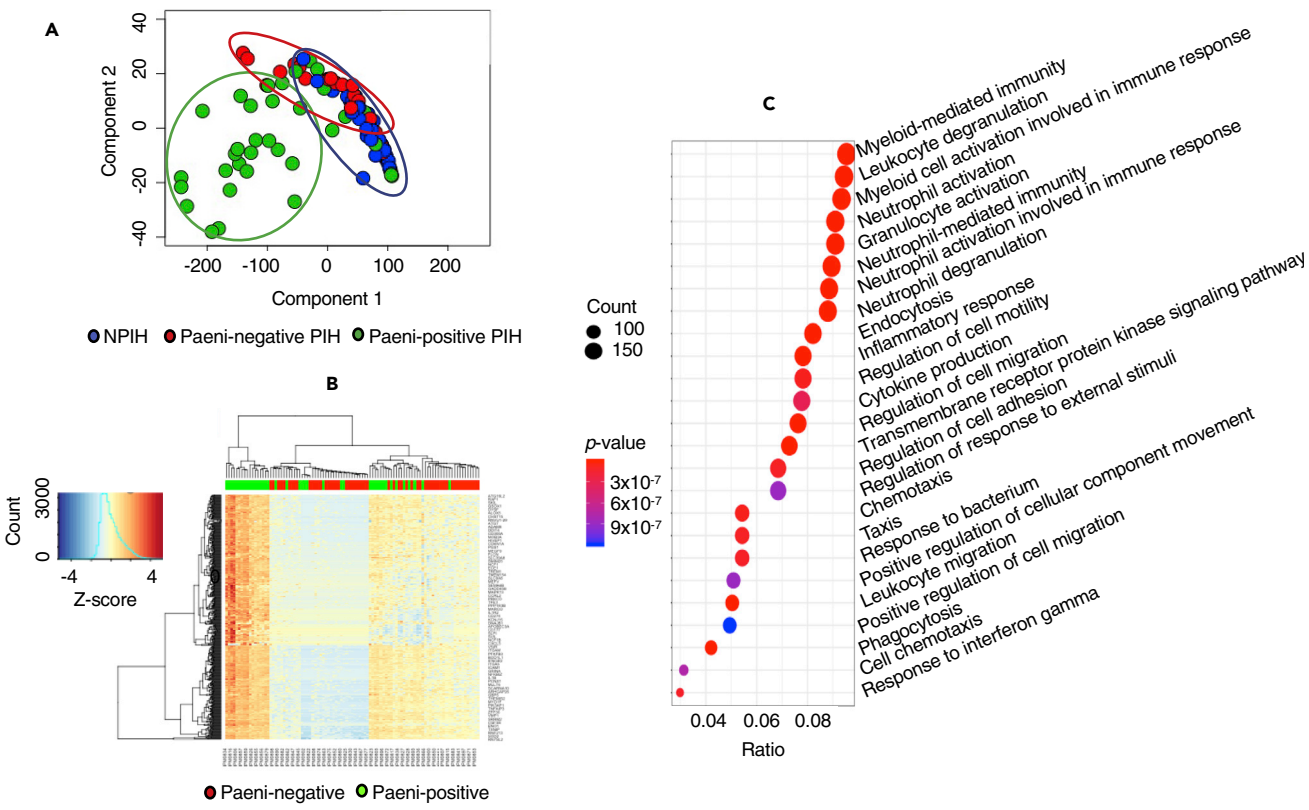


Figure 3. Transcriptomic profile of infants with postinfectious hydrocephalus (PIH) and those without (NPIH), based on 16s rRNA-determined *Paenibacillus* spp. status

(A) Multidimensional scaling of normalized gene expression demonstrating clustering of *Paenibacillus* spp.-positive (Paeni-positive) infants from the remaining groups. The abscissa and ordinate of the scatterplot of individual participants represent the first and second components, respectively. Each oval (not drawn to scale) encircles the majority of patients belonging to the color-matched group, with blue as infants in the NPIH group, red as Paeni-negative PIH infants, and green as Paeni-positive infants.

(B) Heatmap of the 500 most differentially expressed genes between the Paeni-positive and Paeni-negative infants, with the dendrogram demonstrating hierarchical clustering on Euclidean distance of gene identified in infants based on *Paenibacillus* spp. status.

(C) Gene ontology analyses of differentially expressed genes of infants based on 16s rRNA-determined *Paenibacillus* spp. status. There was enrichment for functions associated with neuroinflammation, extracellular matrix structure, and cell-cell adhesion among Paeni-positive infants compared with Paeni-negative infants. The abscissae (ratio) of the dot plots correspond to the number of proteins per total number of proteins, and the size of each circle reflects the relative number of proteins expressed that are enriched for the corresponding function (ordinates).

Paeni-positive samples, the lncRNAs with the largest fold-changes (upregulation) were FAM157C, SPACA6P, LINC00243, LINC01128, and SEMA3B. No lncRNAs were significantly downregulated in Paeni-positive samples. The nature of how these lncRNAs are associated with disease and *Paenibacillus* spp. positivity will be of future interest.

Role of CMV co-infection

Adjusting for CMV status at the transcriptome level did not significantly alter the results of the gene ontology enrichment analysis based on *Paenibacillus spp.* positivity. Sixty-four genes were differentially expressed based on CMV status, and those genes were enriched for functions related to host response to virus (Figure 4A). Principal-component analysis (PCA) using RNA abundance of all genes was not able to cluster samples by CMV status (Figure 4B). Adjusting for CMV status on proteomic data did not change the differentially abundant proteins (Figure 4C).

Network correlation analysis was employed to identify sets of genes that share a pattern of expression among patient RNA data. First, a subset of patients was identified using hierarchical clustering based on expression of genes that were differentially expressed between Paeni-positive and Paeni-negative

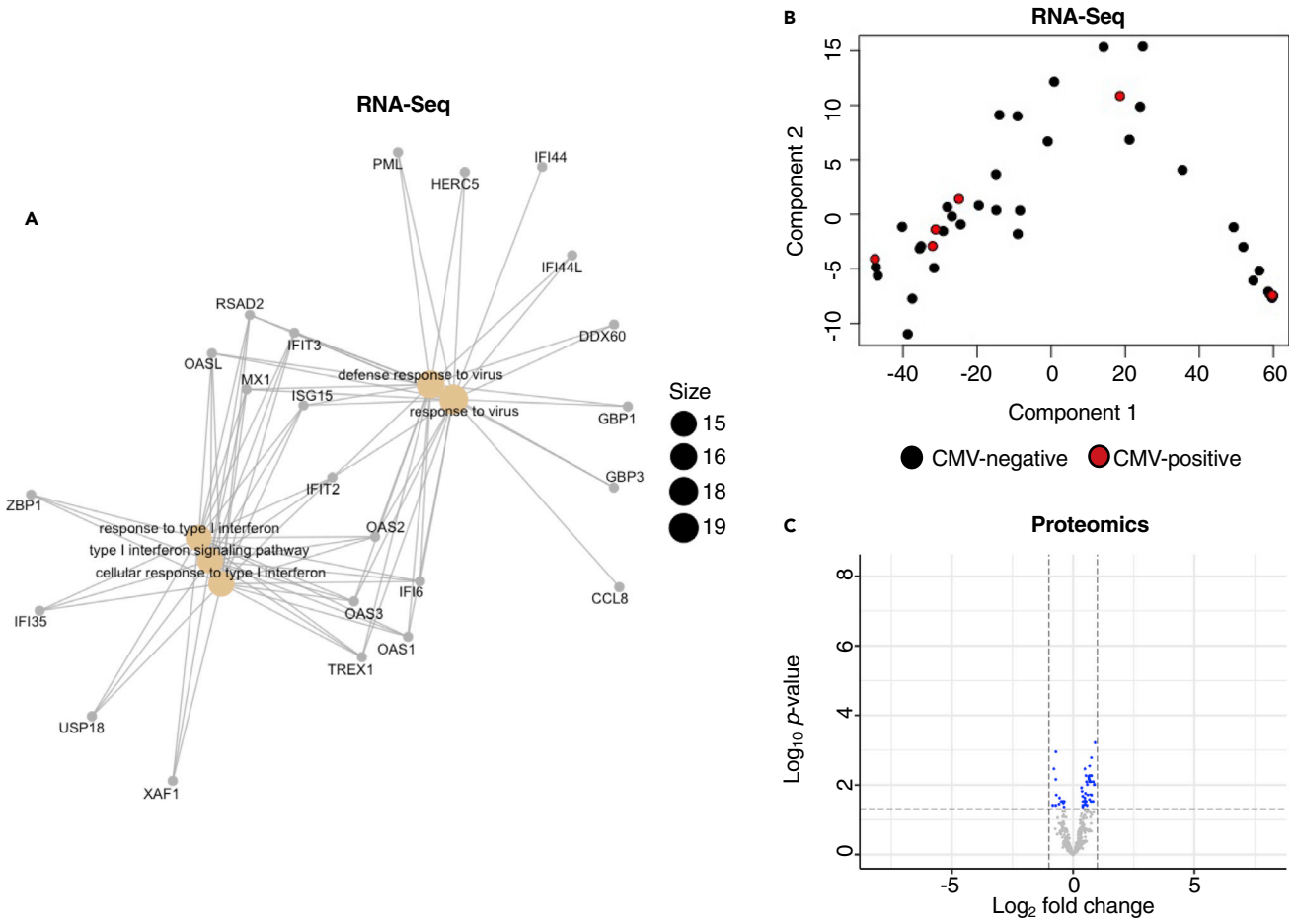


Figure 4. Role of CMV co-infection

(A) An interactome of the 64 genes that were differentially expressed based on CMV status, enriched for functions related to host response to virus.

(B) Principal-component analysis plot demonstrating that using RNA abundance of all genes was not able to cluster samples by CMV status. The abscissa and ordinate of the scatterplot of individual participants represent the first and second components, respectively.

(C) Volcano plot demonstrating there are no differentially expressed proteins based on CMV status alone (blue points) that had a fold change greater than 1 or met statistical significance criteria. Genes with enrichment below the statistically significant threshold are displayed in gray. Differential expression of proteins based on *Paenibacillus* spp. status was similar with or without adjusting for CMV status (Figure 2).

patients. Then, using Weighted Correlation Network Analysis (WGCNA), two gene sets (Modules 1 and 2) were identified as having similar patterns of expression within this subset of 33 patients (Table S2). These modules were assessed for network correlation with four clinical variables: CSF and peripheral blood WBC counts, *Paenibacillus spp.* status, and CSF CMV status (Figure 5A). Module 1 contained all but 27 of the 2,205 genes differentially expressed with respect to *Paenibacillus spp.* status. This module correlated positively with CSF cell count and negatively with *Paenibacillus spp.* status. Genes in Module 1 were enriched for functions related to host immune response (Figure 5B). Module 2 was positively correlated with *Paenibacillus spp.* status, but the 27 genes included in this module were not enriched for any gene ontologies, indicating no specific functional enrichment for the small number of genes within this module.

Marker genes of bulk RNA sequencing

Patient bulk RNA expression was analyzed for immune cell-type signatures using reference single-cell RNA sequencing data (Figure 5C). The *R* package xCell identifies cell type proportions within bulk RNA sequencing data based on enrichment of marker gene expression (Aran et al., 2017). Upon hierarchical clustering, two groups were evident: a mixture of PIH and NPIH samples with a predominance of hematologic cells and PIH-only samples with T helper and natural killer (NK) cell populations.

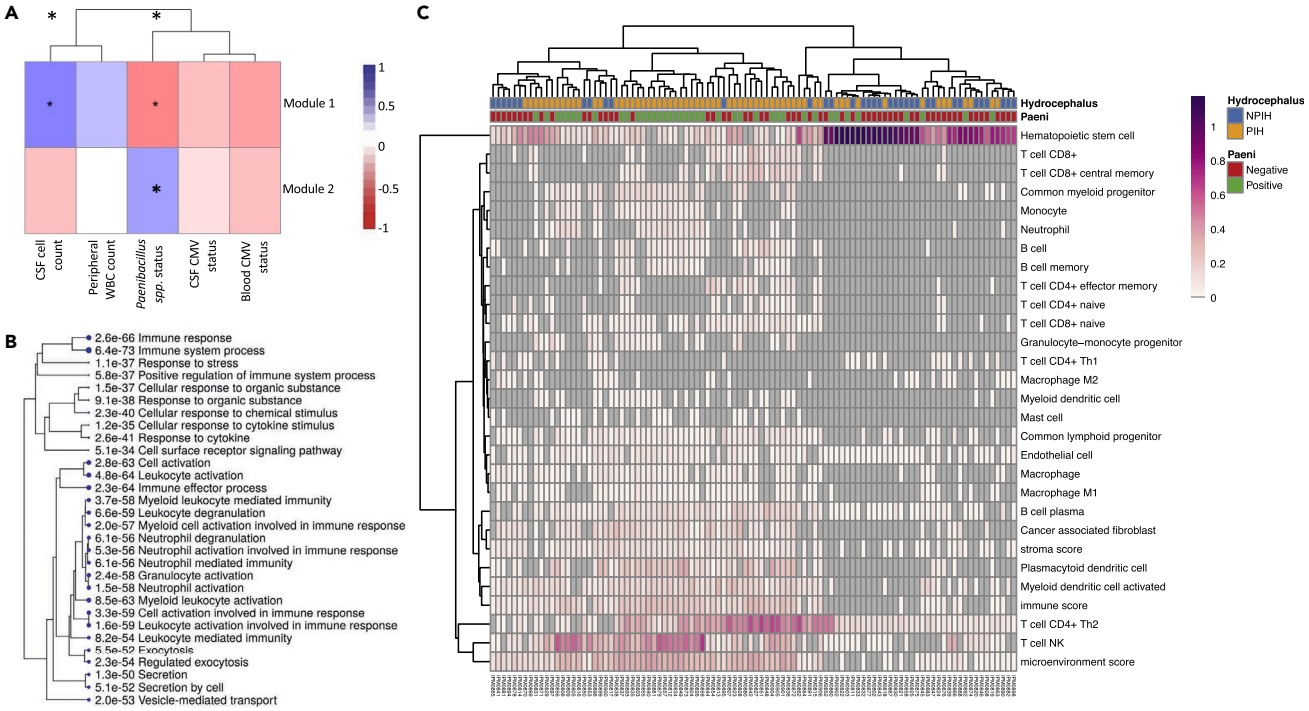


Figure 5. Weighted correlation network analysis (WGCNA) and single-cell RNA deconvolution

(A) Gene modules within RNA-seq data for PIH and NPIH cohorts identified by WGCNA. Module 1 positively correlated with CSF cell count and negatively with *Paenibacillus* spp. status and was enriched for host immune responses, including leukocyte and neutrophil functions. (B) Module 2 was positively correlated with *Paenibacillus* spp. status, but the associated genes had no specific functional enrichment.

(C) Deconvolution of bulk RNA data into immune cell populations expressed on a scale of 0 to 1 that is calculated for each patient. There was hierarchical clustering of a mixture of NPIH and PIH samples with a hematologic predominance, and PIH-only samples with T helper and NK cell populations.

Proteogenomic integration

Integration of proteomics and transcriptomics data was performed following dimension reduction and feature selection of differentially expressed proteins and genes. Unsupervised gene ontology analyses of the concatenated data recapitulated enrichment for functions related to the immune system, metabolism, response to oxidative stress, cell-cell junction interactions, and extracellular matrix structure in association with peptidase activity (Figure 6A).

There were 33 genes detected by both proteomics and RNA-seq as significant; infants with Paeni-positive PIH had consistently differing counts of those markers than did the Paeni-negative infants and infants with NPIH (Figure 6B). Pathway enrichment analysis of those genes demonstrated a predominance of functions associated with the immune system, particularly those involved with interleukin (IL)-4, IL-12, IL-13, interferon, and neutrophil activity, as well as those relating to Janus kinase/signal transducers and activators of transcription (Jak/STAT) pathway. In addition, there was enrichment for processes involving response to platelet-activating factors and host recognition of microbes including antigen presentation and Class I major histocompatibility complex (MHC) antigen processing (Figure 7, Table S3).

DISCUSSION

Disease pathogenesis involves the combinatorial interaction of the proteome, transcriptome, and environment (Rédei, 2008; Heintzman and Ren, 2006). By matching deep-scale proteomics (i.e., high gene coverage) with transcriptomics (Nesvizhskii, 2014; Zhang et al., 2014), the interplay within the host immune response can be optimally reduced into network models representing disease (Ruggles et al., 2017).

In this study, genes with differential mRNA or protein abundance among patients with PIH were enriched for functions related to neuroinflammation, reaction to oxidative stress, cell-cell junction structure, and

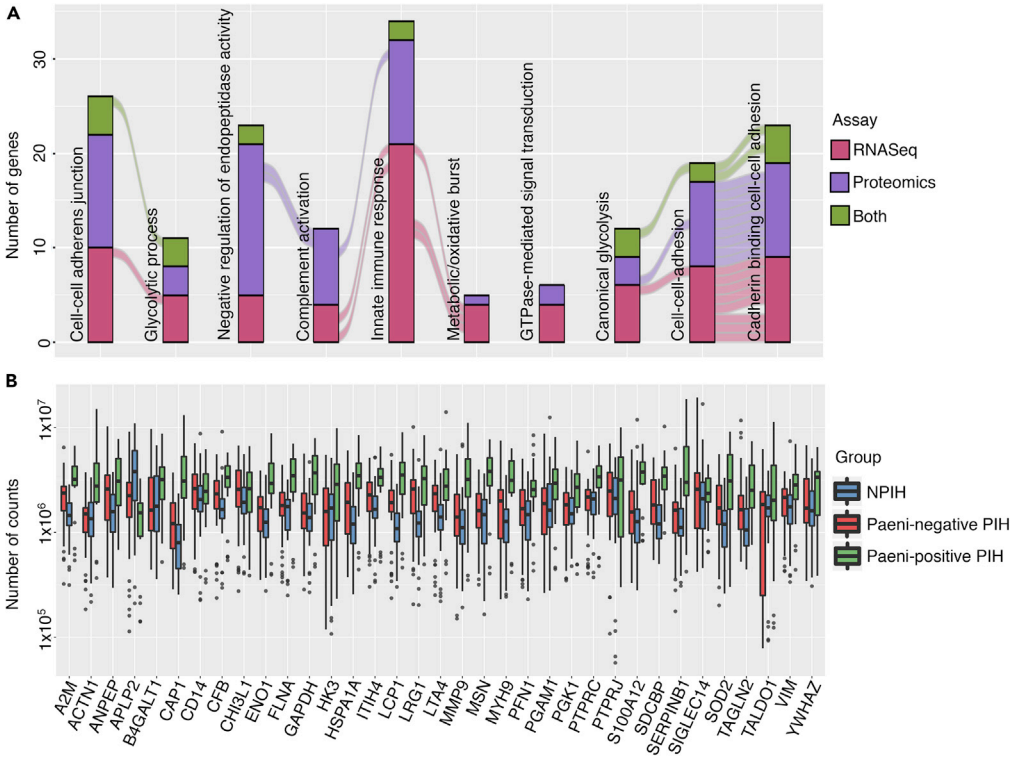


Figure 6. Proteogenomic integration of proteins and genes expressed in RNA-seq and/or proteomics among infants with postinfectious and non-postinfectious hydrocephalus stratified by cerebrospinal fluid 16s rRNA *Paenibacillus* spp. status

(A) Alluvial plot demonstrating the most prominent gene ontological (GO) functions and interactions for PIH pathophysiology between *Paenibacillus* spp.-positive and *Paenibacillus* spp.-negative infants. Each ontological clustering occupies a column in the diagram and is horizontally connected to preceding and succeeding significance clustering by stream fields, representing similar gene involvement. Each stacked bar is color-coded based on the assay being assessed, with red representing RNA-seq data, purple for proteomics, and green for genes common to both RNA-seq and proteomics. The ordinate shows the number of genes represented in each cluster.

(B) Box and whisker plot of the 33 genes that were differentially expressed based on *Paenibacillus* spp. status. Counts (ordinate) of each gene (abscissa) are shown for each group, with blue representing NPIH, red for Paeni-negative PIH, and green for Paeni-positive PIH infants.

extracellular matrix organization. Combining proteomics and RNA-seq results narrowed the spectrum of responses to the innate immune system, including neutrophil activity and signaling via IL-4, IL-12, IL-13, interferon, and Jak/STAT pathways, in addition to platelet-activating factors. Furthermore, there was enrichment for factors involved with microbe recognition such as Class I MHC antigen-presenting complex. Our findings are consistent with a major role for up-regulation of genes and proteins associated with neutrophil activation. Nevertheless, in addition to the direct role in fighting the infection, we note that there appears to be substantial concomitant parenchymal loss in these patients, and we cannot rule out innate-immunity-mediated tissue damage caused by this immune response. One of the difficulties of ascribing infection causality to the presence of pathogen genomes is the ubiquitous problem of environmental and reagent contamination in low-biomass samples, the sensitivity of sequencing techniques, the potential artifacts and biases contributed by sequence amplification when investigating bacterial 16S, as well as the presence of microbes within nominally sterile body spaces that may not be causing active infection. Our findings demonstrated a patterned innate immune response associated with the presence of *Paenibacillus* 16S DNA that strengthens the causal association of *Paenibacillus* with infection. In light of the recent confluence of evidence suggesting that dysregulated neuroinflammation propagates inflammatory hydrocephalus (including PIH and post-hemorrhagic, PHH post-hemorrhagic) (Karimy et al., 2020; Warf, 2005), the pathways we identified are potential targets for adjunctive treatments to reduce the hazards of neuroinflammation and risk of hydrocephalus following neonatal sepsis.

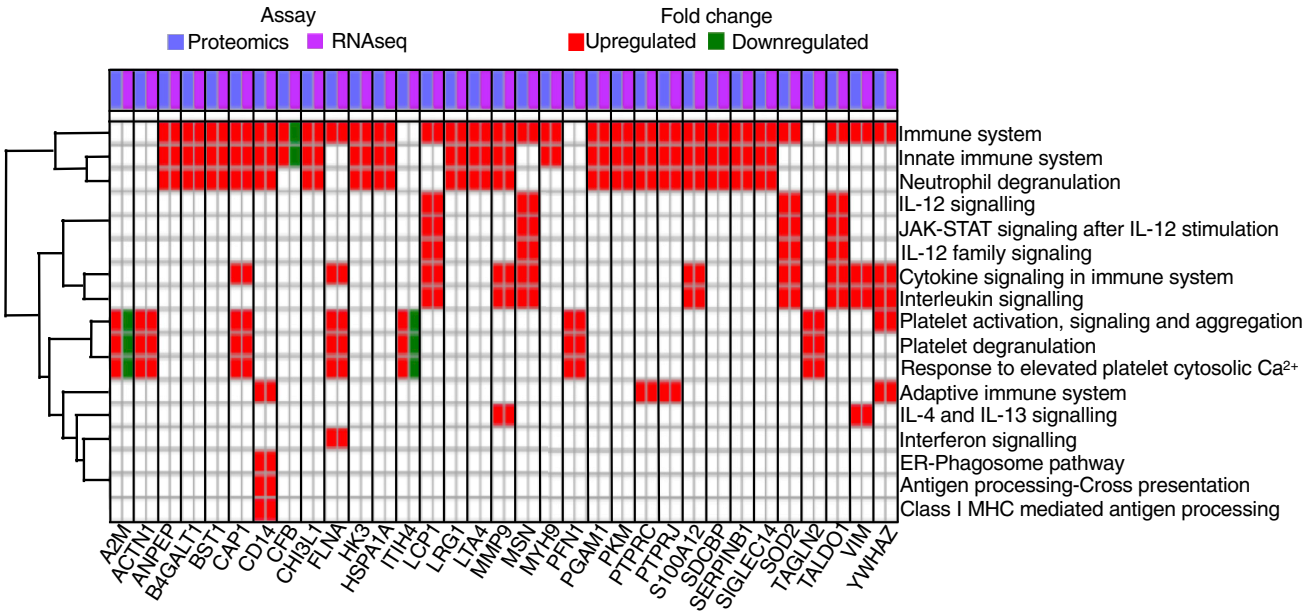


Figure 7. Pathway analysis of 33 genes that were differentially expressed in both RNA-seq and proteomics among infants with postinfectious and non-postinfectious hydrocephalus stratified by cerebrospinal fluid 16s rRNA *Paenibacillus* spp. status

Corresponding pairs of upregulated (red boxes) and downregulated (green boxes) proteins (blue columns) and genes (purple columns) in the *Paenibacillus* spp.-positive group that were identified with proteomics and RNA-seq, respectively, are listed on the abscissa. The 33 common genes demonstrated predominant involvement of the immune system, particularly the innate system and those associated with neutrophil-mediated activity, interleukins, interferon, and the Janus kinase/signal transducers and activators of transcription (JAK-STAT) pathway (ordinate). Differential expression was defined as $\log_2$ fold change of >1 or < -1 at an alpha significance level of 0.05.

Host immune responses in PIH

In addition to ventriculomegaly (enlargement of the cerebral ventricles), the PIH phenotype is characterized by CSF fluid loculations, debris within fluid spaces, ectopic calcification, and brain abscesses (Ciurea et al., 2005), suggesting that severe inflammation occurs locally during the antecedent neonatal sepsis. Typically, once a pathogen breaches a host's endothelial and epithelial barriers, immune responses are activated. Although there is a rapid accumulation of immunologic competence after birth, the innate immune system is the primary active defense for infants less than 3 months of age, since they lack the antigenic experience that informs acquired immunity. Consequently, the neonatal host immune response involves antigen-independent immune components such as neutrophils, phagocytes, NK cells, and antigen-presenting cells (Bell, 2003; Kawai and Akira, 2006; Villarino et al., 2017; PrabhuDas et al., 2011; Ygberg and Nilsson, 2012). Our RNA-seq immune cell signature analyses support the role of NK cells in the host response of patients with PIH. The toll-like receptor (TLR) immune system pathway is commonly implicated in post-inflammatory (including infectious or hemorrhagic) hydrocephalus (Karimy et al., 2017, 2020; Sokol et al., 2016; Lattke et al., 2012); however, the cytokine- and interferon-induced Jak/STAT pathway is considered the most efficient form of innate immunity, especially with intracellular pathogens (Villarino et al., 2017; O'Shea et al., 2015). Protein mutations in the Jak/STAT pathway have been implicated in inadequate inflammation mediation (Villarino et al., 2017; O'Shea et al., 2015). For example, impaired Jak function in severe combined immunodeficiency is associated with susceptibility to infections due to the absence of NK, B, or T cells (Noguchi et al., 1993, 2008). At the protein and transcript levels, we observed differential expression of factors associated with interleukins (IL-4, IL-12 and IL-13) and interferon activity, both of which are obligate mediators of the Jak/STAT pathway (Villarino et al., 2017; O'Shea et al., 2015). Upregulation of MHC Class I in our cohort could be an indication of presentation to cytotoxic CD8+ cells and IL-12 activity (Adiko et al., 2015).

Inflammation and barrier integrity in PIH

The activated immune pathways found in PIH support the perspective that the neonatal immune response to infection or hemorrhage leads to ependymal gliosis or denudation, scarring of CSF conduits, and disruption of CSF physiology to cause hydrocephalus. Although the host-immune response is necessary for

fighting infection and clearing microbial pathogens, prolonged or overstimulated immune activation may be detrimental (Tchessalova et al., 2018); indeed, recent evidence from both experimental (Guerra et al., 2015a; Ortloff et al., 2013; Rodriguez et al., 2012; Jimenez et al., 2009; Paez et al., 2007; Batiz et al., 2005) and clinical (Ortega et al., 2016; Guerra et al., 2015a; Sival et al., 2011) studies demonstrates that host immune response associated with ependymal cell-cell junction protein disruption is a critical pathogenetic mechanism of hydrocephalus.

We also observed platelet-activating factors and response to reactive oxidative species. Elevations in platelet-activating factors have been associated with compromise of the blood-brain barrier (Brailoiu et al., 2018), and their signaling pathways are an important link between inflammatory and thrombotic processes, including in sepsis (de Oliveira et al., 2017; Yost et al., 2010; Pun et al., 2009). Reactive oxygen species set off a cascade of cellular excitotoxicity and secondary brain injury that impairs brain oxygenation and perpetuates cerebral vascular dysfunction (Lehner et al., 2011). Such vascular pathology is consistent with our finding of a differential expression of factors involved in platelet activity including PFN1, ITIH4, and A2M.

Potential pathologic overlap with NPIH

Inflammation is a component of other forms of hydrocephalus (Habiyaremye et al., 2017; Karimy et al., 2017), which informed our decision to use noninfectious hydrocephalic infants as controls. In a recent review, Karimy et al. discussed evidence that PIH shares common host immune pathways with PHH (Karimy et al., 2020). In our PIH cohort, we identified upregulation of IL-12 signaling, which is typically involved in augmentation of CD8+ T cell cytotoxicity. PHH is also associated with significantly elevated levels in other cytokines including IL-1, IL-10, CCL-3, and CCL-13 (Habiyaremye et al., 2017). Neuroinflammation has been linked to the ependymal barrier damage found in humans and several experimental models of hydrocephalus (McAllister et al., 2017). Such ependymal damage was shown to result from cleavage of cell adhesion proteins (McAllister et al., 2017; Castaneyra-Ruiz et al., 2018). Indeed, many forms of human congenital hydrocephalus result from genetic alterations of proteins involved in cell-cell junctional integrity including N-cadherin, connexin, and L1CAM (Guerra et al., 2015b), which are critical for the differentiation of the ventricular and subventricular zone neural stem cells into mature ependyma (Jin et al., 2020). In addition, a recent exome sequencing of 381 infants with congenital hydrocephalus identified a predominance of *de novo* mutations in genes associated with neural stem cell differentiation (Jin et al., 2020). Experimentally, a series of studies on *hyh* mice (Jimenez et al., 2009; Batiz et al., 2005, 2006; Perez-Figares et al., 1998; Wagner et al., 2003) that develop perinatal aqueductal stenosis also support the concept of a defect in cell junction complexes as an underlying cause of hydrocephalus. Although our use of NPIH as controls focused on the acute inflammatory responses that lead to PIH, there are linkages to common mechanisms in other forms of infant hydrocephalus.

Role of immunomodulation for PIH prevention

It would be premature to suggest that our findings support the use of anti-inflammatory or immunosuppressive agents during neonatal sepsis to mitigate the risk of developing hydrocephalus. There have been many attempts to use corticosteroids in treating infants with acute bacterial meningitis; but except for some notable exceptions such as *Haemophilus influenza* (Schaad et al., 1993; Syrogiannopoulos et al., 1994; King et al., 1994; Wald et al., 1995; Kilpi et al., 1995) and pneumococcal meningitis (Kanra et al., 1995; Kilpi et al., 1995; King et al., 1994; Schaad et al., 1993; Syrogiannopoulos et al., 1994; Wald et al., 1995), there have been many failures with some devastating outcomes (Kanra et al., 1995; Kilpi et al., 1995; King et al., 1994; Lebel et al., 1988, 1989; Odio et al., 1991; Syrogiannopoulos et al., 1994) and this remains an area of controversy (Schaad et al., 1995; McIntyre et al., 1997; Coyle, 1999; Goodman et al., 2002). Our work points to critical intervention pathways and suggests that selective and targeted modulation of aspects of the immune response might be studied when treating brain infections to prevent hydrocephalus. For example, in addition to pro-inflammatory immune factors, we found differential expression of IL-4 and IL-13, which generally have anti-inflammatory effects. IL-4 and IL-13 can share a common receptor (Hilton et al., 1996; Jiang et al., 2000; Callard et al., 1996) and typically work synergistically to decrease inflammation by counteracting the activity of pro-inflammatory cytokines such as IL-12 (Minty et al., 1993; Mori et al., 2016). IL-4 and IL-13 are also considered neuroprotective as they can induce death of microglial cells that mediate neuronal damage (Won et al., 2013; Deboy et al., 2006; Yang et al., 2002). However, IL-4 and IL-13 can potentiate oxidative stress-related injury to neurons (Nam et al., 2012;

Park et al., 2008). Thus, one possibility is that the balance of IL-12 and IL-4/IL-13 within the CSF contributes to risk of PIH, implying that their targeted modulation is a potential therapeutic target.

Viral co-infection in PIH

Identifying CMV in the CSF of a subset of infants in our cohort (Paulson et al., 2020) raises the important question of whether the host immune response was in part attributable to CMV. Although there were differentially expressed transcripts detected in CSF consistent with the host response to central nervous system viral infection in CMV-positive infants, we did not detect any functional proteins associated with a response to CMV nor was there any effect of CMV status on the differential expression of host immune markers at the protein level. This may be due to our small sample size (8/100 CSF and 27/100 blood) of CMV-positive infants, or the ability of the virus to modulate protein synthesis in the host (Marshall and Geballe, 2009). Alternatively, CSF CMV status using PCR underestimates CMV burden as intracellular CMV may be difficult to detect in CSF samples with low cell counts, and because CMV cycles through latent and active states in affected individuals (Cheeran et al., 2009). It remains unclear whether CMV was a risk factor for developing PIH or whether the presence of this virus will have long-term effects in these infants.

Conclusions

Inflammation following neonatal infection is a dominant cause of childhood hydrocephalus. An integrated proteogenomic strategy identified gene pathways involving the innate immune system, cell-cell junction structure, platelet activation, and microbial recognition in African infants with PIH. These findings enable the development of preventive hydrocephalus risk reduction strategies during the treatment of neonatal sepsis.

Limitations of the study

Limitations to this study include small cohort size for subset analyses such as CMV-positive participants and low RNA abundance in NPIH samples. Because NK cell activity, IL-12 signaling, and JAK/STAT pathway may indicate viral immune response, it is possible that co-infection with CMV was more common in our patients than our conservative approach to CMV detection permitted. The observation of Th2 cell activation by IL-4/IL-13 needs confirmation in an independent cohort. In addition, attribution of RNA expression to specific cells cannot be completed for the current samples, limiting the ability to confirm the role of specific immune cells in PIH. Furthermore, the majority of activated gene networks identified were pathogen-stratified based on *Paenibacillus* spp. presence.

Resource availability

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Steven J. Schiff (steven.j.schiff@gmail.com).

Materials availability

This study did not generate new unique reagents.

Data and code availability

The code generated during this study is available at our public Github repository https://github.com/Schiff-Lab/iScience_Paper_Isaacs_Morton_2021.

RNA sequencing and sample metadata are available in the NCBI archive under project ID #PRJNA605220. Proteomic data are available in the Proteome Exchange under project ID #PDX024842.

METHODS

All methods can be found in the accompanying Transparent Methods supplemental file.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2021.102351.

ACKNOWLEDGMENTS

The expert technical assistance of Petra Erdmann-Gilmore, Rose Connors, Yiling Mi, and Alan Davis at the Washington University Proteomics Shared Resource (WU-PSR), where the proteomic experiments were performed, is gratefully acknowledged. We also appreciate the efforts of all staff at the genomics core of the Institute of Personalized Medicine at Penn State University, where the majority of the genomic studies were performed. We are grateful to all our colleagues and collaborators in Uganda who were involved in many stages of this study. We thank our funding agencies who supported this work: The National Institutes of Health Director's Pioneer Award 5DP1HD086071-05 to S.J.S., Hydrocephalus Association 2018 Discovery Science Award to R.R.T., and Vanier Canada Graduate Scholarship 396212 to A.M.I. The WU-PSR is supported in part by the Washington University Institute of Clinical and Translational Sciences (NCATS UL1 TR000448), the Mass Spectrometry Research Resource (NIGMS P41 GM103422; R24GM136766), and the Siteman Comprehensive Cancer Center Support Grant (NCI P30 CA091842). None of the organizations listed had any role in the study design, data collection, data analysis, data interpretation, writing, or decision to submit the report for publication.

AUTHOR CONTRIBUTIONS

Study conceptual design and methodology were by S.J.S., J.N.P., D.D.L., R.R.T., J.R.B., A.M.I., M.H., S.U.M., and B.C.W. Formal data analysis and validation were by A.M.I., S.U.M., M.M., Q.Z., C.H., S.J.S., and L.Z. Investigations were performed by A.M.I., C.H., J.O., R.M., E.K., S.J.S., and J.R.B. Resources were provided by D.M.M., E.K., P.S., J.O., R.M., B.C.W., J.R.B., R.R.T., D.D.L., J.N.P., and S.J.S. Data curation was by J.N.P., M.M., C.H., L.Z., and Q.Z. Writing (original draft) was by A.M.I., S.U.M., and J.N.P. Writing (review and editing) were by A.M.I., S.U.M., J.N.P., and S.J.S., with contributions from Q.Z., C.H., L.Z., D.M., S.S., J.E.E., E.M., P.S., J.O., R.M., M.H., and J.R.B. Visualization was by A.M.I., S.J.S., M.M., and J.N.P. Supervision was provided by J.R.B., R.R.T., D.D.L., J.N.P., and S.J.S. Project administration was by C.H., E.K., and S.A.S. Project administration was by A.M.I., S.U.M., C.H., D.M.M., J.R.B., J.N.P., and S.J.S. Acquisition of financial support was by S.J.S., D.D.L., A.M.I., and R.R.T. Final manuscript review and approval was by all authors.

DECLARATION OF INTERESTS

Dr. Limbrick receives research funds and/or research equipment for unrelated projects from Medtronic, Inc. and Microbot Medical, Inc. Dr. Limbrick has received philanthropic equipment contributions for humanitarian relief work from Karl Storz, Inc. and Aesculap, Inc. The authors have no personal, financial, or institutional interest in any of the materials or devices described in this article.

Received: November 10, 2020
Revised: February 24, 2021
Accepted: March 19, 2021
Published: April 23, 2021

REFERENCES

Adiko, A.C., Babdor, J., Gutierrez-Martinez, E., Guermonprez, P., and Saveanu, L. (2015). Intracellular transport routes for MHC I and their relevance for antigen cross-presentation. Front. Immunol. *6*, 335.

Aran, D., Hu, Z., and Butte, A.J. (2017). xCell: digitally portraying the tissue cellular heterogeneity landscape. Genome Biol. *18*, 220.

Batiz, F., Paez, P., Jimenez, A.J., Rodriguez, S., Perez-Figares, J.M., and Rodr¸Guez, E.M. (2005). Clinical and neuropathological evolution of the hydrocephalus developed by the mutant mouse hyh. Cerebrospinal Fluid Res. https://doi.org/10.1186/1743-8454-2-S1-S9.

Batiz, L.F., Paez, P., Jimenez, A.J., Rodriguez, S., Wagner, C., Perez-Figares, J.M., and Rodriguez, E.M. (2006). Heterogeneous expression of hydrocephalic phenotype in the hyh mice carrying a point mutation in [alpha]-SNAP. Neurobiol. Dis. *23*, 152–168.

Bell, E. (2003). Innate immunity - TLR signalling. Nat. Rev. Immunol. *3*, 692.

Brailoiu, E., Barlow, C.L., Ramirez, S.H., Abood, M.E., and Brailoiu, G.C. (2018). Effects of platelet-activating factor on brain microvascular endothelial cells. Neuroscience *377*, 105–113.

Callard, R.E., Matthews, D.J., and Hibbert, L. (1996). IL-4 and IL-13 receptors: are they one and the same? Immunol. Today *17*, 108–110.

Castaneyra-Ruiz, L., Morales, D.M., Mcallister, J.P., Brody, S.L., Isaacs, A.M., Strahle, J.M., Dahiya, S.M., and Limbrick, D.D. (2018). Blood exposure causes ventricular zone disruption and glial activation in vitro. J. Neuropathol. Exp. Neurol. *77*, 803–813.

Cheeran, M.C., Lokensgard, J.R., and Schleiss, M.R. (2009). Neuropathogenesis of congenital cytomegalovirus infection: disease mechanisms and prospects for intervention. Clin. Microbiol. Rev. *22*, 99–126.

Chen, Z.W., Fuchs, K., Sieghart, W., Townsend, R.R., and Evers, A.S. (2012). Deep amino acid sequencing of native brain GABAA receptors using high-resolution mass spectrometry. Mol. Cell Proteomics *11*, M111 011445.

Ciurea, A.V., Coman, T.C., and Mircea, D. (2005). Postinfectious Hydrocephalus in Children. *Pediatric Hydrocephalus* (Springer).

Coyle, P.K. (1999). Glucocorticoids in central nervous system bacterial infection. Arch. Neurol. *56*, 796–801.

de Oliveira, Y.P.A., Pontes-De-Carvalho, L.C., Couto, R.D., and Noronha-Dutra, A.A. (2017). Oxidative stress in sepsis. Possible production of free radicals through an erythrocyte-mediated positive feedback mechanism. Braz. J. Infect. Dis. *21*, 19–26.

Deboy, C.A., Xin, J., Byram, S.C., Serpe, C.J., Sanders, V.M., and Jones, K.J. (2006). Immune-mediated neuroprotection of axotomized mouse facial motoneurons is dependent on the IL-4/STAT6 signaling pathway in CD4(+) T cells. Exp. Neurol. *201*, 212–224.

Dewan, M.C., Rattani, A., Mekary, R., Glancz, L.J., Yunusa, I., Baticulon, R.E., Fieggen, G., Wellons, J.C., Park, K.B., and Warf, B.C. (2018). Global hydrocephalus epidemiology and incidence: systematic review and meta-analysis. J. Neurosurg. 1–15.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21.

Goodman, S., Sprung, C.L., and International Sepsis, F. (2002). The International Sepsis Forum's controversies in sepsis: corticosteroids should be used to treat septic shock. Crit. Care *6*, 381–383.

Guerra, M., Henzi, R., Ortloff, A., Lictin, N., Vio, K., Jiminez, A.J., Dominguez-Pinos, M.D., Gonzalez, C., Bioch, M.C.J., Hinostroza, F., et al. (2015a). A cell junction pathology of neural stem cells Is associated to ventricular zone disruption, hydrocephalus and neurogenesis abnormalities. J. Neuropathol. Exp. Neurol. *74*, 653–671.

Guerra, M.M., Henzi, R., Ortloff, A., Lichtin, N., Vio, K., Jimenez, A.J., Dominguez-Pinos, M.D., Gonzalez, C., Jara, M.C., Hinostroza, F., et al. (2015b). Cell junction pathology of neural stem cells is associated with ventricular zone disruption, hydrocephalus, and abnormal neurogenesis. J. Neuropathol. Exp. Neurol. *74*, 653–671.

Habiyaremye, G., Morales, D.M., Morgan, C.D., Mcallister, J.P., Crevecoeur, T.S., Han, R.H., Gabir, M., Baksh, B., Mercer, D., and Limbrick, D.D.,, JR. (2017). Chemokine and cytokine levels in the lumbar cerebrospinal fluid of preterm infants with post-hemorrhagic hydrocephalus. Fluids Barriers CNS *14*, 35.

Heintzman, N.D., and Ren, B. (2006). The gateway to transcription: identifying, characterizing and understanding promoters in the eukaryotic genome. Cell Mol. Life Sci. *64*, 386–400.

Hilton, D.J., Zhang, J.G., Metcalf, D., Alexander, W.S., Nicola, N.A., and Willson, T.A. (1996). Cloning and characterization of a binding subunit of the interleukin 13 receptor that is also a component of the interleukin 4 receptor. Proc. Natl. Acad. Sci. U S A. *93*, 497–501.

Isaacs, A.M., Riva-Cambrin, J., Yavin, D., Hockley, A., Pringsheim, T.M., Jette, N., Lethebe, B.C., Lowerison, M., Dronyk, J., and Hamilton, M.G. (2018). Age-specific global epidemiology of hydrocephalus: systematic review, metanalysis and global birth surveillance. PLoS One *13*, e0204926.

Jiang, H., Harris, M.B., and Rothman, P. (2000). IL-4/IL-13 signaling beyond JAK/STAT. J. Allergy Clin. Immunol. *105*, 1063–1070.

Jimenez, A.J., Garcia-Verdugo, J.M., Gonzalez, C.A., Batiz, L.F., Rodriguez-Perez, L.M., Paez, P., Soriano-Navarro, M., Roales-Bujan, R., Rivera, P., Rodriguez, S., et al. (2009). Disruption of the neurogenic niche in the subventricular zone of postnatal hydrocephalic hyh mice. J. Neuropathol. Exp. Neurol. *68*, 1006–1020.

Jin, S., Dong, W., Kundishora, A., Panchagnula, S., Moreno-De-Luca, A., Furey, C., Allocco, A., Walker, R., Nelson-Williams, C., Smith, H., et al. (2020). Exome sequencing implicates genetic disruption of prenatal neuro-gliogenesis in sporadic congenital hydrocephalus. Nat. Med. *26*, 1754–1765.

Kanra, G.Y., Ozen, H., Secmeer, G., Ceyhan, M., Ecevit, Z., and Belgin, E. (1995). Beneficial effects of dexamethasone in children with pneumococcal meningitis. Pediatr. Infect Dis. J. *14*, 490–494.

Karimy, J.K., Reeves, B.C., Damisah, E., Duy, P.Q., Antwi, P., David, W., Wang, K., Schiff, S.J., Limbrick, D.D.,, JR., Alper, S.L., et al. (2020). Inflammation in acquired hydrocephalus: pathogenic mechanisms and therapeutic targets. Nat. Rev. Neurol. *16*, 285–296.

Karimy, J.K., Zhang, J., Kurland, D.B., Theriault, B.C., Duran, D., Stokum, J.A., Furey, C.G., Zhou, X., Mansuri, M.S., Montejo, J., et al. (2017). Inflammation-dependent cerebrospinal fluid hypersecretion by the choroid plexus epithelium in posthemorrhagic hydrocephalus. Nat. Med. *23*, 997–1003.

Kawai, T., and Akira, S. (2006). TLR signaling. Cell Death Differ. *13*, 816–825.

Kilpi, T., Peltola, H., Jauhiainen, T., and Kallio, M.J. (1995). Oral glycerol and intravenous dexamethasone in preventing neurologic and audiologic sequelae of childhood bacterial meningitis. The Finnish Study Group. Pediatr. Infect Dis. J. *14*, 270–278.

King, S.M., Law, B., Langley, J.M., Heurter, H., Bremner, D., Wang, E.E., and Gold, R. (1994). Dexamethasone therapy for bacterial meningitis: Better never than late? Can J. Infect Dis. *5*, 210–215.

Koskinen, V.R., Emery, P.A., Creasy, D.M., and Cottrell, J.S. (2011). Hierarchical clustering of shotgun proteomics data. Mol. Cell Proteomics *10*, M110 003822.

Kulkarni, A.V., Schiff, S.J., Mbabazi-Kabachelor, E., Mugamba, J., Ssenyonga, P., Donnelly, R., Levenbach, J., Monga, V., Peterson, M., Macdonald, M., et al. (2017). Endoscopic treatment versus shunting for infant hydrocephalus in Uganda. N. Engl. J. Med. *377*, 2456–2464.

Lattke, M., Magnutzki, A., Walther, P., Wirth, T., and baumann, B. (2012). Nuclear factor kappaB activation impairs ependymal ciliogenesis and links neuroinflammation to hydrocephalus formation. J. Neurosci. *32*, 11511–11523.

Lebel, M.H., Freij, B.J., Syrogiannopoulos, G.A., Chrane, D.F., Hoyt, M.J., Stewart, S.M., Kennard, B.D., Olsen, K.D., and Mccracken, G.H.,, JR. (1988). Dexamethasone therapy for bacterial meningitis. Results of two double-blind, placebo-controlled trials. N. Engl. J. Med. *319*, 964–971.

Lebel, M.H., Hoyt, M.J., Waagner, D.C., Rollins, N.K., Finitzo, T., and Mccracken, G.H.,, JR. (1989). Magnetic resonance imaging and dexamethasone therapy for bacterial meningitis. Am. J. Dis. Child *143*, 301–306.

Lehner, C., Gehwolf, R., Tempfer, H., Krizbai, I., Hennig, B., Bauer, H.C., and Bauer, H. (2011). Oxidative stress and blood-brain barrier dysfunction under particular consideration of matrix metalloproteinases. Antioxid. Redox Signaling *15*, 1305–1323.

Li, B., and Dewey, C.N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. Bmc Bioinformatics *12*.

Marshall, E.E., and Geballe, A.P. (2009). Multifaceted evasion of the interferon response by cytomegalovirus. J. Interferon Cytokine Res. *29*, 609–619.

McAlister, G.C., Huttlin, E.L., Haas, W., Ting, L., Jedrychowski, M.P., Rogers, J.C., Kuhn, K., Pike, I., Grothe, R.A., Blethrow, J.D., and Gygi, S.P. (2012). Increasing the multiplexing capacity of TMTs using reporter ion isotopologues with isobaric masses. Anal Chem. *84*, 7469–7478.

McAllister, J.P., 2nd, Williams, M.A., Walker, M.L., Kestle, J.R., Relkin, N.R., Anderson, A.M., Gross, P.H., Browd, S.R., and Hydrocephalus Symposium Expert, P. (2015). An update on research priorities in hydrocephalus: overview of the third National Institutes of Health-sponsored symposium "Opportunities for hydrocephalus research: pathways to Better outcomes. J. Neurosurg. *123*, 1427–1438.

McAllister, J.P., Guerra, M.M., Ruiz, L.C., Jimenez, A.J., Dominguez-Pinos, D., Sival, D., Den Dunnen, W., Morales, D.M., Schmidt, R.E., Rodriguez, E.M., and Limbrick, D.D. (2017). Ventricular zone disruption in human neonates with intraventricular hemorrhage. J. Neuropathol. Exp. Neurol. *76*, 358–375.

McIntyre, P.B., Berkey, C.S., King, S.M., Schaad, U.B., Kilpi, T., Kanra, G.Y., and Perez, C.M. (1997). Dexamethasone as adjunctive therapy in bacterial meningitis. A meta-analysis of randomized clinical trials since 1988. JAMA *278*, 925–931.

Mertins, P., Tang, L.C., Krug, K., Clark, D.J., Gritsenko, M.A., Chen, L., Clauser, K.R., Clauss, T.R., Shah, P., Gillette, M.A., et al. (2018). Reproducible workflow for multiplexed deep-scale proteome and phosphoproteome analysis of tumor tissues by liquid chromatography-mass spectrometry. Nat. Protoc. *13*, 1632–1661.

Minty, A., Chalon, P., Derocq, J.M., Dumont, X., Guillemot, J.C., Kaghad, M., Labit, C., Leplatois, P., Liauzun, P., Miloux, B., et al. (1993). Interleukin-13 is a new human lymphokine regulating inflammatory and immune responses. Nature *362*, 248–250.

Morales, D.M., Townsend, R.R., Malone, J.P., Ewersmann, C.A., Macy, E.M., Inder, T.E., and Limbrick, D.D.,, JR. (2012). Alterations in protein regulators of neurodevelopment in the cerebrospinal fluid of infants with posthemorrhagic hydrocephalus of prematurity. Mol. Cell Proteomics *11*, M111 011973.

Mori, S., Maher, P., and Conti, B. (2016). Neuroimmunology of the interleukins 13 and 4. Brain Sci. *6*, 18.

Nam, J.H., Park, K.W., Park, E.S., Lee, Y.B., Lee, H.G., Baik, H.H., Kim, Y.S., Maeng, S., Park, J., and Jin, B.K. (2012). Interleukin-13/-4-induced oxidative stress contributes to death of hippocampal neurons in abeta1-42-treated hippocampus in vivo. Antioxid. Redox Signal *16*, 1369–1383.

Nesvizhskii, A.I. (2014). Proteogenomics: concepts, applications and computational strategies. Nat. Methods *11*, 1114–1125.

Noguchi, M., Yi, H.F., Rosenblatt, H.M., Filipovich, A.H., Adelstein, S., Modi, W.S., Mcbride, O.W., and Leonard, W.J. (1993). Interleukin-2 receptor gamma chain mutation results in X-linked severe combined immunodeficiency in humans. Cell *73*, 147–157.

Noguchi, M., Yi, H.F., Rosenblatt, H.M., Filipovich, A.H., Adelstein, S., Modl, W.S., Mcbride, O.W., and Leonard, W.J. (2008). Interleukin-2 receptor gamma chain mutation results in x-linked severe combined immunodeficiency in humans. J. Immunol. *181*, 5817–5827.

O'shea, J.J., Schwartz, D.M., Villarino, A.V., Gadina, M., Mcinnes, I.B., and Laurence, A. (2015). The JAK-STAT pathway: impact on human disease and therapeutic intervention. Annu. Rev. Med. *66*, 311–328.

Odio, C.M., Faingezicht, I., Paris, M., Nassar, M., Baltodano, A., Rogers, J., Saez-Llorens, X., Olsen, K.D., and Mccracken, G.H.,, JR. (1991). The beneficial effects of early dexamethasone administration in infants and children with bacterial meningitis. N. Engl. J. Med. *324*, 1525–1531.

Ortega, E., Munoz, R.I., Luza, N., Guerra, F., Guerra, M., Vio, K., Henzi, R., Jaque, J., Rodriguez, S., Mcallister, J.P., and Rodriguez, E. (2016). The value of early and comprehensive diagnoses in a human fetus with hydrocephalus and progressive obliteration of the aqueduct of Sylvius: Case Report. BMC Neurol. *16*, 45.

Ortloff, A.R., Vio, K., Guerra, M., Jaramillo, K., Kaehne, T., Jones, H., Mcallister, J.P., 2nd, and Rodriguez, E. (2013). Role of the subcommissural organ in the pathogenesis of congenital hydrocephalus in the HTx rat. Cell Tissue Res *352*, 707–725.

Paez, P., Batiz, L.F., Roales-Bujan, R., Rodriguez-Perez, L.M., Rodriguez, S., Jimenez, A.J., Rodriguez, E.M., and Perez-Figares, J.M. (2007). Patterned neuropathologic events occurring in hyh congenital hydrocephalic mutant mice. J. Neuropathol. Exp. Neurol. *66*, 1082–1092.

Park, K.W., Baik, H.H., and Jin, B.K. (2008). Interleukin-4-induced oxidative stress via microglial NADPH oxidase contributes to the death of hippocampal neurons in vivo. Curr. Aging Sci. *1*, 192–201.

Paulson, J.N., Chen, C.Y., Lopes-Ramos, C.M., Kuijjer, M.L., Platig, J., Sonawane, A.R., Fagny, M., Glass, K., and Quackenbush, J. (2017). Tissue-aware RNA-Seq processing and normalization for heterogeneous and sparse data. BMC Bioinformatics *18*, 437.

Paulson, J.N., Williams, B.L., Hehnly, C., Mishra, N., Sinnar, S.A., Zhang, L., Ssentongo, P., Mbabazi-Kabachelor, E., Wijetunge, D.S.S., Von Bredow, B., et al. (2020). Paenibacillus infection with frequent viral coinfection contributes to postinfectious hydrocephalus in Ugandan infants. Sci. Transl Med. *12*, eaba0565.

Perez-Figares, J.M., Jimenez, A.J., Perez-Martin, M., Fernandez-Llebrez, P., Cifuentes, M., Riera, P., Rodriguez, S., and Rodriguez, E.M. (1998). Spontaneous congenital hydrocephalus in the mutant mouse hyh. Changes in the ventricular system and the subcommissural organ. J. Neuropathol. Exp. Neurol. *57*, 188–202.

PrabhuDas, M., Adkins, B., Gans, H., King, C., Levy, O., Ramilo, O., and Siegrist, C.A. (2011). Challenges in infant immunity: implications for responses to infection and vaccines. Nat. Immunol. *12*, 189–194.

Pun, P.B.L., Lu, J., and Moochhala, S. (2009). Involvement of ROS in BBB dysfunction. Free Radic. Res. *43*, 348–364.

Rédei, G.P. (2008). Encyclopedia of Genetics, Genomics, Proteomics & Informatics (Springer).

Rodriguez, E.M., Guerra, M.M., Vio, K., Gonzalez, C., Ortloff, A., Batiz, L.F., Rodriguez, S., Jara, M.C., Munoz, R.I., Ortega, E., et al. (2012). A cell junction pathology of neural stem cells leads to abnormal neurogenesis and hydrocephalus. Biol. Res. *45*, 231–242.

Rohlwink, U.K., Figaji, A., Wilkinson, K.A., Horswell, S., Sesay, A.K., Deffur, A., Enslin, N., Solomons, R., Van Toorn, R., Eley, B., et al. (2019). Tuberculous meningitis in children is characterized by compartmentalized immune responses and neural excitotoxicity. Nat. Commun. *10*, 3767.

Ruggles, K.V., Krug, K., Wang, X., Clauser, K.R., Wang, J., Payne, S.H., Fenyo, D., Zhang, B., and Mani, D.R. (2017). Methods, tools and current perspectives in proteogenomics. Mol. Cell Proteomics *16*, 959–981.

Schaad, U.B., Kaplan, S.L., and Mccracken, G.H.,, JR. (1995). Steroid therapy for bacterial meningitis. Clin. Infect Dis. *20*, 685–690.

Schaad, U.B., Lips, U., Gnehm, H.E., Blumberg, A., Heinzer, I., and Wedgwood, J. (1993). Dexamethasone therapy for bacterial meningitis in children. Swiss Meningitis Study Group. Lancet *342*, 457–461.

Simon, T.D., Riva-Cambrin, J., Srivastava, R., Bratton, S.L., Dean, J.M., Kestle, J.R., and Hydrocephalus Clinical Research, N. (2008). Hospital care for children with hydrocephalus in the United States: utilization, charges, comorbidities, and deaths. J. Neurosurg. Pediatr. *1*, 131–137.

Sinnar, S.A., and Schiff, S.J. (2020). The problem of microbial dark matter in neonatal sepsis. Emerg. Infect Dis. *26*, 2543–2548.

Sival, D.A., Guerra, M., Den Dunnen, W.F., Batiz, L.F., Alvial, G., Castaneyra-Perdomo, A., and Rodriguez, E.M. (2011). Neuroependymal denudation is in progress in full-term human foetal spina bifida aperta. Brain Pathol. *21*, 163–179.

Sokol, B., Wasik, N., Jankowski, R., Holysz, M., Wieckowska, B., and Jagodzinski, P. (2016). Soluble toll-like receptors 2 and 4 in cerebrospinal fluid of patients with acute hydrocephalus following aneurysmal subarachnoid haemorrhage. PLoS One *11*, e0156171.

Syrogiannopoulos, G.A., Lourida, A.N., Theodoridou, M.C., Pappas, I.G., Babilis, G.C., Economidis, J.J., Zoumboulakis, D.J., Beratis, N.G., and Matsaniotis, N.S. (1994). Dexamethasone therapy for bacterial meningitis in children: 2- versus 4-day regimen. J. Infect Dis. *169*, 853–858.

Tchessalova, D., Posillico, C.K., and Tronson, N.C. (2018). Neuroimmune activation drives multiple brain states. Front. Syst. Neurosci. *12*, 39.

Villarino, A.V., Kanno, Y., and O'shea, J.J. (2017). Mechanisms and consequences of Jak-STAT signaling in the immune system. Nat. Immunol. *18*, 374–384.

Wagner, C., Batiz, L.F., Rodriguez, S., Jimenez, A.J., Paez, P., Tome, M., Perez-Figares, J.M., and Rodriguez, E.M. (2003). Cellular mechanisms involved in the stenosis and obliteration of the cerebral aqueduct of hyh mutant mice developing congenital hydrocephalus. J. Neuropathol. Exp. Neurol. *62*, 1019–1040.

Wald, E.R., Kaplan, S.L., Mason, E.O., Jr., Sabo, D., Ross, L., Arditi, M., Wiedermann, B.L., Barson, W., Kim, K.S., Yogov, R., et al. (1995). Dexamethasone therapy for children with bacterial meningitis. Meningitis Study Group. Pediatrics *95*, 21–28.

Warf, B.C. (2005). Hydrocephalus in Uganda: the predominance of infectious origin and primary management with endoscopic third ventriculostomy. J. Neurosurg. *102*, 1–15.

Warf, B.C., and East African Neurosurgical Research, C. (2010). Pediatric hydrocephalus in East Africa: prevalence, causes, treatments, and strategies for the future. World Neurosurg. *73*, 296–300.

Won, S.Y., Kim, S.R., Maeng, S., and Jin, B.K. (2013). Interleukin-13/Interleukin-4-induced oxidative stress contributes to death of prothrombinkringle-2 (pKr-2)-activated microglia. J. Neuroimmunol *265*, 36–42.

Yang, M.S., Park, E.J., Sohn, S., Kwon, H.J., Shin, W.H., Pyo, H.K., Jin, B., Choi, K.S., Jou, I., and Joe, E.H. (2002). Interleukin-13 and -4 induce death of activated microglia. Glia *38*, 273–280.

Ygberg, S., and Nilsson, A. (2012). The developing immune system - from foetus to toddler. Acta Paediatr. *101*, 120–127.

Yost, C.C., Weyrich, A.S., and Zimmerman, G.A. (2010). The platelet activating factor (PAF) signaling cascade in systemic inflammatory responses. Biochimie *92*, 692–697.

Zhang, B., Wang, J., Wang, X., Zhu, J., Liu, Q., Shi, Z., Chambers, M.C., Zimmerman, L.J., Shaddox, K.F., Kim, S., et al. (2014). Proteogenomic characterization of human colon and rectal cancer. Nature *513*, 382–387.

iScience, Volume *24*

Supplemental information

Immune activation during *Paenibacillus* brain infection in African infants with frequent cytomegalovirus co-infection

Albert M. Isaacs, Sarah U. Morton, Mercedeh Movassagh, Qiang Zhang, Christine Hehnly, Lijun Zhang, Diego M. Morales, Shamim A. Sinnar, Jessica E. Ericson, Edith Mbabazi-Kabachelor, Peter Ssenyonga, Justin Onen, Ronnie Mulondo, Mady Hornig, Benjamin C. Warf, James R. Broach, R. Reid Townsend, David D. Limbrick Jr., Joseph N. Paulson, and Steven J. Schiff

TRANSPARENT METHODS

Study Cohort

Participants were recruited at the CURE Children's Hospital of Uganda (CCHU), a pediatric neurosurgery hospital that serves as the largest referral center for patients with hydrocephalus in Uganda. Inclusion criteria for PIH: a) age 3 months or less, b) weight greater than 2.5 kg, c) mothers at least 18 years old to give informed consent, and d) no history of hydrocephalus at birth, and either: i) a history of febrile illness and/or seizures preceding the onset of hydrocephalus, or ii) alternative findings such as imaging and endoscopic results indicative of prior ventriculitis including septations, loculations, or deposits of debris within the ventricular system. Inclusion criteria for NPIH: a) age 3 months or less, b) weight greater than 2.5 kg, c) absence of findings consistent with PIH or congenital origin of hydrocephalus, d) mothers at least 18 years old to give informed consent, and e) findings of non-infectious origin of hydrocephalus on computed tomography (CT) scan or at endoscopy including structural cause (obstruction of the Aqueduct of Sylvius, Dandy-Walker cyst, aneurysm, cavernous malformation, or other congenital malformation) or evidence of hemorrhage (bloody CSF). Exclusion criteria for hydrocephalus study: a) prior surgery on the nervous system (shunt, third ventriculostomy, or myelomeningocele closure), or b) evidence of communication of nervous system with skin such as meningocele, encephalocele, dermal sinus tract, or fistula. Clinical patient information including demographics, computed tomography (CT) neuroimaging, CSF cell counts, and CSF culture results were obtained from review of the medical records.

Ethics approval and consent to participate

Ethics oversight was provided by the CCHU Institutional Review Board (IRB), Mbarara University of Science and Technology Research Ethics Committee, the Ugandan National Council on Science and Technology, and the Pennsylvania State University (PSU) IRB. Verbal informed consent was obtained from the mother of each subject.

CSF Samples

Sterile CSF samples were acquired directly from the cerebral ventricles of all patients at the time of surgery (shunt insertion or endoscopic treatment) for treatment of hydrocephalus. After collection, aliquots of each sample were alternately placed into either a genomic preservative (DNA/RNA Shield, Zymo Research, Irvine CA) to preserve DNA and RNA, or fresh frozen to preserve protein content, and maintained in liquid nitrogen or in a -80-degree Celsius freezer prior to cryogenic shipment, for final -80-degree Celsius storage until proteomics and/or RNA-Seq was performed.

Proteomics characterization

Sample preparation and mass spectrometry

A standard validated proteomics workflow was utilized (Mertins et al., 2018). High abundance proteins in CSF (albumin, IgG, α_1-antitrypsin, IgA, transferrin and haptoglobin) were first removed using a reproducible affinity method (Morales et al., 2012). The depleted samples were then enzymatically digested with trypsin after protein denaturation, reduction and alkylation. Peptides were prepared for isobaric labeling using robotic solid phase extraction (Mertins et al., 2018, Chen et al., 2012). A ten-channel isobaric peptide mass tagging (TMT-10) was used. The purified peptides were labeled and pooled in sets of ten with one label used for tagging a reference pool for all sets of nine samples. The labeled reference pool was used for normalization and merging multiple ten plex data sets. The labeled peptides were analyzed by high-resolution liquid chromatography-mass spectrometry (LC-MS) using a Q-Exactive Plus instrument coupled to an EASY-nanoLC 1000 system (Thermo-Fisher Scientific). The pooled samples in batches of ten CSF patient samples were loaded onto a 75 µm i.d. × 25 cm Acclaim® PepMap 100 RP column (Thermo-Fisher Scientific). Peptide separations were started with 95% mobile phase A (0.1% FA) for 5 min and increased to 30% B (100% MeCN , 0.1% FA) over 180 min, followed by a 25-min gradient to 45% B, a 5-min gradient to 95% B and washed at 90% B for 7 min, with a flow rate of 300 nL/min. Full-scan mass spectra were then acquired using the Orbitrap mass analyzer in the mass-to-charge ratio (m/z) of 375 to 1400 and with a mass

resolving power set to 70,000. Fifteen data-dependent high-energy collisional dissociation (HCD) were performed with a mass resolving power set to 35,000, a fixed first m/z 100, an isolation width of 0.7 m/z, and the normalized collision energy setting of 32. The maximum injection time was 50 ms for parent-ion analysis and 105 ms for product-ion analysis. Target ions already selected for mass spectrometry (MS)/MS were dynamically excluded for 30 sec. An automatic gain control target value of 3×10^6 ions was used for full MS scans and 1×10^5 ions for MS/MS scans. Peptide ions with charge states of one or greater than six were excluded from MS/MS interrogation.

Protein identification

MS raw data were converted to peak lists using Proteome Discoverer (version 2.1.0.81, Thermo-Fisher Scientific) with the integration of reporter-ion intensities of TMT 10-plex at a mass tolerance of ±3.15 mDa (Werner et al., 2014). MS/MS spectra with charges +2, +3 and +4 were analyzed using Mascot search engine (Perkins et al., 1999) (Matrix Science, London, UK; version 2.6.2). Mascot was set up to search against a SwissProt database of human (version June 2016, 20,237 entries) and common contaminant proteins (cRAP, version 1.0 Jan. 1st, 2012, 116 entries), with trypsin/P set as the digestion enzyme and allowing for a maximum of 4 missed cleavages. The searches were performed with a fragment ion mass tolerance of 0.02 Da and a parent ion tolerance of 20 ppm. Carbamidomethylation of cysteine was specified in Mascot as a fixed modification. Deamidation of asparagine, formation of pyro-glutamic acid from N-terminal glutamine, acetylation of protein N-terminus, oxidation of methionine, and pyro-carbamidomethylation of N-terminal cysteine were specified as variable modifications. Peptide spectrum matches (PSM) were filtered at 1% false-discovery rate (FDR) by searching against a reversed database and the ascribed peptide identities were accepted. The uniqueness of peptide sequences among the database entries was determined using the principle of parsimony. Protein identities were inferred using a greedy set cover algorithm from Mascot and the identities containing ≥ 2 Occam's razor peptides were accepted (Koskinen et al., 2011).

Protein relative quantification

The processing, quality assurance and analysis of TMT data were performed with proteoQ (version 1.0; https://github.com/qzhang503/proteoQ). Reporter-ion intensities under 10-plex TMT channels were first obtained from Mascot, followed by removal of PSM entries from shared peptides or those with intensity values lower than 1E3. Intensity of PSMs were converted to logarithmic ratios at base two, relative to the average intensity of reference samples which bookmarked all 10-plex TMT sets. Under each TMT channel, Dixon's outlier removals were carried out recursively for peptides with greater than two identifying PSMs. The median of the ratios of PSM that could be assigned to the same peptide was first taken to represent the ratios of the incumbent peptide. Peptide identifications that could be assigned to more than one protein were excluded from protein quantification. The median of the ratios of peptides were then taken to represent the ratios of the incumbent protein. To align protein ratios under different TMT channels, likelihood functions were first estimated for the log-ratios of proteins using finite mixture modelling, assuming two-component Gaussian mixtures (Benaglia et al., 2009). The ratio of distributions was then aligned by maximum likelihood of log-ratios centered at zero for each sample. Scaling normalization was performed to standardize the log-ratios of proteins across samples. To discount the influence of outliers from either log-ratios or reporter-ion intensities, values between the 5^{th} and 95^{th} percentile of log-ratios and 5^{th} and 95^{th} percentile of intensity were used for standard deviation calculations.

Transcriptomics Characterization

RNA Extraction

RNA was extracted from shield samples using TRIzol™ LS Reagent and Direct-zol™ RNA MiniPrep Plus kit (Zymo Research, R2072). Two aliquots of 250 µl of CSF and shield mixture were added to 750 µl of TRIzol™ LS Reagent (Invitrogen, 10296010). After bead lysis, 200 µl of chloroform was added, incubated at room temperature for 15 minutes then centrifuged at 12,000 xg for 15 minutes at 4°C. Once spun down the aqueous layers from the two aliquots were put into the same tube. Then each organic layer was re-

suspended in 400 µl of water then centrifuged with the same parameters again for 5 minutes. The now 1.6 mL of aqueous phases are combined with 1.6 mL of 100% ethanol and then put through the Zymo-Spin™ IC Column twice and then proceeded from step two of the manufacturer's protocol.

RNA-Seq

The Illumina® TruSeq® Stranded RNA Sequencing kit was used following the manufacturer's protocol, which encompassed rRNA depletion and fragmentation, first strand cDNA synthesis, second strand cDNA synthesis, adenylate 3' ends, adapters ligation, and DNA fragment enrichment. Due to the nature of low cell counts in CSF, varying concentrations of RNA were extracted from the samples (Supplemental Table 1). Majority of the RNA concentrations fell below the recommended input of 100 ng so a standard 10 µl of total RNA was used for input ranging from 10-150 ng of RNA. Using biotinylated, target-specific oligos combined with Ribo-Zero ribosomal RNA (rRNA) removal beads of divalent cations and high temperatures (60 – 100^{o}C). Ribosomal RNA was then depleted from the total RNA and the remaining RNA was purified, fragmented, and primed for cDNA synthesis. The cleaved RNA fragments were then copied into first strand cDNA using reverse transcriptase (Part # 18064-014) and random hexamers. A second strand cDNA synthesis was obtained by removing the RNA template, synthesizing a replacement strand, and incorporating dUTP in place of dTTP to generate the ds (double stranded) cDNA, using DNA Polymerase I and RNaseH. A single 'A' nucleotide was then added to the 3' ends of the blunt fragments and a corresponding single 'T' nucleotide was added on the 3' end of the adapter. Multiple indexing adapters were then ligated to the ends of the ds cDNA fragments to prepare them for hybridization. Polymerase chain reaction (PCR) was used to enrich the DNA fragments which was done without dUTP to ensure only the first strand was enriched. The libraries were quantified by fluorometry and size- and quality-checked on an Agilent Technologies 2100 Bioanalyzer. Libraries were then pooled to 4 nM, diluted and sequenced on the NovaSeq 6000 using the S2 flow cell with 2x100 reads aiming for 25 million reads per sample.

Transcriptomics data processing

We used Trimmomatic (v0.38) to trim and crop paired-end RNA sequence adaptors and low sequencing score reads as quality control with default parameters. After filtering for quality control, paired-end RNA-Seq were mapped to the human reference genome hg38 using STAR (Dobin et al., 2013). Gene and isoform expression levels were estimated from mapped RNA-Seq data (Li and Dewey, 2011). Read counts from all samples were aggregated into a matrix for further analysis. Quality measures are included in Supplemental Table 1. Quality control was assessed through scatterplots between log RNA input and log library size where we observed a weak Spearman correlation (0.39). The average and median number of reads was 36 million and 19 million respectively. Gene expression sets from the RNASeq matrix data were normalized (Paulson et al., 2017), and those with expression of at least 1 count per million in a minimum of 18 samples were filtered for inclusion in downstream analysis.

Data Analysis

All analyses were performed with *R* v.3.5™ (http://www.R-project.org) utilizing the following packages: ComBat (Johnson et al., 2007), clusterProfiler (Yu et al., 2012), DESeq2 (Love et al., 2014), edgeR (Robinson and Smyth, 2007), Limma (Ritchie et al., 2015), RSEM (Li and Dewey, 2011), STAR (Dobin et al., 2013), and YARN (Paulson et al., 2017). Continuous demographic variables were evaluated using the non-parametric Mann-Whitney test (2-group comparisons) following Shapiro-Wilk's test for normality. Fisher's exact test was performed for categorical variables. Analyses are made available in our public Github repository https://github.com/Schiff-Lab/iScience_Paper_Isaacs_Morton_2021.

Transcript and protein data analyses

Following quantile normalization and correction of batch effects (Paulson et al., 2017, Johnson et al., 2007), Principal component analysis (PCA) of gene expression and protein $\log_2$ ratios were performed. Log fold change values of the gene and protein abundances between groups were calculated by comparing mean

expression levels using *edgeR* (Robinson and Smyth, 2007) and *Limma* (Szklarczyk et al., 2019) respectively. Genewise Negative Binomial Generalized Linear Model and Empirical Bayes Moderation were carried out to ensure variances were no longer dependent on mean gene or protein expression levels, respectively. The number of differentially expressed genes or proteins were examined with significance set at a false discovery rate controlled for multiple testing comparisons using the method of Benjamini and Hochberg (Benjamini and Hochberg, 1995) with adjusted p-value cutoff of 0.05 and absolute log fold change of 1. Gene and proteins that were differentially expressed in multiple comparisons were extracted and examined from smallest to largest adjusted *p*-value with their associated log fold changes. *clusterProfiler* was then used to map differentially expressed genes and proteins to their respective Gene ontology (GO) terms to assess for enrichment. Clinical factors that correlate with the expression of gene sets were determined using the R package WGCNA. Average expression of the 2205 genes differentially expressed within Paenibacillus positive patient samples were used to cluster the samples, and a height cutoff was applied to the height of the hierarchical clustering graph to separate a cluster of 33 samples with the highest gene expression values. Modules of genes with expression values that correlated significantly with clinical factors among those 33 samples were identified.

Marker genes of bulk RNA sequencing

Filtered RNA counts generated in expression analysis were normalized as transcripts per million bases (TPM) values. The R package ImmuneDeconv was used to deconvolute the per-patient bulk RNA data into immune cell populations. Identity with immune cell populations, expressed on a scale of 0-1, was calculated for each patient.

Proteogenomic integration

Integration of proteomics and transcriptomics data was performed following dimension reduction and feature selection of differentially expressed proteins and genes. GO enrichment using Database for Annotation, Visualization and Integrated Discovery (DAVID v6.8) (Huang et al., 2009) and reactome over-

representation pathway analysis were performed by concatenating differentially expressed proteins and genes with a set absolute log2 fold change of 1 and a false discovery rate of 0.05. The integrated data were examined for gene/pathway/GO enrichment. Physical and functional protein-protein interactions of the genes common to both proteomics and RNA-Seq were assessed using the STRING (Szklarczyk et al., 2019) online tools. Gene-gene relationships between differentially expressed genes were quantified by assigning a composite score to each pairwise comparison. Each composite score reflects chromosome neighborhood, gene fusion, phylogenetic co-occurrence, homology, co-expression, experimentally determined interactions, database annotation and automated text mining between pairs of genes, and the higher the score (i.e., closer to 1), the stronger the predicted gene-gene relationship (Szklarczyk et al., 2019).

BENAGLIA, T., CHAUVEAU, D., HUNTER, D. R. & YOUNG, D. 2009. mixtools: AnRPackage for Analyzing Finite Mixture Models. *Journal of Statistical Software,* 32.

BENJAMINI, Y. & HOCHBERG, Y. 1995. Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B-Statistical Methodology,* 57**,** 289-300.

CHEN, Z. W., FUCHS, K., SIEGHART, W., TOWNSEND, R. R. & EVERS, A. S. 2012. Deep amino acid sequencing of native brain GABAA receptors using high-resolution mass spectrometry. *Mol Cell Proteomics,* 11**,** M111 011445.

DOBIN, A., DAVIS, C. A., SCHLESINGER, F., DRENKOW, J., ZALESKI, C., JHA, S., BATUT, P., CHAISSON, M. & GINGERAS, T. R. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics,* 29**,** 15-21.

HUANG, D. W., SHERMAN, B. T. & LEMPICKI, R. A. 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols,* 4**,** 44-57.

JOHNSON, W. E., LI, C. & RABINOVIC, A. 2007. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics,* 8**,** 118-127.

KOSKINEN, V. R., EMERY, P. A., CREASY, D. M. & COTTRELL, J. S. 2011. Hierarchical clustering of shotgun proteomics data. *Mol Cell Proteomics,* 10**,** M110 003822.

LI, B. & DEWEY, C. N. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *Bmc Bioinformatics,* 12.

LOVE, M. I., HUBER, W. & ANDERS, S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol,* 15**,** 550.

MERTINS, P., TANG, L. C., KRUG, K., CLARK, D. J., GRITSENKO, M. A., CHEN, L., CLAUSER, K. R., CLAUSS, T. R., SHAH, P., GILLETTE, M. A., PETYUK, V. A., THOMAS, S. N., MANI, D. R., MUNDT, F., MOORE, R. J., HU, Y., ZHAO, R., SCHNAUBELT, M., KESHISHIAN, H., MONROE, M. E., ZHANG, Z., UDESHI, N. D., MANI, D., DAVIES, S. R., TOWNSEND, R. R., CHAN, D. W., SMITH, R. D., ZHANG, H., LIU, T. & CARR, S. A. 2018. Reproducible workflow for multiplexed

deep-scale proteome and phosphoproteome analysis of tumor tissues by liquid chromatography-mass spectrometry. *Nat Protoc,* 13**,** 1632-1661.

MORALES, D. M., TOWNSEND, R. R., MALONE, J. P., EWERSMANN, C. A., MACY, E. M., INDER, T. E. & LIMBRICK, D. D., JR. 2012. Alterations in protein regulators of neurodevelopment in the cerebrospinal fluid of infants with posthemorrhagic hydrocephalus of prematurity. *Mol Cell Proteomics,* 11**,** M111 011973.

PAULSON, J. N., CHEN, C. Y., LOPES-RAMOS, C. M., KUIJJER, M. L., PLATIG, J., SONAWANE, A. R., FAGNY, M., GLASS, K. & QUACKENBUSH, J. 2017. Tissue-aware RNA-Seq processing and normalization for heterogeneous and sparse data. *BMC Bioinformatics,* 18**,** 437.

PERKINS, D. N., PAPPIN, D. J., CREASY, D. M. & COTTRELL, J. S. 1999. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis,* 20**,** 3551-67.

RITCHIE, M. E., PHIPSON, B., WU, D., HU, Y., LAW, C. W., SHI, W. & SMYTH, G. K. 2015. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res,* 43**,** e47.

ROBINSON, M. D. & SMYTH, G. K. 2007. Moderated statistical tests for assessing differences in tag abundance. *Bioinformatics,* 23**,** 2881-7.

SZKLARCZYK, D., GABLE, A. L., LYON, D., JUNGE, A., WYDER, S., HUERTA-CEPAS, J., SIMONOVIC, M., DONCHEVA, N. T., MORRIS, J. H., BORK, P., JENSEN, L. J. & MERING, C. 2019. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Research,* 47**,** D607-D613.

WERNER, T., SWEETMAN, G., SAVITSKI, M. F., MATHIESON, T., BANTSCHEFF, M. & SAVITSKI, M. M. 2014. Ion coalescence of neutron encoded TMT 10-plex reporter ions. *Anal Chem,* 86**,** 3594-601.

YU, G. C., WANG, L. G., HAN, Y. Y. & HE, Q. Y. 2012. clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *Omics-a Journal of Integrative Biology,* 16**,** 284-287.

Table S1, RNA library characteristics, Related to Figure 3

Sample	TotalRead	UniqeMapRead	UniqeMapPercent(%)	MultipleMapRead	MultipleMapPercent(%)	MaptooManyRead	MaptooManyPercent(%)	RNA Input (mg)
IPM0802	73526702	2841876	3.87%	70542999	95.94%	87357	0.12%	0.119
IPM0803	22584637	1782920	7.89%	20739603	91.83%	10603	0.05%	0.553
IPM0804	33165936	1007234	3.04%	32088912	96.75%	23731	0.07%	0.488
IPM0805	17824517	4475434	25.11%	13251675	74.35%	11779	0.07%	2.04
IPM0806	19087508	3056512	16.01%	15978508	83.71%	8484	0.04%	0.262
IPM0807	19658293	597234	3.04%	19016068	96.73%	7678	0.04%	0.105
IPM0808	16833340	476522	2.83%	16310219	96.89%	5657	0.03%	0.105
IPM0809	17923971	11607911	64.76%	6109956	34.09%	20394	0.11%	3.772
IPM0811	6004333	146842	2.45%	5817240	96.88%	3484	0.06%	0.13
IPM0812	12567292	11285465	89.80%	1142589	9.09%	12796	0.10%	3.2
IPM0813	26557668	879111	3.31%	25586048	96.34%	13603	0.05%	0.009
IPM0814	20895259	500503	2.40%	20365054	97.46%	11197	0.05%	0.026
IPM0815	19745837	544408	2.76%	19170512	97.09%	9245	0.05%	0.031
IPM0816	14086428	7255898	51.51%	6677533	47.40%	15624	0.11%	0
IPM0817	21290339	535267	2.51%	20713563	97.29%	8755	0.04%	0.013
IPM0818	24442848	583117	2.39%	23775939	97.27%	38002	0.16%	0.023
IPM0819	57585259	1564100	2.72%	55777291	96.86%	142644	0.25%	0.031
IPM0820	64126950	1582687	2.47%	62355478	97.24%	104014	0.16%	0.033
IPM0821	47374670	1372740	2.90%	45878057	96.84%	77321	0.16%	0.002
IPM0822	38705264	1073660	2.77%	37526650	96.95%	57081	0.15%	0.018
IPM0823	20302717	717891	3.54%	19535396	96.22%	21250	0.10%	0.204
IPM0824	26648882	643800	2.42%	25921753	97.27%	38259	0.14%	0.027
IPM0825	31791867	813828	2.56%	30911375	97.23%	31466	0.10%	0.005
IPM0826	220730424	7850396	3.56%	212520643	96.28%	243993	0.11%	0.006
IPM0827	487299694	17028124	3.49%	469539303	96.36%	434591	0.09%	0.084
IPM0828	495674080	15472269	3.12%	479291366	96.69%	539846	0.11%	0.257
IPM0829	257872145	8337968	3.23%	248888739	96.52%	367106	0.14%	0.043
IPM0830	52238718	1602630	3.07%	50496988	96.67%	86414	0.17%	5.996
IPM0831	50418659	1569256	3.11%	48749241	96.69%	58030	0.12%	9.212
IPM0832	53019251	1596825	3.01%	51296766	96.75%	80243	0.15%	3.524
IPM0833	62256881	1971793	3.17%	60152942	96.62%	85274	0.14%	1.898
IPM0834	16878011	14157443	83.88%	2406262	14.26%	26924	0.16%	14.725
IPM0835	21024466	2507397	11.93%	18458548	87.80%	6010	0.03%	4.519
IPM0836	20431295	2479531	12.14%	17906335	87.64%	6190	0.03%	9.805
IPM0837	12600089	698303	5.54%	11867411	94.19%	7257	0.06%	14.27
IPM0838	4803516	419410	8.73%	4348770	90.53%	2545	0.05%	7.229
IPM0839	12322992	319129	2.59%	11964574	97.09%	8121	0.07%	13.049
IPM0840	10357436	1284482	12.40%	9035195	87.23%	5194	0.05%	0.807
IPM0841	9895503	308278	3.12%	9569011	96.70%	3844	0.04%	0.986
IPM0842	11095736	440858	3.97%	10636410	95.86%	4290	0.04%	7.048
IPM0843	62687436	1692351	2.70%	60915334	97.17%	32240	0.05%	0.234
IPM0844	43269168	1112682	2.57%	42074839	97.24%	23927	0.06%	0.12
IPM0845	61618439	2067065	3.35%	59446731	96.48%	54361	0.09%	0.188
IPM0846	19946502	505495	2.53%	19403119	97.28%	7961	0.04%	0.139
IPM0847	13766300	287727	2.09%	13453084	97.72%	5234	0.04%	0.233
IPM0848	14290972	363880	2.55%	13900381	97.27%	5986	0.04%	0.197
IPM0849	14214873	6127174	43.10%	7769890	54.66%	17782	0.13%	6.656
IPM0850	20608401	559572	2.72%	20022795	97.16%	6695	0.03%	0.257
IPM0851	19319677	560000	2.90%	18740626	97.00%	5862	0.03%	0.224
IPM0852	12455639	315176	2.53%	12106890	97.20%	4126	0.03%	0.157
IPM0853	18306919	527085	2.88%	17740737	96.91%	9150	0.05%	1.212
IPM0854	16221337	522477	3.22%	15665628	96.57%	8055	0.05%	0.214
IPM0855	21137657	4283404	20.26%	16672221	78.87%	20589	0.10%	6.578
IPM0856	22063704	980652	4.44%	21056874	95.44%	10908	0.05%	1.703
IPM0857	20324726	1752364	8.62%	18537771	91.21%	10195	0.05%	2.966
IPM0858	42916606	1187214	2.77%	41668154	97.09%	32262	0.08%	0.07
IPM0859	50069240	3545788	7.08%	46440091	92.75%	27543	0.06%	3.983
IPM0860	30454577	668050	2.19%	29748938	97.68%	14023	0.05%	0.493
IPM0861	21752609	1005150	4.62%	20704878	95.18%	10049	0.05%	1.539
IPM0862	8869254	126835	1.43%	8729065	98.42%	6492	0.07%	0.36
IPM0863	13619910	261754	1.92%	13335003	97.91%	8879	0.07%	0.04
IPM0864	11574071	216112	1.87%	11341552	97.99%	5520	0.05%	0.084
IPM0865	10647826	248229	2.33%	10383391	97.52%	4824	0.05%	0.177
IPM0866	6687209	138236	2.07%	6533061	97.69%	2428	0.04%	0.087
IPM0867	13933453	245334	1.76%	13655152	98.00%	7811	0.06%	0.156
IPM0868	3725277	37399	1.00%	3659738	98.24%	1409	0.04%	0.078
IPM0869	7809393	185361	2.37%	7603645	97.37%	2726	0.03%	0.17
IPM0870	3616290	76017	2.10%	3525802	97.50%	1248	0.03%	0.111
IPM0871	5025825	153013	3.04%	4858187	96.66%	1844	0.04%	0.065

IPM0872	8866570	910183	10.27%	7930074	89.44%	4590	0.05%	0.229
IPM0873	8181686	2128832	26.02%	6016069	73.53%	5810	0.07%	3.201
IPM0874	15168527	415418	2.74%	14708024	96.96%	5063	0.03%	0.046
IPM0875	9240454	154332	1.67%	9048712	97.92%	2780	0.03%	0.04
IPM0876	7069623	75361	1.07%	6963338	98.50%	1784	0.03%	0.004
IPM0877	12469488	255390	2.05%	12185838	97.73%	5484	0.04%	0.081
IPM0878	14071291	305383	2.17%	13720389	97.51%	8432	0.06%	0.296
IPM0879	12598603	1300202	10.32%	11271515	89.47%	4095	0.03%	1.044
IPM0880	14373494	608152	4.23%	13741497	95.60%	4006	0.03%	0.076
IPM0881	17958639	574899	3.20%	17353431	96.63%	10576	0.06%	0.131
IPM0882	13894618	268993	1.94%	13595112	97.84%	4968	0.04%	0.099
IPM0883	5583577	106984	1.92%	5447718	97.57%	1683	0.03%	0.81
IPM0890	37451770	998022	2.66%	36418078	97.24%	18832	0.05%	0.209
IPM0891	31058853	794103	2.56%	30232085	97.34%	14243	0.05%	0.155
IPM0892	16580033	397188	2.40%	16155252	97.44%	11823	0.07%	0.138
IPM0893	16070169	717622	4.47%	15313143	95.29%	6376	0.04%	0.207
IPM0894	16195763	443370	2.74%	15725136	97.09%	9178	0.06%	0.094
IPM0895	27386159	1030445	3.76%	26302107	96.04%	14005	0.05%	0.182
IPM0896	18743065	1191990	6.36%	17519566	93.47%	8297	0.04%	0.294
IPM0897	22476344	601354	2.68%	21840642	97.17%	10989	0.05%	0.05
IPM0898	21477768	571675	2.66%	20875224	97.19%	9674	0.05%	0.079
IPM0899	23432707	1182933	5.05%	22201706	94.75%	10229	0.04%	0.157
IPM0900	16683802	579733	3.47%	16069798	96.32%	7011	0.04%	0.085
IPM0901	17240772	1190723	6.91%	16014873	92.89%	6713	0.04%	0.136
IPM0902	3285	432	13.15%	2844	86.58%	0	0.00%	0.632
IPM0903	2442	391	16.01%	2044	83.70%	0	0.00%	0.104
IPM0904	18414293	2746153	14.91%	15610730	84.78%	8723	0.05%	0.081
IPM0905	19674860	8676628	44.10%	10916860	55.49%	9816	0.05%	0.722
IPM0906	21474264	613609	2.86%	20822712	96.97%	12631	0.06%	0.024

Table S2, Genes with Modules, Related to Figure 5

Module 1 gene list												Module 2 gene list
ABCA1	B4GALT1	COQ6	EXO5	HLA-B	LINC02649	NDRG1	PNISR	RSAD2	SPTY2D1	TTC9	A2M
ABCA3	B4GALT5	CORO1A	EXOC8	HLA-C	LITAF	NDST1	PNKD	RSC1A1	SQOR	TTLL3	ALAS2
ABCB6	BAALC	CORO1C	EXOSC5	HLA-DQA1	LMNB1	NDUFA6-DT	PNPLA6	RSRC2	SQSTM1	TTLL4	ATP1B1
ABCG1	BACH1	CORO7	EYA3	HLA-E	LMO7	NDUFAF5	PNRC1	RSRP1	SRC	TUBB2B	COL20A1
ABHD2	BACH2	COTL1	F11R	HLA-F	LMTK2	NDUFAF8	POGLUT3	RTF2	SRD5A3	TUG1	CPED1
ABHD5	BASP1	CPD	FADS1	HM13	LOC400499	NDUFS4	POLR2A	RTN3	SREBF1	TUT7	EPHA1
ABR	BAZ1A	CPEB4	FAM111B	HMBS	LONP1	NEAT1	POLR2I	RTN4	SRGN	TXNDC5	EPS8
ABTB1	BAZ2A	CPPED1	FAM117B	HMGA1	LPAR2	NECAB3	POLR3B	RTN4IP1	SRRM2	TXNIP	ERICH3
ACAA1	BAZ2B	CPQ	FAM120C	HMGN3	LPAR5	NECAP1	POM121	RUBCNL	SSH2	TYK2	FCGBP
ACACB	BBS12	CR1	FAM126B	HNRNPA2B1	LPCAT1	NEDD9	POM121C	RXRA	SSR2	TYMP	GPR34
ACADVL	BCAN	CREB1	FAM157A	HNRNPC	LRCH4	NEIL2	PON2	RXYLT1	SSU72	TYROBP	HBA1
ACAP2	BCAS4	CREB5	FAM160A2	HNRNPF	LRFN4	NEU1	POR	RYBP	SSX2IP	TYW3	HBA2
ACAT1	BCKDHB	CREBBP	FAM193B	HNRNPH2	LRG1	NF1	PORCN	RYR1	ST14	UACA	HBB
ACBD4	BCKDK	CREBRF	FAM210A	HNRNPK	LRMDA	NFAM1	POU2F2	S100A11	ST20	UBA1	HBG1
ACIN1	BCL10	CRISPLD2	FAM210B	HNRNPM	LRP1	NFAT5	PPCDC	S100A12	ST3GAL1	UBAP1	HBG2
ACOT2	BCL2A1	CRK	FAM214B	HNRNPU	LRP10	NFE2L2	PPFIA1	S100A6	ST3GAL2	UBC	HLA-DQA1
ACP2	BCL3	CROT	FAM49B	HOMER2	LRP3	NFIA	PPIF	S100A8	ST8SIA4	UBE2D1	HTRA1
ACSL1	BCL6	CRTC2	FAM53C	HOPX	LRPAP1	NFIL3	PPIL3	S100A9	STAC3	UBE2D3	IGFBP7
ACSL4	BDH1	CRYBG1	FAM89A	HPCAL1	LRRC25	NFIX	PPM1F	SAFB	STAG3	UBE2R2	LPL
ACSL5	BDH2	CSF1	FAM91A1	HPF1	LRRFIP1	NFKB1	PPM1H	SAFB2	STAT1	UBE2W	LSAMP
ACTB	BEST1	CSF2RB	FANCE	HPSE	LRRK2	NFKB2	PPP1CB	SAMD4B	STAT2	UBE4A	RNA5-8SN1
ACTG1	BEX3	CSF3R	FANCF	HS3ST3B1	LSP1	NFKBIA	PPP1R10	SAMD8	STAT3	UBR4	RNA5-8SN2
ACTN1	BHLHE40	CSGALNACT2	FASTKD3	HSD17B11	LST1	NFKBID	PPP1R12C	SAMD9	STAT5A	UBXN2B	RNA5-8SN3
ACTN4	BID	CSK	FBRS	HSD17B4	LTA4H	NFKBIE	PPP1R15A	SAMD9L	STAT5B	UNC119	RNA5-8SN5
ACTR2	BIN1	CSNK1A1	FBXL5	HSF2	LUCAT1	NFKBIZ	PPP1R15B	SAMSN1	STAT6	UNC13D	SEMA5A
ACTR3	BIRC3	CSNK1D	FBXO32	HSP90AA1	LXN	NFU1	PPP1R18	SAPCD2	STEAP4	UPB1	SIGLEC8
ACVR1B	BIRC6	CSNK1G2	FBXO38	HSP90B1	LYN	NFXL1	PPP1R3B	SARS2	STK10	UPF1	TREM2
ACY1	BIVM	CSRNP1	FBXO4	HSPA1A	LYPLAL1	NIBAN1	PPP1R3G	SASH3	STK35	UPP1	
ADAM10	BMP1	CSTB	FBXW7	HSPA1B	LYST	NIBAN2	PPP2R2D	SAT1	STK4	USB1	
ADAM17	BMS1P1	CTBS	FBXW8	HSPA5	LYVE1	NID1	PPP4C	SBDSP1	STK40	USF3	
ADAM8	BOD1L1	CTDP1	FCAR	HSPA6	LYZ	NINJ1	PPP4R2	SBF2	STMN3	USP15	
ADAMTSL4-AS1	BOLA3	CTDSP1	FCER1G	HSPH1	LZTFL1	NIPBL	PRADC1	SBNO1	STON2	USP3	
ADAR	BPGM	CTNNBIP1	FCGR1B	HTR7	M6PR	NLRC4	PRAM1	SBNO2	STRADB	USP32	
ADCY3	BPHL	CTPS2	FCGR2A	ICAM1	MACROH2A1	NLRC5	PRDM11	SCAF1	STX11	USP34	
ADCY6	BRAF	CTSA	FCGR2B	ICAM3	MAEA	NMI	PRDM2	SCAF11	STX3	USP4	
ADGRB1	BRD2	CTSB	FCGR3A	ID3	MAFB	NMT2	PREX1	SCAMP5	STX4	USP54	
ADGRE1	BRD4	CTSC	FCGR3B	IDS	MAFF	NOD2	PRIM1	SCARF1	STX6	USP8	
ADGRE2	BRI3	CTSD	FCN1	IDUA	MAFG	NOP10	PRIMPOL	SCARNA10	STXBP2	USP9X	
ADGRE3	BRIP1	CTSF	FCRL1	IER2	MAFK	NOTCH1	PRKAR1A	SCARNA13	STXBP4	VASP	
ADGRE4P	BRPF3	CTSK	FDXR	IER3	MAGEF1	NOTCH2	PRKAR1B	SCIN	SULF2	VAV1	
ADGRE5	BRWD3	CTSL	FECH	IER5	MAGEH1	NOTCH2NLA	PRKCB	SCN3A	SUN2	VCAN	
ADGRG3	BST1	CTSS	FERMT3	IFI16	MALAT1	NPC1	PRKCD	SCN9A	SUPT3H	VCPIP1	
ADGRL1	BTBD19	CTSZ	FFAR2	IFI30	MAML1	NPEPPS	PRKCQ-AS1	SDCBP	SUPT5H	VDR	
ADM	BTG2	CUX1	FGD3	IFIH1	MAML2	NPIPB3	PRKCZ	SDF4	SUPT6H	VEGFA	
ADORA2A	BTN2A1	CUX2	FGFR1	IFIT1	MAN1A1	NPIPB5	PRKD2	SEC14L1	SUSD6	VHL	
ADPGK	BTN2A2	CWC25	FGFRL1	IFIT2	MAP1B	NPLOC4	PRLR	SEC16A	SVIL	VIM	
AFAP1	BTN3A1	CXCL1	FGR	IFIT3	MAP1LC3B	NR1H3	PRMT9	SEC22B	SVIP	VMP1	
AFF1	BYSL	CXCL16	FILIP1L	IFITM1	MAP1S	NR3C1	PROK2	SEC24A	SYK	VNN1	
AFF4	BZW1	CXCL2	FKBP14	IFITM2	MAP2	NR4A2	PROS1	SEC61A1	SYNPO	VNN2	
AFMID	C10orf88	CXCL5	FKBP15	IFITM3	MAP2K3	NR4A3	PRR13	SECTM1	SZT2	VNN3	
AFTPH	C11orf95	CXCL8	FKBP9	IFNAR1	MAP3K10	NRBF2	PRR14L	SEL1L	TAB2	VOPP1	
AGAP3	C12orf29	CXCR1	FLCN	IFNGR1	MAP3K11	NRBP1	PRRC2A	SELL	TACO1	VPS13B	
AGBL5	C15orf39	CXCR2	FLII	IFNGR2	MAP3K2	NRDC	PRRC2C	SEMA4A	TAF7	VPS13D	
AGO2	C15orf48	CXCR4	FLNA	IFNLR1	MAP3K8	NRXN1	PRUNE2	SEMA4B	TAF9B	VPS18	
AGO4	C16orf72	CYB5R4	FLOT1	IGF2R	MAP4K4	NSDHL	PRXL2A	SEMA4D	TAGAP	VPS37C	
AGTPBP1	C17orf75	CYBA	FLOT2	IGFBP5	MAP7D1	NSMAF	PRXL2B	SEMA6B	TAGLN2	VPS4B	
AGTRAP	C1orf198	CYBB	FLT3LG	IGIP	MAPK13	NT5C2	PSAP	SEPTIN3	TALDO1	VPS8	
AHCTF1	C1orf50	CYBC1	FLYWCH2	IGLL5	MAPK14	NT5DC1	PSAT1	SEPTIN7P2	TANC2	VPS9D1	
AHNAK	C1orf54	CYFIP2	FMNL1	IGSF6	MAPK6	NUCB1	PSD	SERF2	TANK	VSIR	
AHR	C1RL	CYP27A1	FN1	IKBKB	MAPK8IP1	NUDT12	PSEN1	SERINC1	TAOK1	WAC	
AIF1	C2CD3	CYP2U1	FNBP1L	IL10RB	MAPKAPK2	NUDT15	PSMG1	SERP1	TAP1	WARS1	
AIFM3	C3AR1	CYSTM1	FNDC3B	IL11RA	MARCHF6	NUFIP2	PSTPIP1	SERPINA1	TAP2	WAS	
AK1	C4orf3	CYTH4	FNIP1	IL17RA	MARCHF7	NUMB	PSTPIP2	SERPINB1	TAPBP	WASHC2A	
AK4	C5AR1	CYTIP	FNIP2	IL18BP	MARCHF9	NUP35	PTAFR	SERPINB9	TATDN1	WASHC2C	
AK6	C6orf62	DAAM1	FOS	IL18RAP	MARCKS	NUP58	PTBP3	SERPING1	TBC1D14	WASHC4	
AKAP10	C6orf89	DANCR	FOSL2	IL1A	MARCO	NUP98	PTEN	SETD5	TBC1D2	WBP2	
AKAP13	C9orf72	DAPK1	FOXO3	IL1B	MARF1	NXF1	PTGDR	SETD6	TBC1D22A	WDFY3	
AKAP17A	C9orf85	DAPP1	FOXRED2	IL1R1	MARK3	NYNRIN	PTGER4	SETX	TBC1D9B	WDR1	
AKAP6	CA5BP1	DARS1	FPR1	IL1R2	MARS2	OAS3	PTGFRN	SF1	TBK1	WDR26	
AKAP7	CABLES1	DAZAP2	FPR2	IL1RAP	MAST3	OAZ2	PTGR1	SF3A1	TBL1X	WDR53	
AKIRIN2	CACNB1	DBNL	FRAT2	IL1RN	MAU2	OCEL1	PTGS2	SF3B1	TBL1XR1	WFS1	
AKNA	CADM1	DCANP1	FRMD3	IL2RG	MBD6	OCIAD2	PTK2B	SF3B2	TCAF2	WIPF1	
ALDH4A1	CAMKK1	DCP2	FTH1	IL4I1	MBLAC2	OCRL	PTK7	SFXN2	TCEAL9	WIPF2	
ALDH5A1	CAMKK2	DCUN1D3	FUOM	IL4R	MBOAT7	OGA	PTP4A2	SFXN4	TCIRG1	WNK1	
ALDH6A1	CAMTA2	DDIT3	FURIN	IL6R	MBP	OGFR	PTPDC1	SGK1	TCTN1	WSB1	
ALDOA	CANT1	DDIT4	FUS	IL6ST	MBTPS2	OLFML2B	PTPN1	SGSH	TDP2	WTAP	
ALDOC	CANX	DDX17	FUT10	ILRUN	MCEMP1	OLFML3	PTPN12	SGTB	TECPR2	WWC3	
ALKBH4	CAP1	DDX39B	FYB1	IMPDH1	MCF2L	OPHN1	PTPN2	SH2D3C	TEF	WWOX	
ALKBH7	CAPS	DDX3X	FZD1	INO80B	MCF2L2	OPLAH	PTPN23	SH3BGRL3	TESC	WWP2	
ALKBH8	CAPZA1	DDX58	G0S2	INPP5D	MCL1	ORAI2	PTPN6	SH3BP2	TET2	XBP1	
ALOX5	CAPZA2	DDX6	GAB2	IP6K1	MCOLN1	ORM1	PTPRC	SH3BP4	TET3	XIAP	
ALOX5AP	CARD16	DDX60L	GABARAP	IPMK	MCTP1	OS9	PTPRE	SH3BP5	TFB1M	XPA	

ALPK1	CARD8	DEDD2	GABARAPL2	IQGAP1	MCTP2	OSBP	PTPRJ	SH3D21	TFE3	XPO6
ALPL	CARMIL1	DENND2B	GADD45B	IQSEC1	MDM2	OSBPL2	PTRH1	SH3PXD2B	TFEC	XRN1
ALS2CL	CASC3	DENND3	GAK	IQSEC2	MDM4	OSBPL8	PTRHD1	SHC3	TFRC	YEATS4
AMDHD2	CASP1	DENND4B	GALNS	IRAK1	ME1	OSER1	PTS	SHISA5	TG	YIPF3
AMIGO2	CASP4	DENND5A	GAPDH	IRAK2	MECP2	OSM	PTTG1IP	SHKBP1	TGFB1	YPEL3
AMOTL1	CASP5	DEXI	GAPT	IRAK3	MED13L	OXSR1	PXN	SHMT1	TGM2	YTHDF3
AMPD2	CASS4	DGAT2	GBE1	IRF1	MED15	P2RX1	PYCR1	SHOC2	TGOLN2	YWHAZ
AMPD3	CAST	DHRS4	GBP1	IRF2	MED28	P2RX4	QKI	SIAE	THBD	ZBED6
ANAPC10	CBL	DHRS4-AS1	GBP2	IRF2BP2	MEFV	P2RY13	QPCT	SIGLEC10	THBS1	ZBTB17
ANK1	CCDC112	DHX34	GBP4	IRF7	MEGF9	P2RY2	QSOX1	SIGLEC11	THEM4	ZBTB42
ANK2	CCDC138	DIAPH1	GBP5	IRF9	MELTF	P4HA1	R3HDM4	SIGLEC14	THEM6	ZBTB7B
ANKAR	CCDC141	DICER1	GCA	IRS1	METTL2A	P4HB	RAB10	SIGLEC16	THEMIS2	ZC3H11A
ANKDD1A	CCDC15	DIP2B	GCC1	IRS2	METTL5	PACRGL	RAB11FIP1	SIGLEC5	THOC3	ZC3H12A
ANKH	CCDC57	DIP2C	GCH1	ISG15	MFSD14A	PACSIN2	RAB1A	SIGLEC9	THUMPD2	ZC3H18
ANKHD1	CCDC71L	DIPK1A	GDF11	ISG20	MFSD2A	PAG1	RAB20	SIK3	TICAM1	ZCCHC2
ANKRD11	CCDC86	DISP1	GDI1	ISOC1	MGA	PAK1	RAB21	SIPA1	TIFA	ZCCHC24
ANKRD12	CCDC88B	DLG5	GGA1	IST1	MGAM	PAK2	RAB24	SIPA1L1	TIGD2	ZDHHC14
ANKRD13A	CCDC93	DMAC1	GGA3	ITCH	MGAT1	PAPOLA	RAB3D	SIPA1L2	TIGD5	ZDHHC18
ANKRD22	CCL18	DMXL2	GGT7	ITGA3	MGLL	PAQR7	RAB4A	SIRPA	TIMM21	ZDHHC5
ANKRD33B	CCL2	DNAAF2	GGTA1P	ITGA5	MICALL2	PARP10	RAB5A	SIRPB1	TIMM9	ZEB2
ANKRD46	CCL3	DNAJB1	GIGYF1	ITGAM	MIDN	PARP14	RAB5C	SIRPB2	TIMP1	ZFAND3
ANKRD9	CCL3L1	DNAJB6	GIGYF2	ITGAX	MIER1	PARP2	RAB7A	SIRT7	TKT	ZFAND5
ANKS6	CCL4	DNAJC19	GK	ITGB2	MIF	PARP6	RAB7B	SKAP1	TLCD4	ZFC3H1
ANKZF1	CCL4L2	DNAJC3	GLG1	ITIH4	MIPEP	PARP9	RAB8B	SKAP2	TLE3	ZFP30
ANOS1	CCND1	DNAJC5	GLT1D1	ITM2B	MIR210HG	PATL1	RABGEF1	SKIL	TLE4	ZFP36
ANP32A	CCNG2	DNAL1	GLUL	ITPR2	MKNK1	PAXBP1-AS1	RAC2	SLA	TLN1	ZFP36L1
ANPEP	CCNK	DNLZ	GMFG	ITPRIP	MKNK2	PBX2	RAD21	SLAIN1	TLN2	ZFYVE16
ANTXR1	CCNL1	DNM2	GMIP	ITSN2	MLLT6	PBX3	RAD51C	SLAMF7	TLR1	ZFYVE26
ANTXR2	CCNL2	DNM3	GNA15	IVNS1ABP	MMAA	PCBD1	RAF1	SLAMF8	TLR2	ZHX2
ANXA3	CCR1	DOCK2	GNAI2	JAK1	MME	PCBP1	RALB	SLC11A1	TLR4	ZIK1
ANXA5	CCRL2	DOCK4	GNAI3	JAK2	MMP12	PCBP1-AS1	RALGAPA2	SLC11A2	TLR6	ZKSCAN2
AP1G1	CD14	DOCK5	GNB1	JAK3	MMP14	PCBP2	RALGDS	SLC12A6	TLR8	ZMIZ1
AP3D1	CD163	DOCK8	GNB2	JAML	MMP19	PCDH12	RANBP2	SLC12A9	TM7SF2	ZMIZ2
AP5B1	CD177	DOK3	GNG2	JAZF1	MMP2	PCF11	RANGRF	SLC15A3	TM9SF4	ZMYND19
AP5Z1	CD22	DOT1L	GNG7	JCHAIN	MMP25	PCGF6	RAP1B	SLC16A1-AS1	TMBIM1	ZNF10
APBA1	CD24	DPH5	GNLY	JMJD1C	MMP9	PCNX1	RAP2C	SLC16A3	TMBIM4	ZNF117
APBA3	CD274	DPH6	GNPDA1	JMJD6	MNDA	PCNX3	RAPGEF1	SLC16A6	TMBIM6	ZNF132
APBB1IP	CD276	DPY30	GNPDA2	JUN	MNT	PDGFC	RARA	SLC19A1	TMCC1	ZNF16
APBB3	CD300A	DPYD	GNS	JUNB	MOB1A	PDIA3	RARG	SLC19A2	TMCC2	ZNF160
APLF	CD300E	DRAM1	GOLGA2P7	KANK2	MOB3A	PDK1	RASAL2	SLC20A1	TMCC3	ZNF207
APLP2	CD36	DSC2	GON7	KAT5	MOB3C	PDLIM7	RASGRP4	SLC24A1	TMEM127	ZNF217
APOBEC3A	CD37	DSE	GORAB	KAT6A	MON1B	PDXK	RASSF2	SLC25A13	TMEM140	ZNF23
APOBR	CD44	DTD2	GOT1	KATNBL1	MON2	PDZD11	RASSF3	SLC25A23	TMEM154	ZNF24
APOC2	CD48	DTWD2	GPAT4	KBTBD6	MOSPD1	PECAM1	RASSF4	SLC25A37	TMEM159	ZNF267
APOL6	CD53	DTX2	GPATCH2L	KCNAB2	MPC1	PELI1	RASSF5	SLC25A38	TMEM165	ZNF274
AQP9	CD55	DTX3	GPBP1	KCNJ15	MPEG1	PER1	RBCK1	SLC25A43	TMEM176B	ZNF281
ARAP1	CD58	DTX3L	GPBP1L1	KCNJ2	MPP6	PEX10	RBFA	SLC26A11	TMEM184B	ZNF286A
ARAP3	CD63	DUSP1	GPCPD1	KCNK13	MR1	PEX12	RBM23	SLC2A1	TMEM201	ZNF292
AREL1	CD68	DUSP16	GPHN	KCTD17	MRPL40	PEX3	RBM25	SLC2A3	TMEM242	ZNF316
ARF1	CD82	DUSP2	GPI	KCTD3	MRPL58	PFKFB2	RBM33	SLC2A6	TMEM254	ZNF32
ARFIP1	CD93	DUSP3	GPM6B	KDM2A	MRPS17	PFKFB3	RBM39	SLC35D2	TMEM259	ZNF324B
ARHGAP25	CDC42	DUSP6	GPN3	KDM3A	MRPS28	PFKFB4	RBM47	SLC35E1	TMEM273	ZNF329
ARHGAP26	CDC42EP3	DYRK1A	GPR108	KDM3B	MRPS31P5	PFKL	RBM5	SLC35E3	TMEM42	ZNF331
ARHGAP27	CDC42SE1	DYSF	GPR65	KDM4B	MRPS33	PFN1	RBPJ	SLC38A10	TMEM64	ZNF333
ARHGAP30	CDC42SE2	E2F3	GPR84	KDM5C	MRVI1	PGAM1	RC3H1	SLC39A14	TMEM71	ZNF35
ARHGAP33	CDK18	EAF1	GPRASP1	KDM6A	MS4A4E	PGAP1	RDH10	SLC39A8	TMEM86B	ZNF354C
ARHGAP45	CDK5R1	EBAG9	GPSM1	KDM6B	MSANTD2	PGAP6	REC8	SLC3A2	TMF1	ZNF358
ARHGAP9	CDKN1A	ECE1	GPSM2	KDM7A	MSL1	PGBD2	RECK	SLC43A1	TMOD3	ZNF382
ARHGDIA	CDKN2AIPNL	ECI1	GPSM3	KDM8	MSMO1	PGBD4	REL	SLC43A2	TMSB4X	ZNF395
ARHGDIB	CDV3	ECI2	GRAMD1A	KHNYN	MSN	PGK1	RELA	SLC43A3	TMTC4	ZNF438
ARHGEF1	CEACAM1	EDEM1	GRASP	KIAA0040	MSRB1	PGM2L1	RELB	SLC45A4	TNFAIP2	ZNF462
ARHGEF2	CEACAM3	EEF1AKMT3	GRB2	KIAA0232	MTF1	PGS1	RELT	SLC5A3	TNFAIP3	ZNF470
ARHGEF9	CEBPB	EEF1E1	GRINA	KIAA0513	MTFR1	PHACTR1	RERE	SLC6A6	TNFAIP6	ZNF48
ARID3A	CELF1	EEF2KMT	GRK2	KIF1A	MTHFS	PHC2	RESF1	SLC7A7	TNFRSF10C	ZNF527
ARID3B	CENPL	EFCAB13	GRK6	KIF24	MTMR11	PHF12	RFFL	SLC9A7P1	TNFRSF10D	ZNF542P
ARID4B	CEP170	EFHD2	GRN	KIF5C	MTMR3	PHF20L1	RFK	SLC9A8	TNFRSF14	ZNF546
ARL4A	CEP170B	EGLN1	GSAP	KIFC3	MTMR4	PHF21A	RGL2	SLC9B2	TNFRSF1A	ZNF550
ARMCX2	CEP57L1	EGLN2	GSDMD	KLF10	MTSS2	PI3	RGPD2	SLCO4A1	TNFRSF1B	ZNF568
ARMH1	CEP76	EGR3	GSK3A	KLF6	MTURN	PIAS1	RGS19	SLK	TNFRSF8	ZNF570
ARNT2	CERK	EHBP1L1	GSPT2	KLHL2	MVP	PICALM	RHBDD2	SLPI	TNFSF10	ZNF573
ARPC1B	CERT1	EHD1	GSTCD	KLHL20	MX1	PIGU	RHBDF2	SLU7	TNFSF14	ZNF593
ARPC2	CFAP298	EIF1	GSTM2	KLHL21	MX2	PIGW	RHOA	SLX1B-SULT1A4	TNIP1	ZNF608
ARPC3	CFL1	EIF3A	GTF2H4	KMT2B	MXD1	PIK3AP1	RHOBTB2	SMAD6	TNIP2	ZNF618
ARPC4	CFL2	EIF4G1	GTF2H5	KMT2C	MYADM	PIK3CB	RHOG	SMAP2	TNR	ZNF626
ARPC5	CFLAR	EIF4H	GTPBP1	KMT2D	MYD88	PIK3CD	RHOQ	SMCHD1	TNRC18	ZNF641
ARRB2	CFP	EIF5	GTPBP2	KMT2E	MYEF2	PIK3R3	RICTOR	SMCR8	TNS1	ZNF646
ARRDC2	CHCHD10	ELAC1	GTPBP8	KMT5C	MYH9	PIK3R5	RILPL2	SMG1	TNS3	ZNF704
ARRDC3	CHD2	ELF1	GUK1	KPNA4	MYL12A	PILRA	RIPOR1	SMG1P1	TOM1	ZNF710
ARSG	CHEK2	ELF2	GUSBP1	KPNB1	MYL6	PIM1	RIPOR2	SMG1P3	TOR1AIP2	ZNF710-AS1
ARV1	CHFR	ELF4	GUSBP1	KRAS	MYLIP	PIM2	RIT1	SMG7	TOX4	ZNF718
ASAH1	CHI3L1	ELL	GUSBP3	KRBA2	MYO1B	PIM3	RLF	SMIM25	TP53BP1	ZNF738
ASF1A	CHIT1	ELMSAN1	H2AC19	KREMEN1	MYO1F	PIN4	RLIM	SMYD3	TP53I3	ZNF74
ASGR1	CHMP1A	EMD	H2BC18	KRT23	MYO1G	PISD	RMND5A	SNAP23	TPD52L2	ZNF773
ATF5	CHMP2A	EMILIN2	H3-3A	LAMB3	MYO5B	PITPNA	RMRP	SNCA	TPI1	ZNF777

ATG10	CHMP3	EML2	H3-3B	LAMP1	MYO9B	PJA2	RN7SK	SNN	TPM2	ZNF792
ATG16L2	CHMP4B	EML4	H3C6	LAMP2	MZB1	PKD1P6-NPIPP1	RN7SL1	SNORA48	TPM3	ZNF799
ATG2A	CHN2	ENAH	H4C1	LAPTM5	N4BP1	PKM	RN7SL2	SNORA73A	TPM4	ZNF804A
ATG3	CHRNB1	ENDOD1	HAAO	LARP4B	N4BP2L2	PKN1	RNASEK	SNORA73B	TPP1	ZNF813
ATG7	CHST11	ENO1	HADH	LASP1	N6AMT1	PLA2G4A	RNF13	SNORD17	TPRA1	ZNF846
ATIC	CHST14	ENTPD4	HAL	LAT2	NAA20	PLAG1	RNF130	SNORD3A	TRAF1	ZNFX1
ATP11A	CHST15	EP300	HBP1	LBR	NAB1	PLAGL2	RNF144A	SNTA1	TRAF3	ZSCAN12
ATP11B	CHSY1	EPB41L1	HCAR2	LCP1	NABP1	PLAU	RNF144B	SNX10	TRAFD1	ZSCAN2
ATP13A3	CIC	EPB41L2	HCAR3	LCP2	NADK	PLAUR	RNF145	SNX18	TRAK1	ZSCAN5A
ATP1A1	CKAP4	EPB41L5	HCK	LDHA	NAGK	PLCG2	RNF149	SNX20	TRANK1	ZSWIM4
ATP2A3	CKLF	EPG5	HCLS1	LENG8	NAIP	PLCL1	RNF157	SNX21	TREM1	ZSWIM8
ATP2B1	CLCN7	EPHB3	HCP5	LEP	NAMPT	PLEC	RNF19B	SNX24	TREML2	ZXDA
ATP6AP1	CLEC4E	EPHX1	HDAC11	LEPROT	NAPA	PLEK	RNF213	SNX27	TRIAP1	ZYX
ATP6V0B	CLEC7A	EPS15L1	HDAC4	LGALS3	NAPEPLD	PLEKHA2	RNF24	SOCS3	TRIB1	ZZEF1
ATP6V0C	CLIC1	EPSTI1	HDAC7	LGALS8	NARF	PLEKHB1	RNF40	SOD2	TRIM25	
ATP6V0D1	CLIC4	ERBIN	HDDC3	LIAS	NARS2	PLEKHB2	RNF8	SON	TRIM33	
ATP6V0E2	CLK1	ERCC8	HDGFL3	LILRA1	NBEA	PLEKHG3	RNMT	SORBS1	TRIM38	
ATP6V1A	CLK3	ERF	HECA	LILRA2	NBEAL2	PLEKHM1	RNU1-1	SORL1	TRIM56	
ATP6V1B2	CLN3	ERI3	HELB	LILRA5	NBL1	PLEKHM2	RNU1-2	SOWAHC	TRIP12	
ATP6V1C1	CLUAP1	ERN1	HELZ2	LILRA6	NBN	PLEKHO1	RNU1-3	SOX12	TRPC4AP	
ATP6V1E1	CMAS	ERO1A	HEXIM1	LILRB1	NBPF14	PLEKHO2	RNU1-4	SP1	TRPM2	
ATP6V1E2	CMIP	ERV3-1	HGF	LILRB2	NBPF19	PLIN2	RNU4-1	SP100	TRPM7	
ATP6V1G1	CMSS1	ESCO1	HGS	LILRB3	NBR1	PLIN4	RNU4-2	SP110	TRPV1	
ATXN1	CMTM6	ESR1	HHAT	LILRB4	NCAM1	PLIN5	RNVU1-18	SP3	TRUB1	
ATXN7	CNBP	ETAA1	HIF1A	LIMK2	NCF1	PLK3	RNVU1-29	SPAG9	TSC22D3	
AVL9	CNN2	ETF1	HILPDA	LINC00339	NCF1B	PLP1	ROCK1	SPATC1	TSC22D4	
AVPI1	CNN3	ETFRF1	HINT2	LINC00476	NCF1C	PLP2	RP2	SPECC1L-ADORA2A	TSEN2	
AXL	CNOT1	ETS2	HIP1	LINC00528	NCF2	PLPP3	RPA3	SPEN	TSEN34	
AZIN1	CNOT3	ETV6	HIPK3	LINC00865	NCF4	PLPPR2	RPP25	SPG11	TSNARE1	
B2M	COL18A1	EVI2A	HIVEP1	LINC00893	NCOA1	PLSCR1	RPPH1	SPHK1	TSPAN14	
B3GLCT	COL1A1	EVI2B	HIVEP2	LINC00937	NCOR2	PLTP	RPS6KA1	SPI1	TSPAN5	
B3GNT5	COL8A2	EWSR1	HK2	LINC01184	NCS1	PLXNC1	RPUSD2	SPOPL	TTC28	
B3GNT8	COMMD3-BMI1	EXD2	HK3	LINC01270	NCSTN	PMEPA1	RRAGC	SPPL2A	TTC30A	

Table S3, Genes from Proteogenomic Integration, Related to Figure 7

Gene_Name	Gene_Mean	Gene_std	Sample_Type	Assay
A2M	39.2	11.73502423	NPIH	RNA
A2M	135.56	55.5019911	NonPaeni_PIH	RNA
A2M	97.947368	26.94361	Paeni_PIH	RNA
ACTN1	1.2285714	0.41228306	NPIH	RNA
ACTN1	13.12	5.2187001	NonPaeni_PIH	RNA
ACTN1	175.842105	55.506921	Paeni_PIH	RNA
ANPEP	0.3428571	0.14333254	NPIH	RNA
ANPEP	6.88	2.7985271	NonPaeni_PIH	RNA
ANPEP	216.921053	77.077825	Paeni_PIH	RNA
APLP2	9.8	2.73884462	NPIH	RNA
APLP2	61.72	25.4472921	NonPaeni_PIH	RNA
APLP2	672.026316	211.962453	Paeni_PIH	RNA
B4GALT1	1.5714286	0.42962833	NPIH	RNA
B4GALT1	11.84	4.6847313	NonPaeni_PIH	RNA
B4GALT1	140.236842	40.097342	Paeni_PIH	RNA
CAP1	7.6	2.19494368	NPIH	RNA
CAP1	71.72	29.337323	NonPaeni_PIH	RNA
CAP1	477.342105	133.249663	Paeni_PIH	RNA
CD14	14.6571429	4.56295345	NPIH	RNA
CD14	199.96	76.7045376	NonPaeni_PIH	RNA
CD14	1192.842105	372.150024	Paeni_PIH	RNA
CFB	0.1714286	0.1227057	NPIH	RNA
CFB	0.8	0.2882307	NonPaeni_PIH	RNA
CFB	4.157895	1.329377	Paeni_PIH	RNA
CHI3L1	0.4571429	0.15088552	NPIH	RNA
CHI3L1	3.24	1.2785569	NonPaeni_PIH	RNA
CHI3L1	327.763158	116.682184	Paeni_PIH	RNA
ENO1	12.9428571	4.1950354	NPIH	RNA
ENO1	137.36	54.2521146	NonPaeni_PIH	RNA
ENO1	1457.210526	443.381282	Paeni_PIH	RNA
FLNA	9.8285714	2.77330979	NPIH	RNA
FLNA	164.12	65.5032899	NonPaeni_PIH	RNA
FLNA	3886.131579	1720.563964	Paeni_PIH	RNA
GAPDH	26.0857143	8.6329354	NPIH	RNA
GAPDH	235.04	86.4910752	NonPaeni_PIH	RNA
GAPDH	2387.684211	733.492259	Paeni_PIH	RNA
HK3	0.8571429	0.36545874	NPIH	RNA
HK3	14.08	5.3380838	NonPaeni_PIH	RNA
HK3	102.421053	30.000464	Paeni_PIH	RNA
HSPA1A	4.7428571	2.74825857	NPIH	RNA
HSPA1A	24.08	9.0769226	NonPaeni_PIH	RNA
HSPA1A	979.842105	533.003675	Paeni_PIH	RNA
ITIH4	0.2285714	0.15979011	NPIH	RNA
ITIH4	1.96	0.8457677	NonPaeni_PIH	RNA
ITIH4	5.263158	1.918615	Paeni_PIH	RNA
LCP1	11.3714286	3.31425151	NPIH	RNA
LCP1	179.56	73.4673361	NonPaeni_PIH	RNA
LCP1	2150.421053	590.964767	Paeni_PIH	RNA
LRG1	0.1714286	0.08438117	NPIH	RNA
LRG1	1.88	0.86881	NonPaeni_PIH	RNA
LRG1	407.315789	141.803951	Paeni_PIH	RNA
LTA4H	1.8	0.68405885	NPIH	RNA
LTA4H	32.32	14.0998331	NonPaeni_PIH	RNA
LTA4H	166.078947	37.400004	Paeni_PIH	RNA
MMP9	0.3714286	0.16468206	NPIH	RNA
MMP9	34.32	16.2950412	NonPaeni_PIH	RNA
MMP9	312.447368	107.362986	Paeni_PIH	RNA
MSN	9.2571429	2.67627521	NPIH	RNA
MSN	127.48	49.7424481	NonPaeni_PIH	RNA
MSN	835.184211	233.577103	Paeni_PIH	RNA
MYH9	9.4	2.6473107	NPIH	RNA
MYH9	140.76	52.0221929	NonPaeni_PIH	RNA
MYH9	1286.421053	371.399901	Paeni_PIH	RNA
PFN1	5.3714286	1.62083934	NPIH	RNA
PFN1	100.96	40.9351262	NonPaeni_PIH	RNA
PFN1	538.342105	149.108671	Paeni_PIH	RNA
PGAM1	4.3142857	2.89080958	NPIH	RNA
PGAM1	37.96	15.602414	NonPaeni_PIH	RNA
PGAM1	298.578947	93.172454	Paeni_PIH	RNA

PGK1	6.0285714	2.07956257	NPIH	RNA
PGK1	71.84	30.6036539	NonPaeni_PIH	RNA
PGK1	968	293.330977	Paeni_PIH	RNA
PKM	14.6285714	5.22950811	NPIH	RNA
PKM	132.24	50.8984594	NonPaeni_PIH	RNA
PKM	1279.105263	386.18805	Paeni_PIH	RNA
PTPRC	8.6857143	3.12307628	NPIH	RNA
PTPRC	145	60.2966513	NonPaeni_PIH	RNA
PTPRC	1207.605263	335.684923	Paeni_PIH	RNA
S100A12	0.1142857	0.06632566	NPIH	RNA
S100A12	1.28	0.6691441	NonPaeni_PIH	RNA
S100A12	83.973684	27.020123	Paeni_PIH	RNA
SDCBP	4.3142857	1.67605519	NPIH	RNA
SDCBP	28.52	12.2459563	NonPaeni_PIH	RNA
SDCBP	458.026316	135.968171	Paeni_PIH	RNA
SERPINB1	1.5142857	0.70189249	NPIH	RNA
SERPINB1	15.08	6.6992353	NonPaeni_PIH	RNA
SERPINB1	220.473684	73.837009	Paeni_PIH	RNA
SIGLEC14	0.3714286	0.18723432	NPIH	RNA
SIGLEC14	5.4	2.5660505	NonPaeni_PIH	RNA
SIGLEC14	174.684211	55.188807	Paeni_PIH	RNA
SOD2	8.6285714	2.73689829	NPIH	RNA
SOD2	95.16	40.5393514	NonPaeni_PIH	RNA
SOD2	7109.078947	2415.092354	Paeni_PIH	RNA
TAGLN2	1.3714286	0.47490302	NPIH	RNA
TAGLN2	17.52	6.7153372	NonPaeni_PIH	RNA
TAGLN2	178.526316	53.84761	Paeni_PIH	RNA
TALDO1	2	0.63245553	NPIH	RNA
TALDO1	16.04	6.4844145	NonPaeni_PIH	RNA
TALDO1	158.210526	50.789744	Paeni_PIH	RNA
VIM	15.8857143	6.94602449	NPIH	RNA
VIM	150.12	58.0973274	NonPaeni_PIH	RNA
VIM	880.473684	226.804177	Paeni_PIH	RNA
YWHAZ	5.6285714	2.23290178	NPIH	RNA
YWHAZ	47.32	19.0210992	NonPaeni_PIH	RNA
YWHAZ	344.947368	89.984254	Paeni_PIH	RNA
A2M	133221.7	9061.293	NPIH	Proteomics
A2M	198472.13	15890.586	NonPaeni_PIH	Proteomics
A2M	274404.45	27597.465	Paeni_PIH	Proteomics
ACTN1	116714.06	9776.423	NPIH	Proteomics
ACTN1	119576.72	9089.844	NonPaeni_PIH	Proteomics
ACTN1	297445.89	53791.338	Paeni_PIH	Proteomics
ANPEP	108576.41	9985.235	NPIH	Proteomics
ANPEP	173373.44	17886.847	NonPaeni_PIH	Proteomics
ANPEP	216110.82	27340.908	Paeni_PIH	Proteomics
APLP2	547840.41	62433.879	NPIH	Proteomics
APLP2	328137.7	37983.756	NonPaeni_PIH	Proteomics
APLP2	198439.28	24707.301	Paeni_PIH	Proteomics
B4GALT1	162247.44	19432.395	NPIH	Proteomics
B4GALT1	169430.82	20373.678	NonPaeni_PIH	Proteomics
B4GALT1	247502.43	35288.886	Paeni_PIH	Proteomics
CAP1	60769.19	6210.648	NPIH	Proteomics
CAP1	94136.8	9107.84	NonPaeni_PIH	Proteomics
CAP1	296594.36	51942.821	Paeni_PIH	Proteomics
CD14	269025.58	30322.495	NPIH	Proteomics
CD14	306931.74	26472.734	NonPaeni_PIH	Proteomics
CD14	278283.01	33204.389	Paeni_PIH	Proteomics
CFB	159222.5	19766.775	NPIH	Proteomics
CFB	190851.87	15216.808	NonPaeni_PIH	Proteomics
CFB	268398.88	26505.674	Paeni_PIH	Proteomics
CHI3L1	321831.41	29053.86	NPIH	Proteomics
CHI3L1	434327.09	38137.66	NonPaeni_PIH	Proteomics
CHI3L1	424837.82	58773.803	Paeni_PIH	Proteomics
ENO1	159438.94	14402.893	NPIH	Proteomics
ENO1	196381.01	14636.721	NonPaeni_PIH	Proteomics
ENO1	411495.75	55166.856	Paeni_PIH	Proteomics
FLNA	165793.31	12503.779	NPIH	Proteomics
FLNA	196536.38	14890.969	NonPaeni_PIH	Proteomics
FLNA	378300.76	40310.06	Paeni_PIH	Proteomics
GAPDH	185309.05	17179.248	NPIH	Proteomics
GAPDH	212531.47	20277.813	NonPaeni_PIH	Proteomics

GAPDH	443295.9	54278.46	Paeni_PIH	Proteomics
HK3	218889.04	22577.89	NPIH	Proteomics
HK3	231991.67	26295.58	NonPaeni_PIH	Proteomics
HK3	402744.42	65914.391	Paeni_PIH	Proteomics
HSPA1A	94243.84	9161.338	NPIH	Proteomics
HSPA1A	149336.59	14528.191	NonPaeni_PIH	Proteomics
HSPA1A	268265.79	33231.181	Paeni_PIH	Proteomics
ITIH4	117579.96	11077.351	NPIH	Proteomics
ITIH4	151069.07	12051.139	NonPaeni_PIH	Proteomics
ITIH4	199949.48	20480.571	Paeni_PIH	Proteomics
LCP1	113417.43	9116.376	NPIH	Proteomics
LCP1	195344.4	16039.965	NonPaeni_PIH	Proteomics
LCP1	340143.61	43948.485	Paeni_PIH	Proteomics
LRG1	108031.17	10929.886	NPIH	Proteomics
LRG1	206250.39	22096.464	NonPaeni_PIH	Proteomics
LRG1	251746.12	32363.663	Paeni_PIH	Proteomics
LTA4H	199245.98	22214.737	NPIH	Proteomics
LTA4H	274854.73	21338.75	NonPaeni_PIH	Proteomics
LTA4H	414409.87	73708.092	Paeni_PIH	Proteomics
MMP9	135606.2	21383.917	NPIH	Proteomics
MMP9	172651.95	18211.817	NonPaeni_PIH	Proteomics
MMP9	354630.68	52463.93	Paeni_PIH	Proteomics
MSN	153328.27	14199.509	NPIH	Proteomics
MSN	181385.28	15701.558	NonPaeni_PIH	Proteomics
MSN	399781.27	43917.617	Paeni_PIH	Proteomics
MYH9	142872.64	12546.343	NPIH	Proteomics
MYH9	207808.86	21902.471	NonPaeni_PIH	Proteomics
MYH9	379414.18	40316.424	Paeni_PIH	Proteomics
PFN1	235567.9	21701.309	NPIH	Proteomics
PFN1	280671.85	24642.05	NonPaeni_PIH	Proteomics
PFN1	478063.72	65158.86	Paeni_PIH	Proteomics
PGAM1	175566.46	28071.23	NPIH	Proteomics
PGAM1	167330.31	20852.254	NonPaeni_PIH	Proteomics
PGAM1	246561.2	31372.817	Paeni_PIH	Proteomics
PGK1	207332.41	16847.118	NPIH	Proteomics
PGK1	223880.57	17165.741	NonPaeni_PIH	Proteomics
PGK1	358956.96	49762.833	Paeni_PIH	Proteomics
PKM	261737.79	17946.082	NPIH	Proteomics
PKM	266826.01	19378.661	NonPaeni_PIH	Proteomics
PKM	428380.3	43798.24	Paeni_PIH	Proteomics
PTPRC	808464.82	76515.077	NPIH	Proteomics
PTPRC	831798.58	91176.168	NonPaeni_PIH	Proteomics
PTPRC	1128370.95	228452.191	Paeni_PIH	Proteomics
S100A12	109727.46	21462.527	NPIH	Proteomics
S100A12	117201.9	11805.296	NonPaeni_PIH	Proteomics
S100A12	226202.57	21157.16	Paeni_PIH	Proteomics
SDCBP	50213.4	4283.095	NPIH	Proteomics
SDCBP	77780.78	6725.238	NonPaeni_PIH	Proteomics
SDCBP	130499.8	15048.869	Paeni_PIH	Proteomics
SERPINB1	166738.39	13430.713	NPIH	Proteomics
SERPINB1	224181.28	22991.06	NonPaeni_PIH	Proteomics
SERPINB1	623367.36	118273.928	Paeni_PIH	Proteomics
SIGLEC14	75642.08	13376.91	NPIH	Proteomics
SIGLEC14	85814.32	13300.28	NonPaeni_PIH	Proteomics
SIGLEC14	69099.42	6459.969	Paeni_PIH	Proteomics
SOD2	216602.91	19426.173	NPIH	Proteomics
SOD2	350152.5	32592.274	NonPaeni_PIH	Proteomics
SOD2	592462.98	92008.911	Paeni_PIH	Proteomics
TAGLN2	49330.14	4441.276	NPIH	Proteomics
TAGLN2	88738	13122.68	NonPaeni_PIH	Proteomics
TAGLN2	128991.53	16839.534	Paeni_PIH	Proteomics
TALDO1	394224.77	31213.484	NPIH	Proteomics
TALDO1	397284.78	46011.468	NonPaeni_PIH	Proteomics
TALDO1	763613.88	152020.927	Paeni_PIH	Proteomics
VIM	212003.12	22419.596	NPIH	Proteomics
VIM	236106.62	15654.802	NonPaeni_PIH	Proteomics
VIM	329432.96	37277.74	Paeni_PIH	Proteomics
YWHAZ	196027.15	19719.259	NPIH	Proteomics
YWHAZ	219335.11	23143.789	NonPaeni_PIH	Proteomics
YWHAZ	291817.05	34784.541	Paeni_PIH	Proteomics