Acta Neuropathologica Communications

RESEARCH

Open Access

Distinct cerebrospinal fluid proteomic signatures define clinicopathological subtypes of sporadic Creutzfeldt-Jakob disease and predict patient survival

Giuseppe Mario Bentivenga[1†], Angela Mammana[2†], Dea Gogishvili[3†], Simone Baiardi[1,2], Erica Vittoriosi[1], Andrea Mastrangelo[1,2], Agustina Ranieri[2], Isabel M. Houtkamp[3,4], Kathrin Brockmann[5,6], Sanne Abeln[3,7], Sabina Capellari[1,2] and Piero Parchi[1,2*]

Abstract

Sporadic Creutzfeldt-Jakob disease (sCJD) is a highly heterogeneous neurodegenerative disorder encompassing six major histopathological and molecular subtypes, showing diverse clinical features and prognosis. Currently, no accurate biomarkers are available for the antemortem differentiation and prognostication of sCJD subtypes. We retrospectively analyzed the cerebrospinal fluid (CSF) proteome from 126 sCJD patients belonging to the prevalent MM(V)1, VV2, and MV2K subtypes and 42 non-neurodegenerative controls (CTRL) using proximity extension assay technology, quantifying 797 unique proteins in each sample. Differential expression analysis, machine learning models, and Cox regression analyses were employed to uncover subtype-specific protein signatures and identify diagnostic and prognostic biomarkers. A workflow combining Weighted Gene Co-expression Network Analysis (WGCNA) and Hierarchical HotNet was used for the unsupervised discovery of novel dysregulated biological pathways. A subset of proteins, including HDGF, FOSB, PAG1, APEX1, CCDC80, WASF1, and GPC5, emerged as robust biomarkers for subtype classification (model receiver operating characteristics-area under the curve 0.93). WASF1, CCDC80, and GPC5 were among the most informative biomarkers distinguishing MM(V)1 from V2 strain-related subtypes (VV2 and MV2K). Survival analysis identified ten proteins, most notably CCDC80 (HR 1.49, 95% CI 1.17–1.91, $p=0.001$), as independent prognostic biomarkers in prion disease. WGCNA identified six modules of co-regulated proteins underlying distinct biological processes. Some, such as response to toxic substances, nervous system development, and chemotaxis, were altered across all subtypes compared to the CTRL group (all $p<0.01$), while others showed subtype-specific dysregulation. Specifically, the intracellular signal transduction module was selectively altered in V2 strain-related subtypes, while the epithelium morphogenesis one was dysregulated in MM(V)1 (all $p<0.001$). In conclusion, this study reveals distinct and shared CSF proteomic signatures across sCJD

†Giuseppe Mario Bentivenga, Angela Mammana and Dea Gogishvili contributed equally to this study.

*Correspondence:
Piero Parchi
piero.parchi@unibo.it

Full list of author information is available at the end of the article

subtypes, identifying novel diagnostic and prognostic biomarkers and dysregulated molecular pathways. These findings enhance our understanding of prion disease heterogeneity and suggest new pathogenic mechanisms.

Keywords Prion, Creutzfeldt–Jakob disease, Biomarker, Cerebrospinal fluid, Proteomic, Proximity extension assay, Dementia, Machine learning, Co-expression module, Strain

Introduction

Sporadic Creutzfeldt-Jakob disease (sCJD), the most common human prion disease, is a highly heterogeneous neurodegenerative disorder related to prion protein (PrP) misfolding. The current sCJD classification recognises six main clinicopathological subtypes depending on the polymorphic codon 129 genotype (encoding methionine, M or valine, V) of the prion protein gene (*PRNP*), the type (1 or 2) of misfolded PrP (PrP^{Sc}) accumulating in the brain, and distinctive histopathological characteristics [43, 44]. Each sCJD subtype arises from a specific combination of codon 129 genotype and PrP^{Sc} type, with three key exceptions: (1) MM1 and MV1 are phenotypically identical and grouped as MM(V)1; (2) MM2 includes two subtypes with distinct lesion profiles: MM2-Cortical (MM2C) and MM2-Thalamic (MM2T); (3) MV2 consists of two subtypes-MV2-Kuru (MV2K), marked by kuru plaques, and MV2C, resembling MM2C (together known as MM(V)2C subtype) [4, 43]. Approximately 35% of patients, often with the MM-codon 129 genotype, exhibit a "mixed" phenotype, showing features of "pure" subtypes but dominated by the more prevalent PrP^{Sc} type [12, 45]. The histopathological diversity of sCJD subtypes is recapitulated by the concept of "prion strains" (thought to be enciphered by different PrP^{Sc} conformers), i.e., human isolates that, when inoculated into syngenic hosts, produce diseases with varying characteristics, such as incubation periods and brain regional distribution of lesions [4, 70]. Transmission studies revealed five distinct prion strains among the six sCJD subtypes, with the VV2 and MV2K sharing transmission properties and now collectively referred to as the V2 strain [4, 9, 19]. In addition to molecular and neuropathological features, the six main subtypes of sCJD differ significantly in clinical presentation. MM(V)1, the most common (65% of cases), presents as a multisystemic neurological syndrome with an average disease duration of 4 months. VV2 (15–20%) is characterised by cerebellar and subcortical impairment, lasting approximately 6 months. MV2K (10–15%) is marked by subcortical involvement with a longer average duration of 17 months. The remaining three subtypes are much rarer (about 5% collectively). They are characterised by prominent cortical symptoms (MM(V)2 C and VV1) or show a peculiar clinical phenotype known as sporadic fatal insomnia (MM2T) [70]. Despite the growing clinical need to distinguish sCJD subtypes ante mortem for prognostic and epidemiological purposes, a precise subtype-specific diagnosis is currently possible only at autopsy due to the partial overlap of clinical characteristics. Although brain magnetic resonance imaging (MRI) and cerebrospinal fluid (CSF) markers such as total tau (t-tau), 14-3-3, and neurofilament light assays have some predictive value, the accuracy of in vivo diagnosis of clinicopathological subtypes of sCJD remains suboptimal [10, 21, 25]. Furthermore, the development of targeted therapies depends on a deeper understanding of the molecular mechanisms and biological pathways that drive the heterogeneity of sCJD [69]. Since CSF reflects biochemical alterations in the brain, it is increasingly used in proteomic studies to identify molecular signatures of neurodegenerative disorders [17]. Moreover, this approach holds promise for identifying novel biomarkers for diagnosis, prognosis, and potential therapeutic targets [14, 15, 22]. In this study, we employed high-throughput proteomics with the proximity extension assay (PEA) technology [2] to analyse CSF samples from a large, well-characterised cohort of sCJD patients representing the three most common disease subtypes. Our comprehensive approach aimed to address three critical challenges in sCJD research and clinical practice: (i) identifying biomarkers to differentiate sCJD subtypes, (ii) identifying prognostic biomarkers for disease monitoring, (iii) understanding the molecular basis of disease heterogeneity through the identification of subtype-specific protein signatures.

Materials and methods

Patient selection

The study cohort comprised 168 CSF samples from 126 patients with sCJD and 42 non-neurodegenerative controls (CTRL). sCJD samples were submitted from 2000 to 2024 to the Neuropathology Laboratory (NP-Lab) of the Institute of Neurological Sciences of Bologna, a major referral center for prion disease in Italy. CTRL samples were recruited at the NP-Lab (n = 18) and the University Hospital of Tuebingen biobank (n = 24). The sCJD group included 57 patients with a neuropathological diagnosis and 69 with a clinical diagnosis of probable sCJD according to the current criteria [25, 43], who were all positive by CSF prion Real-Time Quaking-Induced Conversion seeding amplification assay (RT-QuIC) [31]. sCJD cases with a definite (i.e., neuropathological) diagnosis were classified into subtypes according to Parchi et al. [43]. Among these, 11 patients showing a mixed subtype, i.e., MM1 + 2 C (n = 10) and MV2K + 2 C (n = 1), were classified based on the dominant histotype

according to published criteria [45]. For the biomarker analysis by molecular subtype, we merged the patients with definite sCJD with those with a probable diagnosis and a high level of certainty for a given subtype, as previously described [7]. Further details regarding the classification of patients with probable sCJD are reported in the Supplementary methods (see Additional file 1). Overall, the cohort included 42 MM(V)1, 42 VV2, and 42 MV2K. The presence of significant Alzheimer's disease (AD) and Lewy body disease (LBD) co-pathologies was ruled out through immunohistochemistry for amyloid beta (Aβ), tau, and α-synuclein (αSyn) [40] for patients with available brain tissue, and through a negative CSF biomarker profile ($Aβ_{42}/Aβ_{40}$ ratio, phospho-tau181, and αSyn RT-QuIC) for the others [3, 8]. The CTRL group included individuals with no clinical evidence of neurodegenerative disorders, in addition to a negative AD CSF biomarker profile and αSyn RT-QuIC assay [3, 8].

CSF biomarker analyses

CSF samples were obtained in the morning by lumbar puncture (LP) at the L3/L4 or L4/L5 intervertebral level, centrifuged in case of blood contamination, divided into aliquots, and stored in polypropylene tubes at − 80 °C until analysis. For the AD core biomarker measurements, t-tau, p-tau, $Aβ_{42}$, and $Aβ_{40}$ were measured by automated chemiluminescent enzyme immunoassay on the Lumipulse G600II platform (Fujirebio, Gent, Belgium). Pathological values for defining the A/T status were determined using validated in-house cut-off values ($Aβ_{42}/Aβ_{40}$ ratio < 0.68 and a p-tau181 >62 pg/mL) [3]. CSF αSyn RT-QuIC was performed in all participants without a neuropathological diagnosis, as reported [8]. In all sCJD patients, we performed second-generation CSF prion RT-QuIC, as described [31].

CSF proteome profiling

The PEA technology was used to quantify 828 proteins in each CSF sample using nine validated multiplex antibody-based protein panels (Cardiovascular II, Cell Regulation, Immune Response, Inflammation, Metabolism, Neuro Exploratory, Neurology, Oncology II, and Organ Damage; Olink Proteomics, Uppsala, Sweden). Further details regarding panel characteristics and validation data are available on the manufacturer's webpage [42]. Each panel assays 92 unique proteins (the full list is provided in Additional file 2); however, 27 proteins were measured in several panels (replicates), resulting in a total of 797 distinct proteins measured per sample. All proteomic measurements were performed at the NP-Lab by personnel blinded to clinical diagnosis. Briefly, samples were randomised across plates (two per panel) containing appropriate intra- and inter-plate quality controls. Each plate included 8 bridging samples (two per diagnostic group), which were used to normalise across plates to adjust for potential batch effects, using the R package Olink Analyze version 4.2.0 [41]. Missing values in the dataset were imputed using the k-nearest neighbours (KNN) method (k = 3) from the scikit-learn Python library [46]. In cases where a protein was measured in more than one panel, only one of the replicates was kept for further statistical analysis.

Statistical analyses requirements

Statistical analyses were performed using Python 3.12.2 and R 4.4.0 with the following packages: scikit-learn (1.4.2) [46], NumPy (1.26.4) [24], Pandas (2.2.3) [59], SciPy (1.13.1) [63], SHAP (0.46.0) [37, 38], lifelines (0.30.0) [16], Matplotlib (3.9.2) [27], Seaborn (0.13.2) [64], NetworkX (3.2.1) [23], scikit-posthocs (0.11.2) [54], statsmodels (0.14.2) [50], UMAP-learn (0.5.6) [39], Plotly (5.22.0) [47], WGCNA (1.73) [30], dplyr (1.1.4) [67], ggplot2 (3.5.2) [66], tidyr (1.3.1) [68], and ropls (1.26.4) [61].

Descriptive statistics

The normality of continuous variables was assessed using Shapiro-Wilk and Kolmogorov-Smirnov tests. Normally distributed variables were compared using t-test or one-way ANOVA, while non-normally distributed continuous variables were analysed using the Mann–Whitney U or Kruskal-Wallis test, depending on the number of groups. Categorical variables were compared using chi-square tests when expected frequencies were sufficient (≥ 5 in at least 80% of the contingency table cells) or Fisher's exact test when they were not. Post-hoc tests were performed when the analyses involved > 2 groups. Spearman's rank correlation coefficients were computed to determine significant associations between individual protein levels and clinical variables (i.e., timespan LP-death, disease stage). For all analyses, $p < 0.05$ was considered statistically significant.

Uniform manifold approximation and projection

Protein expression data were visualised using Uniform Manifold Approximation and Projection (UMAP) for dimensionality reduction (umap-learn) [39]. UMAP parameters were set to n_neighbors = 15, min_dist = 0.1, and n_components = 2.

Differential expression analyses

To identify protein biomarkers distinguishing between clinical subtypes, we performed differential expression analysis (DEA) using nested generalised linear models. For each protein, we compared a base model that included age and gender as covariates to a full model additionally including clinical group membership. Significance was assessed using two-sided F-tests comparing

these nested models [14, 32]. *P*-values were adjusted for multiple testing using the Benjamini-Hochberg false discovery rate (FDR) method, with a significance threshold set at $q \leq 0.05$. For each pairwise comparison between clinical groups, we calculated log2 fold changes (log2FC) to quantify the magnitude and direction of protein level changes. Beta coefficients from the full models were also recorded to provide effect size estimates, controlling for covariates. Volcano plots were generated to visualise the relationship between statistical significance and effect size. The overlap between differentially expressed proteins (DEPs) across subtypes was plotted using Venn diagrams. Proteins showing significant differences across all clinical group pairwise comparisons were further analysed using trajectory plots to examine their expression patterns across groups.

Hierarchical clustering

Hierarchical clustering was performed to analyse protein expression patterns across sCJD subtypes. Specifically, we selected only the top 50 DEPs ranked by lowest *q*-value (FDR-corrected *p*-value) for pairwise comparison (MM(V)1 vs. VV2, MM(V)1 vs. MV2K, VV2 vs. MV2K). Protein expression values were normalized to the total protein expression per sample by dividing each protein measurement by the sum of all protein measurements within that sample. Subsequently, z-score normalization was applied across each protein feature. Spearman correlation matrices were computed for both proteins and samples, followed by distance transformation (1 - correlation). Hierarchical clustering was conducted using the average linkage method on condensed distance matrices. The results were visualised through a heatmap.

Multi-class classification

Machine learning models were implemented in Python using scikit-learn library [46]. Proteomics data were used to train Random Forest (RF) multi-class classification models (with default parameters, n_estimators = 100) to distinguish CJD subtypes. Three RF classifiers were analysed: one trained on all available protein features (individual protein measurements), another using a reduced set of protein features, and a third (baseline) model with shuffled labels as a negative control. For the reduced feature set, to prevent data leakage during the feature selection process within each individual cross-validation fold, we first performed the train/test split. Then we conducted feature selection exclusively on the training set of that specific fold. Therefore, feature selection was performed five separate times, once for each training set. Specifically, feature selection was performed using ANOVA F-tests (SelectKBest with f_classif scoring) to identify the top 50 most discriminative proteins within each fold. Model performance was assessed using 5-fold stratified cross-validation, with performance metrics including accuracy, precision, recall, F1-score, and receiver operating characteristics-area under the curve (ROC-AUC). Due to the limited sample size, we opted for cross-validation rather than a single held-out test set to ensure robust performance estimation across all available data. For feature importance analysis, we implemented SHapley Additive exPlanations (SHAP) analysis using TreeExplainer, which quantifies the contribution of each protein marker to individual predictions and reveals complex interaction patterns. SHAP values were aggregated across all folds to generate a comprehensive view of feature importance both globally and at the class level. A confusion matrix was used to display the averaged classification results across all folds, with the results expressed as percentages.

Decision tree classifier

To create a more clinically applicable and explainable diagnostic model, we trained a simplified, fully interpretable decision tree classifier. Unlike our previous RF models, this approach prioritizes transparency and ease of use in clinical settings. We constrained the tree complexity (max depth = 4, min samples per leaf = 5) to prevent overfitting and ensure interpretability. The model incorporated only the top 10 protein biomarkers previously identified as most important in our multi-class classification analysis, significantly reducing the number of measurements needed for classification. To assess model performance, we implemented 5-fold stratified cross-validation, allowing comparison with the RF classifier. A confusion matrix was generated to illustrate classification accuracy across subtypes. Following validation, we retrained the model on the entire dataset to maximize the use of available data for the final clinical decision tool. This final model represents the decision rules that would be implemented in practice, utilizing all available training data. Gini feature importance scores were calculated to understand each biomarker's relative contribution to the classification decisions. The trained tree structure was visualized for readability.

One-versus-rest classification

To identify subtype-specific protein signatures, we employed a binary classification approach using RF models. For each disease subtype, we performed a one-versus-rest classification, where samples from the target subtype were compared against all other subtypes combined. To address the class imbalance inherent in this approach (42 samples of target subtype versus 84 samples from other subtypes), we implemented class weighting in our RF classifier. Feature selection and model validation were performed using a 5-fold stratified cross-validation framework to ensure robust and unbiased results. In each

fold, we selected the top 50 features using ANOVA F-test based solely on the training data to prevent data leakage. We assessed feature stability across folds, considering features that appeared in at least 60% of the folds (3 out of 5) as stable markers for each subtype. The RF classifier was constructed with 200 trees and class weights inversely proportional to class frequencies. Model performance was evaluated using multiple metrics including accuracy, balanced accuracy, precision, recall, F1-score, and ROC-AUC. To interpret the biological significance of the models, we analysed feature importance using SHAP values.

Orthogonal Partial Least Squares–Discriminant Analysis

To provide complementary insights to our RF/SHAP value-based multi-class classification approach, we performed Orthogonal Partial Least Squares–Discriminant analysis (OPLS-DA) as implemented in the ropls R package [61]. Specifically, for each pairwise comparison between subtypes (i.e., MM(V)1 vs. VV2, VV2 vs. MV2K, and MM(V)1 vs. MV2K), OPLS-DA models were fitted using leave-one-out cross-validation to assess model robustness. Variable Importance in Projection (VIP) scores were extracted and ranked to identify the top discriminating proteins. The top 10 proteins were visualized using bar plots.

Survival analyses

Time-to-event was defined as the duration from LP to death or akinetic mutism, measured in months. The latter was used in place of time to death exclusively when the revision of medical charts indicated the adoption of life-extending treatments (e.g., enteral/parenteral nutrition, tracheostomy), as reported [7]. Four patients were excluded from the survival analyses due to the lack of information on disease duration. Furthermore, the disease stage was calculated as the ratio between the time from disease onset to LP and the overall survival (timespan onset-death) [7]. The biomarkers' prognostic performance was assessed using Cox proportional hazard models, with each biomarker alone (univariate analysis) or by including as covariates other known prognostic factors in prion disease (i.e., codon 129 genotype, age at sampling, and time from onset to sample collection) (multivariate analysis), as reported [7]. Survival analysis results are presented as hazard ratios (HRs) and 95% confidence intervals (CI). The assumption of proportional hazard was assessed by Schoenfeld residuals. All patients in the analysis had complete follow-up data with no censoring. The protein expression data were analysed in their original scale without transformation. Statistical significance was set at $p < 0.05$, and all reported p-values are two-sided. The Kaplan–Meier estimate was used to calculate the cumulative time-dependent probability of death.

Weighted Gene Co-expression Network Analysis (WGCNA)

We used the WGCNA R package [30] to identify protein co-expression modules based on scaled protein abundance levels. An optimal soft-thresholding power was determined using scale-free topology criteria. A topological overlap matrix was constructed, and blockwise module detection was performed with hierarchical clustering and the dynamic tree cut method, which adaptively determines module boundaries in the hierarchical clustering tree. Each module was assigned a colour, and functional enrichment analysis was performed on each module.

DEA was performed to assess log2FC in protein abundance across disease subtypes compared to controls. Mean log2FC was computed for each protein within each subtype, and proteins were mapped to co-expression modules. Module-level summaries were generated by averaging log2FC values per module and diagnostic group. One-sample t-tests were conducted to determine whether module-level log2FCs significantly deviated from zero. *P*-values were adjusted using the FDR method.

Hierarchical HotNet

As most modules comprised >50 proteins, Hierarchical HotNet was used to identify the most significantly dysregulated proteins in each module, as reported [11, 49]. Specifically, we used correlations between protein abundance levels within each module as edges and the p-values of the DEA between sCJD and CTRL as scores in the network. Subnetworks were visualized in Python using the library NetworkX [23].

Functional enrichment analysis

Functional enrichment was performed using Metascape, selecting Gene Ontology Biological Processes as ontology source [71]. To control for potential bias in our protein panel selection, we used a custom background set comprising all proteins included in our measurement panel instead of the default genome-wide background. Terms with a p-value < 0.01, a minimum count of 3, and an enrichment factor >1.5 are collected and grouped into clusters based on their membership similarities. More specifically, p-values are calculated based on the cumulative hypergeometric distribution, and q-values are calculated using the Benjamini-Hochberg procedure to account for multiple testing. Kappa scores are used as the similarity metric when performing hierarchical clustering on the enriched terms, and sub-trees with a similarity of >0.3 are considered a cluster. The most statistically significant term (lowest p-value) within a cluster is chosen to represent the cluster.

Results

Demographic characteristics

The study included 168 participants, i.e., 126 sCJD patients (42 each of MM(V)1, VV2, and MV2K subtypes) and 42 controls. The groups were matched for age (mean range: 64.4–67.6 years, $p = 0.295$), sex distribution (female proportion range: 38.1–54.8%, $p = 0.372$), and disease stage (median range 0.5–0.6, $p = 0.230$). Time from onset to death was longest in MV2K (median 15.0 months), intermediate in VV2 (5.5 months), and shortest in MM(V)1 (2.8 months) ($p < 0.001$). Further details on the participants' demographic, clinical, and diagnostic data for the four cohorts are reported in Table 1 and Additional file 3.

Distinct proteomic profiles across sCJD subtypes

To investigate potential proteomic differences among diagnostic groups, we performed a UMAP analysis of proteomic data. We identified clear clustering patterns distinguishing between sCJD cases and controls and among sCJD subtypes (Fig. 1A). Specifically, the control samples formed a distinct cluster, while the sCJD subtypes showed partially overlapping but distinguishable patterns, suggesting subtype-specific molecular signatures.

When performing the DEA, we found 201 DEPs between sCJD and CTRL (146 up- and 55 downregulated) (Fig. 1B). The five most upregulated biomarkers were involved in microtubule assembly and stability (MAPT, $\beta = 4.58$), neurofilament components (NEFL, $\beta = 3.54$), cell signalling (CRKL, $\beta = 2.72$), and glucose metabolism (ENO2, $\beta = 1.47$; and RBKS, $\beta = 1.67$). When stratifying the analysis by clinicopathological subtype, we found 164 DEPs for MM(V)1, 185 for VV2, and 110 for MV2K, compared to CTRL. Specifically, we identified 20 proteins selectively changed in MM(V)1, 38 in VV2, and 13 in MV2K (Additional file 1-Supplementary Fig. 1A-D). In addition, we explored DEPs between the sCJD subtypes. We found 71 biomarkers differentially expressed between MM(V)1 and MV2K, 103 between VV2 and MV2K, and 49 between MM(V)1 and VV2 (Fig. 1C). Notably, only 20 biomarkers were significantly differentially expressed in all pairwise comparisons. All were up-regulated compared to CTRL, showed higher values in VV2, and a smaller increase in MM(V)1 and MV2K (Additional file 1-Supplementary Fig. 1E-F). All differential expression analysis results are shown in Additional file 4.

The hierarchical clustering analysis of protein expression data revealed distinct patterns across sCJD subtypes (Fig. 2). Notably, MV2K and VV2 cases formed well-defined, distinct clusters, while MM(V)1 cases showed a more heterogeneous clustering pattern.

Table 1 Demographic characteristics

	MM(V)1 ($n = 42$)	VV2 ($n = 42$)	MV2K ($n = 42$)	CTRL ($n = 42$)	*p*-value
Female, n/N (%)	18/42 (42.9)	16/42 (38.1)	22/42 (52.4)	23/42 (54.8)	0.372
Age at LP	66.3 ± 8.5	66.4 ± 8.6	67.6 ± 6.2	64.4 ± 7.7	0.295
Time onset-LP	1.1 (1.0–2.0)	3.0 (2.5-4.0)	8.0 (3.6–11.8)	N/A	< 0.001
Time LP-death	1.2 (0.9-2.0)	2.0 (1.0-3.1)	6.0 (3.5–11.0)	N/A	< 0.001
Time onset-death	2.8 (2.0-3.5)	5.5 (5.0–7.0)	15.0 (12.0–17.0)	N/A	< 0.001
Disease stage	0.5 (0.4–0.7)	0.6 (0.4–0.8)	0.5 (0.3–0.7)	N/A	0.230

Age at LP is expressed as mean (standard deviation) in years. The timespans between onset, LP, and death are reported as median (interquartile range) in months, and the disease stage as median (interquartile range). Abbreviations: CTRL, healthy controls; LP, lumbar puncture; N/A, not applicable

Machine learning identifies potential biomarkers for sCJD subtype differentiation

We trained three RF classification models to identify protein biomarkers that could distinguish sCJD subtypes. The RF model using a selected set of protein features demonstrated superior performance across all metrics compared to the complete feature set and random baseline models (Fig. 3A). This optimised model achieved high ROC-AUC values (0.93) and maintained consistent performance across other metrics. A detailed analysis of the model classification performance revealed high accuracy in distinguishing between subtypes (Fig. 3B). The confusion matrix demonstrated misclassification primarily in MM(V)1 cases. SHAP analysis identified the top 20 proteins contributing to subtype classification (Fig. 3C). Among these, almost all proteins (including top markers such as HDGF, FOSB, PAG1, and APEX1) showed higher levels in VV2, intermediate in MM(V)1, and low in MV2K. Notable exceptions include GPC5 and CCDC80, which exhibited higher levels in V2-strain-related subtypes (VV2 and MV2K) compared to MM(V)1, and WASF1, which showed an opposite trend, i.e., lower levels in V2-strain-related subtypes compared to MM(V)1 (Fig. 3D and Additional file 1-Supplementary Fig. 2). All proteins except three (NEFL, CCDC80, and GPC5) showed a significant negative correlation with the timespan LP-death across the entire sCJD cohort (Additional file 5-Supplementary Table 4). Moreover, no significant associations between protein levels and disease stage were observed in any sCJD clinicopathological subtype (Additional file 6-Supplementary Table 5). When performing UMAP analysis on the top 20 protein classifiers, all three sCJD subtypes showed clear clustering patterns (Additional file 1-Supplementary Fig. 3A). OPLS-DA results were highly consistent with the RF for multiclass classification. Specifically, the top 10 proteins with the highest VIP scores in the pairwise comparisons

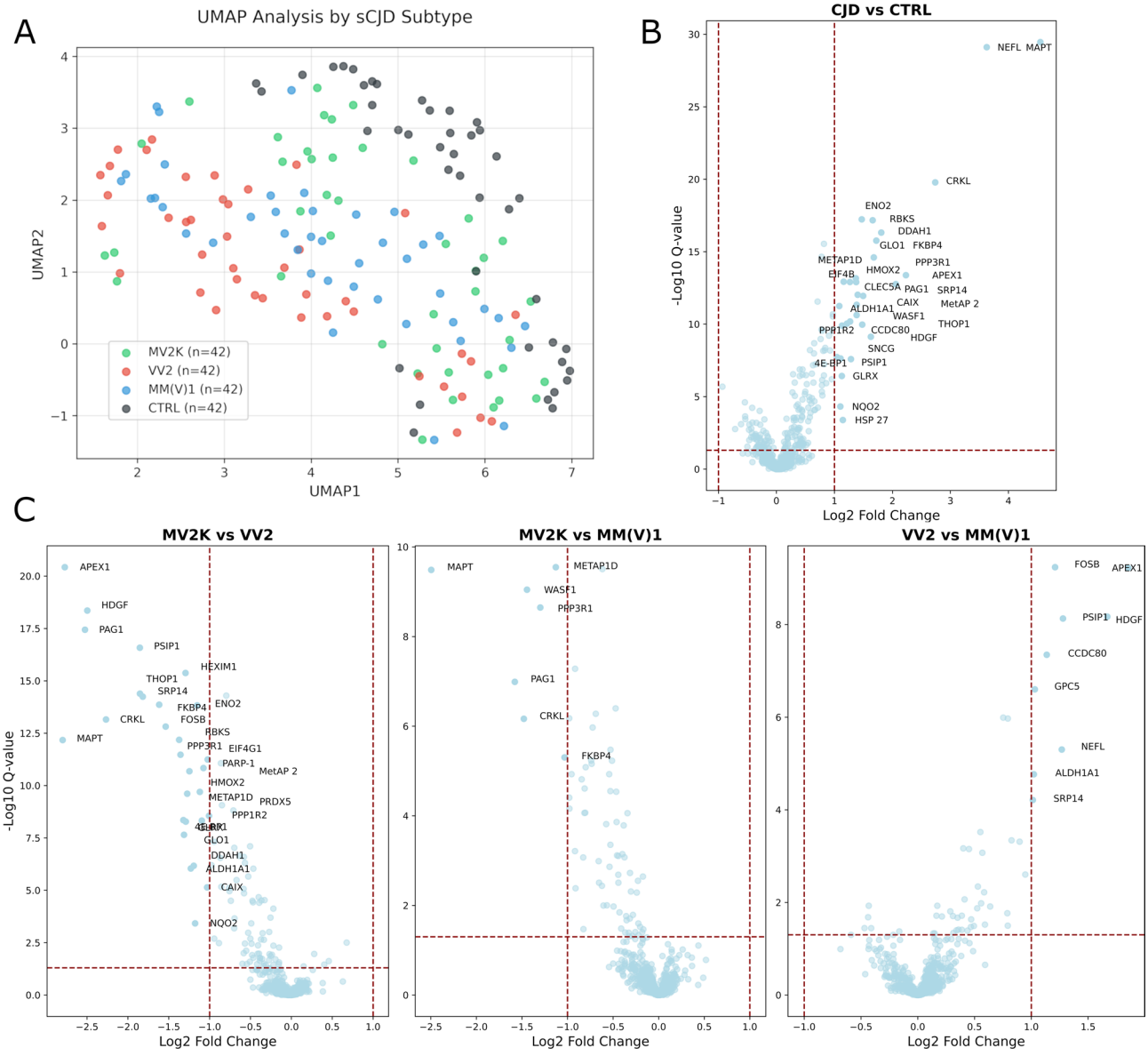

Fig. 1 Distinct proteomic profiles across sCJD subtypes. (**A**) UMAP projection of protein expression profiles showing distinct clustering patterns among sCJD subtypes and controls (CTRL, gray; MM(V)1, blue; MV2K, green; VV2, red). The distribution of samples in the UMAP space suggests subtype-specific protein expression signatures, with notable separation between controls and sCJD cases. (**B**) Volcano plot displaying DEPs between sCJD and control samples. Notable proteins with high significance and large positive log2FC include MAPT and NEFL at the top right, followed by CRKL, ENO2, and RBKS. (**C**) Volcano plot showing DEPs between sCJD subtypes. Abbreviations: CTRL, non-neurodegenerative controls; DEPs, differentially expressed proteins; sCJD, sporadic Creutzfeldt-Jakob disease; UMAP, Uniform Manifold Approximation and Projection

(MM(V)1 vs. VV2, MM(V)1 vs. MV2K, and VV2 vs. MV2K) were included among the top 20 proteins contributing to subtype classification in 90%, 90%, and 70% of cases, respectively (see Additional file 1-Supplementary Fig. 4). Further details on the biological functions, subcellular localization, and patterns of cellular and tissue-specific expression of the top 20 protein classifiers are provided in Additional file 7.

Next, to propose an interpretable and scalable model based on a reduced set of biomarkers that could classify sCJD subtypes, we trained a Decision Tree classifier, including the top ten proteins contributing to subtype classification. Among these, PAG1, FOSB, GPC5, MAPT, METAP1D, and WASF1 emerged as the most critical protein markers for sCJD subtype classification. This model, although underperforming in classifying VV2, correctly classified most of MM(V)1 and MV2K cases (Additional file 1-Supplementary Fig. 5A-C).

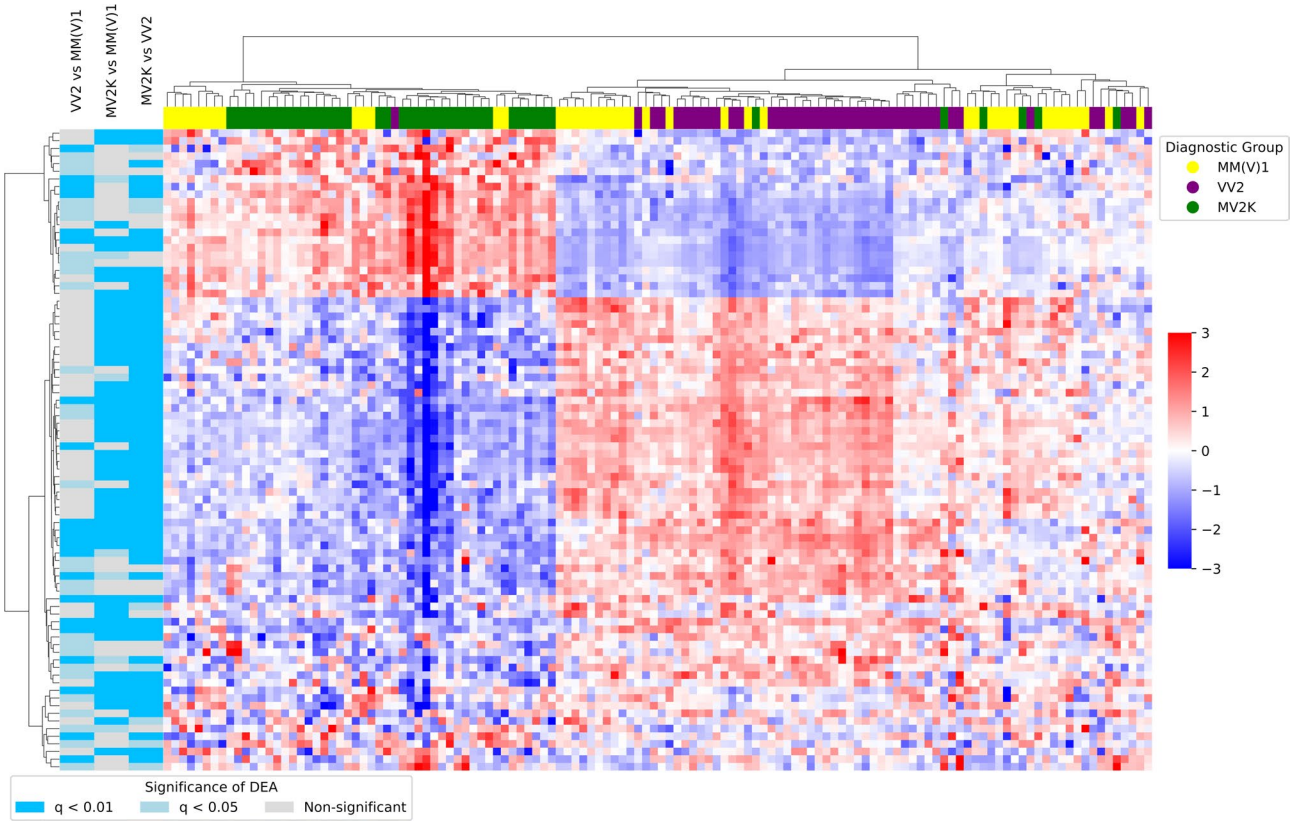

Fig. 2 Hierarchical clustering of the top 50 DEPs in inter-subtype pairwise comparisons. The samples are represented on the x-axis and coloured according to the sCJD subtype (MM(V)1, yellow; VV2, purple; MV2K, green). Normalized protein abundance levels are plotted on the y-axis. Q-values of the differential expression in each pairwise comparison are indicated to the left of the figure. Shades of turquoise indicate statistically significant values in the DEA. Abbreviations: DEA, differential expression analysis; sCJD, sporadic Creutzfeldt-Jakob disease

Machine learning reveals biomarkers differentiating M1 and V2 strain-related subtypes

We applied a one-versus-rest classification approach to identify specific molecular signatures and biomarkers for each subtype. The VV2 subtype was the easiest to classify with an accuracy of 0.89 (± 0.04) and ROC-AUC of 0.92 (± 0.04), suggesting a highly distinctive protein profile (Fig. 4A-B). The model identified 51 stable proteins that consistently distinguished VV2 from other subtypes across cross-validation folds. The MV2K subtype demonstrated similar strong discrimination with an accuracy of 0.86 (± 0.02) and ROC-AUC of 0.89 (± 0.06), with 49 stable protein markers. In contrast, despite the high overall accuracy (0.83 ± 0.06) and ROC-AUC (0.93 ± 0.06), the MM(V)1 subtype exhibited lower recall (0.48 ± 0.15) compared to other subtypes, suggesting a greater difficulty in identifying MM(V)1 cases. Interestingly, MM(V)1 was characterised by a more focused set of 40 stable protein markers. Of note, GPC5, CCDC80, WASF1, and, to a lesser extent, NEFL emerged as the most critical protein markers in differentiating MM(V)1 and V2-strain-related subtypes.

Functional enrichment analyses on the identified protein profiles are shown in Fig. 4C. The MV2K showed significant enrichment in pathways related to the regulation of membrane potential, peptidyl-amino acid modification, DNA metabolic process, generation of precursor metabolites and energy, and protein-RNA complex assembly (all $p < 0.001$). Similarly, VV2 demonstrated enrichment in the regulation of membrane potential, chromosome organization, regulation of double-strand break repair, RNA splicing, and response to zinc ions ($p < 0.001$). The MM(V)1 subtype showed significant enrichment in axo-dendritic transport ($p < 0.001$).

Survival analysis

Next, we assessed whether the 20 top protein markers exhibited any prognostic value (Fig. 5A). In univariate Cox regression, 17 out of 20 were significantly associated with survival. However, when adjusting for factors known to predict prognosis in prion disease (i.e., age at LP, sex, codon 129 genotype, timespan between onset and LP), only 11 proteins correlated with survival, i.e., CCDC80, FOSB, APEX1, THOP1, FKBP4, ARHGEF12, PAG1, PRDX3, PPP3R1, MAPT, WASF1, with HR

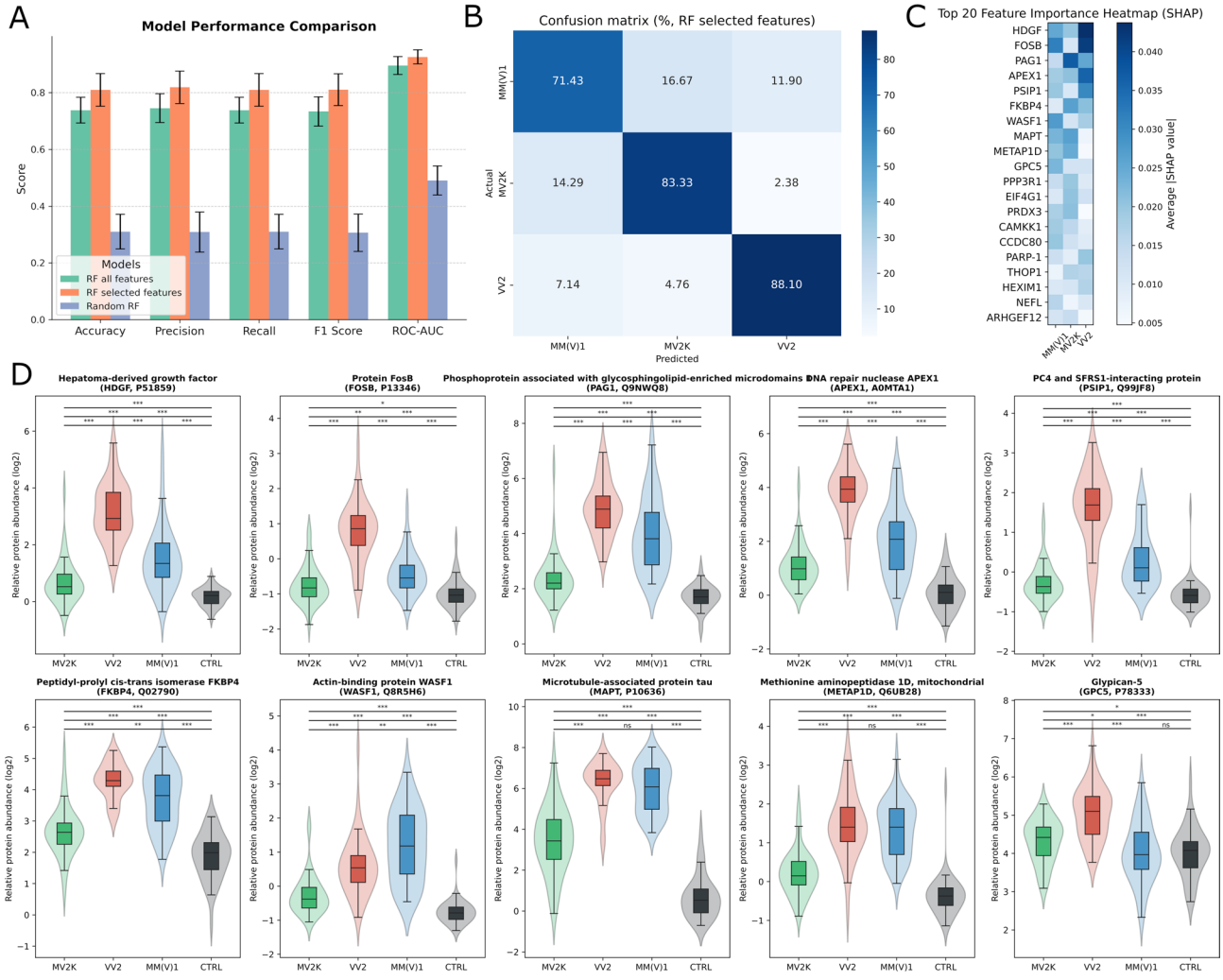

Fig. 3 Machine learning identifies potential biomarkers for sCJD subtype differentiation (**A**) Model performance was assessed using 5-fold stratified cross-validation, with performance metrics including accuracy, precision, recall, F1-score, and ROC-AUC. (**B**) A confusion matrix heatmap (RF selected features model) displays the averaged classification results across all folds, with results expressed as percentages. (**C**) A class-specific heatmap derived from SHAP values (RF selected features model) for visualising the top 20 most important features across folds. (**D**) Differential levels of the ten most important proteins for differentiating CJD subtypes. Box plots with violin plots show the distribution of relative protein abundance for the biomarkers with the highest SHAP importance values. Statistical significance between groups is indicated by asterisks (* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, ns: not significant) based on Mann-Whitney U tests. Abbreviations: RF, Random Forest; ROC-AUC, receiver operating characteristic-area under the curve; sCJD, sporadic Creutzfeldt-Jakob disease; SHAP, SHapley Additive exPlanations

ranging 1.16–1.80 (all $p < 0.05$). Notably, when adjusting also for the clinicopathological subtype, 10 proteins were significantly associated with survival: CCDC80 (HR 1.49, 95% CI 1.17–1.91, $p = 0.001$), PARP-1 (HR 1.32, 95% CI 1.00-1.75, $p = 0.049$), FOSB (HR 1.27, 95% CI 1.02–1.58, $p = 0.031$), APEX1 (HR 1.16, 95% CI 1.02–1.32, $p = 0.025$), THOP1 (HR 1.20, 95% CI 1.01–1.44, $p = 0.044$), FKBP4 (HR 1.27, 95% CI 1.04–1.56, $p = 0.020$), PAG1 (HR 1.17, 95% CI 1.01–1.36, $p = 0.032$), PRDX3 (HR 1.75, 95% CI 1.11–2.75, $p = 0.016$), MAPT (HR 1.21, 95% CI 1.06–1.38, $p = 0.005$), WASF1 (HR 1.26, 95% CI 1.01–1.59, $p = 0.044$). Survival curves for these 10 proteins are shown in Fig. 5B and Additional file 1-Supplementary Fig. 6. The UMAP plot of the top 20 protein markers confirmed that these proteins broadly capture both the clinicopathological subtype and the survival (Additional file 1-Supplementary Fig. 3B).

Network-based identification of dysregulated protein modules and molecular hubs in sCJD subtypes

Next, we used the WGCNA algorithm to identify functional groups of co-regulated proteins. Unlike functional enrichment analysis, which relies on pre-annotated functional gene sets and is thus biased toward known pathways, this method does not require prior annotations or sample labels, allowing the discovery of novel functional modules [11, 30, 49]. Specifically, we identified six modules named after their top functional annotation, which

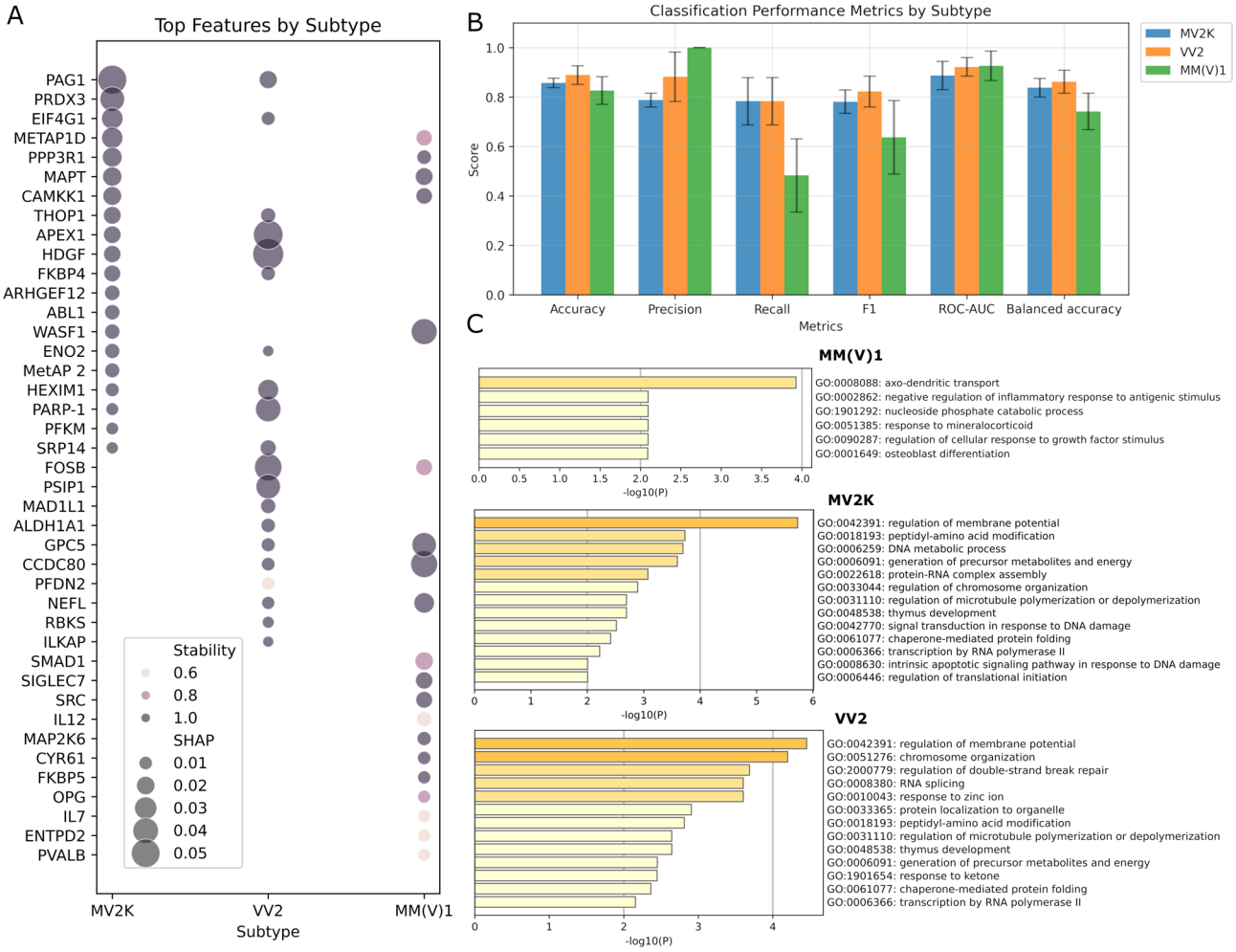

Fig. 4 Feature analysis and protein interaction networks across CJD subtypes. (**A**) Dot plot showing top selected features (proteins) across subtypes, where dot size represents SHAP importance and color intensity indicates selection stability for different folds. (**B**) Model performance was assessed using 5-fold stratified cross-validation, with performance metrics including accuracy, precision, recall, F1-score, and ROC-AUC. (**C**) Gene Ontology biological process enrichment analysis for most important proteins in each sCJD subtype. Abbreviations: RF, Random Forest; ROC-AUC, receiver operating characteristic-area under the curve; sCJD, sporadic Creutzfeldt-Jakob disease; SHAP, SHapley Additive exPlanations

were abbreviated for clarity (Fig. 6A). Module size was highly variable, ranging from 47 proteins (chemotaxis) to 170 (nervous system development). Enrichment analysis was performed on the proteins included in each module to identify the dysregulated biological processes (Fig. 7). In terms of mean log2FC, the toxicity and chemotaxis modules were significantly upregulated across all three sCJD subtypes (toxicity: MM(V)1, $p < 0.001$; VV2, $p < 0.001$; MV2K, $p < 0.001$; chemotaxis: MM(V)1, $p < 0.001$; VV2, $p = 0.003$; MV2K, $p < 0.001$) compared to the CTRL group, whereas the nervous system development module was consistently downregulated (MM(V)1, $p < 0.001$; VV2, $p < 0.001$; MV2K, $p < 0.001$). Notably, the toxicity module exhibited the highest dysregulation in VV2, followed by an intermediate level in MM(V)1, and the least in MV2K. The toxicity module showed significant enrichment in terms related to response to toxic substance, negative regulation of phosphorus metabolic process, pyridine nucleotide metabolic process, chromosome organization, DNA metabolic process, establishment of organelle localization, response to metal ion, and chaperone-mediated protein folding ($p < 0.001$), while the chemotaxis one was related to positive regulation of positive chemotaxis and interleukin-6 production ($p < 0.001$). The nervous system development module was enriched in terms related to the regulation of nervous system development, neuron projection morphogenesis, synapse organization, regulation of growth, developmental growth involved in morphogenesis, and cell-cell adhesion. Interestingly, the epithelium morphogenesis module showed selective upregulation in MM(V)1 ($p < 0.001$), while the intracellular signal transduction one was specifically upregulated in V2-strain-related subtypes (VV2, $p < 0.001$; MV2K, $p < 0.001$). The epithelium

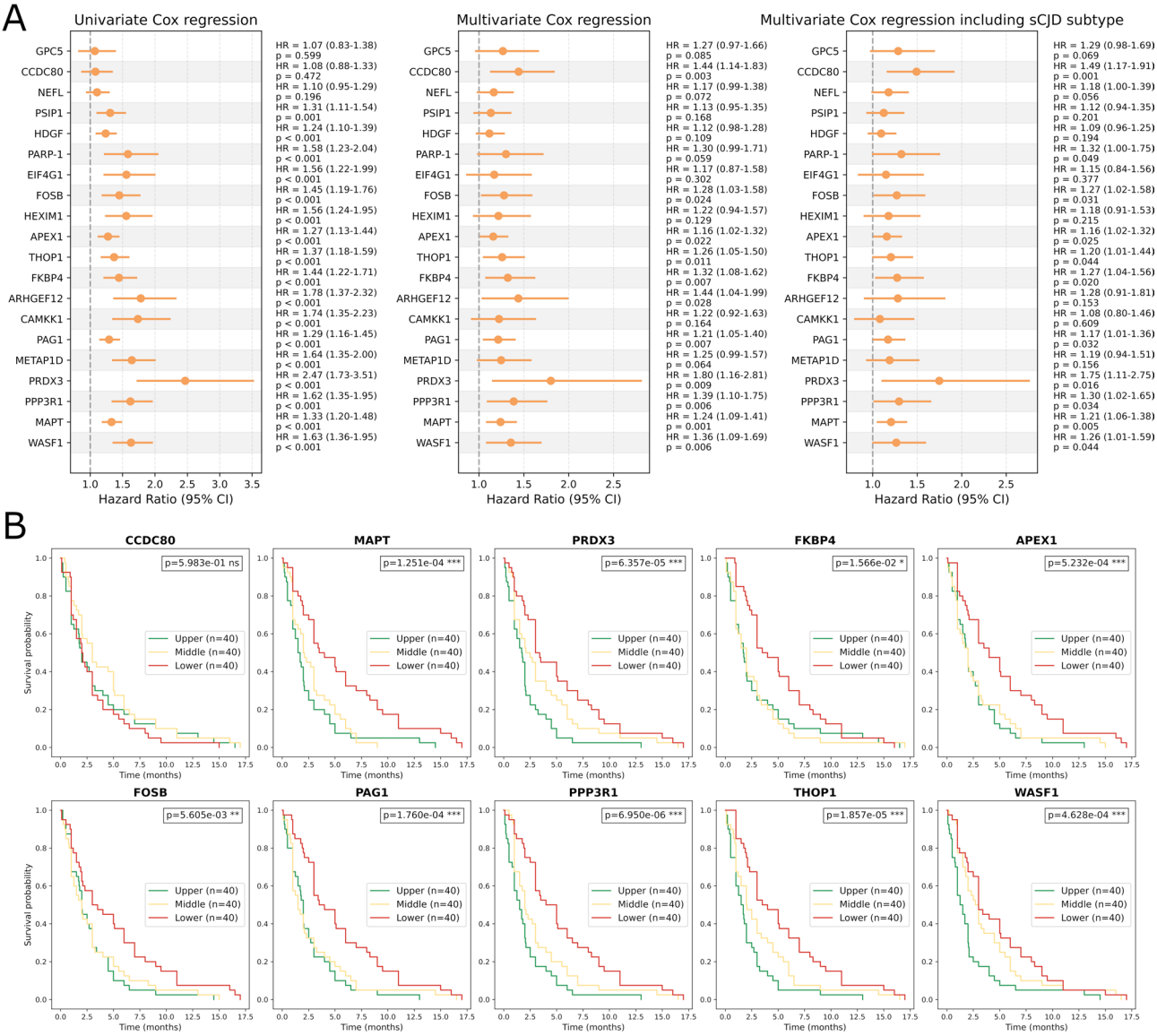

Fig. 5 Survival analysis results. (**A**) Forest plots showing the results of Cox regression analyses. All multivariate Cox regression analyses included sex, age at LP, time from onset to LP, and codon 129 genotype as covariates. In the third plot the sCJD clinicopathological subtype was also included as a covariate. (**B**) Kaplan–Meier survival curves displaying time from sampling to death in the whole sCJD cohort as a function of biomarker levels. Time is shown in months from disease onset, and survival probability is shown on the y-axis. Biomarker concentrations are binned in tertiles. *p*-values were calculated using the log-rank test. All patients were deceased at the time of the analysis. Only the top 10 prognostic biomarkers are shown. Abbreviations: LP, lumbar puncture; sCJD, sporadic Creutzfeldt-Jakob disease

morphogenesis module was mostly enriched in terms related to epithelium morphogenesis, angiogenesis regulation, and brain-derived neurotrophic factor receptor signaling pathway ($p < 0.001$), while the intracellular signal transduction one was enriched in the negative regulation of intracellular signal transduction processes ($p < 0.001$). The remaining "N/A" module, which appeared to be selectively dysregulated in VV2 ($p = 0.012$), showed no significant enrichment in GO Biological Process.

Subnetwork analysis revealed the most highly and significantly differentially abundant proteins associated with each module (Fig. 6B). Interestingly, the toxicity subnetwork consisted of 19 proteins, all upregulated in sCJD compared to controls, among which some of them (i.e., MAPT, PAG1, APEX1, and WASF1) were identified as relevant biomarkers in previous analyses. Key molecular hubs in epithelium morphogenesis and intracellular signal transduction modules included CD48, ERBB2, ICOSLG, LYPD3, SEZ6L, TNFRSF4, TNFSF13, WFDC2, and WIF-1, and BANK1, BTC, MAEA, PRKRA, and SMAD1, respectively.

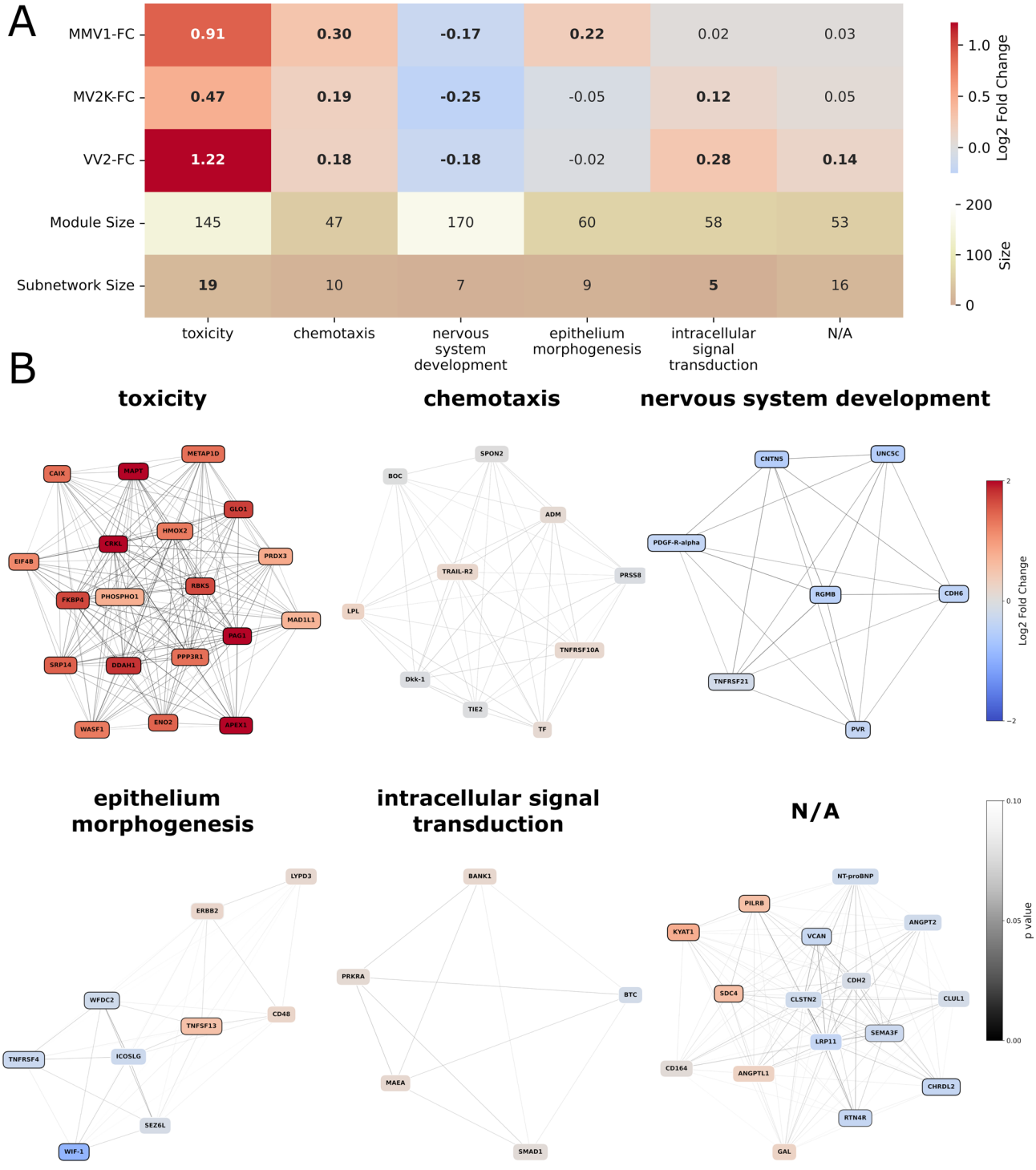

Fig. 6 Network-based identification of dysregulated protein modules in sCJD subtypes. (**A**) Heatmap showing the mean log2FC protein abundance levels for each module. Red and blue indicate an upregulated or downregulated mean log2FC protein abundance level within the module compared to CTRL, respectively. Bold values indicate that the mean log2FC was significantly different from zero. The two bottom rows indicate the size of the modules and subnetworks. (**B**) Subnetworks of the most highly and significantly dysregulated proteins for each module. Red nodes indicate a positive log2FC, and blue nodes indicate a negative log2FC. The darkness of the borders is related to the p-value. The thickness of the lines connecting the nodes represents the weight of the edges. Both the log2FC and the p-value are referred to the DEA between the whole sCJD cohort and CTRL. Abbreviations: CTRL, non-neurodegenerative controls; DEA, differential expression analysis; log2FC, log2 fold changes; sCJD, sporadic Creutzfeldt-Jakob disease

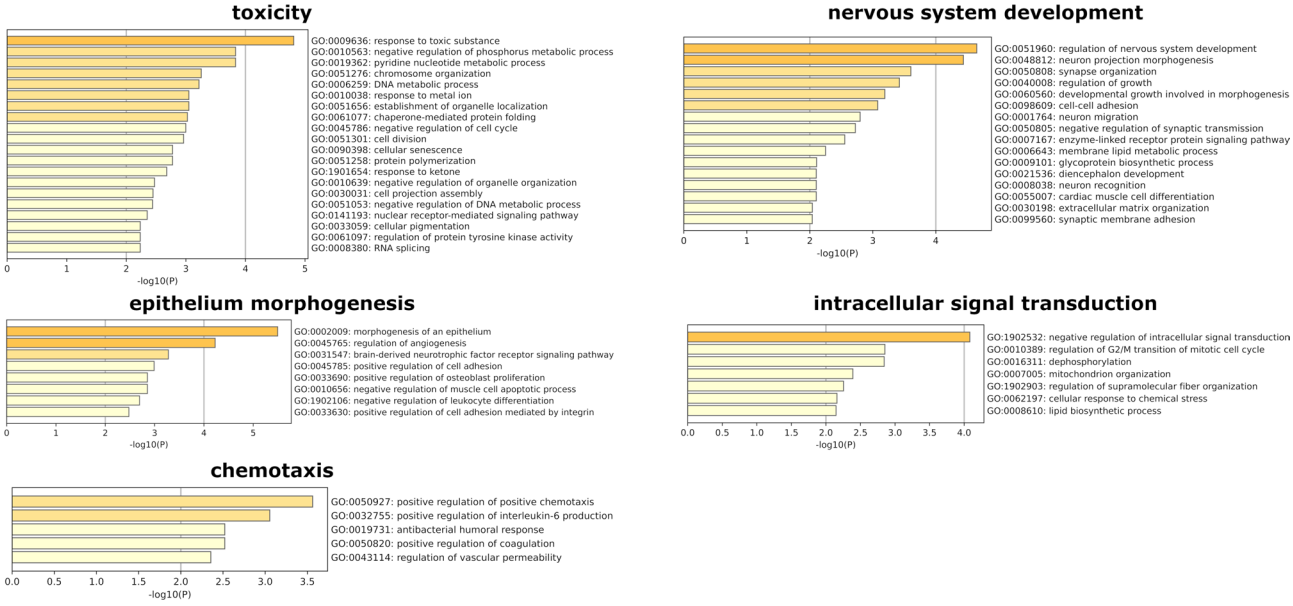

Fig. 7 GO biological process enrichment analysis for all modules. Functional enrichment was performed using Metascape, selecting GO biological processes as ontology source. Bars show the biological pathways enriched in each module. The corresponding GO number and associated biological process are indicated on the right. Darker colors indicate greater enrichment significance. The N/A module is not shown, as it exhibited no significant enrichment. Abbreviations: GO, Gene Ontology

Discussion

This study provides new evidence of distinct CSF proteomic signatures across the most common sCJD clinicopathological subtypes, translating this finding into the discovery of novel biomarkers for in vivo subtype differentiation and prognostication. As the study's main result, we demonstrated that it is possible to differentiate sCJD subtypes with high accuracy (AUC 0.93) based on a reduced set of proteins. Top biomarker candidates include proteins related to transcriptional regulation and DNA repair (e.g., FOSB, PSIP1, HEXIM1, PARP-1, APEX1), cytoskeletal structure and organization (e.g., MAPT, NEFL, WASF1, ARHGEF12), cell signalling (e.g., HDGF, PAG1, GPC5, CAMKK1, PPP3R1), protein metabolism (e.g., FKBP4, EIF4G1, THOP1, METAP1D), stress response (e.g., PRDX3), and extracellular matrix organization (e.g., CCDC80). Among them, only MAPT and NEFL have so far demonstrated diagnostic and prognostic utility for sCJD subtypes, as well as a pathogenic role [1, 4, 7, 36], while evidence for all other proteins is lacking. Upon more detailed analysis, our machine learning model showed strong performance in classifying the VV2 and MV2K subtypes. In contrast, its accuracy was lower for MM(V)1, most likely due to the substantial neuropathological heterogeneity within this group, including a significant number of mixed cases (MM(V)1 + 2C). The robustness of our RF/SHAP value-based approach for biomarker identification was further supported by the consistent results obtained with a different supervised machine learning model, namely OPLS-DA.

Ultimately, despite requiring further validation, we also report that even a small subset of these potential biomarkers (PAG1, FOSB, GPC5, MAPT, METAP1D, and WASF1) might prove accurate for subtype differentiation.

Regarding the distribution of these protein levels across sCJD subtypes, we identified three main patterns. In the first and most common pattern, exemplified by MAPT, CSF levels are highest in VV2, intermediate in MM(V)1, and lowest in MV2K. The second pattern, observed in GPC5 and CCDC80 and best represented by NEFL, is characterized by higher levels in V2-strain-related subtypes (VV2 and MV2K) compared to MM(V)1. Notably, these three proteins were the only ones among the top 20 identified biomarker candidates that did not exhibit a significant negative correlation with the timespan between LP and death, suggesting that their levels do not merely reflect disease severity. The third pattern, which applies exclusively to WASF1, shows elevated levels in MM(V)1 relative to the other subtypes. The underlying reasons for these distinct patterns remain to be fully elucidated. Almost all of the identified proteins (with the notable exception of CCDC80, which is secreted to the extracellular matrix) have a predominantly intracellular localization in nervous system cells. It is likely that these proteins are released as a consequence of the extensive brain damage occurring in sCJD, and therefore reach higher levels in the more aggressive subtypes (i.e., VV2 and MM(V)1) compared to the more slowly progressive MV2K. However, evidence indicates that CSF protein levels reflect both the extent and speed of brain damage, as well as

the predominant regional and subcellular pathology of a given disorder. For instance, NEFL, a neurodegeneration marker of myelinated axons, is thought to have higher levels in V2-strain-related subtypes due to predominant subcortical involvement [1]. In contrast, MAPT, which is ubiquitously represented in the CNS, primarily reflects the intensity of neurodegeneration, i.e., highest in VV2, intermediate in MM(V)1, and lowest in MV2K [7]. WASF1 is tendentially more expressed in the cortex and hippocampus, while GPC5 and CCDC80 are more abundant in subcortical areas (data available from v24.proteinatlas.org), which may help explain the peculiar distribution of their CSF levels across subtypes [51, 55–58].

Consistent with these observations, GPC5, CCDC80, WASF1, and, to a lesser extent, NEFL emerged as protein markers with potential in distinguishing MM(V)1 from V2-strain-related subtypes. Their distinctive distribution makes them promising for differentiating patients with PrP^{Sc} type 1 or 2 when carrying the same MV genotype (MV1 vs. MV2K), a critical diagnostic challenge in clinical practice.

Next, we evaluated the association between the identified biomarker levels and survival using previously validated prognostic models for sCJD [52]. Given the small size of the clinicopathologically defined subgroups (i.e., 42 per subtype), we decided to include the subtype as a covariate in the model rather than stratifying patients into three separate analyses. This approach allowed us to preserve statistical power while accounting for subtype effects. Notably, statistically significant differences emerged despite the non-ideal context of a small cohort with relatively low survival variability within each molecular subtype, which does not fully reflect the disease's heterogeneity. Specifically, when adjusting for factors known to predict prognosis in prion disease (i.e., age at LP, sex, codon 129 genotype, timespan between onset and LP, and clinicopathological subtype), 10 proteins correlated with survival, namely, CCDC80, PARP-1, FOSB, APEX1, THOP1, FKBP4, PAG1, PRDX3, MAPT, and WASF1. Notably, CCDC80 exhibited one of the strongest associations, overpowering MAPT, which is currently considered one of the best-performing CSF biomarkers for CJD prognostication [7, 52]. Its potential as an independent prognostic biomarker, in addition to its diagnostic value, makes CCDC80 particularly promising for future validation efforts.

Previous studies have shown an association between biomarker levels and disease stage in CJD, with distinct biomarker trajectories depending on the disease subtype [6, 7, 20]. Regarding this aspect, none of the top-identified proteins exhibited significant correlations with the disease stage across any sCJD clinicopathological subtype, suggesting that their concentrations remain relatively stable throughout the disease course. However, given the relatively small sample size and limited within-subgroup variability, this analysis was likely underpowered. Therefore, validation studies in larger cohorts, possibly using specific immunoassays for the relevant proteins we identified, should also address this issue further.

Beyond biomarker research, this study explored the potential of proteomics to enhance our understanding of the pathogenic mechanisms driving the disease and its phenotypic heterogeneity. DEA identified numerous dysregulated proteins in sCJD compared to CTRL. Among these, MAPT and NEFL (two proteins widely recognized as biomarkers in the differential diagnosis of sCJD) stood out as the most significantly altered, further validating the robustness of our proteomic approach. Unsupervised analyses, including UMAP and hierarchical clustering of the top DEPs, revealed distinct clustering patterns among sCJD clinicopathological subtypes, suggesting the existence of subtype-specific proteomic signatures. We applied two complementary approaches to investigate further the biological processes differentiating these subtypes. In the first one, we employed a machine learning model performing a binary classification task (one subtype versus all others) to identify the most predictive proteins for each subtype. Functional enrichment analysis of these proteins uncovered several dysregulated biological pathways across subtypes. For instance, the MV2K and VV2 showed significant enrichment in pathways related to membrane potential regulation and nucleic acid metabolism, while the MM(V)1 displayed significant enrichment in axo-dendritic transport. In the second approach, to overcome the limitations of conventional functional enrichment analyses, which rely on pre-annotated functional groups of proteins, we adopted a workflow combining WGCNA and Hierarchical HotNet. This strategy enables the unsupervised discovery of novel biological relationships and mechanisms not yet well-characterised in existing pathway databases [11, 30, 49]. While some modules were dysregulated across all sCJD subtypes, others displayed subtype-specific alterations. The toxicity module exhibited the highest level of dysregulation in VV2, intermediate in MM(V)1, and the lowest in MV2K. This module was significantly enriched in processes related to the response to toxic substances, negative regulation of phosphorus metabolism, pyridine nucleotide metabolism, chromosome organization, DNA metabolism, organelle localization, response to metal ions, and chaperone-mediated protein folding. Many of the previously identified promising biomarkers (e.g., MAPT, CRKL, PAG1, APEX1, and WASF1) came up as molecular hubs in this module, suggesting a direct pathogenetic role in sCJD. The chemotaxis module, which showed the most pronounced dysregulation in MM(V)1 compared to other subtypes, was associated

with regulatory processes of the immune response, including positive regulation of chemotaxis, interleukin-6 production, antibacterial humoral response, coagulation, and vascular permeability regulation. The nervous system development module, which was significantly downregulated across all subtypes, was primarily involved in neurogenesis, cellular communication and migration, and the maintenance of synaptic integrity. Overall, these findings deepen our understanding and align with a growing body of evidence demonstrating that disruptions in protein homeostasis, neuroinflammation, vascular and synaptic dysfunction, and metal ion metabolism are key pathogenic processes in prion diseases [5, 26, 29, 35, 60, 62, 65]. Notably, the epithelium morphogenesis and intracellular signal transduction modules exhibited distinct strain-specific dysregulation. The former, which is involved in epithelial morphogenesis, angiogenesis regulation, BDNF receptor signalling, cell adhesion, osteoblast proliferation, inhibition of muscle cell apoptosis, suppression of leukocyte differentiation, and integrin-mediated adhesion, was selectively altered in MM(V)1. The intracellular signal transduction module, which plays a crucial role in intracellular signal transduction inhibition, G2/M transition regulation, dephosphorylation, mitochondrial organization, supramolecular fiber organization, cellular response to chemical stress, and lipid biosynthesis, was specifically dysregulated in V2-strain-related subtypes. These results are consistent with our previous reports, which, based on transcriptomic data, highlighted a selective alteration in the processes of dephosphorylation regulation and mitochondrial dysfunction in VV2 compared to MM(V)1, and expand on the characterization of the molecular differences underlying the phenotypic heterogeneity of sCJD [53]. None of the protein hubs identified in these two modules has a known pathogenic role in prion diseases. However, some are recognized for their association with other neurodegenerative disorders, highlighting the existence of shared pathological mechanisms across the neurodegeneration spectrum. For instance, WIF1, a regulator of the Wnt signalling pathway, has been implicated in both AD and Parkinson's disease [18, 34], while SEZ6, a known synaptic regulator, is altered in AD [28]. Additionally, PRKRA has been linked to early-onset generalized dystonia-parkinsonism [13].

Regarding the identification of protein hubs within individual modules, similar implementations of Hierarchical HotNet have been successfully applied in cancer research and neurological disorders [11, 33, 48, 49]. However, it is worth noting that its performance depends on the quality and completeness of protein-protein interaction networks, which may be biased toward well-studied proteins, and is sensitive to parameter selection. Furthermore, the method assumes that disease-relevant proteins are topologically connected, and computational complexity increases with network size, which can hinder reproducibility across datasets [49].

The proteomic approach applied in this study warrants further consideration. Olink proteomics, which employs PEA technology, provides distinct advantages over conventional mass spectrometry (MS)-based methods. Although MS offers comprehensive proteome coverage, it generally requires larger sample volumes, labor-intensive preparation, and complex data analysis, factors that may constrain its use in rare diseases where biological material is scarce. By contrast, PEA leverages dual antibody recognition coupled with DNA-based quantification to enable detection of proteins at femtogram-per-milliliter concentrations from minimal sample input. This enables the highly sensitive, specific, and reproducible quantification of hundreds of low-abundance biomarkers in parallel [2, 22]. Importantly, owing to its antibody-based detection principle, PEA is more readily translatable to clinical applications than MS-based approaches [15]. Given the limited availability of CSF samples in our cohort and the importance of generating clinically actionable insights, PEA represents a particularly well-suited approach for this study.

This study has several limitations. First, biomarker discovery and model training were performed on the same retrospective cohort without independent external validation. The cross-validation approach employed provides internal validation within our dataset, but cannot account for potential biases that might affect model generalisability. Another limitation is that our analysis focuses exclusively on the three main subtypes (collectively accounting for 90–95% of sCJD cases), while rarer variants were not included. This is particularly relevant for the MM(V)2 C subtype, which often presents with cognitive deficits at onset and can be difficult to distinguish from the more common MM1 and MV1 subtypes, which may also manifest solely with cognitive impairment in early phases. Moreover, excluding rare subtypes, which often have longer disease durations, may have biased our analysis of prognostic biomarkers and limit their applicability to the broader sCJD population. Furthermore, we acknowledge that the classification of sCJD subtypes may not have been entirely accurate due to the absence of neuropathological assessment in some patients. However, the risk of patient misdiagnosis or misclassification was substantially reduced through comprehensive evaluation of medical records, including survival data, codon 129 genotyping, and the application of second-generation prion RT-QuIC. As the study's main strength, pathology-specific biomarkers enabled us to control confounding effects from common copathologies such as AD and LBD. Furthermore, the application and integration of various machine learning models enabled us to

interrogate the data from multiple perspectives, yielding complementary insights into the proteomic profile of the disease, the most promising biomarker candidates, and the underlying biological processes. Validation studies in independent, possibly prospective, neuropathological cohorts, including rare subtypes (i.e., MM(V)2 C, MM2T, and VV1) and mixed cases, will be essential to confirm and expand our findings.

Conclusions

In conclusion, this study demonstrates the existence of both shared and subtype-specific CSF proteomic signatures in sCJD, shedding new light on the biological processes dysregulated in prion disease and driving its clinicopathological heterogeneity. We also identified proteins likely involved in sCJD pathogenesis and that may thus be considered therapeutic targets. In addition, we propose a set of biomarkers that accurately distinguish sCJD subtypes in vivo and may help in prognostication. Future studies in independent cohorts should validate our findings.

Abbreviations

αSyn	α-synuclein
Aβ	Amyloid beta
AD	Alzheimer's disease
AUC	Area under the curve
CI	Confidence interval
CSF	Cerebrospinal fluid
CTRL	Non-neurodegenerative controls
DEA	Differential expression analysis
FDR	False discovery rate
GO	Gene Ontology
HR	Hazard ratio
LBD	Lewy body disease
log2FC	log2 fold change
LP	Lumbar puncture
MS	Mass spectrometry
OPLS-DA	Orthogonal Partial Least Squares–Discriminant Analysis
PEA	Proximity extension assay
PrP^{Sc}	Misfolded prion protein
RF	Random Forest
ROC	Receiver operating characteristics
RT-QuIC	Real-time quaking-induced conversion assay
sCJD	Sporadic Creutzfeldt-Jakob disease
SHAP	SHapley Additive exPlanations
t-tau	Total tau
UMAP	Uniform Manifold Approximation and Projection
VIP	Variable Importance in Projection
WGCNA	Weighted Gene Co-expression Network Analysis

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40478-025-02168-9.

Additional file 1

Additional file 2

Additional file 3

Additional file 4

Additional file 5

Additional file 6

Additional file 7

Acknowledgements

The authors wish to thank the patients and their caregivers for supporting the research in neurodegenerative diseases and the Italian neurologists who provided the clinical information.

Author contributions

Conceptualization and design of the study: GMB, AM (Angela Mammana), DG, SA and PP. Drafting/revision of the manuscript for content, including medical writing for content GMB, DG and PP. Proximity extension assay experiments: AM (Angela Mammana), and EV. Major role in the acquisition, analysis, or interpretation of data: GMB, AM (Angela Mammana), DG, SB, EV, AM (Andrea Mastrangelo), AR, IMH, KB, SA, SC, and PP. Statistical analysis: GMB, DG, IMH, SA. Supervision: SA and PP. Critical review of the manuscript and approval of the final version: all authors.

Funding

P.P. is supported by JPND (01ED2407A), the Ministero della Salute (Ricerca Corrente), and the #NextGenerationEU (NGEU) funded by the Ministry of University and Research (MUR), National Recovery and Resilience Plan (NRRP), project MNESYS (PE0000006). D.G., I.M.H., and S.A. are supported by JPND (01ED2407A), ZonMW, Alzheimer Nederland, and Health ~ Holland. The sponsors had no role in study design, in the collection, analysis and interpretation of data, in the writing of the report, and in the decision to submit the article for publication.

Data availability

The datasets used and analyzed during the current study are available from the corresponding author upon reasonable request. All scripts used for data analysis are publicly available at GitHub using the following weblink: https://github.com/Deagogishvili/sCJD-CSF-proteomics.

Declarations

Ethics approval and consent to participate

The study was conducted according to the revised Declaration of Helsinki and Good Clinical Practice guidelines and approved by the local ethics committee (approval number AVEC:18025, 113/2018/OSS/AUSLBO). Written informed consent was given by study participants or the next of kin.

Consent for publication

Not applicable.

Competing interests

GMB, AM (Angela Mammana), DG, SB, EV, AM (Andrea Mastrangelo), AR, IMH, KB, SC, and PP declare no competing interests. Outside the submitted work: SA reports a patent pending; SA is in a consortium agreement with Cergentis BV as part of the TargetSV project; SA is in a consortium agreement with Olink and Quanterix as part of the NORMAL project.

Author details

[1]Department of Biomedical and Neuromotor Sciences (DiBiNeM), University of Bologna, Bologna, Italy
[2]IRCCS, Istituto delle Scienze Neurologiche di Bologna, Bologna, Italy
[3]AI Technology for Life, Department of Computing and Information Sciences, Department of Biology, Utrecht University, Utrecht, The Netherlands
[4]Neurochemistry Laboratory, Department of Laboratory Medicine, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam Neuroscience, Amsterdam, The Netherlands
[5]Department of Neurodegenerative Diseases, Hertie Institute for Clinical Brain Research, Eberhard Karls University Tübingen, Tübingen, Germany
[6]German Center for Neurodegenerative Diseases (DZNE), Tübingen, Germany
[7]Department of Computer Science, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

Received: 9 June 2025 / Accepted: 20 October 2025
Published online: 15 December 2025

References

1. Abu-Rumeileh S, Baiardi S, Polischi B, Mammana A, Franceschini A, Green A et al (2019) Diagnostic value of surrogate CSF biomarkers for Creutzfeldt–Jakob disease in the era of RT-QuIC. J Neurol 266:3136–3143. https://doi.org/10.1007/s00415-019-09537-0
2. Assarsson E, Lundberg M, Holmquist G, Björkesten J, Thorsen SB, Ekman D et al (2014) Homogenous 96-plex PEA immunoassay exhibiting high sensitivity, specificity, and excellent scalability. PLoS ONE 9:e95192. https://doi.org/10.1371/journal.pone.0095192
3. Baiardi S, Quadalti C, Mammana A, Dellavalle S, Zenesini C, Sambati L et al (2022) Diagnostic value of plasma p-tau181, NfL, and GFAP in a clinical setting cohort of prevalent neurodegenerative dementias. Alzheimers Res Ther 14:153. https://doi.org/10.1186/s13195-022-01093-6
4. Baiardi S, Rossi M, Capellari S, Parchi P (2019) Recent advances in the histomolecular pathology of human prion disease. Brain Pathol 29:278–300. https://doi.org/10.1111/bpa.12695
5. Bartoletti-Stella A, Corrado P, Mometto N, Baiardi S, Durrenberger PF, Arzberger T et al (2019) Analysis of RNA expression profiles identifies dysregulated vesicle trafficking pathways in Creutzfeldt-Jakob disease. Mol Neurobiol 56:5009–5024. https://doi.org/10.1007/s12035-018-1421-1
6. Bentivenga GM, Baiardi S, Mastrangelo A, Zenesini C, Mammana A, Polischi B et al (2023) Diagnostic and prognostic value of cerebrospinal fluid SNAP-25 and neurogranin in Creutzfeldt-Jakob disease in a clinical setting cohort of rapidly progressive dementias. Alzheimers Res Ther 15:150. https://doi.org/10.1186/s13195-023-01300-y
7. Bentivenga GM, Gonzalez-Ortiz F, Baiardi S, Kirsebom B-E, Mastrangelo A, Mammana A et al (2024) Clinical value of novel blood-based tau biomarkers in Creutzfeldt–Jakob disease. Alzheimers Dement 21:e14422. https://doi.org/10.1002/alz.14422
8. Bentivenga GM, Mammana A, Baiardi S, Rossi M, Ticca A, Magliocchetti F et al (2024) Performance of a seed amplification assay for misfolded alpha-synuclein in cerebrospinal fluid and brain tissue in relation to Lewy body disease stage and pathology burden. Acta Neuropathol 147:18. https://doi.org/10.1007/s00401-023-02663-0
9. Bishop MT, Will RG, Manson JC (2010) Defining sporadic Creutzfeldt-Jakob disease strains and their transmission properties. Proc Natl Acad Sci U S A 107:12005–12010. https://doi.org/10.1073/pnas.1004688107
10. Bizzi A, Pascuzzo R, Blevins J, Moscatelli MEM, Grisoli M, Lodi R et al (2021) Subtype diagnosis of sporadic Creutzfeldt-Jakob disease with diffusion magnetic resonance imaging. Ann Neurol 89:560–572. https://doi.org/10.1002/ana.25983
11. Bridel C, van Gils JHM, Miedema SSM, Hoozemans JJM, Pijnenburg YAL, Smit AB et al (2023) Clusters of co-abundant proteins in the brain cortex associated with fronto-temporal lobar degeneration. Alzheimers Res Ther 15:1–18. https://doi.org/10.1186/s13195-023-01200-1
12. Cali I, Castellani R, Alshekhlee A, Cohen Y, Blevins J, Yuan J et al (2009) Co-existence of scrapie prion protein types 1 and 2 in sporadic Creutzfeldt-Jakob disease: its effect on the phenotype and prion-type characteristics. Brain 132:2643–2658. https://doi.org/10.1093/brain/awp196
13. Camargos S, Scholz S, Simón-Sánchez J, Paisán-Ruiz C, Lewis P, Hernandez D et al (2008) DYT16, a novel young-onset dystonia-parkinsonism disorder: identification of a segregating mutation in the stress-response protein PRKRA. Lancet Neurol 7:207–215. https://doi.org/10.1016/S1474-4422(08)70022-X
14. del Campo M, Peeters CF, Johnson ECB, Vermunt L, Hok-A-Hin YS, van Nee M et al (2022) CSF proteome profiling reveals protein panels reflecting the pathophysiological diversity of Alzheimer's disease. Nat Aging 2:1040–1053. https://doi.org/10.1038/s43587-022-00300-1
15. del Campo M, Vermunt L, Peeters CFW, Sieben A, Hok-A-Hin YS, Lleó A et al (2023) CSF proteome profiling reveals biomarkers to discriminate dementia with Lewy bodies from Alzheimer´s disease. Nat Commun 14:5635. https://doi.org/10.1038/s41467-023-41122-y
16. Davidson-Pilon (2019) Lifelines: survival analysis in python. J Open Source Softw 4:1317. https://doi.org/10.21105/joss.01317
17. Dayon L, Cominetti O, Affolter M (2022) Proteomics of human biological fluids for biomarker discoveries: technical advances and recent applications. Expert Rev Proteom 19:131–151. https://doi.org/10.1080/14789450.2022.2070477
18. Folke J, Pakkenberg B, Brudek T (2019) Impaired Wnt signaling in the prefrontal cortex of Alzheimer's disease. Mol Neurobiol 56:873–891. https://doi.org/10.1007/s12035-018-1103-z
19. Gambetti P, Cali I, Notari S, Kong Q, Zou W-Q, Surewicz WK (2011) Molecular biology and pathology of prion strains in sporadic human prion diseases. Acta Neuropathol 121:79–90. https://doi.org/10.1007/s00401-010-0761-3
20. Gmitterová K, Heinemann U, Bodemer M, Krasnianski A, Meissner B, Kretzschmar HA et al (2009) 14-3-3 CSF levels in sporadic Creutzfeldt-Jakob disease differ across molecular subtypes. Neurobiol Aging 30:1842–1850. https://doi.org/10.1016/j.neurobiolaging.2008.01.007
21. Gmitterová K, Heinemann U, Krasnianski A, Gawinecka J, Zerr I (2016) Cerebrospinal fluid markers in the differentiation of molecular subtypes of sporadic Creutzfeldt-Jakob disease. Eur J Neurol 23:1126–1133. https://doi.org/10.1111/ene.12991
22. Gogishvili D, Vromen EM, Koppes-den Hertog S, Lemstra AW, Pijnenburg YAL, Visser PJ et al (2023) Discovery of novel CSF biomarkers to predict progression in dementia using machine learning. Sci Rep 13:6531. https://doi.org/10.1038/s41598-023-33045-x
23. Hagberg AA, Schult DA, Swart PJ (2008) Exploring network structure, dynamics, and function using NetworkX. In: Proceedings of the 7th python in science conference 11–15. https://doi.org/10.25080/TCWV9851
24. Harris CR, Millman KJ, van der Walt SJ, Gommers R, Virtanen P, Cournapeau D et al (2020) Array programming with numpy. Nature 585:357–362. https://doi.org/10.1038/s41586-020-2649-2
25. Hermann P, Appleby B, Brandel J-P, Caughey B, Collins S, Geschwind MD et al (2021) Biomarkers and diagnostic guidelines for sporadic Creutzfeldt-Jakob disease. Lancet Neurol 20:235–246. https://doi.org/10.1016/S1474-4422(20)30477-4
26. Hilton KJ, Cunningham C, Reynolds RA, Perry VH (2013) Early hippocampal synaptic loss precedes neuronal loss and associates with early behavioural deficits in three distinct strains of prion disease. PLoS ONE 8:e68062. https://doi.org/10.1371/journal.pone.0068062
27. Hunter JD (2007) Matplotlib: a 2D graphics environment. Comput Sci Eng 9:90–95. https://doi.org/10.1109/MCSE.2007.55
28. Khoonsari PE, Häggmark A, Lönnberg M, Mikus M, Kilander L, Lannfelt L et al (2016) Analysis of the cerebrospinal fluid proteome in Alzheimer's disease. PLoS ONE 11:e0150672. https://doi.org/10.1371/journal.pone.0150672
29. Kushwaha R, Li Y, Makarava N, Pandit NP, Molesworth K, Birukov KG et al (2023) Reactive astrocytes associated with prion disease impair the blood brain barrier. Neurobiol Dis 185:106264. https://doi.org/10.1016/j.nbd.2023.106264
30. Langfelder P, Horvath S (2008) WGCNA: an R package for weighted correlation network analysis. BMC Bioinform 9:559. https://doi.org/10.1186/1471-2105-9-559
31. Lattanzio F, Abu-Rumeileh S, Franceschini A, Kai H, Amore G, Poggiolini I et al (2017) Prion-specific and surrogate CSF biomarkers in Creutzfeldt-Jakob disease: diagnostic accuracy in relation to molecular subtypes and analysis of neuropathological correlates of p-tau and Aβ42 levels. Acta Neuropathol 133:559–578. https://doi.org/10.1007/s00401-017-1683-0
32. de Leeuw FA, Peeters CFW, Kester MI, Harms AC, Struys EA, Hankemeier T et al (2017) Blood-based metabolic signatures in Alzheimer's disease. Alzheimers Dement (Amst) 8:196–207. https://doi.org/10.1016/j.dadm.2017.07.006
33. Leiserson MDM, Vandin F, Wu H-T, Dobson JR, Eldridge JV, Thomas JL et al (2015) Pan-cancer network analysis identifies combinations of rare somatic mutations across pathways and protein complexes. Nat Genet 47:106–114. https://doi.org/10.1038/ng.3168
34. L'Episcopo F, Tirolo C, Caniglia S, Testa N, Morale MC, Serapide MF et al (2014) Targeting Wnt signaling at the neuroimmune interface for dopaminergic neuroprotection/repair in Parkinson's disease. J Mol Cell Biol 6:13–26. https://doi.org/10.1093/jmcb/mjt053
35. Li B, Chen M, Zhu C (2021) Neuroinflammation in prion disease. Int J Mol Sci 22:2196. https://doi.org/10.3390/ijms22042196
36. Liberski PP, Budka H (1999) Neuroaxonal pathology in Creutzfeldt-Jakob disease. Acta Neuropathol 97:329–334. https://doi.org/10.1007/s004010050995
37. Lundberg SM, Erion G, Chen H, DeGrave A, Prutkin JM, Nair B et al (2020) From local explanations to global understanding with explainable AI for trees. Nat Mach Intell 2:56–67. https://doi.org/10.1038/s42256-019-0138-9

38. Lundberg SM, Lee S-I (2017) A unified approach to interpreting model predictions. Advances in neural information processing systems. Curran Associates, Inc
39. McInnes L, Healy J, Melville J (2018) UMAP: uniform manifold approximation and projection. J Open Source Softw 3:861. https://doi.org/10.21105/joss.00861
40. Montine TJ, Phelps CH, Beach TG, Bigio EH, Cairns NJ, Dickson DW et al (2012) National Institute on Aging–Alzheimer's association guidelines for the neuropathologic assessment of Alzheimer's disease: a practical approach. Acta Neuropathol 123:1–11. https://doi.org/10.1007/s00401-011-0910-3
41. Nevola K, Sandin M, Guess J, Forsberg S, Cambronero C, Pucholt P et al (2025) OlinkAnalyze: facilitate analysis of proteomic data from Olink. https://doi.org/10.32614/CRAN.package.OlinkAnalyze. https://cran.r-project.org/web/packages/OlinkAnalyze/index.html Accessed 27 March 2025
42. Olink®—Part of Thermo Fisher Scientific https://olink.com. Accessed 27 Mar 2025
43. Parchi P, de Boni L, Saverioni D, Cohen ML, Ferrer I, Gambetti P et al (2012) Consensus classification of human prion disease histotypes allows reliable identification of molecular subtypes: an inter-rater study among surveillance centres in Europe and USA. Acta Neuropathol 124:517–529. https://doi.org/10.1007/s00401-012-1002-8
44. Parchi P, Giese A, Capellari S, Brown P, Schulz-Schaeffer W, Windl O et al (1999) Classification of sporadic Creutzfeldt-Jakob disease based on molecular and phenotypic analysis of 300 subjects. Ann Neurol 46(2):224–233. https://doi.org/10.1002/1531-8249(199908)46:2%3C224::AID-ANA12%3E3.0.CO;2-W
45. Parchi P, Strammiello R, Notari S, Giese A, Langeveld JPM, Ladogana A et al (2009) Incidence and spectrum of sporadic Creutzfeldt–Jakob disease variants with mixed phenotype and co-occurrence of PrP^{Sc} types: an updated classification. Acta Neuropathol 118:659–671. https://doi.org/10.1007/s00401-009-0585-1
46. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O et al (2011) Scikit-learn: machine learning in python. J Mach Learn Res 12:2825–2830
47. Plotly Technologies Inc. Collaborative Data Science. Plotly (2015) https://plotly.com/python/. Accessed 25 April 2025
48. Reyna MA, Haan D, Paczkowska M, Verbeke LPC, Vazquez M, Kahraman A et al (2020) Pathway and network analysis of more than 2500 whole cancer genomes. Nat Commun 11:729. https://doi.org/10.1038/s41467-020-14367-0
49. Reyna MA, Leiserson MDM, Raphael BJ (2018) Hierarchical hotnet: identifying hierarchies of altered subnetworks. Bioinformatics 34:i972–i980. https://doi.org/10.1093/bioinformatics/bty613
50. Seabold S, Perktold J (2010) Statsmodels: econometric and statistical modeling with Python. In: Proceedings of the 9th Python in science conference 92–96. https://doi.org/10.25080/Majora-92bf1922-011
51. Sjöstedt E, Zhong W, Fagerberg L, Karlsson M, Mitsios N, Adori C et al (2020) An atlas of the protein-coding genes in the human, pig, and mouse brain. Science 367:eaay5947. https://doi.org/10.1126/science.aay5947
52. Staffaroni AM, Kramer AO, Casey M, Kang H, Rojas JC, Orrú CD et al (2019) Association of blood and cerebrospinal fluid Tau level and other biomarkers with survival time in sporadic Creutzfeldt-Jakob disease. JAMA Neurol 76:969. https://doi.org/10.1001/jamaneurol.2019.1071
53. Tarozzi M, Baiardi S, Sala C, Bartoletti-Stella A, Parchi P, Capellari S et al (2022) Genomic, transcriptomic and RNA editing analysis of human MM1 and VV2 sporadic Creutzfeldt-Jakob disease. Acta Neuropathol Commun 10:181. https://doi.org/10.1186/s40478-022-01483-9
54. Terpilowski MA (2019) scikit-posthocs: pairwise multiple comparison tests in python. J Open Source Softw 4:1169. https://doi.org/10.21105/joss.01169
55. The Human Protein Atlas https://www.proteinatlas.org/. Accessed 29 Mar 2025
56. The Human Protein Atlas Brain tissue expression of WASF1—Summary. https://www.proteinatlas.org/ENSG00000112290-WASF1/brain. Accessed 29 Mar 2025
57. The Human Protein Atlas Brain tissue expression of CCDC80—Summary. https://www.proteinatlas.org/ENSG00000091986-CCDC80/brain. Accessed 29 Mar 2025
58. The Human Protein Atlas Brain tissue expression of GPC5—Summary https://www.proteinatlas.org/ENSG00000179399-GPC5/brain. Accessed 29 Mar 2025
59. The pandas development team (2024) pandas-dev/pandas: pandas (v2.2.2). https://doi.org/10.5281/zenodo.10957263. Zenodo
60. Thellung S, Corsaro A, Dellacasagrande I, Nizzari M, Zambito M, Florio T (2022) Proteostasis unbalance in prion diseases: mechanisms of neurodegeneration and therapeutic targets. Front Neurosci. https://doi.org/10.3389/fnins.2022.966019
61. Thévenot EA, Roux A, Xu Y, Ezan E, Junot C (2015) Analysis of the human adult urinary metabolome variations with age, body mass index, and gender by implementing a comprehensive workflow for univariate and OPLS statistical analyses. J Proteome Res 14:3322–3335. https://doi.org/10.1021/acs.jproteome.5b00354
62. Toni M, Massimino ML, De Mario A, Angiulli E, Spisni E (2017) Metal dyshomeostasis and their pathological role in prion and prion-like diseases: the basis for a nutritional approach. Front Neurosci 11 :3 https://doi.org/10.3389/fnins.2017.00003
63. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D et al (2020) SciPy 1.0: fundamental algorithms for scientific computing in python. Nat Methods 17:261–272. https://doi.org/10.1038/s41592-019-0686-2
64. Waskom ML (2021) Seaborn: statistical data visualization. J Open Source Softw 6:3021. https://doi.org/10.21105/joss.03021
65. Watt NT, Taylor DR, Kerrigan TL, Griffiths HH, Rushworth JV, Whitehouse IJ et al (2012) Prion protein facilitates uptake of zinc into neuronal cells. Nat Commun 3:1134. https://doi.org/10.1038/ncomms2135
66. Wickham H (2016) ggplot2: elegant graphics for data analysis, 2nd edn. Springer
67. Wickham H, François R, Henry L (2023) dplyr: a grammar of data manipulation. https://cran.r-project.org/package=dplyr. Accessed 25 April 2025
68. Wickham H, Henry L (2023) tidyr: tidy messy data. https://cran.r-project.org/package=tidyr. Accessed 25 April 2025
69. Woerman AL (2021) Strain diversity in neurodegenerative disease: an argument for a personalized medicine approach to diagnosis and treatment. Acta Neuropathol 142:1–3. https://doi.org/10.1007/s00401-021-02311-5
70. Zerr I, Parchi P (2018) Sporadic Creutzfeldt–Jakob disease. Handbook of clinical neurology. Elsevier, Amsterdam, pp 155–174
71. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O et al (2019) Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat Commun 10:1523. https://doi.org/10.1038/s41467-019-09234-6

Publisher's note