Journal of proteome research

pubs.acs.org/jpr

Article

Deep Proteomics Network and Machine Learning Analysis of Human Cerebrospinal Fluid in Japanese Encephalitis Virus Infection

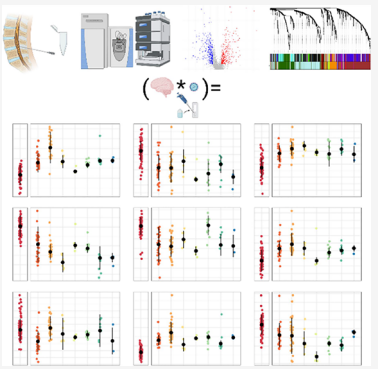
Tehmina Bharucha,* Bevin Gangadharan, Abhinav Kumar, Ashleigh C. Myall, Nazli Ayhan, Boris Pastorino, Anisone Chanthongthip, Manivanh Vongsouvath, Mayfong Mayxay, Onanong Sengvilaipaseuth, Ooyanong Phonemixay, Sayaphet Rattanavong, Darragh P. O'Brien, Iolanda Vendrell, Roman Fischer, Benedikt Kessler, Lance Turtle, Xavier de Lamballerie, Audrey Dubot-Pérès, Paul N. Newton, Nicole Zitzmann, and SEAe Consortium

Cite This: *J. Proteome Res.* 2023, 22, 1614−1629

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Japanese encephalitis virus is a leading cause of neurological infection in the Asia-Pacific region with no means of detection in more remote areas. We aimed to test the hypothesis of a Japanese encephalitis (JE) protein signature in human cerebrospinal fluid (CSF) that could be harnessed in a rapid diagnostic test (RDT), contribute to understanding the host response and predict outcome during infection. Liquid chromatography and tandem mass spectrometry (LC−MS/MS), using extensive offline fractionation and tandem mass tag labeling (TMT), enabled comparison of the deep CSF proteome in JE vs other confirmed neurological infections (non-JE). Verification was performed using data-independent acquisition (DIA) LC−MS/MS. 5,070 proteins were identified, including 4,805 human proteins and 265 pathogen proteins. Feature selection and predictive modeling using TMT analysis of 147 patient samples enabled the development of a nine-protein JE diagnostic signature. This was tested using DIA analysis of an independent group of 16 patient samples, demonstrating 82% accuracy. Ultimately, validation in a larger group of patients and different locations could help refine the list to 2−3 proteins for an RDT. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD034789 and 10.6019/PXD034789.



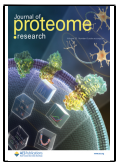
KEYWORDS: *central nervous system infection, neurological infection, encephalitis, flavivirus, Japanese encephalitis virus, diagnosis, clinical proteomics, mass spectrometry, tandem mass tagging, data-independent acquisition, network analysis, machine learning analysis, predictive modeling., Lao PDR*

INTRODUCTION

Japanese encephalitis virus (JEV) is a mosquito-borne flavivirus and a leading cause of neurological infection as Japanese encephalitis (JE) in Asia and the Pacific. It is of considerable public health importance, with recent estimates based on sparse data suggesting 1.5 billion people at risk with 42,000 cases per year.[1,2] It is an emerging disease, with recent evidence of JEV in multiple territories in Australia.[3] Patients may experience devastating socioeconomic consequences; JE predominantly affects children in poor rural areas with a 20−30% case fatality rate and 30−50% of survivors suffer long-term disability.[4] Although no specific treatment is available, several vaccines are available and recommended by the World Health Organization (WHO).[5,6] Although recent efforts have strengthened JEV vaccination programs, still only 15 of 24 endemic countries include JEV vaccine in routine immunization policies, and even then, it is not uniformly nationwide, with vaccine coverage in targeted areas reported to be as low as 39%.[7] JEV is a zoonosis, and sustained vaccine coverage is essential to control disease.

A fundamental limitation in the control of JE is the poor accuracy of existing diagnostic tests, requirement for lumbar puncture and laboratory capacity for diagnosis.[8] Surveillance data suggest that only 11 of 24 countries meet the minimum surveillance standards, equivalent to diagnostic testing in a sentinel site.[7] This is a threat to vaccine implementation, as accessible and accurate diagnostics are essential to understand epidemiology and effectiveness of vaccination, identify associated research knowledge gaps, and facilitate public engagement. This also has implications for appropriate risk-

Received: September 19, 2022
Published: May 23, 2023



https://doi.org/10.1021/acs.jproteome.2c00563
J. Proteome Res. 2023, 22, 1614−1629

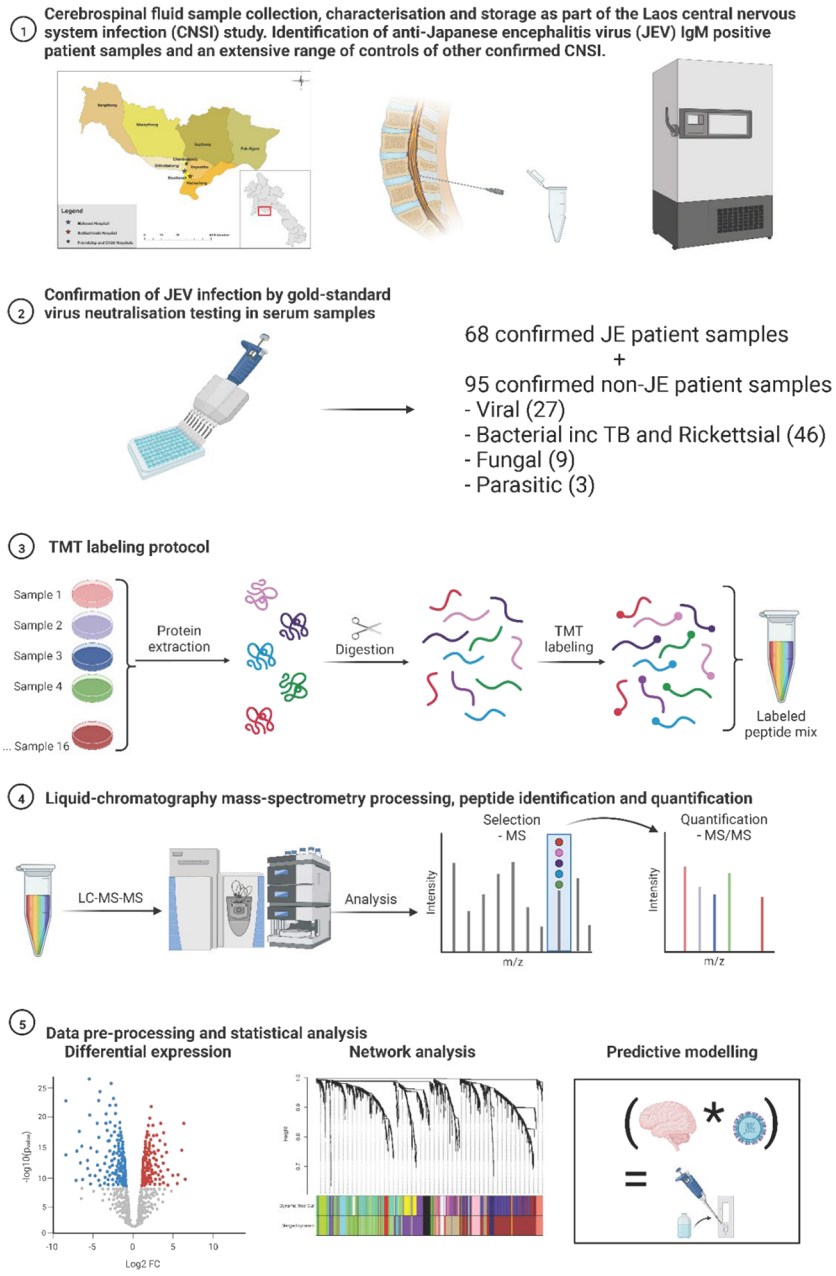


Figure 1. Schematic representation of the study methods. Patient cerebrospinal fluid (CSF) samples were collected during the Laos central nervous system infection study and underwent extensive diagnosis testing. Patients positive by anti-JEV IgM were confirmed by the gold-standard virus seroneutralization. The CSF of 163 patients with confirmed JE (n = 68) and non-JE (n = 95) neurological infections was analyzed by LC−MS/MS, all were tested by tandem mass tagging (TMT) and 16 were also tested by DIA LC-MS/MS. Data were analyzed by conventional univariate analysis, network analysis, and machine learning to identify diagnostic protein biomarkers of JE. The figure was created with BioRender.com.

assessment for travelers. Aside from JEV control, diagnosis is crucial for patients, families, and health-workers, to be able to institute appropriate supportive and rehabilitation care, stop unnecessary antibiotics, or if the test is negative to prompt further investigation.

The gold-standard JEV test is a neutralization assay. However, this requires paired acute and convalescent sera, is laborious, time-consuming, requires specialist skills, high-level isolation facilities for viral cell culture, and may not define the infecting virus in secondary flavivirus infections.[8] The WHO recommended diagnostic test is anti-JEV IgM antibody capture ELISA (MAC-ELISA) of cerebrospinal fluid (CSF). There are limited data from field studies comparing CSF MAC-ELISA with neutralization assays. The manufacturer of the only available commercial kit for clinical diagnosis (InBios) quotes a sensitivity of >90% for well-characterized CSF samples, but sensitivity in the field is as low as 53%.[9] There are also increasingly recognized problems with specificity related to prior vaccination and cross-reactivity with other flaviviruses.[10,11] Reported specificity is >90%; however, a study by our group demonstrated that 13% of patients with anti-JEV IgM detected in CSF by MAC-ELISA had another pathogen detected that may have explained the presentation.[10]

Detection of JEV RNA would be highly specific, but the period of viraemia is brief and hard to capture clinically, often occurring before the onset of neurological symptoms and signs.

https://doi.org/10.1021/acs.jproteome.2c00563
J. Proteome Res. 2023, 22, 1614−1629

RT-qPCR remains insensitive irrespective of the analytical sensitivity or gene targets.[12] For this reason, the application of metagenomics is not likely to significantly improve JEV RNA detection.

Non-targeted, discovery-based liquid chromatography−tandem mass spectrometry (LC−MS/MS) proteomics represents an underused technology for improving diagnosis of JE from the analysis of clinical samples.[13,14] Such an approach is based on the hypothesis that there is a protein signature in the CSF specific for JE and that diagnostic protein biomarkers could be harnessed in an antibody-based point-of-care test. Furthermore, deep proteomics exploration provides insights into disease processes and potential therapeutic targets. Network science and machine learning are two complementary disciplines enabling insights into complex high-dimensional data.[15,16] Networks, composed of nodes and links, are naturally attuned to problems where features have a relational structure[17] and have a track record of success in understanding networks of biological interactions.[18] Weighted correlation network analysis (WGCNA) was developed for the analysis of transcriptomics datasets but is increasingly used in proteomics research, enabling assignment of lengthy lists of proteins into modules with biological insights.[19−21] On the other hand, machine learning can uncover signals in data related to outcome variables and identify predictive markers of disease, a vital exploratory process for constructing diagnostics.[22] Used in conjunction, network science and machine learning provide novel characterization of disease states and can identify robust predictive markers of disease.[23]

Herein, we aimed to test the hypothesis that there is a diagnostic protein signature of JE by performing LC−MS/MS in patient samples recruited as part of the Laos CNS study, incorporating differential expression, network, and machine learning analysis. A subsidiary aim was to utilize the data in the same workflow to evaluate proteins associated with outcome of JE. We first performed a pilot feasibility tandem mass tag labeling (TMT) LC−MS/MS study (n = 15) and then a larger verification TMT LC−MS/MS study (n = 148) including a sample size based on a power calculation. These data were combined in the final analysis (n = 163). The results were verified by data-independent acquisition (DIA) LC−MS/MS in 16 (10%) of the samples. Weighted correlation network analysis (WGCNA) was used to explore the data. For the purposes of feature selection and training a machine learning model for classifying JE vs non-JE patients, the TMT LC−MS/MS data was used excluding the patients analyzed by DIA LC−MS/MS (n = 147).. The model was tested using the DIA LC−MS/MS data (n = 16), providing an independent group of patient samples and alternative methodological analysis.

EXPERIMENTAL SECTION

Patient Samples

A prospective study of central nervous system (CNS) infection has been conducted at Mahosot Hospital, Vientiane, Laos, since 2003. Methods and results from 2003 to 2011 have been described.[24] Patients from 2014 to 2017 were included in the Southeast Asia Encephalitis Project.[25] Inpatients of all ages were recruited for whom diagnostic lumbar puncture was indicated for suspicion of CNS infection because of altered consciousness or neurologic findings and for whom lumbar puncture was not contraindicated. There was no formal definition for CNS infection; patient recruitment was at the discretion of the responsible physician, reflecting local clinical practice. The laboratory also received samples from patients from other hospitals in Vientiane: Friendship, Children's, and Setthathirat Hospitals. Written informed consent was obtained from patients or responsible guardians. Ethical clearance was granted by the Ethical Review Committee of the former Faculty of Medical Sciences, National University of Laos and the Oxford University Tropical Ethics Research Committee. The confirmed etiology was determined by the results of a panel of diagnostic tests which included tests for the direct detection of pathogens in CSF or blood, specific IgM in CSF, seroconversion, or a 4-fold rise in antibody titer between admission and follow-up serum samples.[24] Pathogen detection was confirmed after critical analysis of test results to rule out possible contamination. JEV infection was confirmed, as recommended by the World Health Organization, by detection of anti-JEV IgM by ELISA in CSF or seroconversion in paired serum samples. All anti-JEV IgM positive samples were subsequently confirmed by the gold-standard virus neutralization assay see cited reference 26. Power analysis was performed to estimate the sample size that would be required using different values. A schematic representation of the study design and methods is illustrated in Figure 1.

LC−MS Sample Preparation

CSF samples were diluted 1:5 in 9 M urea and vortexed intermittently at room temperature for 30 min, to solubilize and denature proteins, inactivating any pathogens and rendering the sample acellular. Protein concentration was assessed with a Nanodrop assay ND-1000 spectrophotometer (Thermo Scientific) by measuring the absorbance at 280 nm and normalized by aliquoting different volumes of each sample dependent on the protein concentration, and then the total volume equalized with 7.5 M urea. An equal volume of 100 mM dithiothreitol (DTT) in 50 mM ammonium bicarbonate (AmBic) was added as a reducing agent, and the samples were vortexed and incubated at 56 °C for 45 min. An equal volume of 100 mM iodoacetamide (IAA) in 50 mM AmBic was added as an alkylating agent, vortexed, and incubated at room temperature for 1 h in the dark. 50 mM AmBic was added to each sample to reduce the urea concentration to below 1 M. Digestion was performed with trypsin in a ratio of 1:20 m:m protein:trypsin (Promega, P/N V5072 for the pilot study; V5117 for the larger study); first 75% of the total amount of trypsin added and incubated at 37 °C for 18 h overnight and then the remaining 25% added and incubated at 37.5 °C for 6 h. The samples were frozen at −20 °C to quench the trypsin digestion reaction. A pooled aliquot of each sample was analyzed by label-free LC−MS to verify protein digestion.

Reverse-phase (RP) C18 solid-phase extraction (SPE) was used to desalt the digested proteins, as per the manufacturer's instructions (Waters P/N WAT023590 for the pilot study; Thermo Scientific P/N 60109-001 for the larger study). The total eluate was dried completely using a vacuum concentrator (Savant SpeedVac or Eppendorf concentrator) and for the samples to be labeled by tandem mass tag (TMT), resuspended in 100 mM triethylammonium bicarbonate (TEAB). The samples were vortexed, centrifuged, and sonicated for 3 min, and then this was repeated. The Pierce Quantitative Colorimetric Peptide Assay (Thermo Scientific, UK) was performed as per the manufacturer's instructions. The samples were normalized for peptide concentration with TEAB to make up a final volume of 100 μL required for TMT

https://doi.org/10.1021/acs.jproteome.2c00563
J. Proteome Res. 2023, 22, 1614−1629

labeling. TMT labeling was performed as per the manufacturer's instructions, in two batches of TMT 11-plex (Thermo Scientific, P/N A37724) for the pilot study and 10 batches of 16-plex (Thermo Scientific, P/N A44520) for the larger study. For the larger study, in order to examine technical variability and adjust for batch effects, each batch contained one reference pool and the batches 9 and 10 had two replicate samples. A pooled sample was analyzed by LC−MS to verify labeling efficiency.

Offline High-pH Reverse-Phase Fractionation

For the pilot study, offline high-pH reverse-phase fractionation was performed using a Hypersil Gold column (Thermo Scientific, P/N 25002-202130). The mobile phase A was water adjusted with ammonium hydroxide to pH 10 and B was 10 mM ammonium bicarbonate in 80% acetonitrile (ACN) adjusted with ammonium hydroxide to pH 10 and a flow rate of 300 μL/min. The samples were separated into 91 fractions with each fraction collected every 60 s from the start of the run and using the gradient shown in Supplementary Data (S1 Data). For the larger study, offline high-pH reverse-phase fractionation was performed using an Xbridge BEH C18 column (Waters P/N 186006710). The mobile phase A was water adjusted to pH 10 with ammonium hydroxide and B was 90% ACN adjusted to pH 10 with ammonium hydroxide, at a flow rate of 200 μL/min. Fractions were collected every 60 s from the start of the run (100 fractions) and then concatenated into 44 fractions using the gradient shown in Supplementary Data (S1 Data). The samples analyzed by DIA LC−MS/MS were not processed by offline fractionation.

Liquid Chromatography-Mass Spectrometry

Online peptide desalting was performed with a Dionex Ultimate 3000 nano UHPLC (Thermo Scientific) using 100% of loading mobile phase A = 0.05% TFA in water at a flow rate of 10 μL/min for 4.6 min. The online desalting column (trap column) used was a C18 column (Thermo Scientific P/N 160454). At 4.6 min, the flow from the nanopump was diverted to the trap column in a backward flush direction. For online low-pH reverse-phase fractionation, the trapped peptides were eluted from the column over the gradient time specified in Supplementary Data (S1 Data). For the pilot study, Accucore C18 columns (Thermo Scientific P/N 16126-507569) were used with a nanosource, at a flow rate of 250 nL/min. For the larger study, EASY-Spray PepMap C18 columns (Thermo Scientific P/N ES903) were used with an EASY-Spray source and a flow rate of 300 nL/min. Mobile phase A was 0.1% FA and B was 0.1% FA in 80% ACN. MS was performed with a Q Exactive benchtop hybrid quadrupole-Orbitrap MS (Thermo Scientific); the settings are described in detail in Supplementary Data (S1 Data). For the CSF samples processed by DIA LC−MS/MS, samples were analyzed using a Dionex Ultimate 3000 nano UPLC (Thermo Scientific) coupled to an Orbitrap Fusion Lumos mass spectrometer (Thermo Scientific). Briefly, peptides were trapped on a PepMap C18 trap columns (Thermo) and separated on an EasySpray column (50 cm, P/N ES803, Thermo) over a 60-min linear gradient from 2% buffer B to 35% buffer B (A: 5% DMSO, 0.1% formic acid in water. B: 5% DMSO, 0.1% formic acid in acetonitrile) at a flow rate of 250 nL/min. The instrument was operated in data-independent mode as previously described.[27]

Data Processing and Statistical Analysis

The sample size was estimated using a power calculation based on a *t* test and multiple testing correction, with data from the pilot study and the R package "FDRsampsize".[28]

TMT LC−MS/MS analysis protein identification, quantification, missing value imputation and batch correction: Thermo raw files were imported into Proteome Discoverer v2.5 (Thermo Scientific, UK) for peptide identification using the SEQUEST algorithm[29] searching against the SwissProt *Homo sapiens* and pathogen databases according to the included samples with precursor mass tolerance 10 ppm and fragment mass tolerance 0.02 Da. Carbamidomethylation of cysteine, TMT at N-termini and lysine were set as fixed modifications, and oxidation of methionine was set as a variable modification. False discovery rate (FDR) estimation was performed using the Percolator algorithm.[30] The criteria for protein identification included FDR < 1%, ≥2 peptides per protein, ≥1 unique peptide per protein, ≤2 missed cleavages and ≥6 and ≤144 peptide length (amino acids), coisolation threshold <50%, average S/N threshold >10, and at least two channels with quantification data. Protein quantification was performed in R v 4.1.2 with the package MSstatsTMT.[31] Proteins with >50% missing data were removed, and the data were imputed with the package DreamAI.[32] To incorporate peptide count per protein, jitter was added proportional to 1/median peptide count for each protein. The pilot and larger study data were merged and normalized with the package RobNorm,[33] and then batch correction was performed with the function ComBat[34] in the package sva without modifiers as covariates.[35] The protein list was filtered to remove potential contaminant proteins from the skin or red blood cells, see Supplementary Data S5_contaminants for the list of proteins removed. The effectiveness of batch correction was performed by visualizing the processed data using principal component analysis and hierarchical clustering in MetaboAnalyst v5.0.[36]

TMT LC−MS/MS Data Differential Protein Expression. Differential expression between the protein abundance in the JE vs non-JE patient samples was performed using a *t* test and Benjamini−Hochberg correction for multiple testing.

TMT LC−MS/MS Data Protein Set Enrichment Analysis. Functional analysis of human proteins identified in JE vs non-JE patient samples was performed using the WebGestalt online tool[37] using gene set enrichment analysis (GSEA) and gene ontology.

TMT LC−MS/MS Data Network Analysis. WGCNA was performed using the package WGCNA: constructing a signed weighted coexpression network with a soft power threshold of 12 to produce a power distribution, that is, scale-free topology; applying hierarchical clustering to detect modules of highly interconnected proteins with a minimum module size of five, deepSplit 4 and merge threshold 0.3; classifying intramodular hub proteins as the five proteins with the highest module membership for each module; and then correlating the modules with patient sample data.[38]

Data-Independent Acquisition (DIA) Data Processing. For robustness, final verification was performed on 10% of the samples independently processed via a separate mass spectrometry pipeline using label-free DIA LC−MS/MS. DIA data were analyzed using DIA-NN software (v0.8) with the library-free approach as previously described,[39] using the default settings as recommended. Briefly, for the library-free processing, a library was created from human UniProt SwissProt database (downloaded 24/2/21 containing 20,381

https://doi.org/10.1021/acs.jproteome.2c00563
J. Proteome Res. 2023, 22, 1614−1629

sequences) using deep learning. Trypsin was selected as the enzyme (1 missed cleavage), with carboamidomethylation of C as a fixed modification, oxidation of methionine as a variable modification, and N-term M excision. Identification and quantification of raw data were performed against the in silico library applying 1% FDR at precursor level and match between runs (MBR). The DIA-NN "report.proteingroup" matrix output was further analyzed. Missing values were imputed with half the minimum value for each protein.

Feature selection and predictive modeling: This was performed using the TMT LC−MS/MS data without the samples processed by DIA (n = 147) with the Boruta algorithm (using the random forest classifier) using the package Boruta[40] and with Lasso (least absolute shrinkage and selection operator) regression using the package glmnet.[16,41] A final list of proteins based on the intersect between Boruta and Lasso was selected.[42] Classification of JE vs non-JE was performed with selected proteins using several different machine learning models (random forest, support vector machine, logistic regression, and naïve bayes with the package caret and caretEnsemble).[43] Models were trained using tenfold cross-validation repeated 10 times evaluated on AUC-ROC. The ensemble model was tested with the DIA LC−MS/MS data (n = 16). An analysis of feature importance was performed to identify proteins that best predicted the outcome (alive/died) in JE patients, however due to the small sample size this was considered an exploratory analysis. Feature selection was performed with Boruta and Lasso, and then fivefold cross-validation was performed on the entire TMT LC−MS/MS dataset using different machine learning models. Protein involvement in biological, molecular, and cellular processes was explored using gene ontology using the webserver STRING,[44] the R package WebGestaltR 0.4.4,[45] and tissue expression correlated with the Human Protein Atlas (HPA).[46,47]

RESULTS

Patient Data

Power analysis was performed to estimate the sample size that would be required to compare differential expression of proteins in JE vs non-JE using different values: with 1,000−3,000 biomarkers to be tested, 50−150 finally verified, effect size 0.8, power 90%, FDR < 5%, the total sample size with an equal number of JE cases and non-JE controls of 122. Overall, including the pilot and larger study, 163 patients were included: 68 JE and 95 Non-JE; see Table 1, Supplementary Data S2 and S3.

JE patients were confirmed by the assays with the highest diagnostic confidence; detection of JEV RNA, or detection of anti-JEV IgM in CSF or by seroconversion and confirmed by virus neutralization tests (VNT). Non-JE patients included a range of different categories of infection that are common in the region. None of the patients had dual infections. Details of patient demographics, clinical presentations, laboratory investigations and outcome are reported in Supplementary Data S1 and S2.

Protein Profiling in CSF Reveals Differential Expression in JE

5,070 proteins were identified, including 4,805 human proteins and 265 pathogen proteins, see Supplementary Data S4 for MSstatsTMT output for the pilot and larger studies. The pathogen proteins were bacterial or parasitic proteins. 2,244 human proteins were identified in more than half of the samples included in both the pilot and larger studies. 68 proteins deemed to be contaminants were removed from the list, see Supplementary Data S5, resulting in a filtered list of 2,176 proteins. Data processing and batch correction were effective, see Supplementary Data S7.

268 proteins showed differential expression (167 > 1.2-fold change, FC, and 101 < 0.8 FC) based on the performance of a t test and Benjamini−Hochberg multiple testing correction with p value <0.05, illustrated by the volcano plots in Figure 2.

Molecular Pathways Associated with JE in CSF

2,176 proteins from 163 patient samples were used in protein set enrichment analysis, detailed in Supplementary Data S8, and to build a weighted gene expression network, detailed in Figures 3 and 4 and Supplementary Data S9−S11. A single outlier was identified, see Supplementary Data S9, and removed. Further analysis revealed that this sample had higher overall protein abundances, in spite of peptide normalization prior to TMT labeling and downstream normalization in MSstatsTMT and RobNorm during data processing. 44 modules were identified, and then closely related modules merged into 20 modules, see the tree diagram illustrating the cluster dendrogram and the modules in Figure 3 (top and bottom panels respectively). Module-trait relationships are shown in Figure 4; suggesting that 15 modules were associated with JE (p value <0.05), 9 upregulated (red) and 6 downregulated (green). 10 of the modules included proteins in the top five intramodular proteins, that is, proteins with the highest modular membership, with significant differences in abundance between the JE and non-JE group.

Diagnostic Protein Signature of JE in CSF

Feature Selection. 1,736 proteins were present in both the TMT and DIA LC−MS/MS data. In total, 83 proteins were identified by at least one of the feature selection procedures as important in classifying JE vs non-JE; 68 proteins identified with the Boruta algorithm and 24 with Lasso, see Supplementary Data S12. The proteins were associated with 10 different WGCNA modules, all of which had been identified as associated with JE through WGCNA. Forty seven were upregulated and 36 downregulated in JE in comparison to other neurological infections. Functional enrichment analysis in STRING demonstrated interactions between the proteins, Figure 5. Twenty-two proteins were secreted proteins: Immunoglobulin lambda variable 3−9 (IGLV3−9), immunoglobulin heavy variable 3−74 (IGHV3−74), Golgi membrane protein 1 (GOLM1), Cathepsin L (CTSL), CEA cell adhesion molecule 8 (CEACAM8), phospholipase B domain containing 1 (PLBD1), Cerebellin 1 precursor (CBLN1), secreted phosphoprotein 1 (SPP1), Natriuretic peptide C (NPPC), microtubule-associated protein tau (MAPT), Chitinase 3 like 1 (CHI3L1), ISG15 ubiquitin like modifier (ISG15), Interleukin 18 binding protein (IL18BP), Beta-2-microglobulin (B2M), TNF superfamily member 13b (TNFSF13B), bactericidal permeability increasing protein (BPI), Pentraxin 3 (PTX3), matrix metallopeptidase 9 (MMP9), S100 calcium binding protein A12 (S100A12), Azurocidin 1 (AZU1), Olfactomedin 4 (OLFM4), and matrix metallopeptidase 8 (MMP8). 15 proteins were associated with increased expression in the brain: Brain abundant membrane attached signal protein 1 (BASP1), Aldolase, fructose-bisphosphate C (ALDOC), CBLN1, Metallothionein 3 (MTX3), MAP2, Tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein gamma

https://doi.org/10.1021/acs.jproteome.2c00563
J. Proteome Res. 2023, 22, 1614−1629

Table 1. Summary of Included Patients' Demographics, Clinical Presentations, and Details of Diagnosis

		JE[a]	non-JE viral infections[b]	bacterial infections[c]	Tuberculosis[d]	Rickettsial infections[e]	scrub typhus[f]	fungal infections[g]	parasitic infections[h]	total (N = 163)
age (years)	median	16	22	38	53	48	20	33	25	22
	IQR	10.0, 23.2	6.0, 35.0	20.0, 55.0	32.5, 55.5	32.0, 53.0	11.8, 24.0	27.0, 47.0	24.5, 42.0	12.0, 41.5
sex	female	24 (35.3%)	10 (37.0%)	9 (28.1%)	1 (14.3%)	1 (20.0%)	4 (33.3%)	0 (0.0%)	0 (0.0%)	49 (30.1%)
ethnicity	Lao loum	35 (51.5%)	25 (92.6%)	23 (71.9%)	7 (100.0%)	5 (100.0%)	11 (91.7%)	8 (88.9%)	3 (100.0%)	117 (71.8%)
	Lao sung	16 (23.5%)	1 (3.7%)	3 (9.4%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	20 (12.3%)
	Lao theung	1 (1.5%)	0 (0.0%)	1 (3.1%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	2 (1.2%)
	Lao unspecified	16 (23.5%)	1 (3.7%)	5 (15.6%)	0 (0.0%)	0 (0.0%)	1 (8.3%)	1 (11.1%)	0 (0.0%)	24 (14.7%)
comorbidities	N-Miss	0	0	0	0	0	1	0	0	1
	yes	3 (4.4%)	5 (18.5%)	3 (9.4%)	4 (57.1%)	1 (20.0%)	0 (0.0%)	5 (55.6%)	0 (0.0%)	21 (12.9%)
	HIV	0 (0.0%)	3 (11.1%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	5 (55.6%)	0 (0.0%)	8 (4.9%)
duration of illness (days)	median	4	3	3	6	2	5.5	5	1	4
	IQR	3.8, 5.2	1.0, 5.5	2.0, 5.2	3.0, 7.0	2.0, 4.0	2.8, 8.2	2.0, 7.0	0.5, 4.0	2.0, 6.0
seizures	yes	29 (42.6%)	9 (33.3%)	6 (18.8%)	0 (0.0%)	2 (40.0%)	3 (25.0%)	1 (11.1%)	0 (0.0%)	50 (30.7%)
antibiotics prior to LP	N-Miss	13	1	3	2	0	3	3	0	25
	yes	44 (80.0%)	15 (57.7%)	12 (41.4%)	4 (80.0%)	2 (40.0%)	8 (88.9%)	1 (16.7%)	1 (33.3%)	87 (63.0%)
GCS	N-Miss	7	0	1	0	0	1	0	0	9
	Median	11	14	12	10	10	15	15	15	12
	IQR	8.0, 13.0	9.5, 15.0	10.0, 15.0	8.5, 12.5	7.0, 14.0	14.0, 15.0	12.0, 15.0	15.0, 15.0	9.0, 15.0
blood WCC ($10^6/\mu$L)	N-Miss	9	1	5	0	0	1	0	0	16
	Median	12.8	11.3	12.8	11.5	8.3	13	8.4	7.5	11.7
	IQR	8.4, 17.4	7.8, 12.9	9.4, 17.5	8.6, 12.1	7.0, 13.0	10.0, 13.6	4.8, 14.6	7.4, 8.9	8.1, 15.1
blood CRP (mg/L)	N-Miss	50	11	18	5	1	5	3	1	94
	Median	45	45.1	184.5	20.3	41.6	102.5	33.5	9.8	46.7
	IQR	24.7, 73.1	4.5, 93.1	106.3, 196.5	16.5, 24.0	3.9, 103.4	33.6, 131.3	10.6, 45.5	5.1, 14.6	19.4, 112.6
CSF opening pressure (cm of H_2O)	N-Miss	11	3	0	0	0	1	0	0	15
	Median	23	20	23.5	24	15	23	25.5	27	22
	IQR	18.0, 26.0	16.2, 27.4	11.4, 32.8	18.8, 28.0	11.5, 20.0	20.0, 25.0	12.0, 38.0	22.5, 28.0	16.8, 28.2
CSF color	N-Miss	4	1	3	0	0	4	0	1	13
	clear	50 (78.1%)	18 (69.2%)	6 (20.7%)	6 (85.7%)	3 (60.0%)	7 (87.5%)	6 (66.7%)	1 (50.0%)	97 (64.7%)
	red	1 (1.6%)	3 (11.5%)	1 (3.4%)	0 (0.0%)	1 (20.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	6 (4.0%)
	turbid	13 (20.3%)	4 (15.4%)	20 (69.0%)	1 (14.3%)	0 (0.0%)	0 (0.0%)	3 (33.3%)	1 (50.0%)	42 (28.0%)
	yellow	0 (0.0%)	1 (3.8%)	2 (6.9%)	0 (0.0%)	1 (20.0%)	1 (12.5%)	0 (0.0%)	0 (0.0%)	5 (3.3%)
CSF WCC (10^6/L)	N-Miss	3	3	2	1	2	1	1	0	13
	median	110	147.5	655	192.5	25	60	140	500	150
	IQR	50.0, 215.0	36.2, 255.0	206.2, 1573.8	98.8, 316.2	17.5, 35.0	47.5, 132.5	32.5, 260.0	492.5, 1115.0	51.2, 340.0
CSF % neutrophils	N-Miss	3	3	3	1	2	2	1	0	15
	median	25	29	92.7	30.2	89	67	93.6	72	45.5
	IQR	9.3, 48.4	12.4, 62.8	74.0, 97.2	21.8, 76.6	69.5, 94.5	58.9, 78.0	75.2, 100.0	66.5, 85.0	15.0, 89.6
CSF % lymphocytes	N-Miss	3	3	3	1	2	2	1	0	15
	median	75	71	7.3	69.8	11	33	6.4	28	54.5
	IQR	51.6, 93.3	37.2, 87.6	2.8, 26.0	23.4, 78.2	5.5, 30.5	22.0, 41.1	0.0, 24.8	15.0, 33.5	10.4, 85.0
CSF RCC (10^6/L)	N-Miss	2	0	3	0	0	0	0	0	5

https://doi.org/10.1021/acs.jproteome.2c00563
J. Proteome Res. 2023, 22, 1614−1629

Table 1. continued

		JE[a]	non-JE viral infections[b]	bacterial infections[c]	Tuberculosis[d]	Rickettsial infections[e]	scrub typhus[f]	fungal infections[g]	parasitic infections[h]	total (N = 163)
	median	0	0	0	0	75	14.5	5	75	0
	IQR	0.0, 0.0	0.0, 27.5	0.0, 135.0	0.0, 12.5	5.0, 1005.0	0.0, 37.5	0.0, 18.0	37.5, 85.0	0.0, 19.5
CSF total protein (mg/dL)	N-Miss	6	1	1	0	1	2	0	0	11
	median	60	58.5	111	128	72.5	93.5	39	96	76.5
	IQR	38.2, 98.8	30.5, 117.5	57.5, 259.5	97.5, 270.0	60.0, 75.2	57.8, 139.5	20.0, 88.0	59.0, 108.0	38.8, 120.0
CSF glucose (mmol/L)	N-Miss	4	2	2	0	1	2	0	0	11
	median	3.7	3.3	3	2	5.1	3.4	2.6	2.9	3.4
	IQR	2.9, 4.7	1.9, 4.9	1.7, 5.3	1.4, 4.6	3.7, 5.8	3.1, 6.5	1.8, 2.9	2.6, 4.2	2.3, 4.8
duration of admission (days)	N-Miss	34	3	5	0	1	3	3	0	49
	median	13.5	8.5	11	20	9.5	5	19.5	8	11
	IQR	10.2, 17.0	7.0, 12.2	6.0, 16.0	8.0, 21.5	6.5, 12.0	5.0, 8.0	14.2, 27.8	7.5, 10.0	7.0, 16.0
outcome	N-Miss	26	3	7	0	1	2	5	0	44
	died	9 (21.4%)	2 (8.3%)	7 (28.0%)	1 (14.3%)	0 (0.0%)	1 (10.0%)	0 (0.0%)	0 (0.0%)	20 (16.8%)

[a]JE= Japanese encephalitis"; Miss = missing value. [b]LP = lumbar puncture. [c]GCS = Glasgow Coma Scale. [d]WCC = white cell count. [e]CRP = c-reactive protein. [f]CSF = cerebrospinal fluid. [g]RCD = red cell count. [h]IQR = interquartile range.

(YWHAG), Tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein eta (YWHAH), MARCKS like 1 (MARCKSL1), Secernin 1 (SCRN1), SPP1, MAPT, CHI3L1, Paralemmin (PALM), Reticulon 1 (RTN1), Purkinje cell protein 4 (PCP4), Cytidine/uridine monophosphate kinase 2 (CMPK2), NPPC, Glial fibrillary acidic protein (GFAP), cell cycle exit, and neuronal differentiation 1 (CEND1). Thus, three proteins were secreted and showed an increased expression in the brain: SPP1, MAPT, CHIL3, NPPC, and CBLN1.

JEV has a predilection for the thalamus and substantia nigra of the basal ganglia.[26] One of the proteins were "group enriched" in the thalamus, MMP9, from the HPA database. Four proteins were associated with the GO term "substantia nigra development", associated with BASP1, Glucose-6-phosphate dehydrogenase (G6PD), YWHAH, and 14-3-3 protein epsilon (14-3-3epsilon). The HPA database includes mRNA expression data from 13 brain regions, including the basal ganglia and thalamus; substantia nigra expression on its own is not reported (https://www.proteinatlas.org/humanproteome/brain).

Feature selection identified a final set of nine proteins which together exhibited high predictive performance (Figure 6). When examined using the ensemble model, using 10-fold cross-validation, JE classification demonstrated an AUC-ROC of 98.7 (98.0−99.4), in addition to high sensitivity and specificity—metrics in Table 2 and ROC in Supplementary Data S13 and S14.

Data acquired by DIA LC−MS/MS of 16 samples were used to verify the nine-protein JE diagnostic predictive model. The test metrics are reported in Table 2.

Establishing CSF Molecular Signatures as Predictors of the JE Outcome

Feature Selection. Subgroup analysis was performed using 42 JE samples for which outcome data at hospital discharge (died vs alive) were available. Seven proteins were identified as important in predicting outcome using the Boruta algorithm and two proteins using Lasso, such that two proteins were identified by both Boruta and Lasso, see Supplementary Data S15. In view of the small sample size, the data were not split into a training and test set. These proteins were used to train different models with fivefold CV repeated ten times evaluated on ROC and then combined in an ensemble model with cross-validation scores reported in Table 2, see the list of proteins in Supplementary Data S15 and ROC in Data S16. There were five JE patients in the DIA LC−MS analysis of which 3 had outcome data, and this was considered too small to report test metrics.

CONCLUSIONS

We performed deep untargeted analysis of well-characterized patient CSF samples from a large number of different confirmed neurological infections. To our knowledge, the highest number of proteins in CSF identified to date has been 3,174;[48] thus, this research represents a notable improvement in terms of the numbers of proteins identified, and this serves as a marker of the depth of analysis and prospects for biomarker identification.[49] Offline fractionation into 90 fractions in the pilot study, and 100 fractions concatenated into 44 in the larger study, with two-hour online LC gradients and multiplexing with TMT-16plex contributed to the depth of

https://doi.org/10.1021/acs.jproteome.2c00563
J. Proteome Res. 2023, 22, 1614−1629

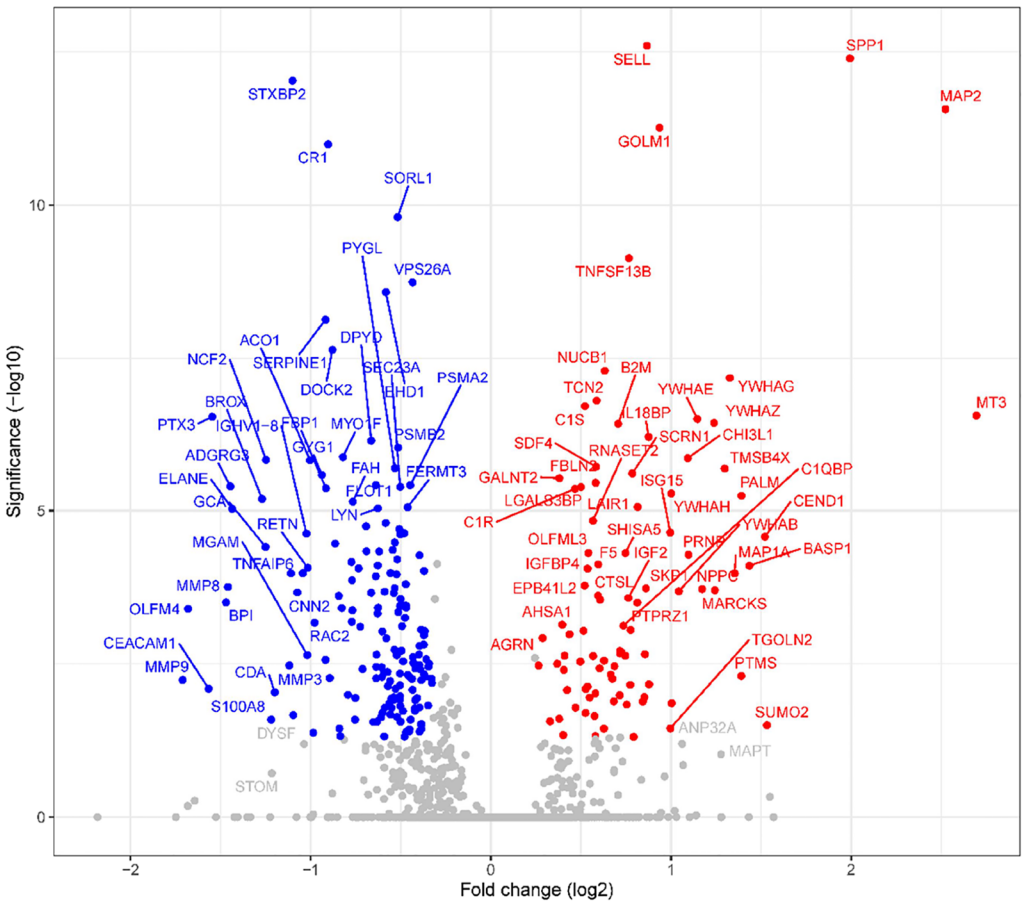


Figure 2. Differential CSF proteome between JE and non-JE neurological infections. Volcano plot of the identified proteins illustrating the statistical significance (*t* test *p* values with multiple testing correction) against the magnitude of change (fold change) for JE vs non-JE neurological infections. Proteins with differential expression are colored and labeled by gene names, upregulated proteins in red, and downregulated proteins in blue.

analysis. Furthermore, the diverse range of neurological infections also augmented the variety of proteins identified.

WGCNA identified 20 clusters of highly correlated proteins and provided insight into the proteins and how they associate with disease mechanisms. The modules were allocated a descriptor, according to gene ontology analysis, as well as the clinical and biological significance of the proteins. For example, one module was associated with IgM (proteins in the module included immunoglobulin heavy constant mu and immunoglobulin J chain) and correlated with JE and *Orientia tsutsugamushi* (OT), as well as the duration of illness. Other important modules associated with upregulation in JE included neuronal damage, antiapoptosis, heat shock response, unfolded protein response, cell adhesion, and macrophage and dendritic cell activation. In contrast, in comparison to other non-JE neurological infections, there was an association with downregulated acute inflammatory response, hepatotoxicity, activation of coagulation, extracellular matrix, and actin regulation.

Predictive modeling using the nine-protein ensemble model enabled classification of JE and non-JE samples with a CV accuracy of 97.0 (95% CI 95.7−98.0) using TMT labeled DDA data, and 81.3% (95% CI 54.4−96.0) in verification with 16 (10%) of the samples by DIA. DIA is a label-free method of analysis, with ongoing improvements in depth and throughput; in this case providing a complementary method to verify the TMT data rather than performing traditional targeted LC−MS/MS proteomics such as parallel reaction monitoring. Three proteins selected as the best disease classifiers were not "significant", i.e., *p* value <0.05 with *t*-test and adjustment for multiple testing, highlighting the limitations of univariate analysis in biomarker identification.[50] Biomarker discovery is a lengthy process, akin to the pharmaceutical pipeline.[13] The work demonstrates important CSF proteins in classifying JE vs non-JE. However, there is no doubt that the protein signature needs to be validated with orthogonal antibody-based methods in additional patient groups. It will also be useful to compare this with protein profiling in other body fluids. This will inform the use of a smaller subset of proteins in an ELISA or rapid diagnostic test to be tested alongside the existing anti-JEV IgM assay.

To date, to our knowledge, two studies have utilized unbiased techniques to examine the CSF proteome in human patients with confirmed JEV infection; while they demonstrate the feasibility of the methods, the patients were not confirmed by seroneutralization and included relatively small numbers of patients (10 and 26 JE patients).[51,52] There have been a handful of studies utilizing ELISA methods to target specific proteins; however, these rarely used power calculations in their experimental design, nor did they include adequate controls.[53−58] Analysis of the transcriptome and proteome in animal models[59−63] and cell culture[53,59,64−69] have been performed; however, the comparability to human CSF and comparison with other neurological infections is limited.

https://doi.org/10.1021/acs.jproteome.2c00563
J. Proteome Res. 2023, 22, 1614−1629

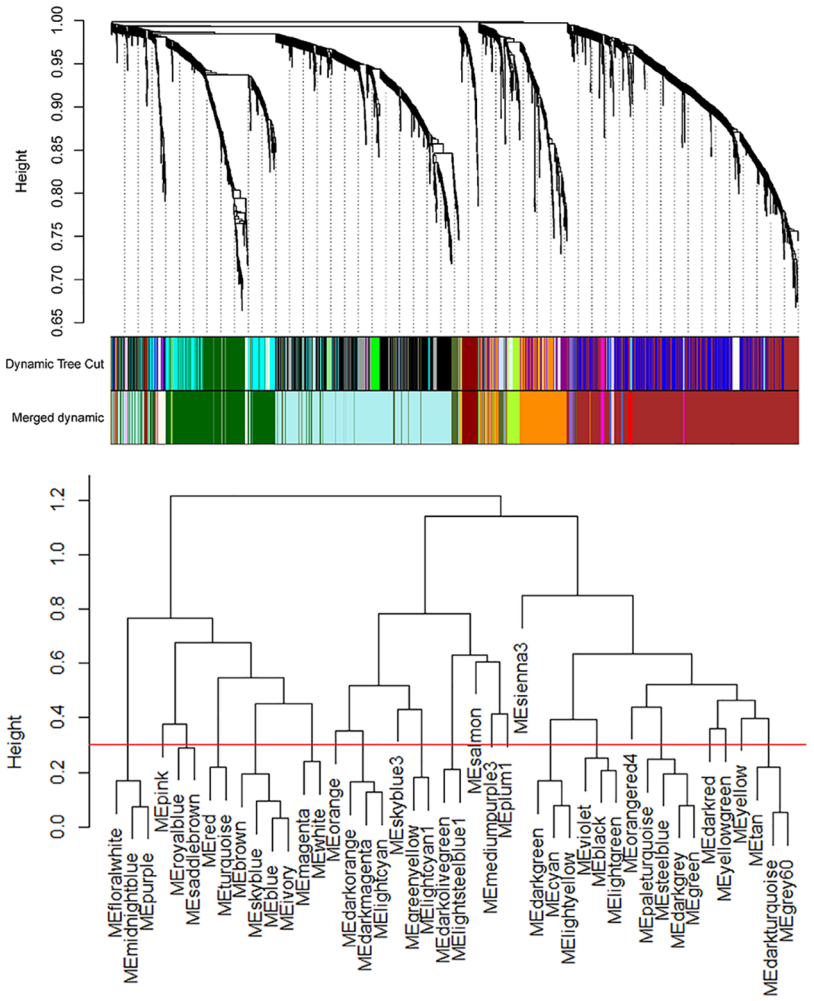


Figure 3. WGCNA identified 20 modules enriched in the CSF of JE patients. Cluster dendrogram (top panel) and clustering of module eigengenes (bottom panel). Expression of 2176 proteins was analyzed using the R package WGCNA, enabling unsupervised construction of correlation-based clusters, called "modules." Modules with close correlation (threshold = 0.3) were merged, producing 20 modules. The red line in the figure indicates the threshold for merging modules together, here the threshold was 0.3.

Furthermore, mRNA expression does not directly correlate with that of the corresponding protein.[70]

As expected, while we included JEV proteins in the search database, we did not identify any JEV proteins. This is compatible with previous publications; non-structural protein 1 is the major secreted protein during flavivirus infections, harnessed widely as a diagnostic biomarker for dengue virus infection, but not a useful diagnostic biomarker for JE.[71] The data provide useful interrogation of the host response to JEV infection. The identified proteins fit well into the existing literature on the host response in JEV and other closely associated flavivirus infections, most importantly West Nile virus infection.[72,73] MAPT and MAP2 are both closely associated microtubule stabilizing proteins specific to neuronal cells.[74] Both proteins were identified in this study as being biomarkers of JE in CSF, and the high levels in comparison to other neurological infections are striking. The association of the former has previously been demonstrated by ELISA, in one of the only studies of this type.[75] The role of actin, microtubule, and intermediate filament cytoskeletal re-organization in flavivirus infection has been described[76] and upregulation of MAPT and MAP2 may represent neuronal damage following transneural spread of JEV. Other proteins that were associated with JE in this study, all within the red WGCNA module, that may reflect neuronal damage include Paralemmin, Calbindin 1, MAP2, Parvalbumin, Secernin 1, and cell cycle exit and neuronal differentiation. The upregulation of ISG15 and ISG20 fit in with the known upregulation of a host of ISGs as part of the innate immune response to a viral infection.[77,78] Additional functional enrichments reflecting different WGCNA modules have previously been described anti-apoptosis,[79] heat shock response,[80,81] unfolded protein response,[82] translation,[83] IgM,[84] cell adhesion and pathogen attachment,[85] endothelial activation,[86] and macrophage activation.[87,88] In comparison to other neurological infections, there was a downregulation in acute phase response proteins and neutrophil enriched proteins, as has been seen by other studies.[89−91] In these, however, the sample size for the analysis of proteins predictive of outcome was less substantial and not supported by an a priori power calculation.

Incomplete coverage and missing data between LC−MS runs is an ongoing issue in the field.[32] It is notable that comparing with other similar studies in the literature, the important proteins may not be exactly the same but are closely related. These issues are now being improved by DIA methods. Further limitations are that the demographics of the cases and

https://doi.org/10.1021/acs.jproteome.2c00563
J. Proteome Res. 2023, 22, 1614−1629

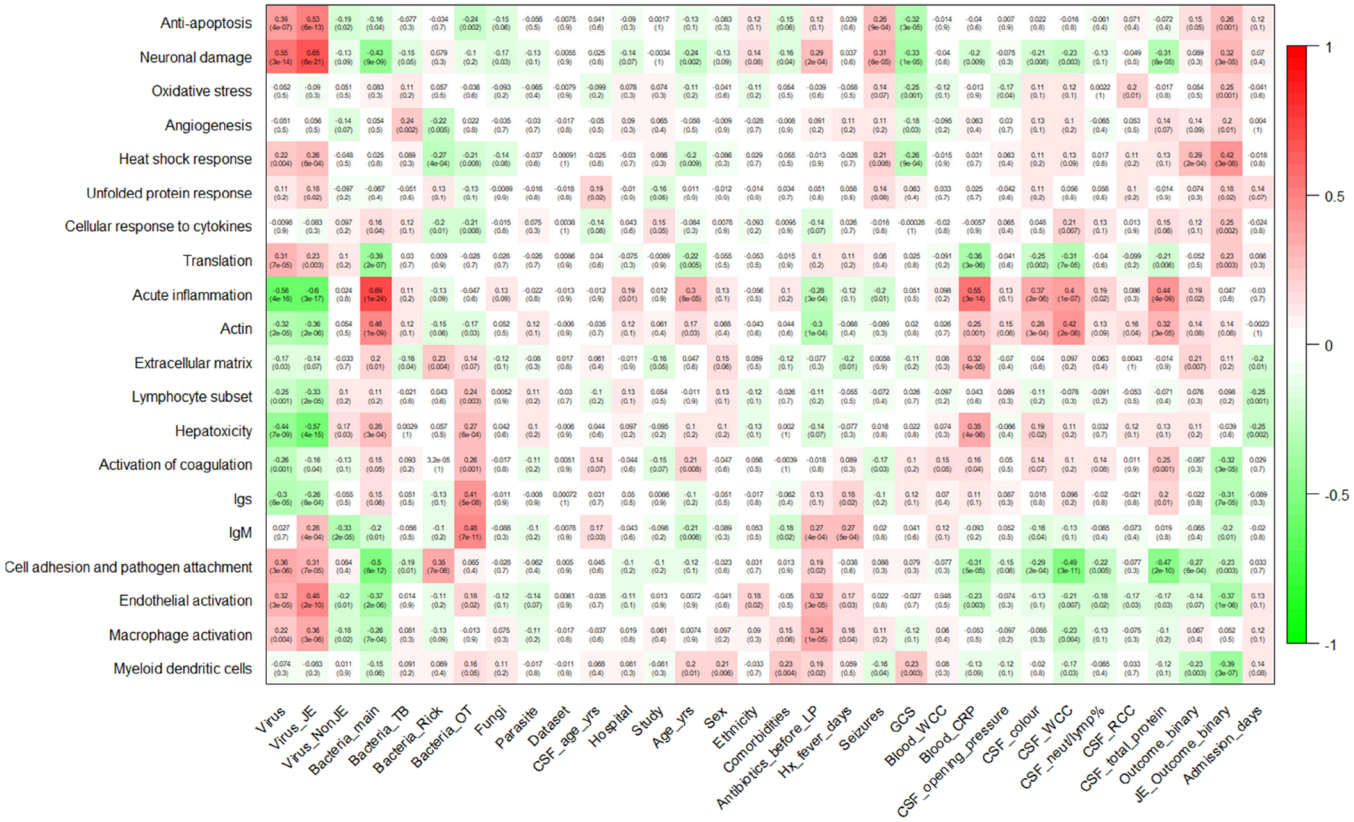


Figure 4. WGCNA module-trait relationships reveal enriched molecular pathways associated with patient clinical conditions. The 20 modules were allocated a descriptor (*y* axis), according to the clinical and biological associations of the proteins. The correlation with module protein expression was analyzed against patient data (*x* axis). Virus_JE = Japanese encephalitis; Virus_NonJE = non-JE viral infections; Bacteria_main = bacterial infections; Bacteria_TB = tuberculosis; Bacteria_Rick = rickettsial infections; Bacteria_OT = scrub typhus; Fungi = fungal infections; Parasite = parasitic infections; CSF = cerebrospinal fluid; LP = lumbar puncture; GCS = Glasgow Coma Score; WCC = white cell count; CRP = c-reactive protein; RCC = red cell count.

Table 2. Predictive Modeling Scores with 95% Confidence Intervals

classification task	data	AUC-ROC	accuracy	sensitivity	specificity	positive predictive value	negative predictive value
JE diagnosis (JE vs non-JE)	training set[a] (*n* = 147)	98.7 (98.0−99.4)	97.0 (95.7−98.0)	99.8 (98.9−100)	94.2 (91.8−96.1)	94.5 (92.2−96.3)	99.8 (98.8−100)
	test set[b] (*n* = 16)	95.5 (86.6−100)	81.3 (54.4−96.0)	100 (47.8−100)	72.7 (39.0−94.0)	62.5 (24.5−91.5)	100 (63.1−100)
JE outcome (dead vs alive)	training set[c] (*n* = 42)	88.5 (84.7−92.2)	86.3 (83.5−88.8)	42.0 (32.2−52.3)	93.7 (91.4, 95.5)	52.5 (41.0−63.8)	90.6 (88.1−92.8)

[a]The training set included patient samples processed by TMT LC−MS/MS; performance metrics are calculated by cross-validation. [b]The test set included patient samples processed by label-free DIA LC−MS/MS. [c]The training set included all the JE patients included in the TMT LC−MS/MS analysis for which outcome data was available; performance metrics are calculated by cross-validation.

controls were not perfectly matched and that we did not include CSF from healthy people in Laos on ethical grounds, or from cohorts from elsewhere on the basis that samples that have undergone different storage conditions may not be comparable. The latter is also the reason that there are no samples from neurological flaviviruses occurring in other geographical areas, such as West Nile virus (WNV) and Zika virus (ZIKV). Furthermore, for the purposes of the objective of finding a diagnostic protein signature of JE, the utmost importance was comparing JE with controls of a wide range of other neurological infections. The analysis of proteins predictive of different categories of infectious etiologies was not sufficiently powered and has not been reported. It is important to keep in mind that the comparison is between different neurological infections in the analysis of proteins that are up- and downregulated.

An RDT to detect JE in less accessible areas is urgently needed. This study demonstrates the feasibility of an unbiased LC−MS approach in the identification of novel protein biomarkers of neurological infections. It also represents a novel workflow for biomarker discovery research involving a large set of patient samples processed by TMT LC−MS/MS and then verified by an alternative DIA LC−MS/MS pipeline. Additional data using antibody-based methods will allow the nine-protein signature to be refined. This could be performed by purchasing or developing ELISA assays and comparing the specific protein abundance in JE and non-JE patients. These data will need to be validated in a larger group of patients, in

https://doi.org/10.1021/acs.jproteome.2c00563
J. Proteome Res. 2023, 22, 1614−1629

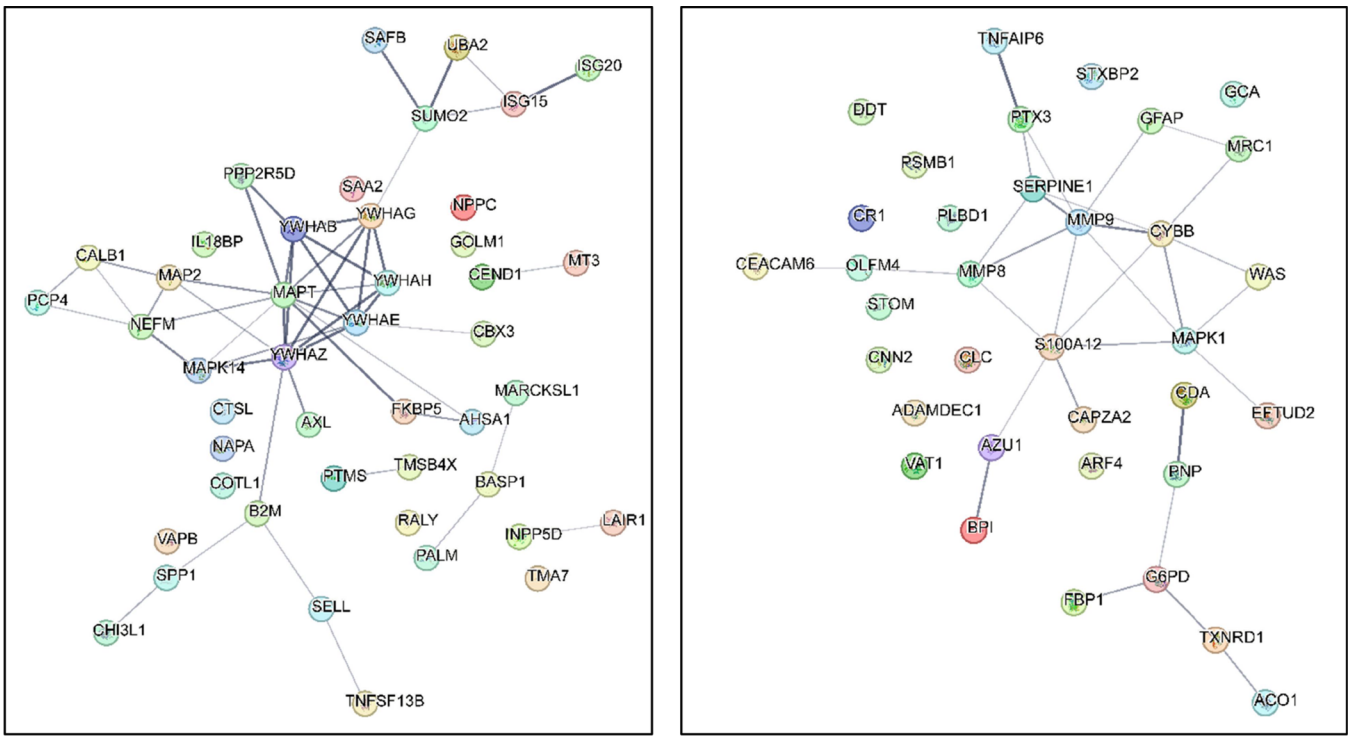


Figure 5. Molecular signatures enriched in JE patient CSF. STRING functional protein association network for upregulated (left) and downregulated (right) proteins. The protein clusters are derived by hierarchically clustering the full STRING network of 83 proteins identified as being important in diagnosing JE using an average linkage algorithm. The smallest clusters generated consist of five proteins and largest used in the enrichment analysis contain 200 proteins. In order to reduce redundancy, clusters with a size difference of less than five proteins toward their child cluster are removed.

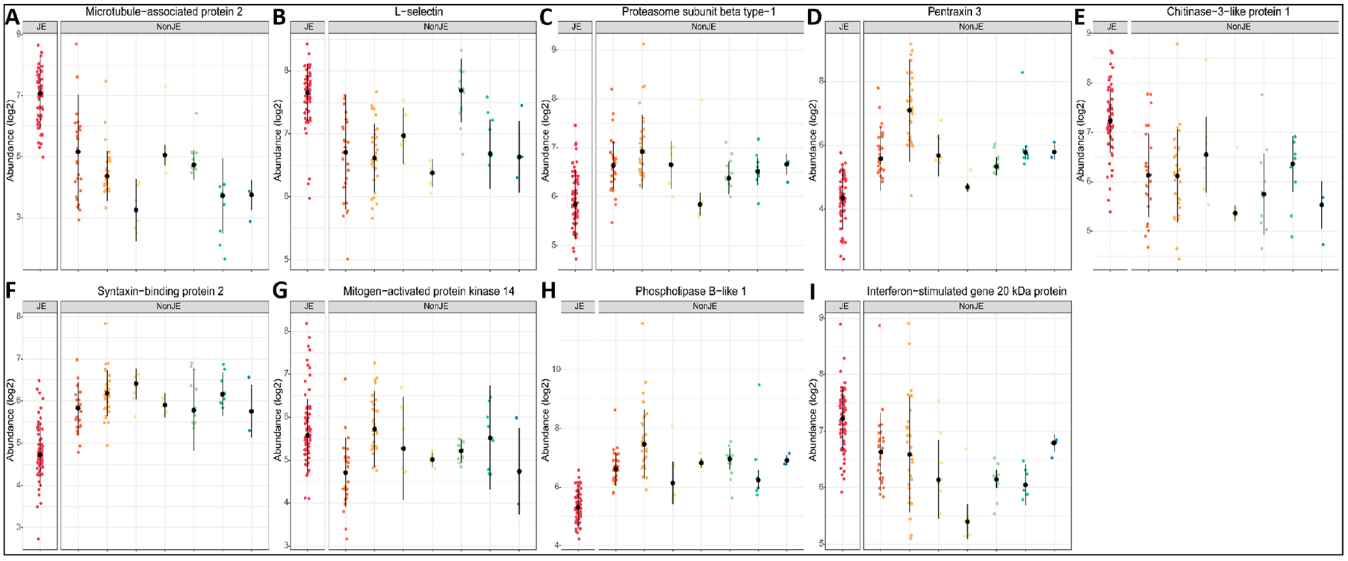


Figure 6. Differential expression across samples in nine proteins as a diagnostic CSF signature of JEV infection. Differential expression across samples in nine proteins as a diagnostic signature of Japanese encephalitis virus infection. Expression profiles of the proteins, microtubule-associated protein 2/P11137/MAP2 (A), L-selectin/P14151/SELL (B), Proteasome subunit beta type-1/P20618/PMSB-1 (C), Pentraxin-3/P26022/PTX3 (D), Chitinase-3-like protein 1/P36222/CH3L1 (E), Syntaxin-binding protein 2/Q15833/STXBP2 (F), Mitogen-activated protein kinase 14/Q16539/MAPK14 (G), Phospholipase B-like 1/Q6P4A8/PLBD1 (H), and Interferon-stimulated gene 20 kDa protein/Q96AZ6/ISG20 (I), are indicated for the categories virus (JE—red and non-JE—dark-orange), bacteria (mainstream—light-orange, TB—yellow, Rickettsia spp.—light-green, *Orientia tsutsugamushi*—dark-green), fungi (turquoise) and parasites (blue).

different locations and in field settings. Ultimately, this could enable the selection of 2−3 proteins for the development of an RDT such as a lateral flow test. We would envisage, as per WHO guidelines, that any testing for JE will be performed contemporaneously with testing for Dengue virus infection and any other endemic flavivirus such as ZIKV and WNV.

https://doi.org/10.1021/acs.jproteome.2c00563
J. Proteome Res. 2023, 22, 1614−1629

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jproteome.2c00563.

S1_Data: LC−MS methods; S2_Patient metadata: Detailed table of the patient metadata; S3_Patient metadata: Summary of included patients' demographics, clinical presentations and details of diagnosis separated into training and test groups; S4_MSstatsTMT output: Results of MSstatsTMT analysis for the pilot (tab 1) and larger (tab 2) study; S5_Contaminants: List of potential protein contaminant proteins that were excluded from downstream analysis; S6_Combined data processed: Normalized and batch corrected data from the pilot and larger study; S7: Assessment of data processing and effectiveness of batch correction; S8_ JE vs non-JE protein set enrichment analysis: Results of protein set enrichment analysis performed using the WebGestalt online tool; S9_WGCNA sample clustering to detect outliers: WGCNA hierarchical clustering of the samples to identify outliers; S10_WGCNA output: Results of WGCNA analysis; S11 WGCNA summary of modules: WGCNA eigengene dendrogram to illustrate groups of correlated modules; S12 Boruta Lasso proteins: List of proteins identified as important in the classification of JE vs non-JE by either Boruta or Lasso algorithms; S13_Predictive modeling to identify a JE diagnostic protein signature; S14_ROC: Receiver operator characteristic curve of the nine-protein model built to classify JE and non-JE patients; S15_Boruta proteins outcome: List of proteins identified as important in predicting the outcome (dead vs alive) of JE patients; and S16 ROC outcome: Receiver operator characteristic curve of the model built to predict the outcome (dead vs alive) of JE patients (ZIP)

■ AUTHOR INFORMATION

Corresponding Author

Tehmina Bharucha − *Department of Biochemistry, University of Oxford, OX1 3QU Oxford, U.K.; Kavli Institute for Nanoscience Discovery, University of Oxford, OX1 3QU Oxford, U.K.; Lao-Oxford-Mahosot Hospital-Wellcome Trust Research Unit (LOMWRU), Microbiology Laboratory, Mahosot Hospital, Vientiane 0100, Lao PDR;* orcid.org/0000-0002-6772-2855; Email: t.bharucha@doctors.org.uk

Authors

Bevin Gangadharan − *Department of Biochemistry, University of Oxford, OX1 3QU Oxford, U.K.; Kavli Institute for Nanoscience Discovery, University of Oxford, OX1 3QU Oxford, U.K.;* orcid.org/0000-0002-7466-2977

Abhinav Kumar − *Department of Biochemistry, University of Oxford, OX1 3QU Oxford, U.K.; Kavli Institute for Nanoscience Discovery, University of Oxford, OX1 3QU Oxford, U.K.*

Ashleigh C. Myall − *Department of Infectious Disease and Department of Mathematics, Imperial College London, London W12 0NN, U.K.*

Nazli Ayhan − *Unité Des Virus Emergents UVE, Aix Marseille Univ, IRD190, INSERM 1207, IHU Méditerranée Infection, Marseille 13005, France*

Boris Pastorino − *Unité Des Virus Emergents UVE, Aix Marseille Univ, IRD190, INSERM 1207, IHU Méditerranée Infection, Marseille 13005, France*

Anisone Chanthongthip − *Lao-Oxford-Mahosot Hospital-Wellcome Trust Research Unit (LOMWRU), Microbiology Laboratory, Mahosot Hospital, Vientiane 0100, Lao PDR*

Manivanh Vongsouvath − *Lao-Oxford-Mahosot Hospital-Wellcome Trust Research Unit (LOMWRU), Microbiology Laboratory, Mahosot Hospital, Vientiane 0100, Lao PDR*

Mayfong Mayxay − *Lao-Oxford-Mahosot Hospital-Wellcome Trust Research Unit (LOMWRU), Microbiology Laboratory, Mahosot Hospital, Vientiane 0100, Lao PDR; Institute of Research and Education Development (IRED), University of Health Sciences, Ministry of Health, Vientiane 43130, Lao PDR; Centre for Tropical Medicine & Global Health, Nuffield Department of Medicine, University of Oxford, Oxford OX3 7LG, U.K.*

Onanong Sengvilaipaseuth − *Lao-Oxford-Mahosot Hospital-Wellcome Trust Research Unit (LOMWRU), Microbiology Laboratory, Mahosot Hospital, Vientiane 0100, Lao PDR*

Ooyanong Phonemixay − *Lao-Oxford-Mahosot Hospital-Wellcome Trust Research Unit (LOMWRU), Microbiology Laboratory, Mahosot Hospital, Vientiane 0100, Lao PDR*

Sayaphet Rattanavong − *Lao-Oxford-Mahosot Hospital-Wellcome Trust Research Unit (LOMWRU), Microbiology Laboratory, Mahosot Hospital, Vientiane 0100, Lao PDR*

Darragh P. O'Brien − *Target Discovery Institute, Centre for Medicines Discovery, Nuffield Department of Medicine, University of Oxford, Oxford OX3 7FZ, U.K.;* orcid.org/0000-0003-4924-7795

Iolanda Vendrell − *Target Discovery Institute, Centre for Medicines Discovery, Nuffield Department of Medicine, University of Oxford, Oxford OX3 7FZ, U.K.; Chinese Academy of Medical Sciences Oxford Institute, Nuffield Department of Medicine, University of Oxford, Oxford OX3 7BN, U.K.*

Roman Fischer − *Target Discovery Institute, Centre for Medicines Discovery, Nuffield Department of Medicine, University of Oxford, Oxford OX3 7FZ, U.K.; Chinese Academy of Medical Sciences Oxford Institute, Nuffield Department of Medicine, University of Oxford, Oxford OX3 7BN, U.K.;* orcid.org/0000-0002-9715-5951

Benedikt Kessler − *Target Discovery Institute, Centre for Medicines Discovery, Nuffield Department of Medicine, University of Oxford, Oxford OX3 7FZ, U.K.; Chinese Academy of Medical Sciences Oxford Institute, Nuffield Department of Medicine, University of Oxford, Oxford OX3 7BN, U.K.;* orcid.org/0000-0002-8160-2446

Lance Turtle − *Institute of Infection, Veterinary and Ecological Sciences, Faculty of Health and Life Sciences, University of Liverpool, Liverpool L69 7BE, U.K.; Tropical and Infectious Disease Unit, Liverpool University Hospitals NHS Foundation Trust (Member of Liverpool Health Partners), Liverpool L69 7BE, U.K.*

Xavier de Lamballerie − *Unité Des Virus Emergents UVE, Aix Marseille Univ, IRD190, INSERM 1207, IHU Méditerranée Infection, Marseille 13005, France*

Audrey Dubot-Pérès − *Lao-Oxford-Mahosot Hospital-Wellcome Trust Research Unit (LOMWRU), Microbiology Laboratory, Mahosot Hospital, Vientiane 0100, Lao PDR; Unité Des Virus Emergents UVE, Aix Marseille Univ, IRD190, INSERM 1207, IHU Méditerranée Infection, Marseille 13005, France; Centre for Tropical Medicine &*

https://doi.org/10.1021/acs.jproteome.2c00563
J. Proteome Res. 2023, 22, 1614−1629

Global Health, Nuffield Department of Medicine, University of Oxford, Oxford OX3 7LG, U.K.

Paul N. Newton − *Lao-Oxford-Mahosot Hospital-Wellcome Trust Research Unit (LOMWRU), Microbiology Laboratory, Mahosot Hospital, Vientiane 0100, Lao PDR; Centre for Tropical Medicine & Global Health, Nuffield Department of Medicine, University of Oxford, Oxford OX3 7LG, U.K.*

Nicole Zitzmann − *Department of Biochemistry, University of Oxford, OX1 3QU Oxford, U.K.; Kavli Institute for Nanoscience Discovery, University of Oxford, OX1 3QU Oxford, U.K.;* orcid.org/0000-0003-1969-4949

SEAe Consortium − *Biology of Infection Unit, Institut Pasteur, 75015 Paris, France*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jproteome.2c00563

Author Contributions

T.B., B.G., N.Z., A.D.P., and P.N.N. conceived the study. The Laos CNS study was completed by A.D.P., P.N.N., X.D.L., M.V., M.M., S.P., A.C., O.S., O.P., and the SEAe collaborators (X.D.L., A.D.P., and P.N.N.). JEV seroneutralization was performed in Marseille by T.B., N.A., B.P. and supervised by A.D.P. and X.D.L. T.B., B.G., and N.Z. developed the methodology for the TMT LC−MS/MS analysis, and T.B. performed the laboratory work with input from A.K., B.G. and D.O. I.V., R.F., and B.K. developed the methodology for the DIA LC−MS/MS analysis, and T.B. and I.V. performed the laboratory work. T.B. performed the data analysis with input from A.M. T.B. wrote the manuscript; all the authors edited successive drafts and approved the final version.

Funding

T.B. was supported by the University of Oxford and the Medical Research Council (grant MR/N013468/1). The work was also supported by the Oxford Glycobiology endowment, the Institute of Research for Development, Aix Marseille University and the European Union's Horizon 2020 research, Fondation Total, Institut Pasteur, International Network Institut Pasteur, Fondation Merieux, Aviesan Sud, Institut national de la santé et de la recherche médicale (Inserm), and innovation programme EVAg (grant agreement 653316). This research was funded in part by the Wellcome Trust. For the purpose of Open Access, the author has applied a CC BY public copyright license to any Author Accepted Manuscript version arising from this submission. A.M. was supported by a scholarship from the Medical Research Foundation National PhD Training Programme in Antimicrobial Resistance Research (MRF-145-0004-TPG-AVISO). D.P.O., I.V., R.F., and B.K. are supported by the Chinese Academy of Medical Sciences (CAMS) Innovation Fund for Medical Science (CIFMS), China (grant number: 2018-I2M-2-002). L.T. is a Wellcome clinical career development fellow, supported by grant number 205228/Z/16/Z, and the NIHR Health Protection Research Unit in emerging and zoonotic infections (NIHR200907) at University of Liverpool in partnership with Public Health England (PHE), in collaboration with Liverpool School of Tropical Medicine and the University of Oxford.

Notes

The authors declare no competing financial interest.

Ethical clearance for the Laos CNS study was granted by the Ethical Review Committee of the former Faculty of Medical Sciences, National University of Laos (now University of Health Sciences) and the Oxford University Tropical Ethics Research Committee, Oxford, UK.

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium[92] via the PRIDE partner repository with the dataset identifier PXD034789 and 10.6019/PXD034789. All other data underlying this article are available in the article and in its online Supplementary Material.

■ ACKNOWLEDGMENTS

We are very grateful to the patients and to Bounthaphany Bounxouei, the former Director of Mahosot Hospital, the late Rattanaphone Phetsouvanh, Director of the Microbiology Laboratory, and the staff of the wards and Microbiology Laboratory of Mahosot Hospital. We also thank Bounnak Saysanasongkham, the former Director of the Department of Healthcare and Rehabilitation, Ministry of Health, and Bounkong Syhavong, the former Minister of Health, Lao PDR for their very kind help and support. We thank the members of the Unité des Virus Émergents (Christine Isnard and Camille Placidi) and the CNR des Arbovirus (Patrick Gravier, Gilda Grard, Isabelle Leparc-Goffart and Mathilde Galla). Finally, we acknowledge useful discussions on the data analysis with Andrew R. Jones at the University of Liverpool, and Damien Ming, Mauricio Barahona and Robert Peach at Imperial College London. The Southeast Asia Encephalitis Project (SEAe) aimed to strengthen and harmonize the identification of causes of encephalitis and microbiological diagnostic capacity of public referral laboratories in Cambodia, Laos, Vietnam, and Myanmar.[25] The consortium set up and conducted the SEAe project; involving writing the protocol, recruiting patients, performing diagnostic testing and analyzing data. A key finding of the SEAe study was that JEV remains a leading etiology of neurological infection in the regions studied, and this has led onto the study reported here, utilizing the SEAe patient samples from Laos to identify diagnostic protein biomarkers of JE. The proteomic study is in line with the objectives of the SEAe project. We are grateful to all the SEAe study researchers, including Philippe Buchy, Em Bunnakea, Julien Cappelle, Mey Channa, Veronique Chevalier, Yoann Crabol, Philippe Dussart, Marc Eloit, Christopher Gorma, Magali Herrant, Nguyen Hien, Chaw Su Hlaing, Jérôme Honnorat, Tran Thi Mai Hung, Tran Thi Thu Huong, Latt Latt Kyaw, Nguyen Van Lam, Denis Laurent, Marc Lecuit, Kyaw Linn, Olivier Lortholary, Aye Mya Min Aye, Philippe Perot, Sommanikhone Phangmanixay, Khounthavy Phongsavath, Phan Huu Phuc, Anne-Laurie Pinto, Patrice Piola, Jean-David Pommier, Bruno Rosset, Ky Santy, Heng Sothy, Arnaud Tarantola, Nguyen Thi Thu Thuy, Htay Tin, Ommar Swe Tin, Pham Nhat An, Dang Duc Anh, Pascal Bonnet, Kimrong Bun, Danoy Chommanam, Viengmon Davong, Patrice Debré, Jean-François Delfraissy, Christian Devaux, Anousone Douangnouvong, Veasna Duong, Benoit Durand, Chanreaksmey Eng, Catherine Ferrant, Didier Fontenille, Lukas Hafner, Le Thanh Hai, Do Thu Huong, Marc Jouan, May July, Magali Lago, Jean-Paul Moatti, Bernadette Murgue, Khin Yi Oo, MengHeng Oum, Khansoudaphone Phakhounthong, Anh Tuan Pham, Do Quyen, Malee Seephonelee, Maud Seguy, Bountoy Sibounheunang, Kanarith Sim, Luong Minh Tan, Cho Thair, Win Thein, Phung Bich Thuy, Hervé Tissot-Dupont, and Malavanh Vongsouvath.

REFERENCES

(1) Moore, S. M. The current burden of Japanese encephalitis and the estimated impacts of vaccination: Combining estimates of the spatial distribution and transmission intensity of a zoonotic pathogen. *PLoS Negl. Trop. Dis.* **2021**, *15*, No. e0009385.

(2) Mulvey, P.; Duong, V.; Boyer, S.; Burgess, G.; Williams, D. T.; Dussart, P.; et al. The Ecology and Evolution of Japanese Encephalitis Virus. *Pathogens* **2021**, *10*, 1534.

(3) Health AGDo. Japanese encephalitis virus 2022 Available from: https://www.health.gov.au/health-alerts/japanese-encephalitis-virus-jev/about.

(4) Mayxay, M.; Douangdala, P.; Vilayhong, C.; Phommasone, K.; Chansamouth, V.; Vongsouvath, M.; et al. Outcome of Japanese Encephalitis Virus (JEV) Infection in Pediatric and Adult Patients at Mahosot Hospital, Vientiane, Lao PDR. *Am. J. Trop. Med. Hyg.* **2020**, *104*, 567−575.

(5) Turtle, L.; Solomon, T. Japanese encephalitis - the prospects for new treatments. *Nat. Rev. Neurol.* **2018**, *14*, 298−313.

(6) Vannice, K. S.; Hills, S. L.; Schwartz, L. M.; Barrett, A. D.; Heffelfinger, J.; Hombach, J.; et al. The future of Japanese encephalitis vaccination: expert recommendations for achieving and maintaining optimal JE control. *npj Vaccines* **2021**, *6*, 82.

(7) Heffelfinger, J. D.; Li, X.; Batmunkh, N.; Grabovac, V.; Diorditsa, S.; Liyanage, J. B.; et al. Japanese Encephalitis Surveillance and Immunization - Asia and Western Pacific Regions, 2016. *MMWR Morb. Mortal. Weekly Rep.* **2017**, *66*, 579−583.

(8) Bharucha, T.; Shearer, F. M.; Vongsouvath, M.; Mayxay, M.; de Lamballerie, X.; Newton, P. N.; et al. A need to raise the bar - A systematic review of temporal trends in diagnostics for Japanese encephalitis virus infection, and perspectives for future research. *Int. J. Infect. Dis.* **2020**, *95*, 444−456.

(9) Robinson, J. S.; Featherstone, D.; Vasanthapuram, R.; Biggerstaff, B. J.; Desai, A.; Ramamurty, N.; et al. Evaluation of three commercially available Japanese encephalitis virus IgM enzyme-linked immunosorbent assays. *Am. J. Trop. Med. Hyg.* **2010**, *83*, 1146−1155.

(10) Dubot-Perés, A.; Sengvilaipaseuth, O.; Chanthongthip, A.; Newton, P. N.; de Lamballerie, X. How many patients with anti-JEV IgM in cerebrospinal fluid really have Japanese encephalitis? *Lancet Infect. Dis.* **2015**, *15*, 1376−1377.

(11) Hills, S.; Van Keulen, A.; Feser, J.; Panella, A.; Letson, B.; Staples, E.; et al. Persistence of IgM Antibodies after Vaccination with Live Attenuated Japanese Encephalitis Vaccine. *Am. J. Trop. Med. Hyg.* **2021**, *104*, 576−579.

(12) Bharucha, T.; Sengvilaipaseuth, O.; Vongsouvath, M.; Vongsouvath, M.; Davong, V.; Panyanouvong, P.; et al. Development of an improved RT-qPCR Assay for detection of Japanese encephalitis virus (JEV) RNA including a systematic review and comprehensive comparison with published methods. *PLoS One* **2018**, *13*, No. e0194412.

(13) Bharucha, T.; Gangadharan, B.; Kumar, A.; de Lamballerie, X.; Newton, P. N.; Winterberg, M.; et al. Mass spectrometry-based proteomic techniques to identify cerebrospinal fluid biomarkers for diagnosing suspected central nervous system infections. A systematic review. *J. Infect.* **2019**, *79*, 407−418.

(14) Dayon, L.; Cominetti, O.; Affolter, M. Proteomics of human biological fluids for biomarker discoveries: technical advances and recent applications. *Expert. Rev. Proteom.* **2022**, *19*, 131−151.

(15) Newman, M. *Networks*, 2nd edition; Oxford University Press: Oxford, 2018.

(16) Hastie, T., Tibshirani, R., Friedman, J. H., Friedman, J. H. *The elements of statistical learning: data mining, inference, and prediction*, 2nd edition; Springer: New York, 2009.

(17) Barabási, A.-L. Network science. *Philos. Trans. R. Soc., A* **1987**, *2013*, No. 20120375.

(18) Barabási, A.-L.; Gulbahce, N.; Loscalzo, J. Network medicine: a network-based approach to human disease. *Nat. Rev. Genet.* **2011**, *12*, 56−68.

(19) Carr, A. V.; Frey, B. L.; Scalf, M.; Cesnik, A. J.; Rolfs, Z.; Pike, K. A.; et al. MetaNetwork Enhances Biological Insights from Quantitative Proteomics Differences by Combining Clustering and Enrichment Analyses. *J. Proteome Res.* **2022**, *21*, 410−419.

(20) Rudolph, J. D.; Cox, J. A Network Module for the Perseus Software for Computational Proteomics Facilitates Proteome Interaction Graph Analysis. *J. Proteome Res.* **2019**, *18*, 2052−2064.

(21) Wu, J. X.; Pascovici, D.; Wu, Y.; Walker, A. K.; Mirzaei, M. Workflow for Rapidly Extracting Biological Insights from Complex, Multicondition Proteomics Experiments with WGCNA and PloGO2. *J. Proteome Res.* **2020**, *19*, 2898−2906.

(22) Mann, M.; Kumar, C.; Zeng, W. F.; Strauss, M. T. Artificial intelligence for proteomics and biomarker discovery. *Cell Syst.* **2021**, *12*, 759−770.

(23) Myall, A. C.; Perkins, S.; Rushton, D.; David, J.; Spencer, P.; Jones, A. R.; et al. An OMICs-based meta-analysis to support infection state stratification. *Bioinformatics* **2021**, *37*, 2347−2355.

(24) Dubot-Perés, A.; Mayxay, M.; Phetsouvanh, R.; Lee, S. J.; Rattanavong, S.; Vongsouvath, M.; et al. Management of Central Nervous System Infections, Vientiane, Laos, 2003-2011. *Emerg. Infect. Dis.* **2019**, *25*, 898−910.

(25) Pommier, J. D.; Gorman, C.; Crabol, Y.; Bleakley, K.; Sothy, H.; Santy, K.; et al. Childhood encephalitis in the Greater Mekong region (the SouthEast Asia Encephalitis Project): a multicentre prospective study. *Lancet Glob. Health* **2022**, *10*, e989−e1002.

(26) Bharucha, T.; Ayhan, N.; Pastorino, B.; Rattanavong, S.; Vongsouvath, M.; Mayxay, M.; et al. Immunoglobulin M seroneutralization for improved confirmation of Japanese encephalitis virus infection in a flavivirus-endemic area. *Trans. R. Soc. Trop. Med. Hyg.* **2022**, *116*, 1032−1042.

(27) Muntel, J.; Kirkpatrick, J.; Bruderer, R.; Huang, T.; Vitek, O.; Ori, A.; et al. Comparison of Protein Quantification in a Complex Background by DIA and TMT Workflows with Fixed Instrument Time. *J. Proteome Res.* **2019**, *18*, 1340−1351.

(28) Pounds, S. 2 FDRsampsize-package. In *Compute Sample Size that Meets Requirements for Average Power*, 2016; Vol. *2*.

(29) Eng, J. K.; McCormack, A. L.; Yates, J. R. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 976−989.

(30) Käll, L.; Canterbury, J. D.; Weston, J.; Noble, W. S.; MacCoss, M. J. Semi-supervised learning for peptide identification from shotgun proteomics datasets. *Nat. Methods* **2007**, *4*, 923.

(31) Huang, T.; Choi, M.; Tzouros, M.; Golling, S.; Pandya, N. J.; Banfai, B.; et al. MSstatsTMT: Statistical Detection of Differentially Abundant Proteins in Experiments with Isobaric Labeling and Multiple Mixtures. *Mol. Cell. Proteom.* **2020**, *19*, 1706−1723.

(32) Ma, W.; Kim, S.; Chowdhury, S.; Li, Z.; Yang, M.; Yoo, S.; et al. DreamAI: algorithm for the imputation of proteomics data. *bioRxiv* **2021**, DOI: 10.1101/2020.07.21.214205.

(33) Wang, M.; Jiang, L.; Jian, R.; Chan, J. Y.; Liu, Q.; Snyder, M. P.; et al. RobNorm: model-based robust normalization method for labeled quantitative mass spectrometry proteomics data. *Bioinformatics* **2021**, *37*, 815−821.

(34) Johnson, W. E.; Li, C.; Rabinovic, A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* **2007**, *8*, 118−127.

(35) Leek, J. T.; Johnson, W. E.; Parker, H. S.; Fertig, E. J.; Jaffe, A. E.; Storey, J. D.; et al. sva: Surrogate variable analysis. *R package version* **2021**, *3*, 882−883.

(36) Pang, Z.; Chong, J.; Zhou, G.; de Lima Morais, D. A.; Chang, L.; Barrette, M.; et al. MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. *Nucleic Acids Res.* **2021**, *49*, W388−W396.

(37) Liao, Y.; Wang, J.; Jaehnig, E. J.; Shi, Z.; Zhang, B. WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. *Nucleic Acids Res.* **2019**, *47*, W199−W205.

(38) Langfelder, P.; Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinform.* **2008**, *9*, 559.

(39) Demichev, V.; Messner, C. B.; Vernardis, S. I.; Lilley, K. S.; Ralser, M. DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. *Nat. Methods* **2020**, *17*, 41−44.
(40) Kursa, M. B.; Rudnicki, W. R. Feature selection with the Boruta package. *J. Stat. Softw.* **2010**, *36*, 1−13.
(41) Friedman, J.; Hastie, T.; Tibshirani, R.; Narasimhan, B.; Tay, K.; Simon, N. et al. Package 'glmnet'. *J. Stat. Softw. 2010a* **2021**, *33* ().
(42) Wen, P.; Dayyani, F.; Tao, R.; Zhong, X. Screening and verification of potential gene targets in esophageal carcinoma by bioinformatics analysis and immunohistochemistry. *Ann. Transl. Med.* **2022**, *10*, 70.
(43) Kuhn, M.; Wing, J.; Weston, S.; Williams, A.; Keefer, C.; Engelhardt, A.; et al. Package 'caret'. *R J.* **2020**, *223*, 7.
(44) Szklarczyk, D.; Gable, A. L.; Lyon, D.; Junge, A.; Wyder, S.; Huerta-Cepas, J.; et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **2019**, *47*, D607−D613.
(45) Mi, H.; Huang, X.; Muruganujan, A.; Tang, H.; Mills, C.; Kang, D.; et al. PANTHER version 11: expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. *Nucleic Acids Res.* **2017**, *45*, D183−D189.
(46) Petryszak, R.; Burdett, T.; Fiorelli, B.; Fonseca, N. A.; Gonzalez-Porta, M.; Hastings, E.; Huber, W.; Jupp, S.; Keays, M.; Kryvych, N.; McMurry, J.; Marioni, J. C.; Malone, J.; Megy, K.; Rustici, G.; Tang, A. Y.; Taubert, J.; Williams, E.; Mannion, O.; Parkinson, H. E.; Brazma, A. Expression Atlas update−a database of gene and transcript expression from microarray- and sequencing-based functional genomics experiments. *Nucleic Acids Res.* **2014**, *42*, D926−D932.
(47) Uhlén, M.; Fagerberg, L.; Hallström, B. M.; Lindskog, C.; Oksvold, P.; Mardinoglu, A.; et al. Tissue-based map of the human proteome. *Science* **2015**, *347*, No. 1260419.
(48) Macron, C.; Lavigne, R.; Núñez Galindo, A.; Affolter, M.; Pineau, C.; Dayon, L. Exploration of human cerebrospinal fluid: A large proteome dataset revealed by trapped ion mobility time-of-flight mass spectrometry. *Data Brief* **2020**, *31*, No. 105704.
(49) Geyer, P. E.; Holdt, L. M.; Teupser, D.; Mann, M. Revisiting biomarker discovery by plasma proteomics. *Mol. Syst. Biol.* **2017**, *13*, 942.
(50) Halder, A.; Verma, A.; Biswas, D.; Srivastava, S. Recent advances in mass-spectrometry based proteomics software, tools and databases. *Drug Discov. Today Technol.* **2021**, *39*, 69−79.
(51) Sengupta, N.; Mukherjee, S.; Tripathi, P.; Kumar, R.; Suryavanshi, A.; Basu, A. Cerebrospinal Fluid Biomarkers of Japanese Encephalitis. *F1000Research* **2015**, *4*, 334.
(52) Yin, R.; Yang, L.; Hao, Y.; Yang, Z.; Lu, T.; Jin, W.; et al. Proteomic landscape subtype and clinical prognosis of patients with the cognitive impairment by Japanese encephalitis infection. *J. Neuroinflamm.* **2022**, *19*, 77.
(53) Baluni, M.; Ghildiyal, S.; Fatima, T.; Tiwari, R.; Upadhyay, S.; Dhole, T. N.; et al. Differential expression of circulating micro RNAs in serum: Potential biomarkers to track Japanese encephalitis virus infection. *J. Med. Virol.* **2022**, *94*, 531−539.
(54) Deval, H.; Alagarasu, K.; Srivastava, N.; Bachal, R.; Mittal, M.; Agrawal, A.; et al. Association of single nucleotide polymorphisms in the CD209, MMP9, TNFA and IFNG genes with susceptibility to Japanese encephalitis in children from North India. *Gene* **2022**, *808*, No. 145962.
(55) Shukla, V.; Shakya, A. K.; Dhole, T. N.; Misra, U. K. Matrix metalloproteinases and their tissue inhibitors in serum and cerebrospinal fluid of children with Japanese encephalitis virus infection. *Arch. Virol.* **2013**, *158*, 2561−2575.
(56) Singh, A.; Kulshreshtha, R.; Mathur, A. Secretion of the chemokine interleukin-8 during Japanese encephalitis virus infection. *J. Med. Microbiol.* **2000**, *49*, 607−612.
(57) Son, H.; Sunwoo, J. S.; Lee, S. K.; Chu, K.; Lee, S. T. Clinical Outcomes of Japanese Encephalitis after Combination Treatment of Immunoglobulin, Ribavirin, and Interferon-α2b. *J. Clin. Neurol.* **2021**, *17*, 428−434.
(58) Tiwari, R.; Ghildiyal, S.; Baluni, M.; Singh, D.; Srivastva, J. K.; Kumar, R.; et al. Association of interleukin-6 (174 G/C) and interleukin-12B (1188 A/C) gene polymorphism with expression and risk of Japanese encephalitis disease in North Indian population. *J. Neuroimmunol.* **2021**, *358*, No. 577630.
(59) Bhaskar, M.; Mukherjee, S.; Basu, A. Involvement of RIG-I Pathway in Neurotropic Virus-Induced Acute Flaccid Paralysis and Subsequent Spinal Motor Neuron Death. *MBio* **2021**, *12*, No. e0271221.
(60) Jhan, M. K.; Chen, C. L.; Shen, T. J.; Tseng, P. C.; Wang, Y. T.; Satria, R. D.; et al. Polarization of Type 1 Macrophages Is Associated with the Severity of Viral Encephalitis Caused by Japanese Encephalitis Virus and Dengue Virus. *Cell* **2021**, *10*, 3181.
(61) Li, Q.; Zhou, D.; Jia, F.; Zhang, L.; Ashraf, U.; Li, Y.; et al. Japanese Encephalitis Virus NS1' Protein Interacts with Host CDK1 Protein to Regulate Antiviral Response. *Microbiol. Spectr.* **2021**, *9*, No. e0166121.
(62) Swarup, V.; Ghosh, J.; Duseja, R.; Ghosh, S.; Basu, A. Japanese encephalitis virus infection decrease endogenous IL-10 production: correlation with microglial activation and neuronal death. *Neurosci. Lett.* **2007**, *420*, 144−149.
(63) Tripathi, A.; Banerjee, A.; Vrati, S. Development and characterization of an animal model of Japanese encephalitis virus infection in adolescent C57BL/6 mouse. *Dis. Models Mech.* **2021**, *14*, No. dmm049176.
(64) Li, M.; Yang, J.; Ye, C.; Bian, P.; Yang, X.; Zhang, H.; et al. Integrated Metabolomics and Transcriptomics Analyses Reveal Metabolic Landscape in Neuronal Cells during JEV Infection. *Virol. Sin.* **2021**, *36*, 1554−1565.
(65) Liu, H.; Zhang, J.; Niu, Y.; Liang, G. The 5′ and 3' Untranslated Regions of the Japanese Encephalitis Virus (JEV): Molecular Genetics and Higher Order Structures. *Front. Microbiol.* **2021**, *12*, No. 730045.
(66) Sharma, K. B.; Chhabra, S.; Aggarwal, S.; Tripathi, A.; Banerjee, A.; Yadav, A. K.; et al. Proteomic landscape of Japanese encephalitis virus-infected fibroblasts. *J. Gen. Virol.* **2021**, *102*, No. 001657.
(67) Wang, P.; Liu, X.; Li, Q.; Wang, J.; Ruan, W. Proteomic analyses identify intracellular targets for Japanese encephalitis virus nonstructural protein 1 (NS1). *Virus Res.* **2021**, *302*, No. 198495.
(68) Xu, P.; Tong, W.; Chen, Y. M. FUSE binding protein FUBP3 is a potent regulator in Japanese encephalitis virus infection. *Virol. J.* **2021**, *18*, 224.
(69) Yang, J.; Li, M.; Yuan, M.; Bian, P.; Dong, Y.; Zhang, H.; et al. Axl(−/−) neurons promote JEV infection by dampening the innate immunity. *Virus Res.* **2022**, *307*, No. 198605.
(70) Kosti, I.; Jain, N.; Aran, D.; Butte, A. J.; Sirota, M. Cross-tissue Analysis of Gene and Protein Expression in Normal and Cancer Tissues. *Sci. Rep.* **2016**, *6*, 24799.
(71) Li, Y. Z.; Counor, D.; Lu, P.; Liang, G. D.; Vu, T. Q.; Phan, T. N. A specific and sensitive antigen capture assay for NS1 protein quantitation in Japanese encephalitis virus infection. *J. Virol. Methods* **2012**, *179*, 8−16.
(72) Fraisier, C.; Papa, A.; Granjeaud, S.; Hintzen, R.; Martina, B.; Camoin, L.; et al. Cerebrospinal fluid biomarker candidates associated with human WNV neuroinvasive disease. *PLoS One* **2014**, *9*, No. e93637.
(73) Ashraf, U.; Ding, Z.; Deng, S.; Ye, J.; Cao, S.; Chen, Z. Pathogenicity and virulence of Japanese encephalitis virus: Neuro-inflammation and neuronal cell damage. *Virulence* **2021**, *12*, 968−980.
(74) Hohmann, T.; Dehghani, F. The Cytoskeleton-A Complex Interacting Meshwork. *Cell* **2019**, *8*, 362.
(75) Dittrich, S.; Sunyakumthorn, P.; Rattanavong, S.; Phetsouvanh, R.; Panyanivong, P.; Sengduangphachanh, A.; et al. Blood-Brain Barrier Function and Biomarkers of Central Nervous System Injury in Rickettsial Versus Other Neurological Infections in Laos. *Am. J. Trop. Med. Hyg.* **2015**, *93*, 232−237.

https://doi.org/10.1021/acs.jproteome.2c00563
J. Proteome Res. 2023, 22, 1614−1629

(76) Foo, K. Y.; Chee, H.-Y. Interaction between Flavivirus and Cytoskeleton during Virus Replication. *Biomed. Res. Int.* **2015**, *2015*, No. 427814.
(77) Schoggins, J. W. Interferon-Stimulated Genes: What Do They All Do? *Ann. Rev. Virol.* **2019**, *6*, 567−584.
(78) Nair, S.; Diamond, M. S. Innate immune interactions within the central nervous system modulate pathogenesis of viral infections. *Curr. Opin. Immunol.* **2015**, *36*, 47−53.
(79) Pan, Y.; Cheng, A.; Wang, M.; Yin, Z.; Jia, R. The Dual Regulation of Apoptosis by Flavivirus. *Front. Microbiol.* **2021**, *12*, No. 654494.
(80) Roby, J. A.; Esser-Nobis, K.; Dewey-Verstelle, E. C.; Fairgrieve, M. R.; Schwerk, J.; Lu, A. Y.; et al. Flavivirus Nonstructural Protein NS5 Dysregulates HSP90 to Broadly Inhibit JAK/STAT Signaling. *Cell* **2020**, *9*, 899.
(81) Wang, Y.; Li, Y.; Ding, T. Heat shock protein 90β in the Vero cell membrane binds Japanese encephalitis virus. *Int. J. Mol. Med.* **2017**, *40*, 474−482.
(82) Lewy, T. G.; Grabowski, J. M.; Bloom, M. E. BiP: Master Regulator of the Unfolded Protein Response and Crucial Factor in Flavivirus Biology. *Yale J. Biol. Med.* **2017**, *90*, 291−300.
(83) Chhajer, H.; Rizvi, V. A.; Roy, R. Life cycle process dependencies of positive-sense RNA viruses suggest strategies for inhibiting productive cellular infection. *J. R. Soc., Interface* **2021**, *18*, No. 20210401.
(84) Burke, D. S.; Nisalak, A.; Ussery, M. A.; Laorakpongse, T.; Chantavibul, S. Kinetics of IgM and IgG Responses to Japanese Encephalitis Virus in Human Serum and Cerebrospinal Fluid. *J. Infect. Dis.* **1985**, *151*, 1093−1099.
(85) Sooryanarain, H.; Ayachit, V.; Gore, M. Activated CD56(+) lymphocytes (NK+NKT) mediate immunomodulatory and anti-viral effects during Japanese encephalitis virus infection of dendritic cells in-vitro. *Virology* **2012**, *432*, 250−260.
(86) Myint, K. S. A.; Kipar, A.; Jarman, R. G.; Gibbons, R. V.; Perng, G. C.; Flanagan, B.; Mongkolsirichaikul, D.; Van Gessel, Y.; Solomon, T. Neuropathogenesis of Japanese encephalitis in a primate model. *PLoS Negl. Trop. Dis.* **2014**, *8*, No. e2980.
(87) Kumar, A.; Kalita, J.; Sinha, R. A.; Singh, G.; Shukla, M.; Tiwari, S.; Dhole, T. N.; Misra, U. K. Impaired Autophagy Flux is Associated with Proinflammatory Microglia Activation Following Japanese Encephalitis Virus Infection. *Neurochem. Res.* **2020**, *45*, 2184−2195.
(88) Wang, Z. Y.; Zhen, Z. D.; Fan, D. Y.; Wang, P. G.; An, J. Transcriptomic Analysis Suggests the M1 Polarization and Launch of Diverse Programmed Cell Death Pathways in Japanese Encephalitis Virus-Infected Macrophages. *Viruses* **2020**, *12*, 356.
(89) Bakochi, A.; Mohanty, T.; Pyl, P. T.; Gueto-Tettay, C. A.; Malmström, L.; Linder, A.; et al. Cerebrospinal fluid proteome maps detect pathogen-specific host response patterns in meningitis. *Elife* **2021**, *10*, No. e64159.
(90) Zatta, M.; Di Bella, S.; Bottazzi, B.; Rossi, F.; D'Agaro, P.; Segat, L.; et al. Determination of pentraxin 3 levels in cerebrospinal fluid during central nervous system infections. *Eur. J. Clin. Microbiol. Infect. Dis.* **2020**, *39*, 665−670.
(91) Behairy Bel, S.; Salama, E. I.; Allam, A. A.; Ali, M. A.; Elaziz, A. M. Lipocalin-2 as a marker of bacterial infections in chronic liver disease: a study in Egyptian children. *Egypt. J. Immunol.* **2011**, *18*, 31−36.
(92) Deutsch, E. W.; Bandeira, N.; Sharma, V.; Perez-Riverol, Y.; Carver, J. J.; Kundu, D. J.; et al. The Proteome Xchange consortium in 2020: enabling 'big data' approaches in proteomics. *Nucleic Acids Res.* **2019**, *48*, D1145−D1152.

https://doi.org/10.1021/acs.jproteome.2c00563
J. Proteome Res. 2023, 22, 1614−1629