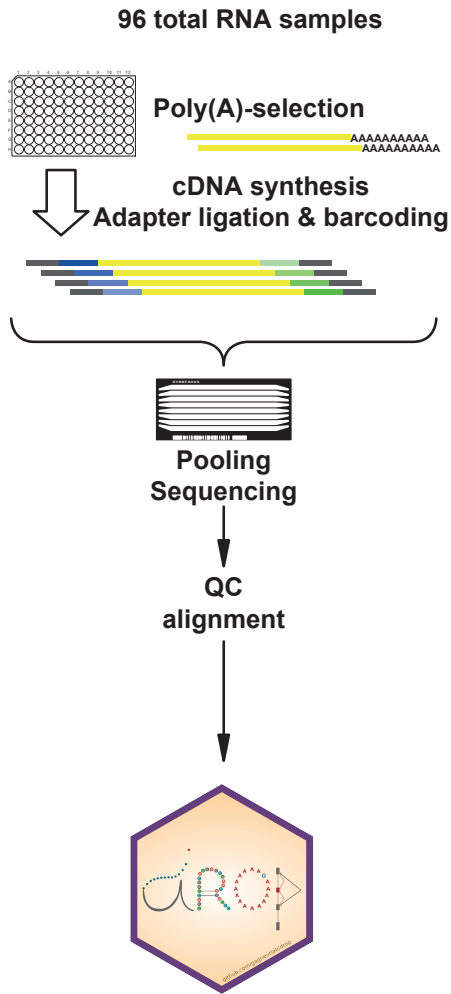
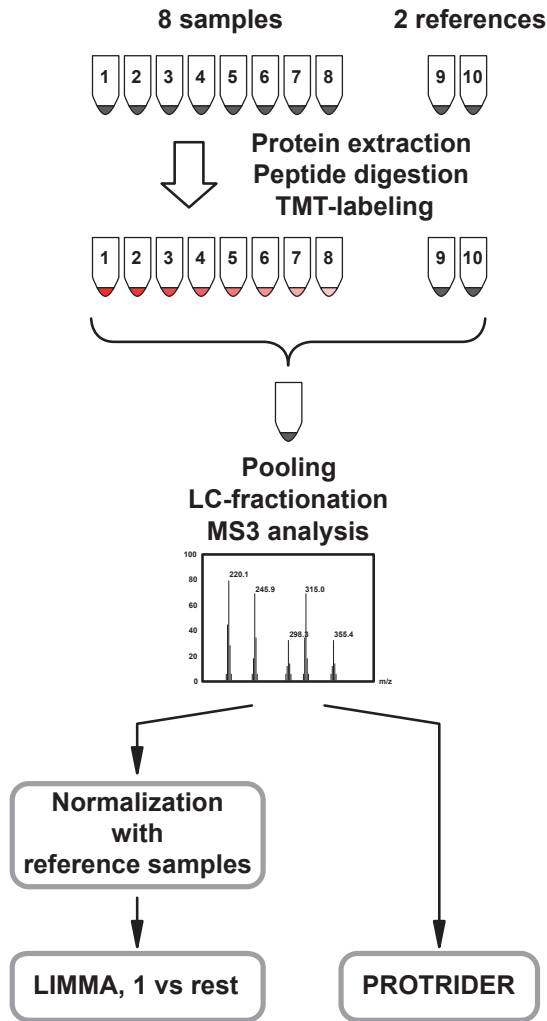
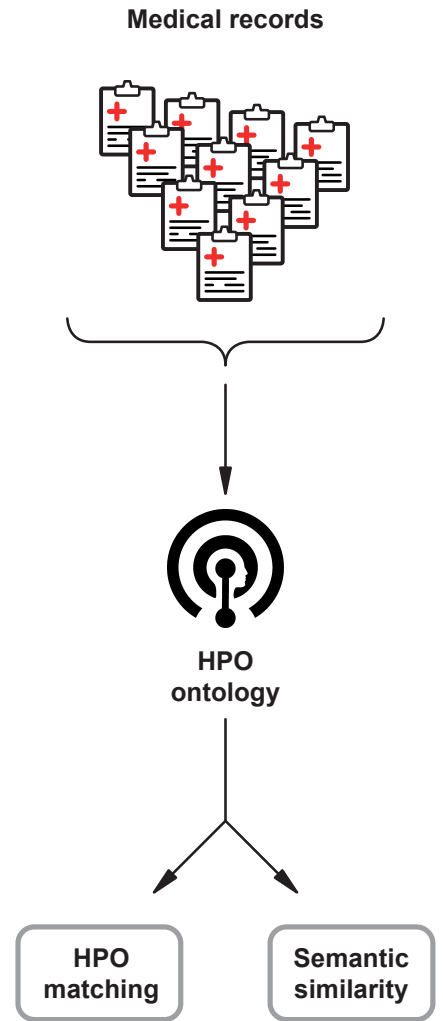


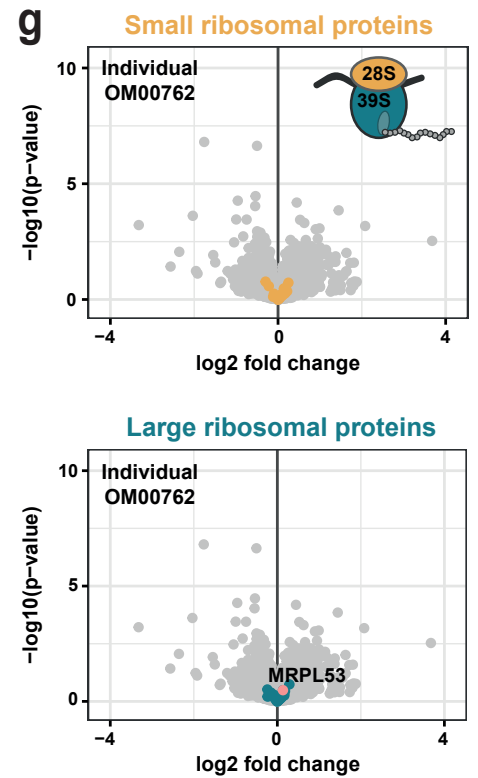
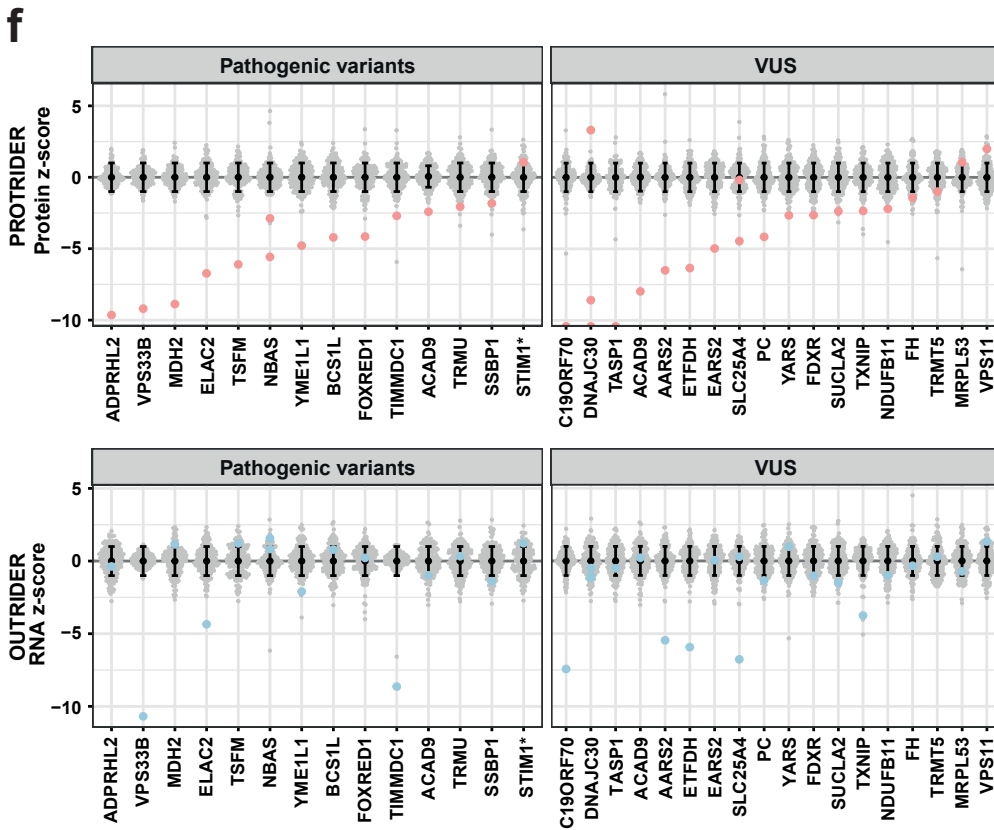
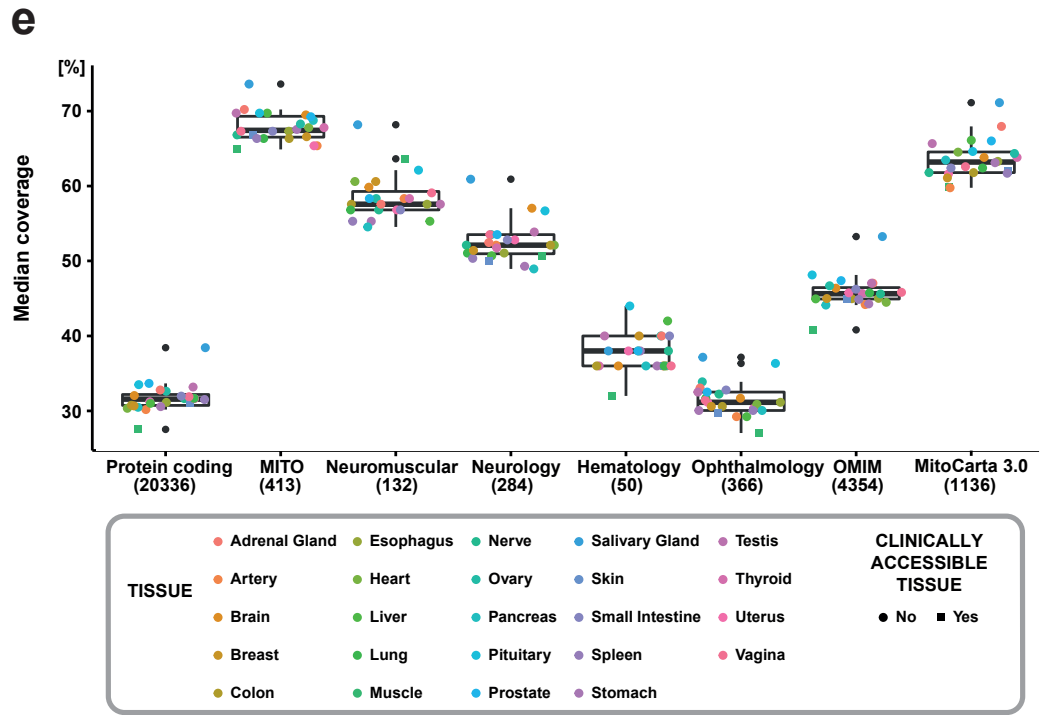
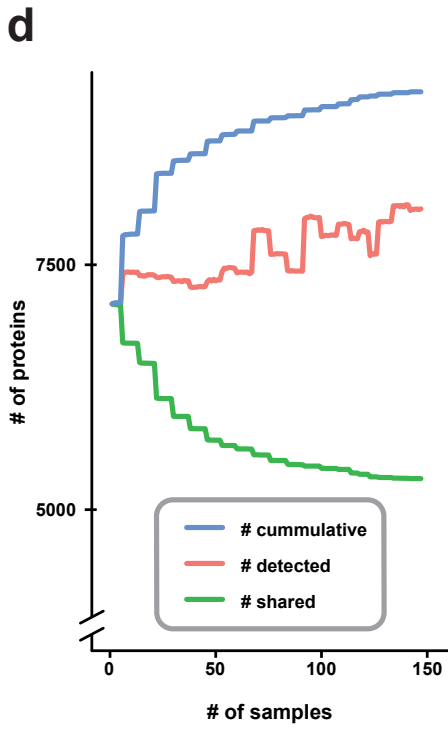
## **Supplementary Information**

The Supplementary information contains:

- Supplementary Figures and Legends
- Case Summaries
- Supplementary References

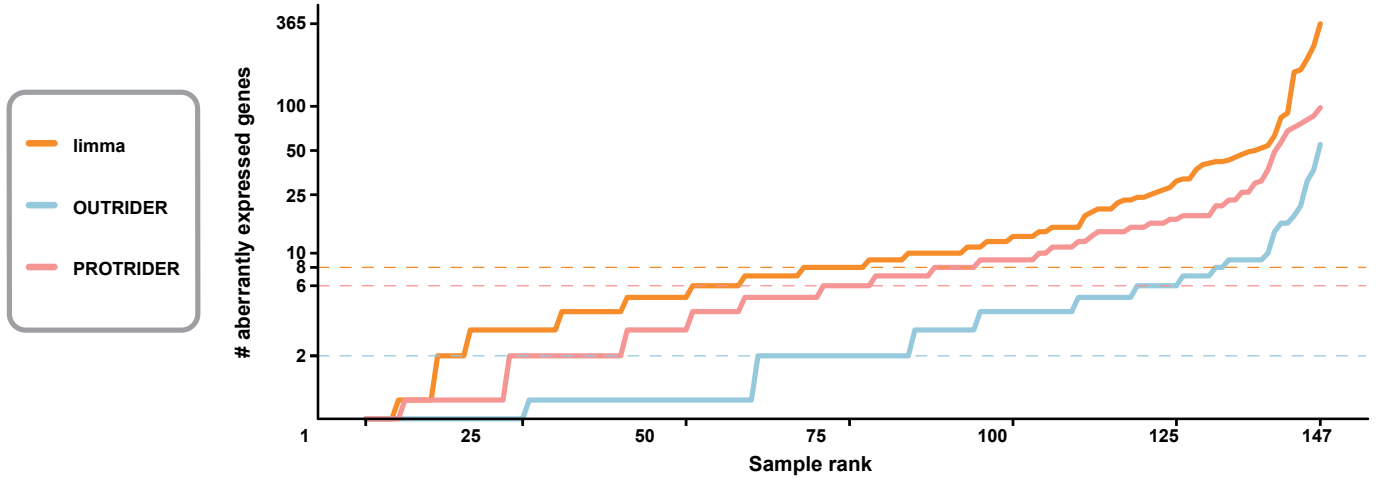
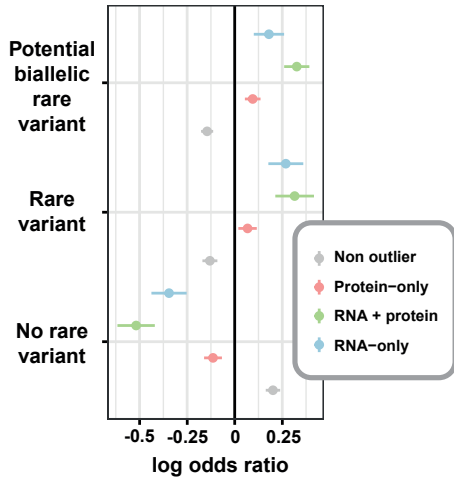
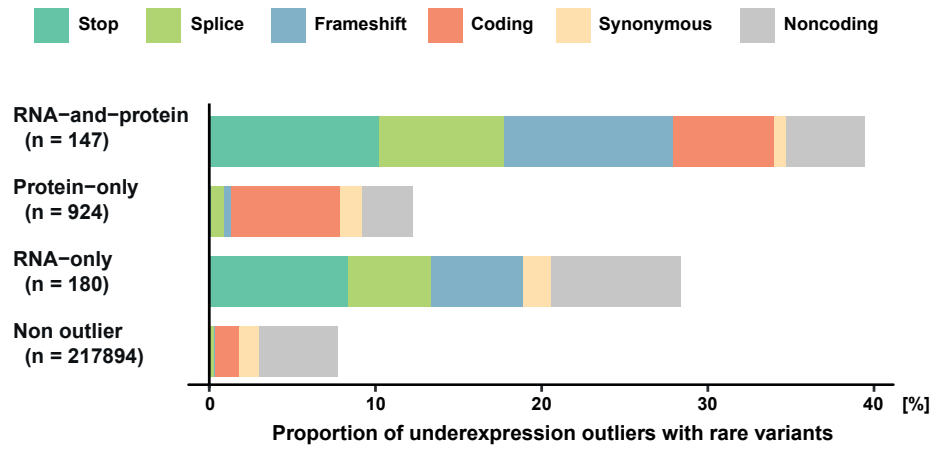
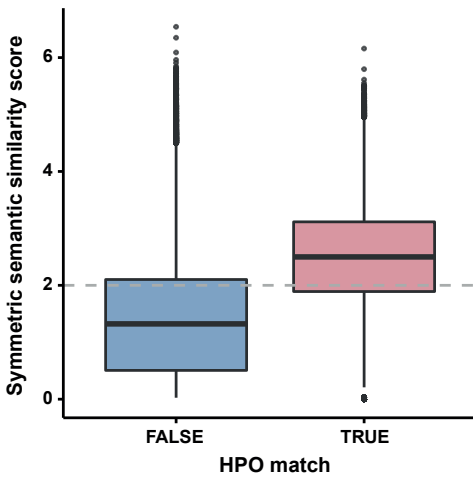
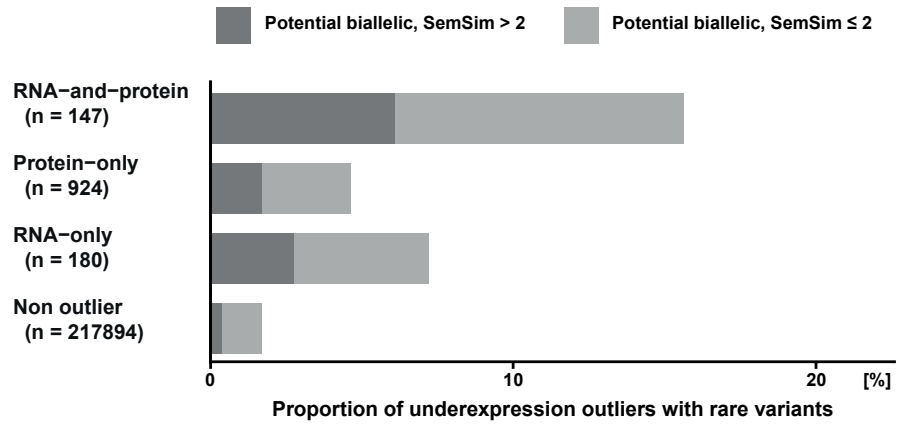
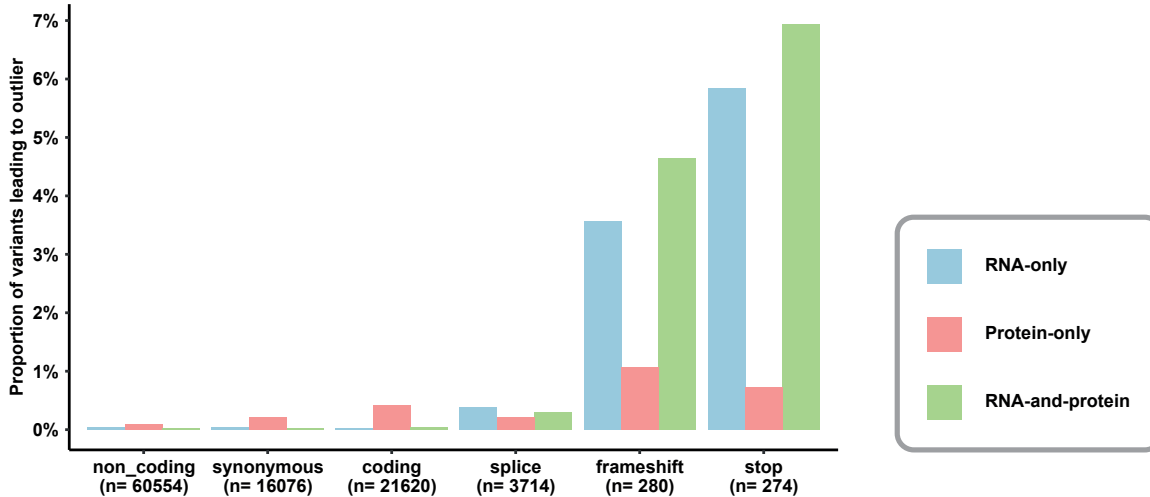
### **Supplementary Figures and Legends**

**a****RNAseq - workflow****b****Proteomics - workflow****c****Phenomics - workflow**



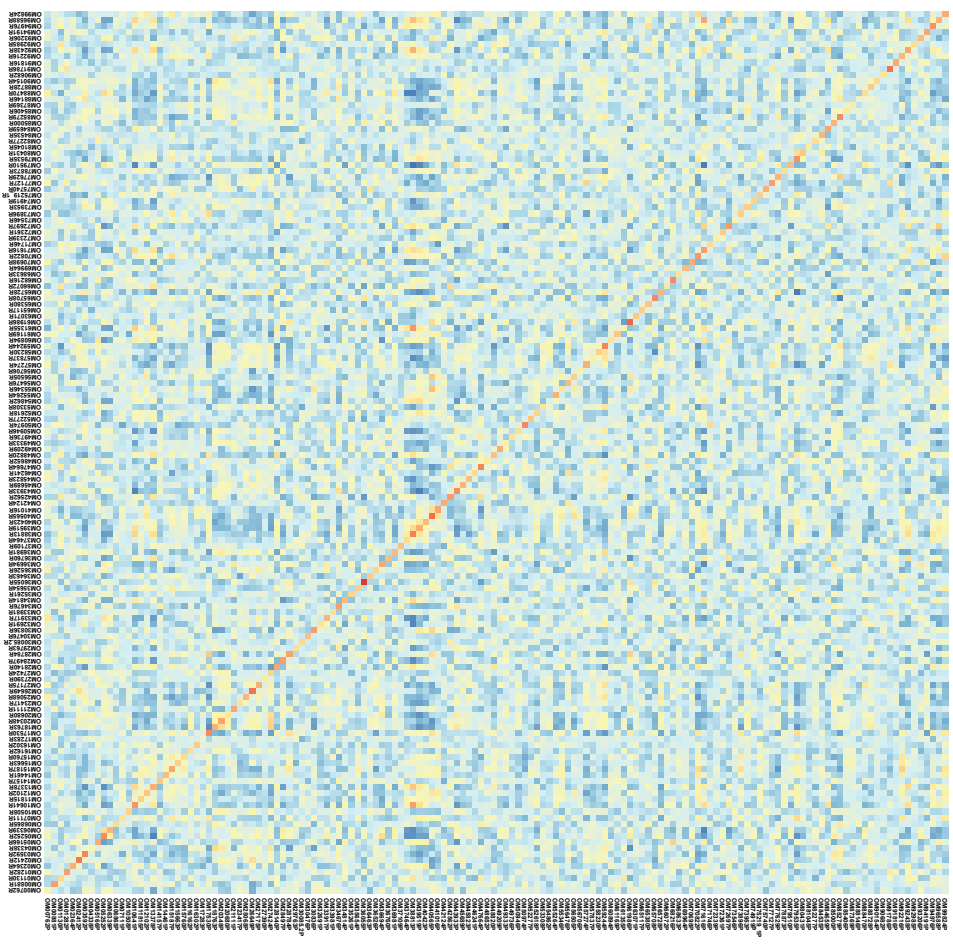
**Supplementary Fig. 1: Transcriptomic, proteomic, and phenomic workflows and the detection of proteins and outliers.** **a**, RNA-seq analysis of patient-derived fibroblasts was performed by strand-specific mRNA sequencing in batches of 96 samples. Discarding lowly expressed genes, approximately 12,000 transcripts were detectable (total 15,652, median 11,432). Data analysis was performed in the DROP pipeline (see **Methods**). **b**, Proteomic analysis of patient-derived fibroblasts was performed by tandem mass tag (TMT) labelled proteomics. Eight patient samples and two reference samples were processed in parallel. Peptides were tagged sample-wise prior to pooling, liquid chromatography (LC) fractionation, and analysis by mass spectrometry (MS3). The reference sample, in duplicate, was included for inter-batch normalization. Approximately 8,000 proteins were quantified per sample (total 9,264, median 7,720). Data analysis was performed by both limma and PROTRIDER (see **Methods**). **c**, Phenomic workflow from extraction of patient phenotype data from the medical records, reported as human phenotype ontology (HPO) terms, to calculation of the symmetric semantic similarity between the patient's phenotype and the disease-gene-associated phenotype reported in OMIM (<https://omim.org>) and determination of a phenotype match to the disease-gene-associated phenotype (see **Methods**). **d**, Number of detected proteins cumulatively (blue), per sample (red), and shared across all samples (green) in our dataset. **e**, Mean coverage of proteins encoded by all protein coding genes, a selection of disease genes for various disease groups, and all OMIM disease genes by tissue as per the publically available GTEx proteomic dataset (Jiang et al., 2020). The center line represents median, box limits correspond to the 25th and 75th percentiles, whiskers extend to the largest value no further than  $1.5 * \text{IQR}$  from the box limits, values beyond the whiskers are outliers and plotted individually. **f**, Protein and RNA z-score distribution for disease-causing genes with known pathogenic variants (left panels) and prioritized variants of uncertain significance (VUS, right panels). The points appear in red/blue for cases known to harbour variants requiring validation and in green for novel cases diagnosed in our downstream systematic approach. Proteins and RNAs that were not detected are depicted on the y-axis at a z-score value of -10. **g**, Normal expression of MRPL53 and both the small and large mitoribosomal subunit in individual OM00762. Data are displayed as a gene-wise protein expression volcano plot of nominal ( $-\log_{10}$ ) p-values against protein intensity  $\log_2$  fold change.



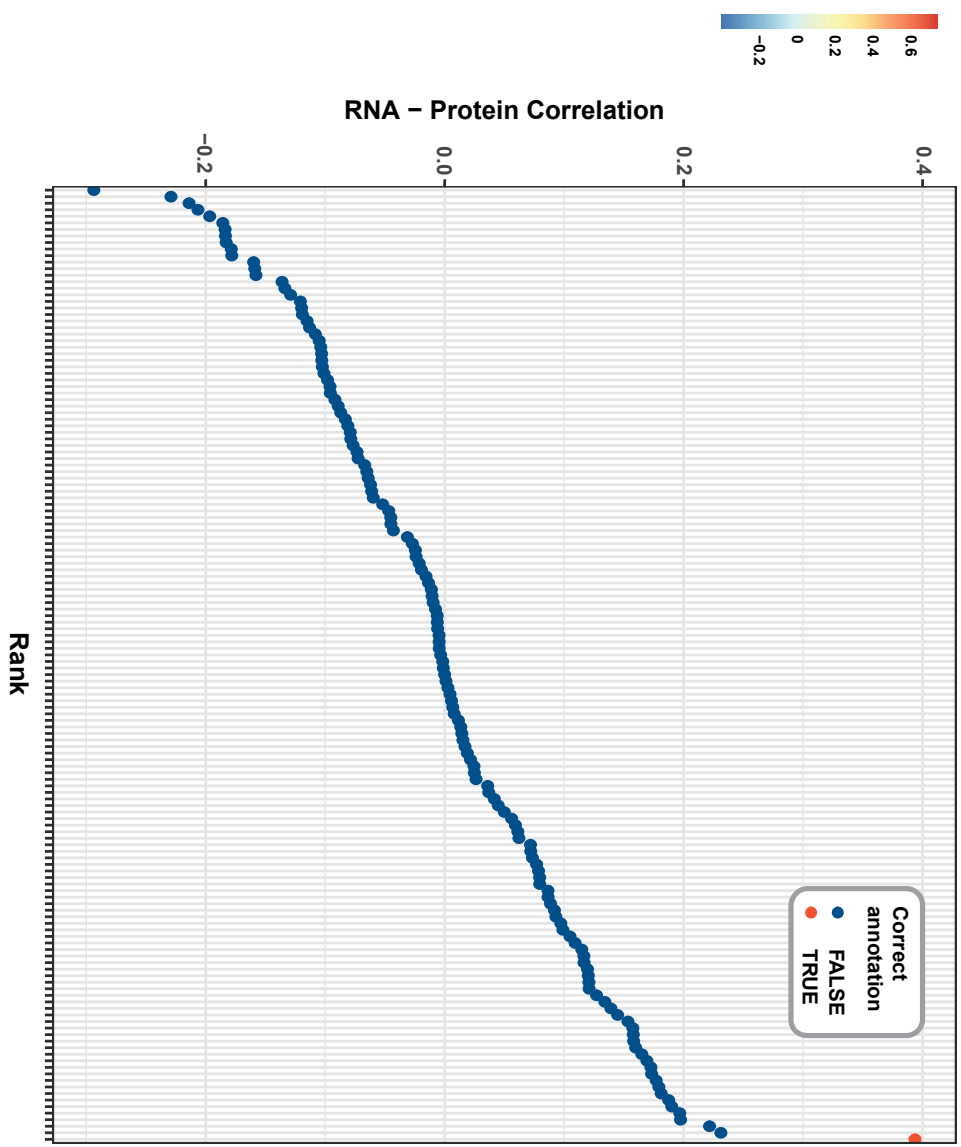
**a****b****c****d****e****f**

**Supplementary Fig. 2: The association of RNA and protein outliers with genetic variants and phenotype.** **a**, Systematic calling of expression outliers over the entire dataset of 143 patients and two healthy controls identifying a median of two aberrantly expressed transcripts (OUTRIDER) and either six (PROTRIDER) or eight (limma) aberrantly expressed proteins per sample. **b**, Enrichment analysis for rare potentially biallelic, rare single, and frequent variants in all classes of outliers. Data are displayed as a log odds ratio with the 95% confidence interval. **c**, Proportion of underexpression outliers with rare variants detected by WES of three different outlier classes (RNA-only, protein-only, and RNA-and-protein) in comparison to non-outliers. 40% of RNA-and-protein outliers, 28% of RNA-only outliers, and 12% of protein-only outliers contain at least one rare variant in the corresponding gene, in comparison to 8% in non-outliers. **d**, Proportion of rare variants in a given category that led to an outlier across outlier classes. **e**, Semantic similarity calculated from patient HPO terms in patients with a phenotype match according to the method defined by Fresard et al., 2019. Semantic similarity threshold of  $\geq 2$  determined to be equivalent to a phenotype match. The center line represents median, box limits correspond to the 25th and 75th percentiles, whiskers extend to the largest value no further than  $1.5 * \text{IQR}$  from the box limits, values beyond the whiskers are outliers and plotted individually. **f**, Proportion of underexpression outliers with rare potentially biallelic variants detected by WES of three different outlier classes (RNA-only, protein-only, and RNA-and-protein) in comparison to non-outliers stratified by semantic similarity score of  $\geq 2$  (phenotype match) and with a semantic similarity score of  $< 2$  or missing (no phenotype match).

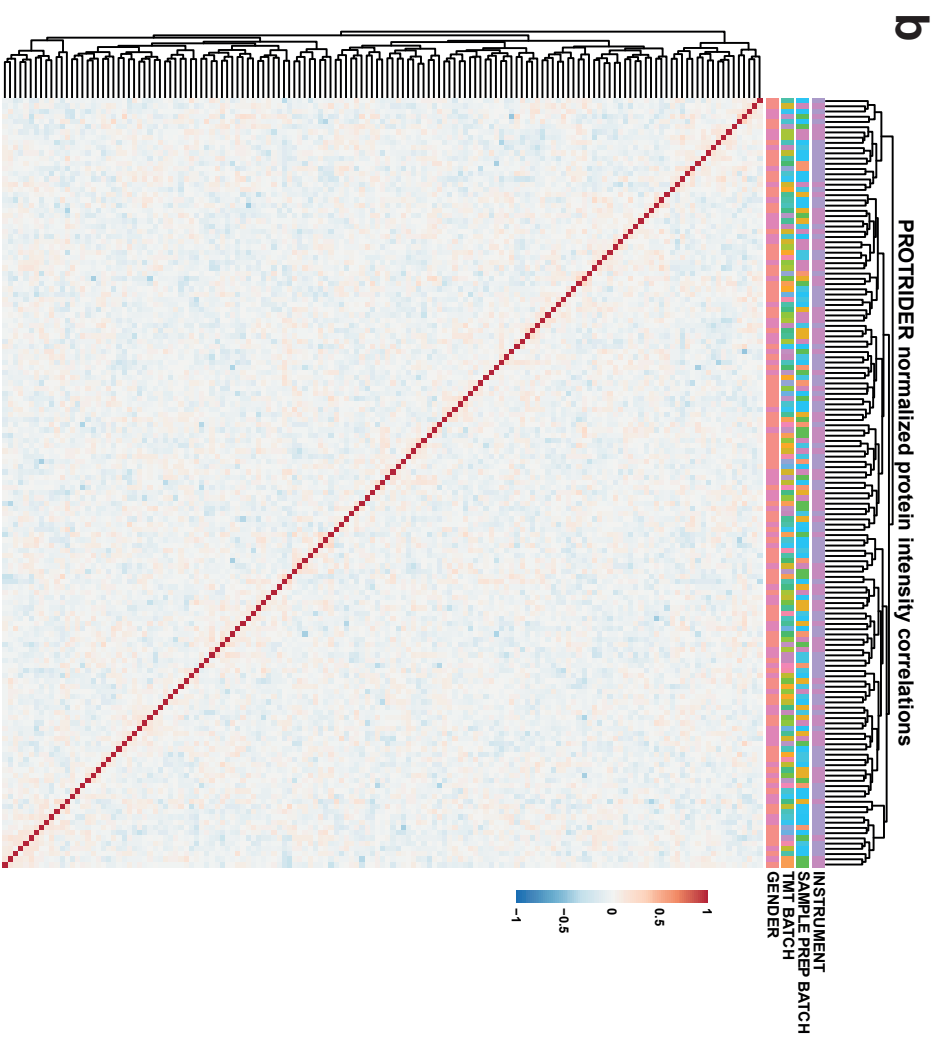
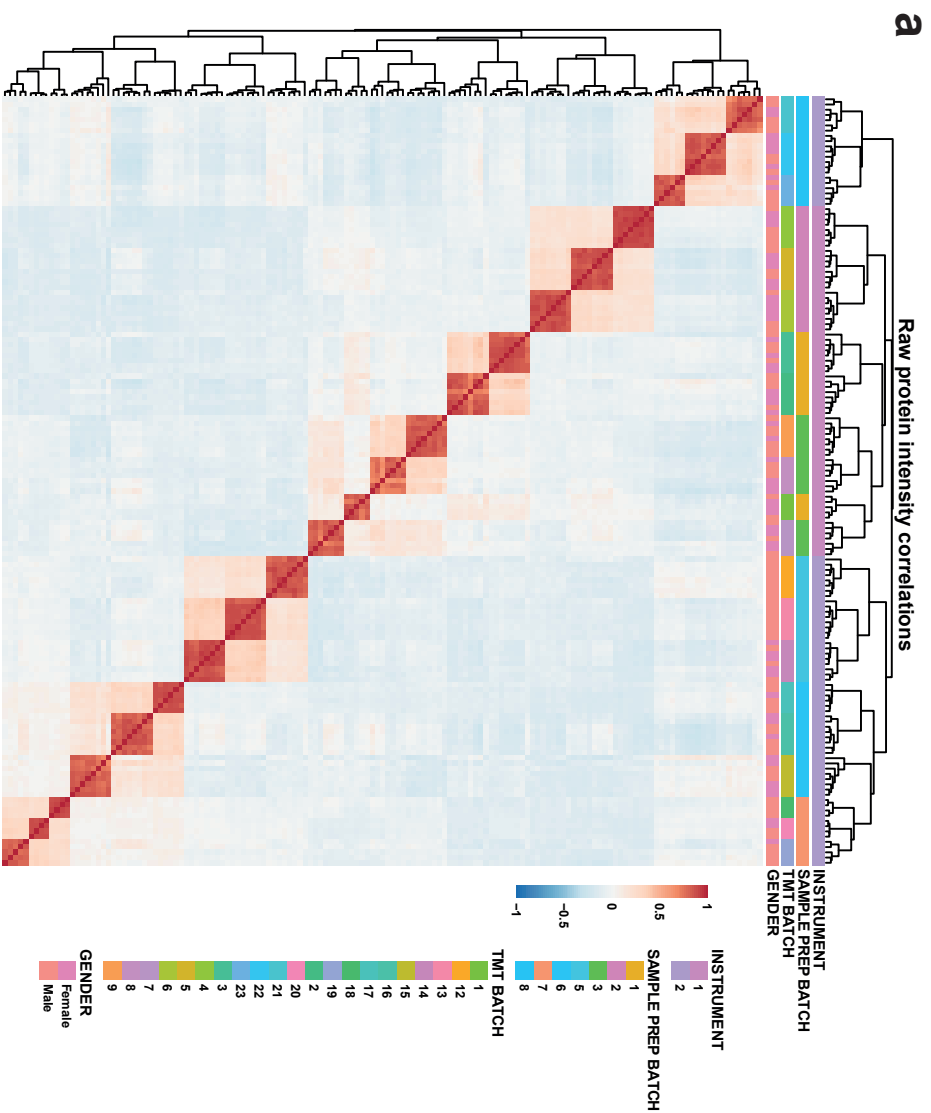
**a** Transcriptome – proteome correlation per sample

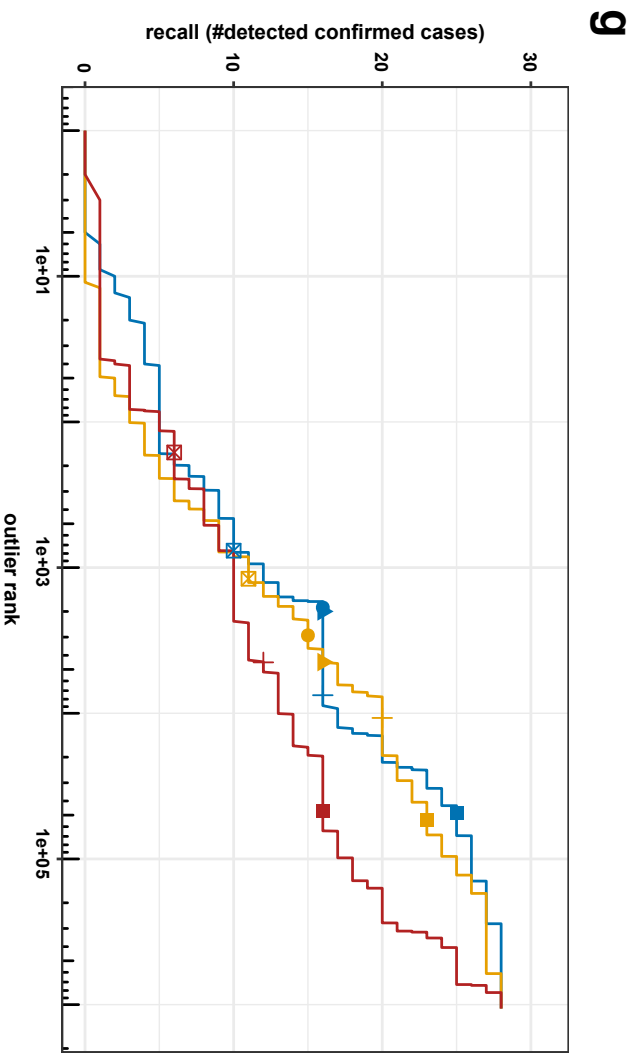
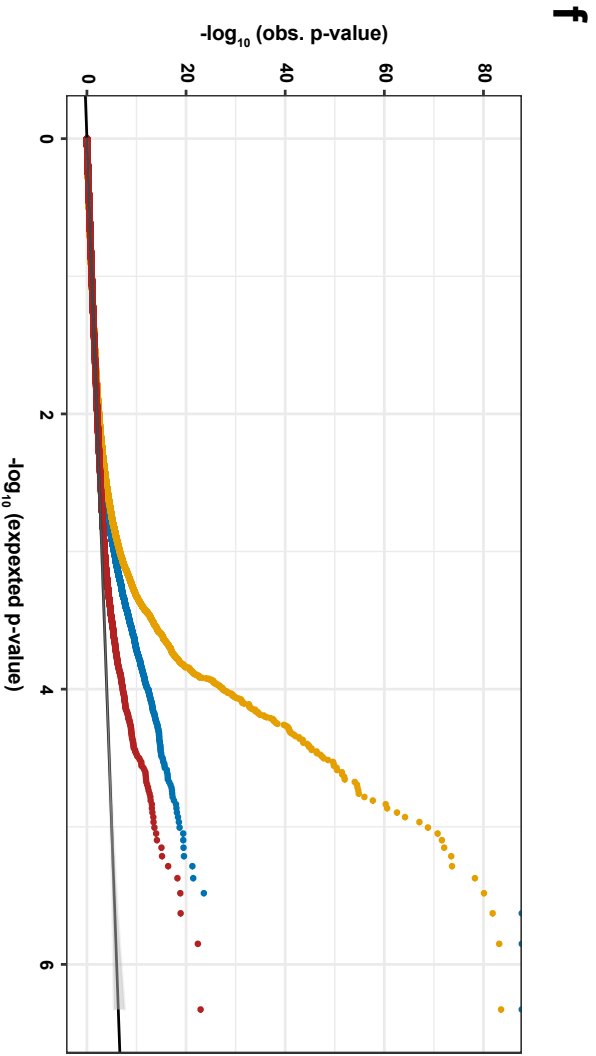
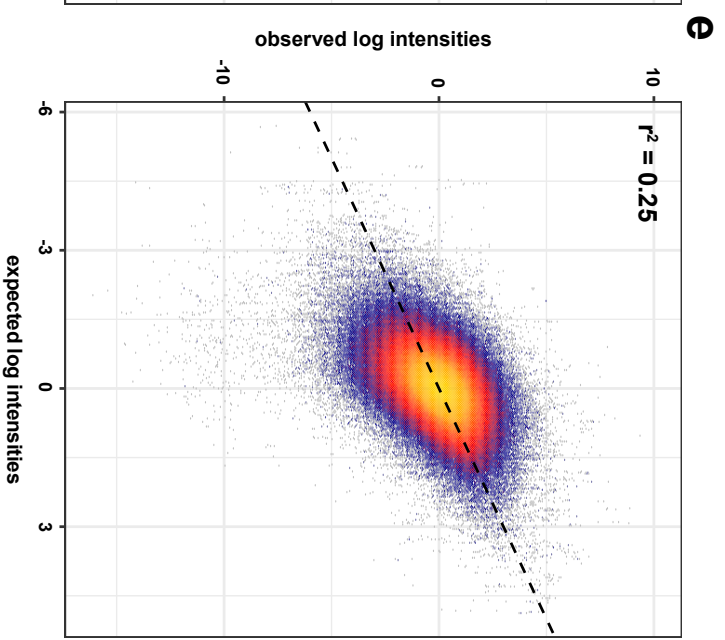
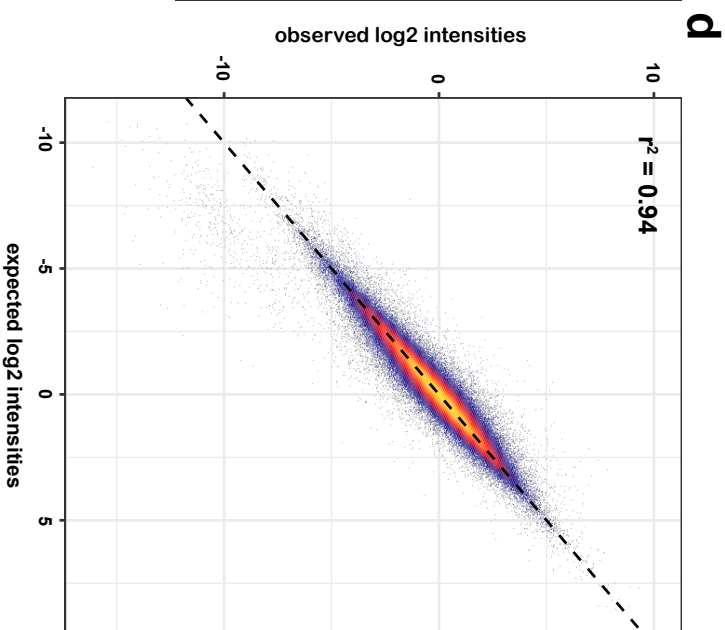
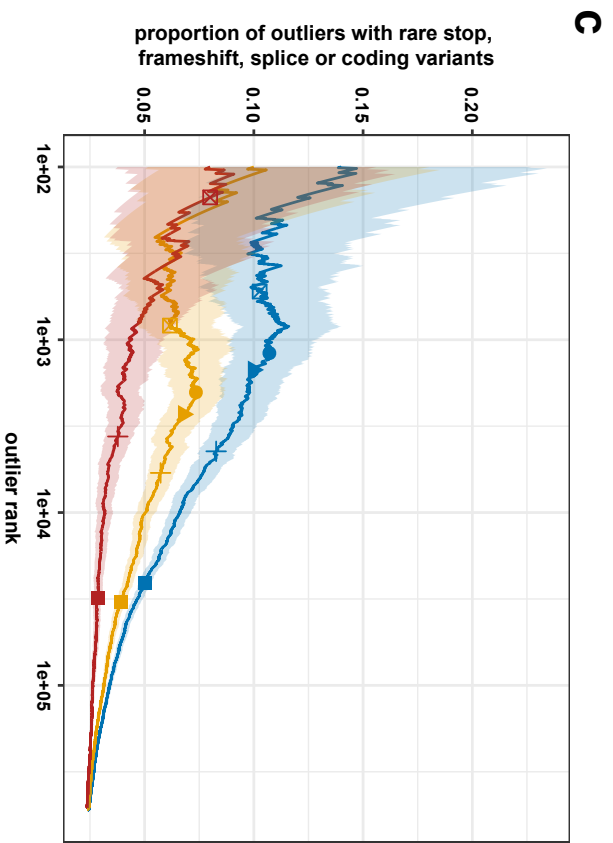


**b**



**Supplementary Fig. 3: RNA-seq-proteomics matching.** **a**, Heatmap of sample-wise transcriptome-proteome Spearman rank correlation. Red indicates positive correlation, blue indicates negative correlation. **b**, Example of Spearman rank correlation for the proteome sample OM35261 to all transcriptomes, confirming the corresponding transcriptome according to the sample annotation (red) to demonstrate the highest correlation coefficient.



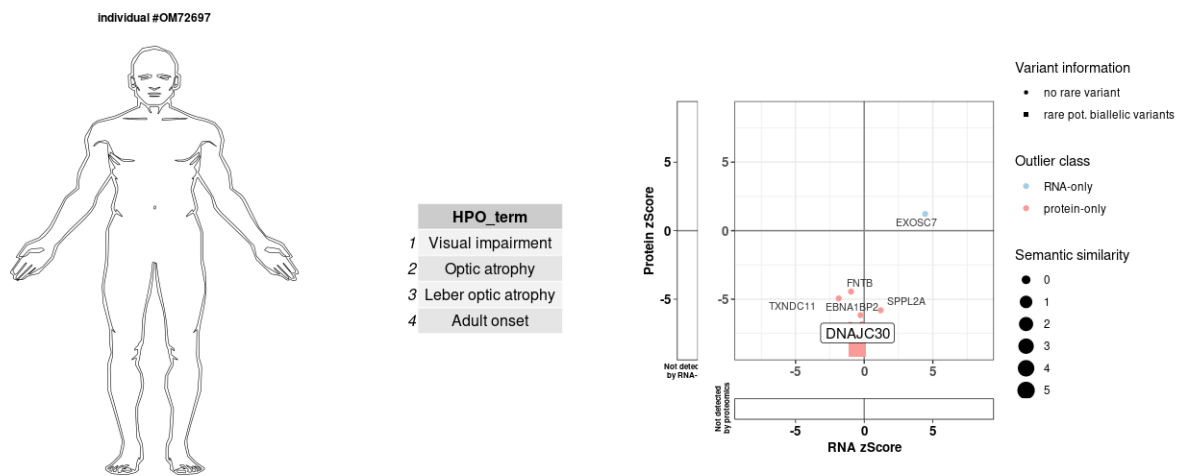


**Supplementary Fig. 4: Comparison of PROTRIDER with alternative approaches for the detection of aberrant protein expression.** Correlation matrix of samples (rows and columns) on **a**, raw log-transformed and protein-centered intensity values and **b**, PROTRIDER normalized log-transformed and protein-centered intensity values. Red indicates positive correlation, blue indicates negative correlation. Colored horizontal tracks display instrument, sample preparation batch, and TMT proteomic batch, and gender of the samples. Two different instruments have been used for the analysis, three batches were processed in parallel indicated by the sample preparation batch, and the samples have been analysed in 23 TMT batches. The dendrogram represents the sample-wise clustering. **c**, Proportion of outlier genes containing a rare (MAF < 0.01) stop, splice, frameshift or coding variant against the ranking of outliers for PROTRIDER (blue), the limma-based approach (yellow) and z-scores on the residuals of a regression of covariates (red). Shapes indicate FDR cutoffs of 0.05 and 0.1 as well as z-score cutoffs of 2, 3, and 5. **d-e**, Observed  $\log_2$  intensities centered per protein (y axis) are plotted against predicted  $\log_2$  intensities (x axis, also centered per protein) for **d**, PROTRIDER and **e**, the covariate regression approach. The color indicates the density of the data points with yellow corresponding to high density and blue to low density.  $r^2$  values describing the proportion of variance explained are shown in the top left corner. **f**, Global quantile-quantile plot of observed negative log-transformed p-values (y-axis) against expected p-values (x-axis) for PROTRIDER (blue), the limma-based approach (yellow) and z-scores on the residuals of a regression of covariates (red). Under the null hypothesis, the data are expected to lie along the diagonal (95% confidence bands shown in gray). **g**, Recall (number of detected confirmed pathogenic outliers, n=28, y axis) is plotted against the ranking of nominal p-values of all protein-sample pairs (x axis) for PROTRIDER (blue), the limma-based approach (yellow) and z-scores on the residuals of a regression of covariates (red). Shapes indicate FDR cutoffs of 0.05 and 0.1 as well as z-score cutoffs of 2, 3 and 5.

## Case Summaries

### Validation cohort

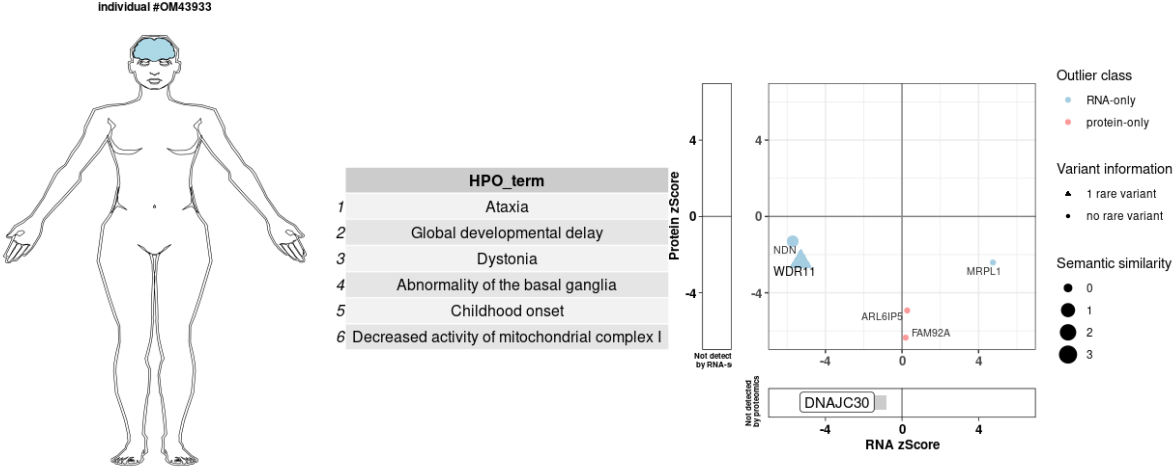
#### Case 1, Sample ID: OM72697, Causal gene: *DNAJC30*



Adult male diagnosed with homozygous missense variant (c.[152A>G], p.[Tyr51Cys]) in *DNAJC30*. Phenotype of the patient matched the published phenotype of the causal gene (Semantic similarity score = 5.95). *DNAJC30* was a significant protein underexpression outlier (FC: 0.01; p-value:  $8.72 \times 10^{-18}$ ; z-score: -8.59). The protein was detected in the TMT-batch but not in this sample, therefore the missing value was imputed as described in the methods.

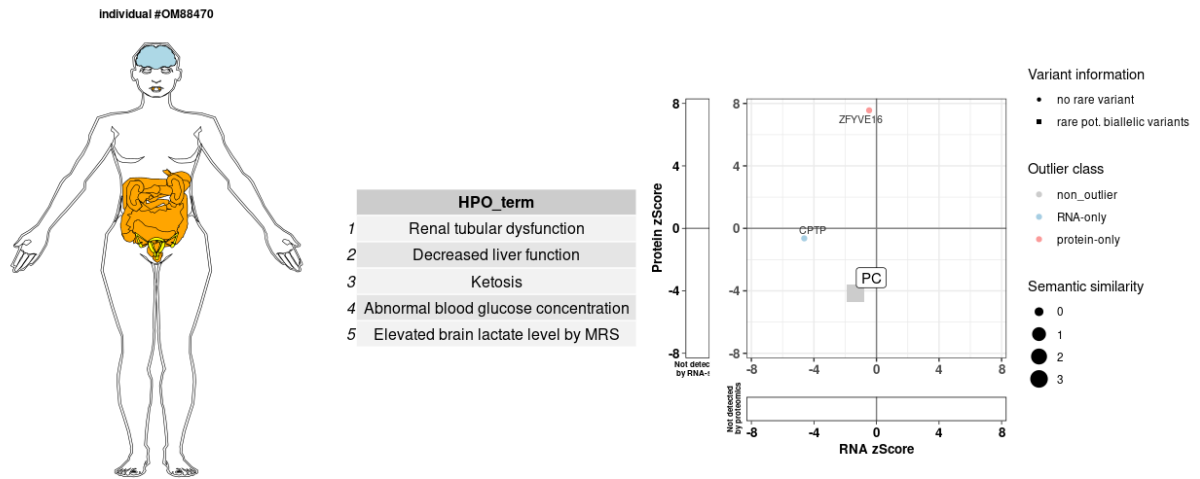


**Case 2, Sample ID: OM43933, Causal gene: *DNAJC30***



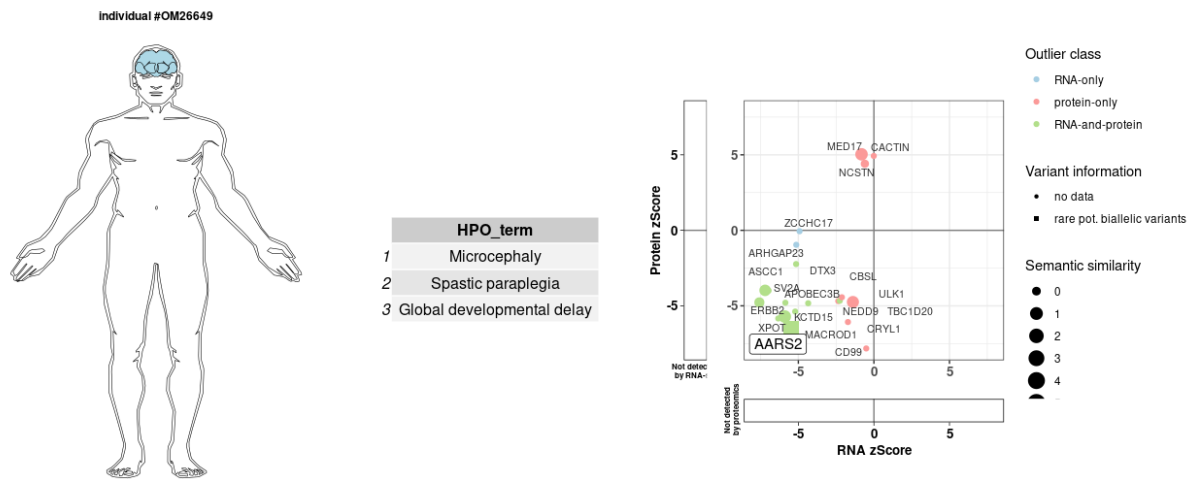
Infant female diagnosed with homozygous missense variant (c.[152A>G], p.[Tyr51Cys]) in *DNAJC30*. Phenotype of the patient matched the published phenotype of the causal gene (Semantic similarity score = 6.54). *DNAJC30* was a significant protein underexpression outlier (FC: 0.15; p-value: 1.66\*10<sup>-15</sup>; z-score: -7.96). The protein was detected in the replicate sample grown in galactose medium.

**Case 3, Sample ID: OM88470, Causal gene: *PC***



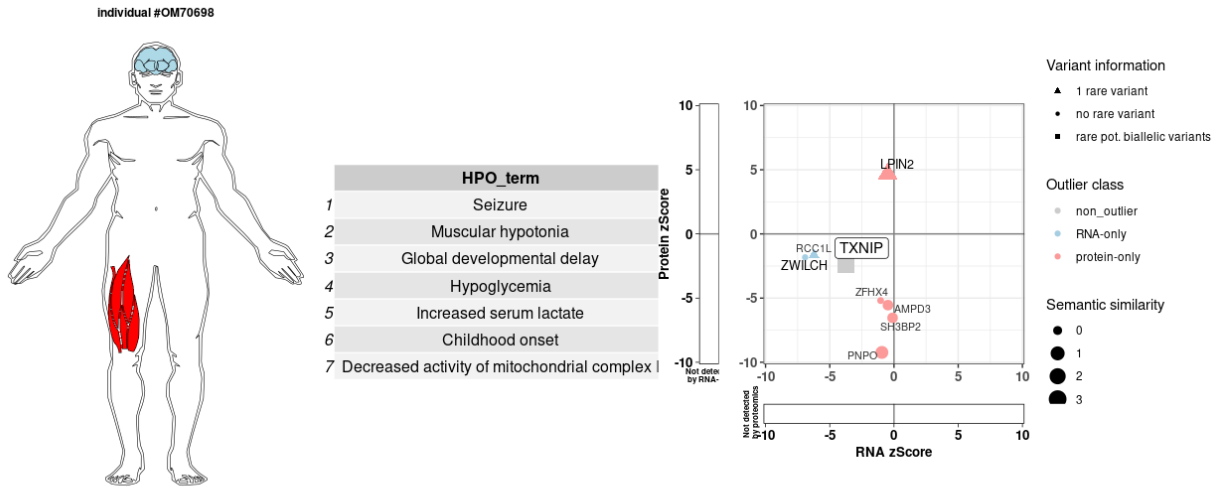
Paediatric female diagnosed with compound heterozygous missense variants (c.[1217G>A], p.[Arg406His]; c.[1891C>T], p.[Arg631Trp]) in *PC*. Phenotype of the patient matched the published phenotype of the causal gene (Semantic similarity score = 3.61). *PC* protein was underexpressed with nominal significance (FC: 0.52; p-value:  $3.19 \times 10^{-5}$ ; z-score: -4.16).

## Case 4, Sample ID: OM26649, Causal gene: *AARS2*



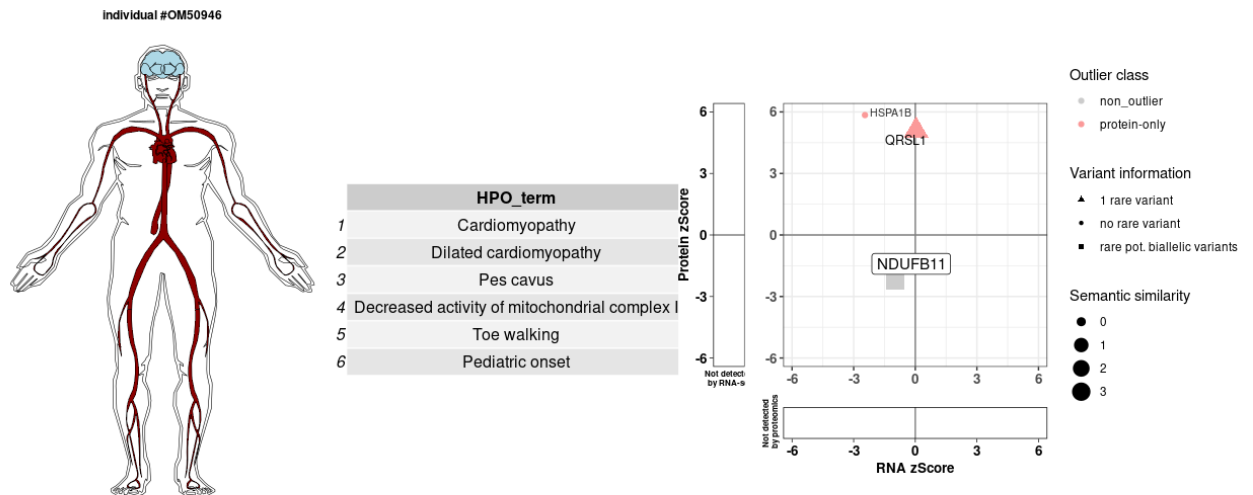
Paediatric male diagnosed with homozygous missense variant (c.[959C>T], p.[Ala320Val]) in *AARS2*. Phenotype of the patient matched the published phenotype of the causal gene (Semantic similarity score = 3.18). *AARS2* was significantly RNA-and-protein underexpression outlier (RNA: FC: 0.54; p-value:  $5.49 \times 10^{-7}$ ; z-score: -5.46 / Protein: FC: 0.20; p-value:  $7.91 \times 10^{-11}$ ; z-score: -6.50).

**Case 5, Sample ID: OM70698, Causal gene: *TXNIP***



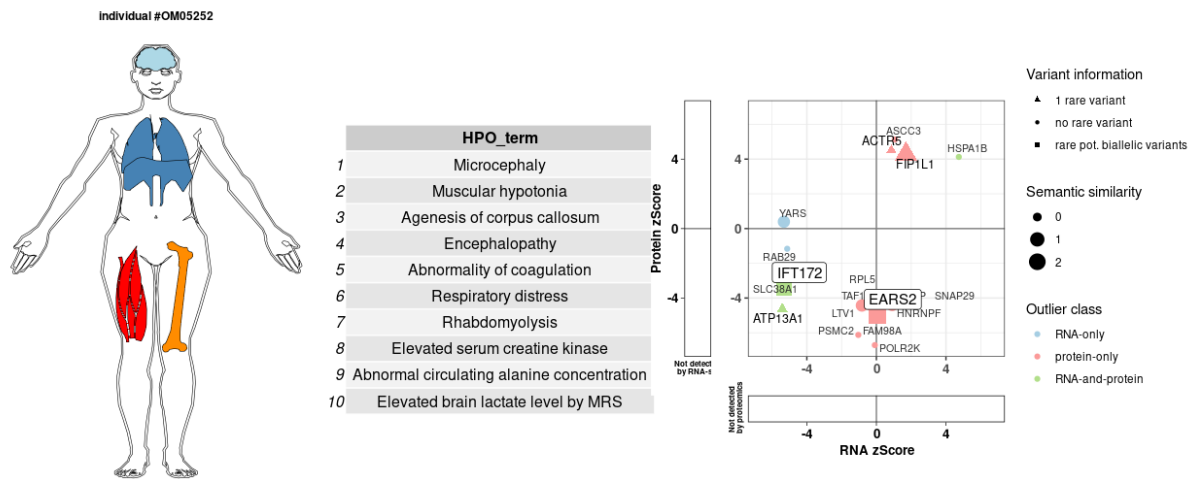
Paediatric male diagnosed with homozygous indel variant causing frameshift (c.[642\_643insT], p.[Ile215Tyrfs\*59]) in *TXNIP*. Phenotype of the patient matched the published phenotype of the causal gene (Semantic similarity score = 3.05). *TXNIP* was not published as a disease-associated gene when the case was initially investigated. Moreover, once reported, the phenotype similarity was questionable given that only one family has been published. *TXNIP* was underexpressed with nominal significance on both RNA and protein levels (RNA: FC: 0.18; p-value:  $9.51 \times 10^{-4}$ ; z-score: -3.75 / Protein: FC: 0.35; p-value:  $1.87 \times 10^{-2}$ ; z-score: -2.35).

**Case 6, Sample ID: OM50946, Causal gene: *NDUFB11***



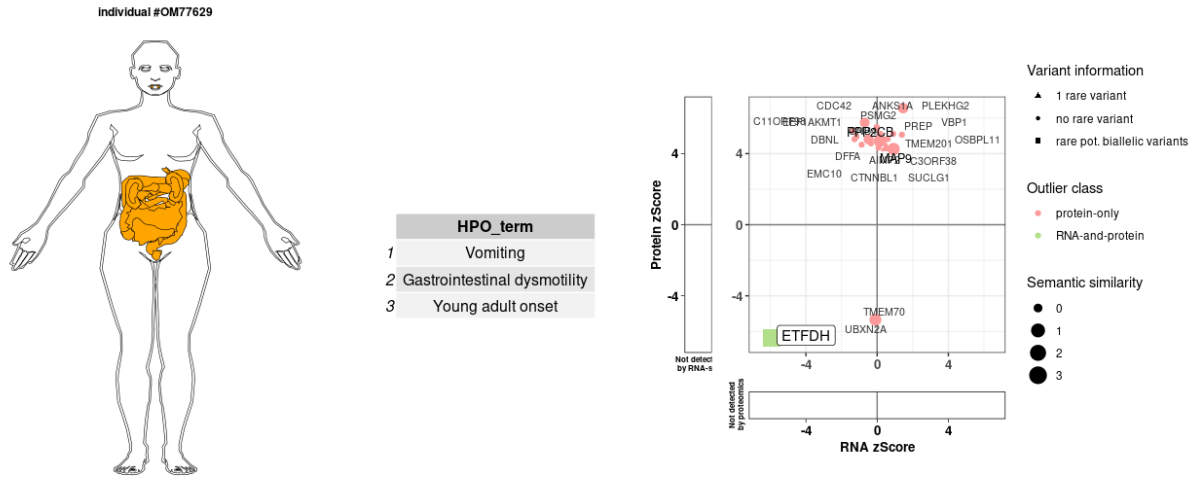
Paediatric male diagnosed with hemizygous missense variant (c.[379T>C], p.[Trp127Arg]) in *NDUFB11*. Phenotype of the patient matched the published phenotype of the causal gene (Semantic similarity score = 4.36). *NDUFB11* protein was underexpressed with nominal significance (Protein: FC: 0.68; p-value:  $2.78 \times 10^{-2}$ ; z-score: -2.20). Furthermore a global reduction of respiratory chain complex I was observed.

**Case 7, Sample ID: OM05252, Causal gene: *EARS2***



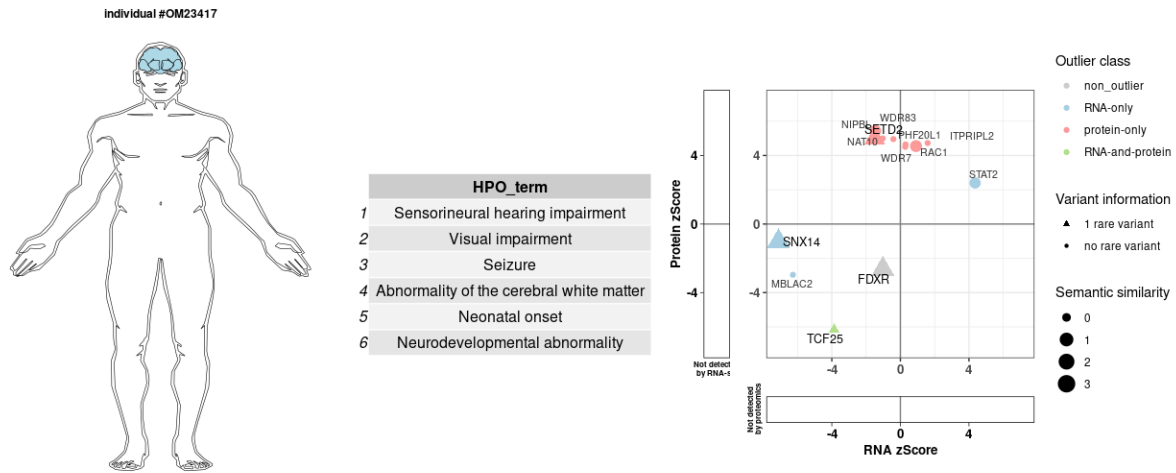
Paediatric female diagnosed with homozygous missense variant (c.[911C>A], p.[Pro304His]) in *EARS2*. Phenotype of the patient matched the published phenotype of the causal gene (Semantic similarity score = 2.96). *EARS2* was a significant protein underexpression outlier (FC: 0.47; p-value:  $6.57 \times 10^{-7}$ ; z-score: -4.97).

**Case 8, Sample ID: OM77629, Causal gene: *ETFDH***



Adult female diagnosed with heterozygous indel variant (c.[685G>T; 687\_688delAA; 691T>C];[=] p.[Ala229Ser; Thr230Ilefs\*2; Phe231Val];[=]) in *ETFDH*. Phenotype of the patient matched the published phenotype of the causal gene (Semantic similarity score = 3.53). *ETFDH* was significant RNA-and-protein underexpression outlier (RNA: FC: 0.60; p-value: 2.71\*10<sup>-8</sup>; z-score: -5.93 / Protein: FC: 0.48; p-value: 2.22\*10<sup>-10</sup>; z-score: -6.34).

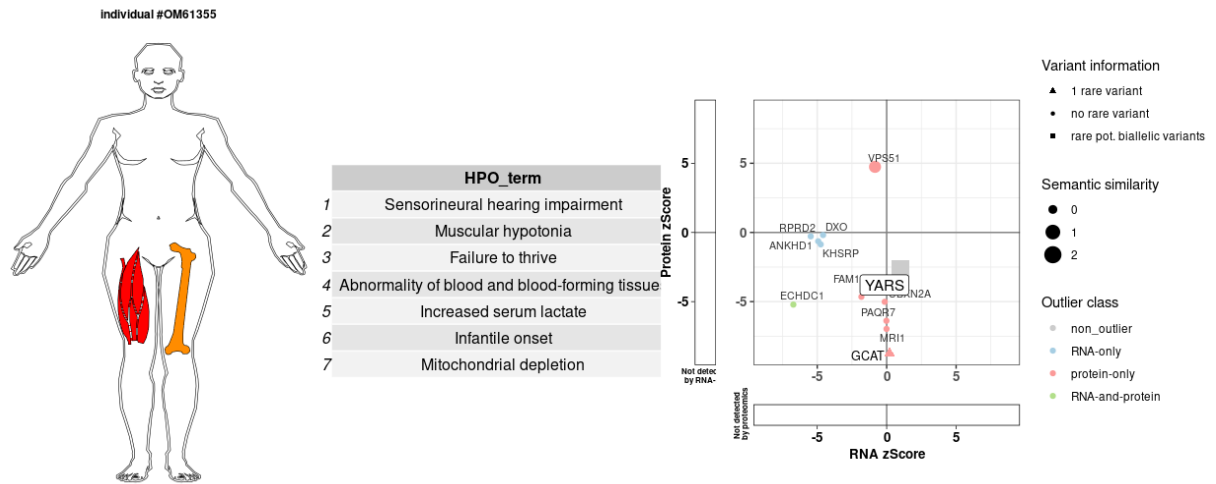
**Case 9, Sample ID: OM23417, Causal gene: *FDXR***



Neonatal male diagnosed with compound heterozygous missense variants (c.[683G>T], p.[Arg228Leu]; c.[1058G>A], p.[Cys353Tyr]) in *FDXR*. Phenotype of the patient matched the published phenotype of the causal gene (Semantic similarity score = 3.53). *FDXR* protein was underexpressed with nominal significance (FC: 0.63; p-value:  $8.31 \times 10^{-3}$ ; z-score: -2.64).

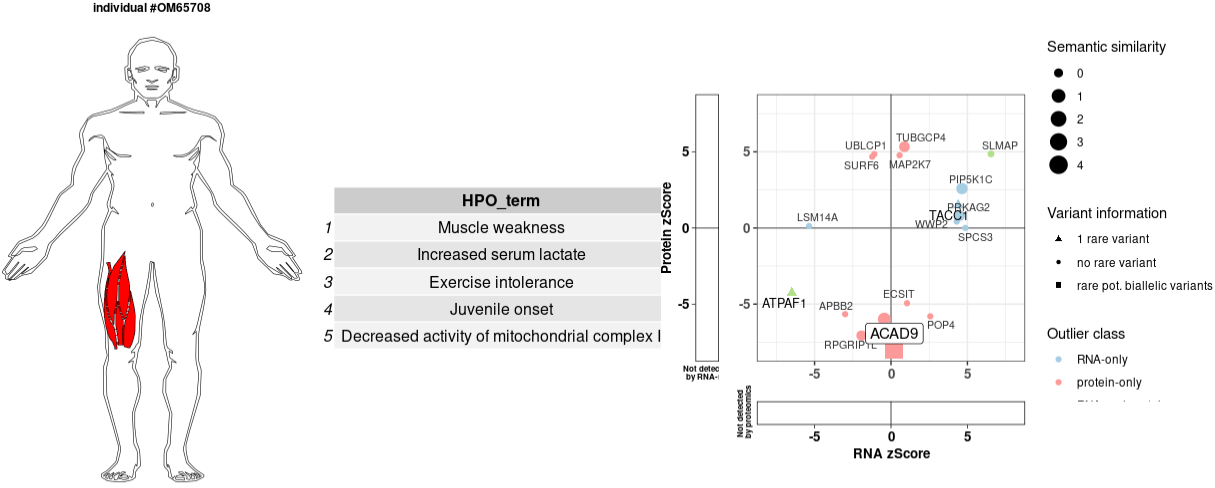


## Case 10, Sample ID: OM61355, Causal gene: *YARS*



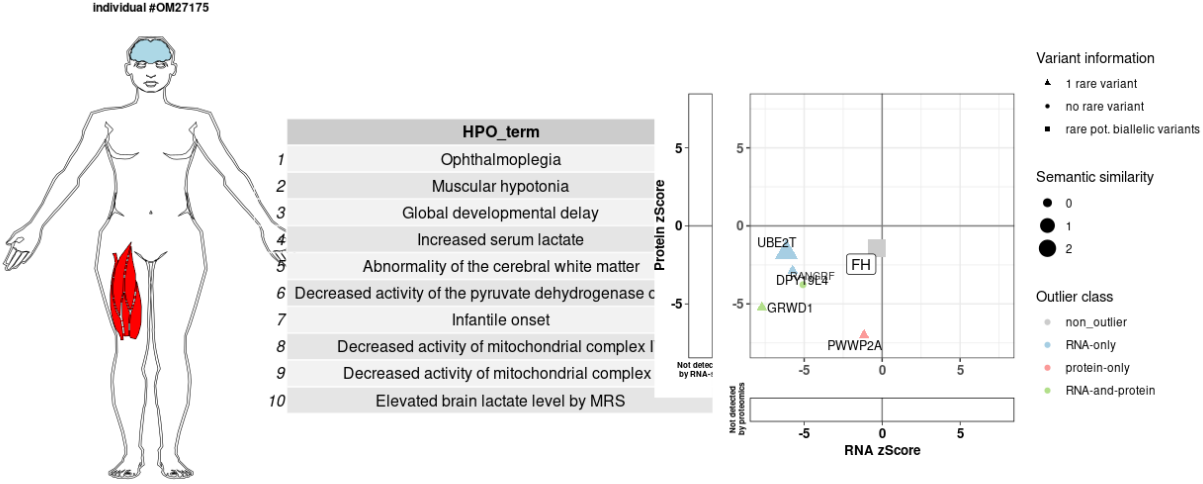
Infant female diagnosed with homozygous missense variant (c.[176T>C], p.[Ile59Thr]) in *YARS*. Phenotype of the patient matched the published phenotype of the causal gene (Semantic similarity score = 2.78). *YARS* protein was underexpressed with nominal significance (FC: 0.81; p-value:  $7.87 \times 10^{-3}$ ; z-score: -2.66).

**Case 11, Sample ID: OM65708, Causal gene: *ACAD9***



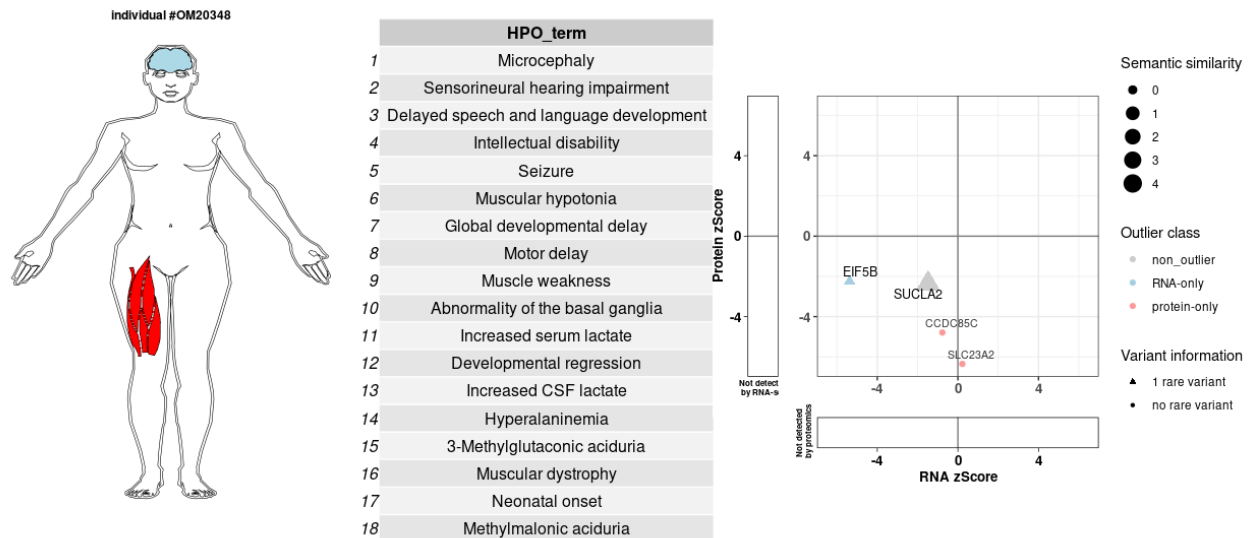
Paediatric male diagnosed compound heterozygous start loss and missense variants (c.[1A>G], p.[?]; c.[261T>G], p.[Ile87Met]) in *ACAD9*. Phenotype of the patient matched the published phenotype of the causal gene (Semantic similarity score = 4.37). *ACAD9* was a significant protein underexpression outlier (FC: 0.29; p-value:  $1.51 \times 10^{-15}$ ; z-score: -7.98). *ECSIT* protein, direct interaction partner of *ACAD9*, also found significantly reduced (FC: 0.37; p-value:  $7.89 \times 10^{-7}$ ; z-score: -4.94).

**Case 12, Sample ID: OM27175, Causal gene: FH**



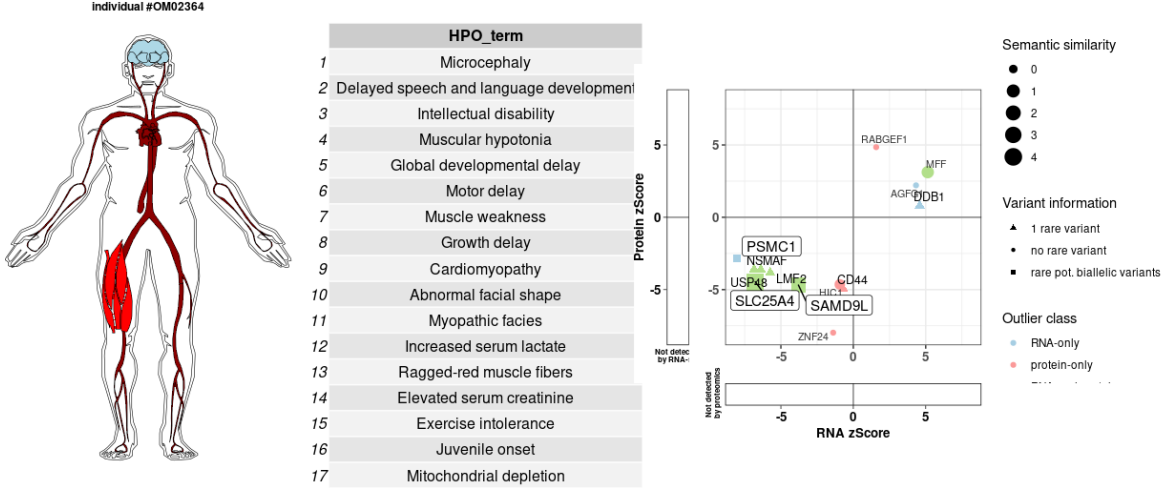
Infant female diagnosed with homozygous missense variant (c.[1048C>T], p.[Arg350Trp]) in *FH*. Phenotype of the patient matched the published phenotype of the causal gene (Semantic similarity score = 2.76). FH protein was underexpressed with nominal significance (FC: 0.88; p-value: 0.15; z-score: -1.44).

**Case 13, Sample ID: OM20348, Causal gene: *SUCLA2***



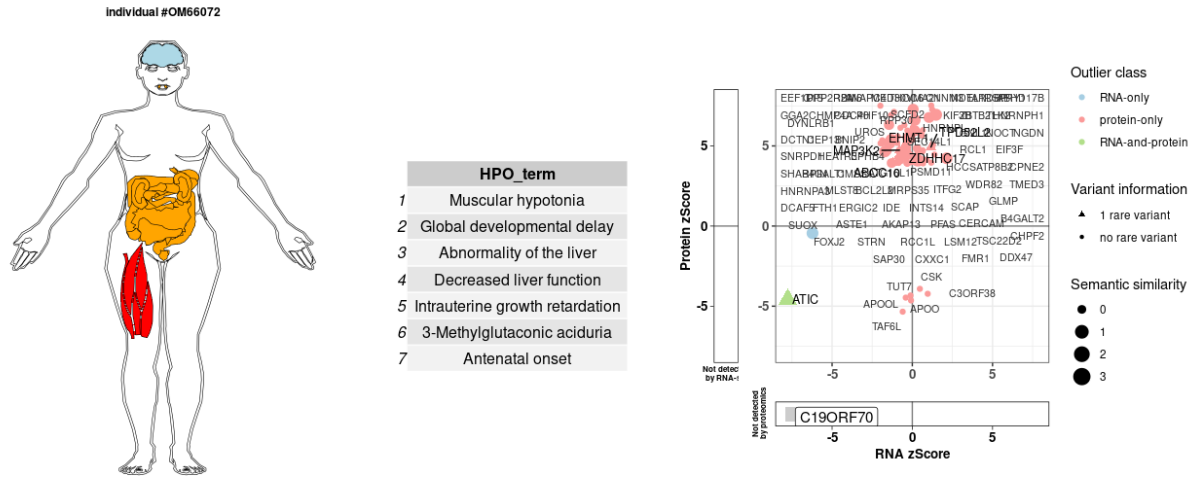
Infant female diagnosed with compound heterozygous missense variants (c.[670C>T], p.[Arg224Cys]; c.[805A>G], p.[Met269Val]) in *SUCLA2*. Phenotype of the patient matched the published phenotype of the causal gene (Semantic similarity score = 4.7). *SUCLA2* protein was underexpressed with nominal significance (FC: 0.81; p-value:  $1.77 \times 10^{-2}$ ; z-score: -2.37). Furthermore a global reduction of Succinyl-CoA synthetase protein complex was observed.

**Case 14, Sample ID: OM02364, Causal gene: SLC25A4**



Paediatric male diagnosed with homozygous missense variant (c.[598G>A], p.[Gly200Arg]) in *SLC25A4*. Phenotype of the patient matched the published phenotype of the causal gene (Semantic similarity score = 4.29). *SLC25A4* was a significant RNA-and-protein underexpression outlier (RNA: FC: 0.24; p-value: 6.39\*10<sup>-9</sup>; z-score: -6.77 / Protein: FC: 0.35; p-value: 8.28\*10<sup>-6</sup>; z-score: -4.46).

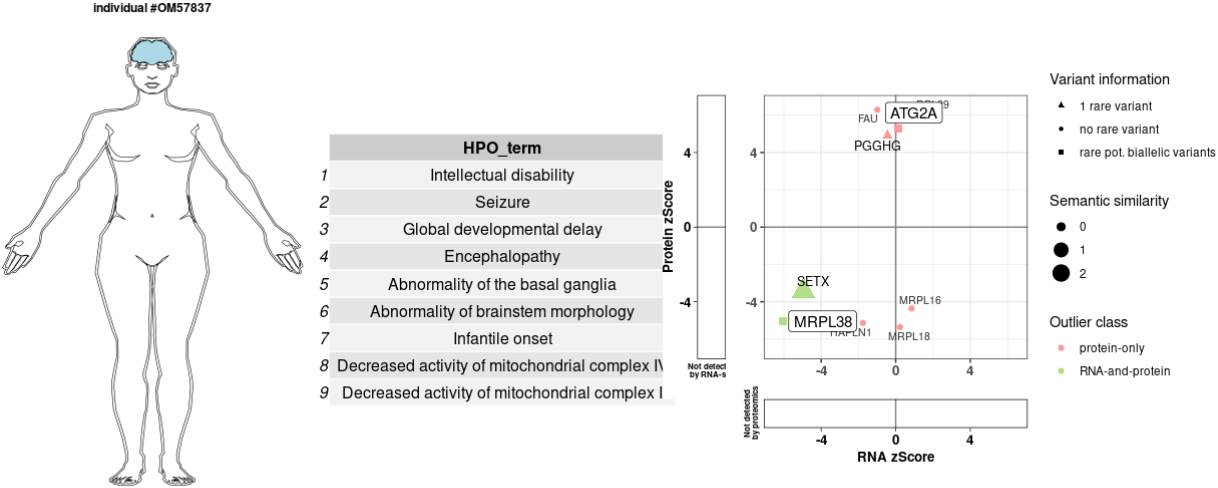
**Case 21 - validated with RNA-seq, Sample ID: OM66072, Causal gene: *C19ORF70***



Paediatric female diagnosed with compound heterozygous frameshift and intronic variants (c.[143delT], p.[Val48Alafs\*42]; c.[29+272G>C], p.[Phe11Trpfs\*5]) in *c19orf70*. Phenotype of the patient matched the published phenotype of the causal gene (Semantic similarity score = 4.49). *c19orf70* transcript was underexpressed with nominal significance and aberrant splicing was detected as well (RNA: FC: 0.59, p-value:  $7.41 \times 10^{-13}$ ; z-score: -7.43). Protein not detected in corresponding TMT-batch. Furthermore a global reduction of the MICOS complex was observed.



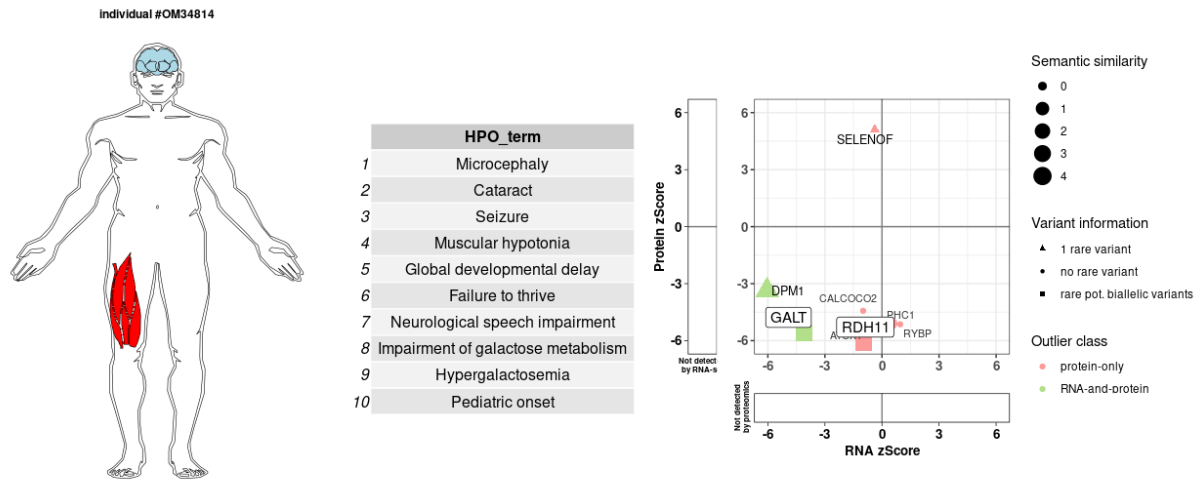
**Case 2, Sample ID: OM57837, Causal gene: MRPL38**



Infant female diagnosed with compound heterozygous 5'UTR-deletion and missense variants (c.[1-174del127], p.[?]; c.[770C>G], p.[Pro257Arg]) in *MRPL38*. *MRPL38* was a significant RNA-and-protein underexpression outlier (RNA: FC: 0.66; p-value: 8.47\*10<sup>-10</sup>; z-score: -6.05 // protein: FC: 0.66; p-value: 8.47\*10<sup>-10</sup>; z-score: -6.05). Missense variant was present in 78% of RNA reads. Furthermore reduction of large mitochondrial ribosomal subunit was observed.

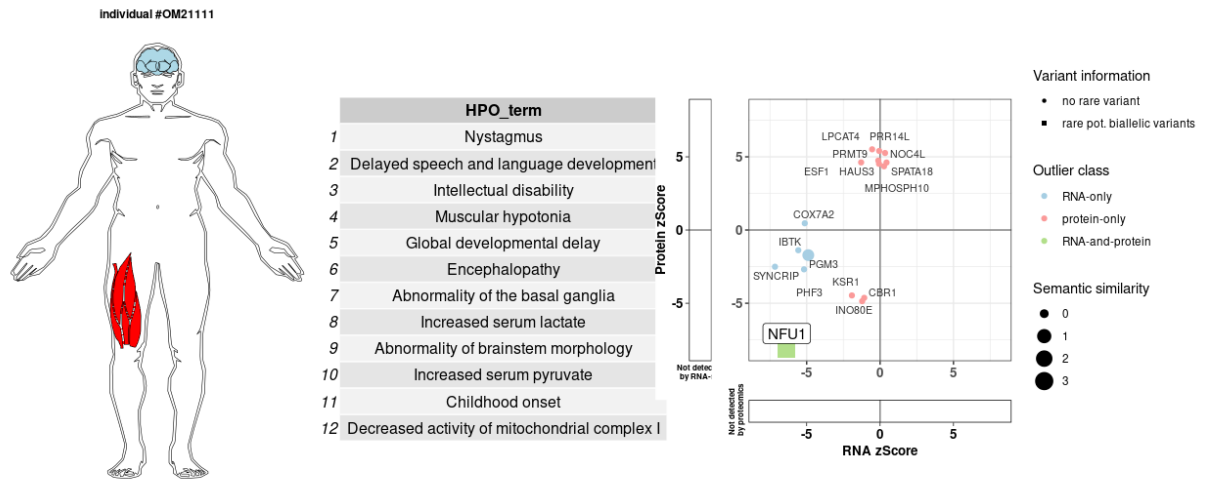


**Case 3, Sample ID: OM34814, Causal gene: *RDH11***



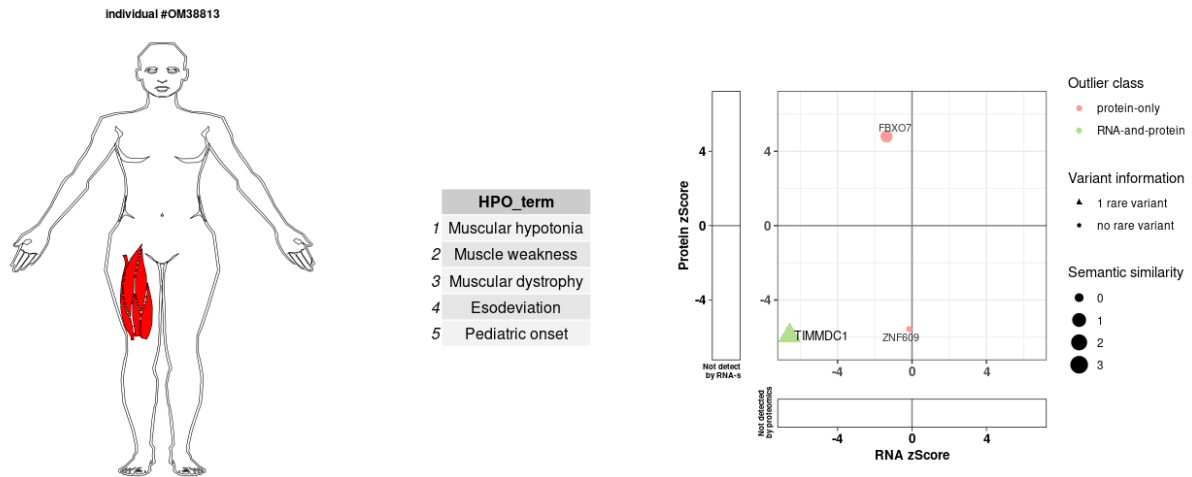
Paediatric male diagnosed with homozygous missense variant (c.[637A>C], p.[Thr213Pro]) in *RDH11*. Phenotype of the patient matched the published phenotype of the causal gene (Semantic similarity score = 2.96). *RDH11* was a significant protein-only underexpression outlier (FC: 0.35; p-value:  $9.89 \times 10^{-10}$ ; z-score: -6.11). Furthermore reduction of *RDH11*-*SELENOF* protein complex was observed.

## Case 4, Sample ID: OM21111, Causal gene: *NFU1*



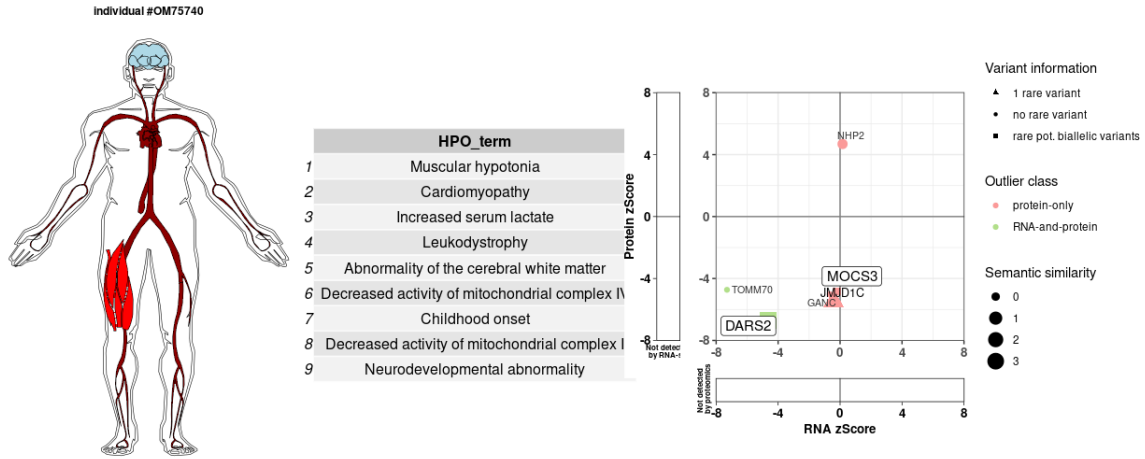
Infant male diagnosed with compound heterozygous missense and CNV variants (c.[362T>C], p.[Val121Ala]; c.[del exon5], p.[?]) in *NFU1*. Phenotype of the patient matched the published phenotype of the causal gene (Semantic similarity score = 3.26). *NFU1* was a significant RNA-and-protein underexpression outlier. Whole exon deletion (exon 5 plus spanning region) resulted in significant RNA underexpression (FC: 0.63 p-value:  $1.18 \times 10^{-9}$ ; z-score: -6.36) plus monoallelic expressed missense variant resulted in protein underexpression (FC: 0.18; p-value:  $4.00 \times 10^{-16}$ ; z-score: -8.14).

**Case 5, Sample ID: OM38813, Causal gene: *TIMMDC1***



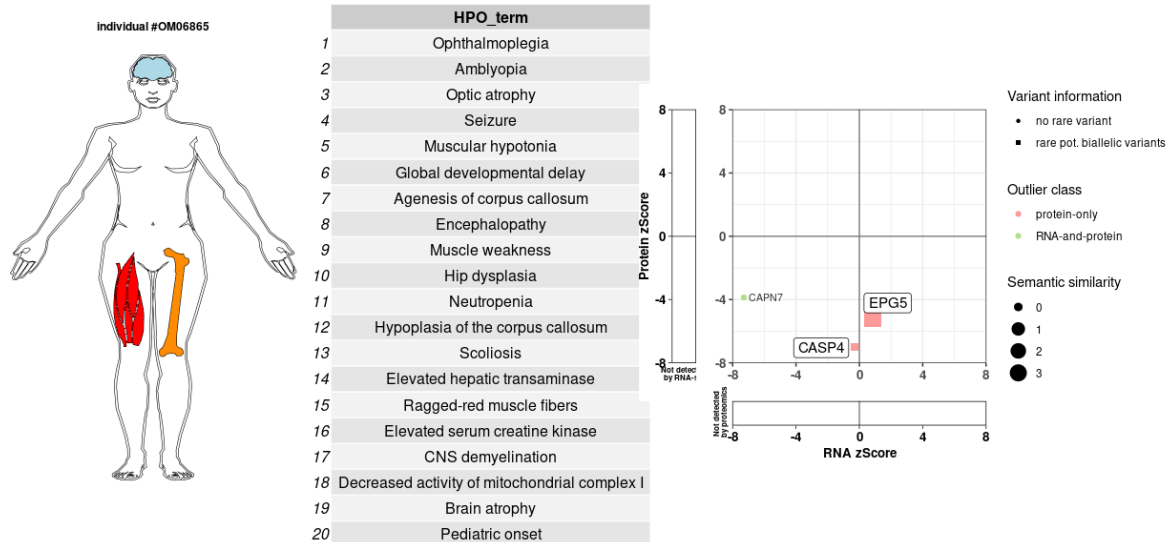
Paediatric male diagnosed with homozygous deep intronic variant (c.[597-1340A>G], p.[Gly199ins5\*]) in *TIMMDC1*. Phenotype of the patient matched the published phenotype of the causal gene (Semantic similarity score = 3.53). Homozygous deep intronic variant caused splice defect (cryptic exon), resulting in frameshift and RNA-and-protein underexpression outlier (RNA: FC: 0.24; p-value:  $3.66 \times 10^{-11}$ ; z-score: -6.58 / Protein: FC: 0.27; p-value:  $3.07 \times 10^{-9}$ ; z-score: -5.93).

**Case 6, Sample ID: OM75740, Causal gene: *DARS2***



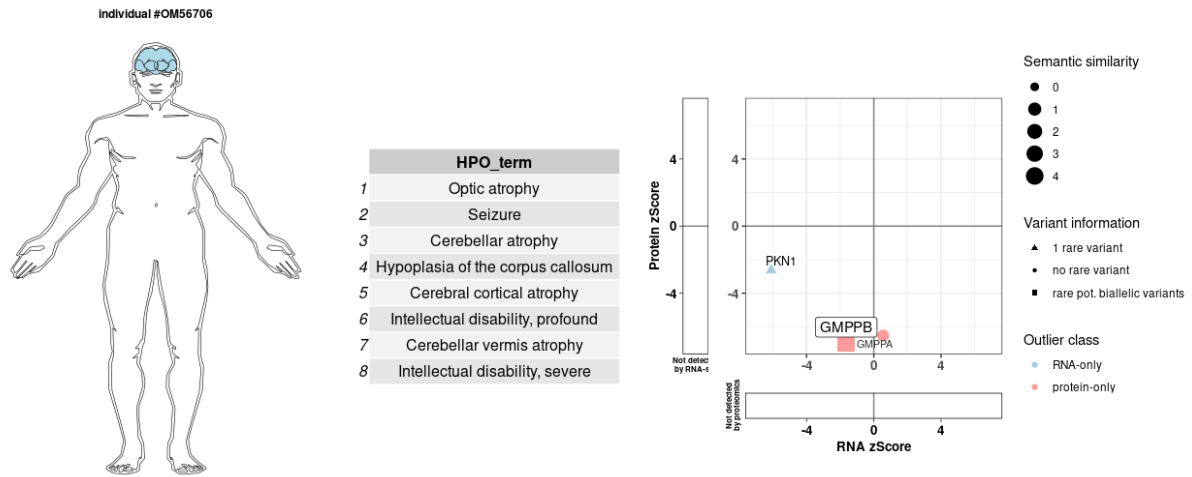
Infant male diagnosed with compound heterozygous nearsplice (1 rare + 1 frequent) and splice (rare) variants (c.[228-20T>C; 228-12C>G], p.[?]; c.[228-20T>C; 492+2T>C], p.[?]) in *DARS2*. Phenotype of the patient matched the published phenotype of the causal gene (Semantic similarity score = 4.03). *DARS2* was a significant RNA-and-protein underexpression outlier (RNA: FC: 0.71; p-value:  $5.95 \times 10^{-5}$ ; z-score: -4.66 / Protein: FC: 0.37; p-value:  $1.85 \times 10^{-11}$ ; z-score: -6.72) due to partial skipping of exons 3 and 5.

**Case 7, Sample ID: OM06865, Causal gene: *EPG5***



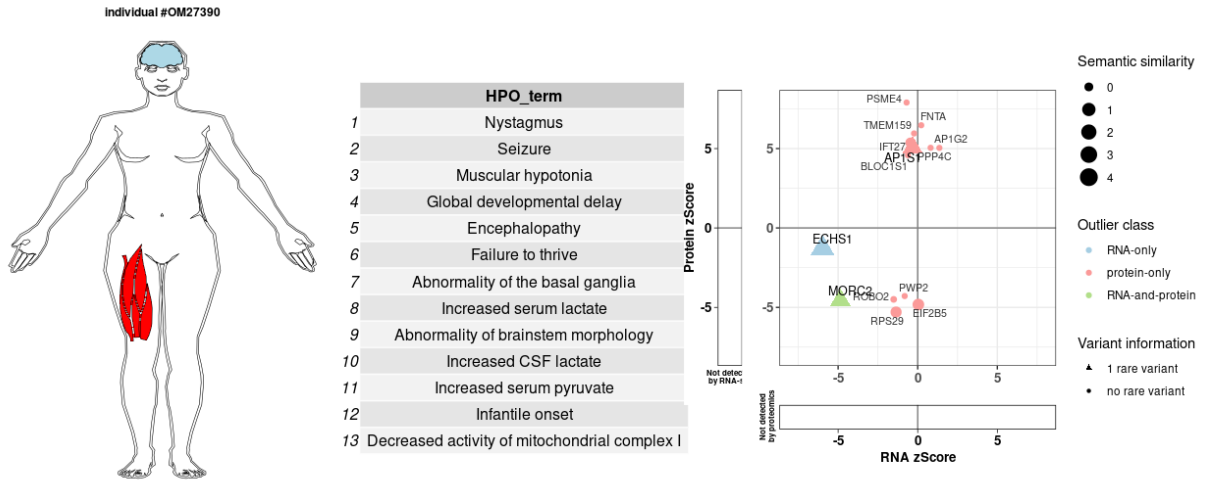
Paediatric female diagnosed with homozygous missense variant (c.[5548T>C], p.[Cys1850Arg]) in *EPG5*. Phenotype of the patient matched the published phenotype of the causal gene (Semantic similarity score = 3.73). *EPG5* was a significant protein-only underexpression outlier (FC: 0.42; p-value:  $2.13 \times 10^{-7}$ ; z-score: -5.19).

**Case 8, Sample ID: OM56706, Causal gene: *GMPPB***



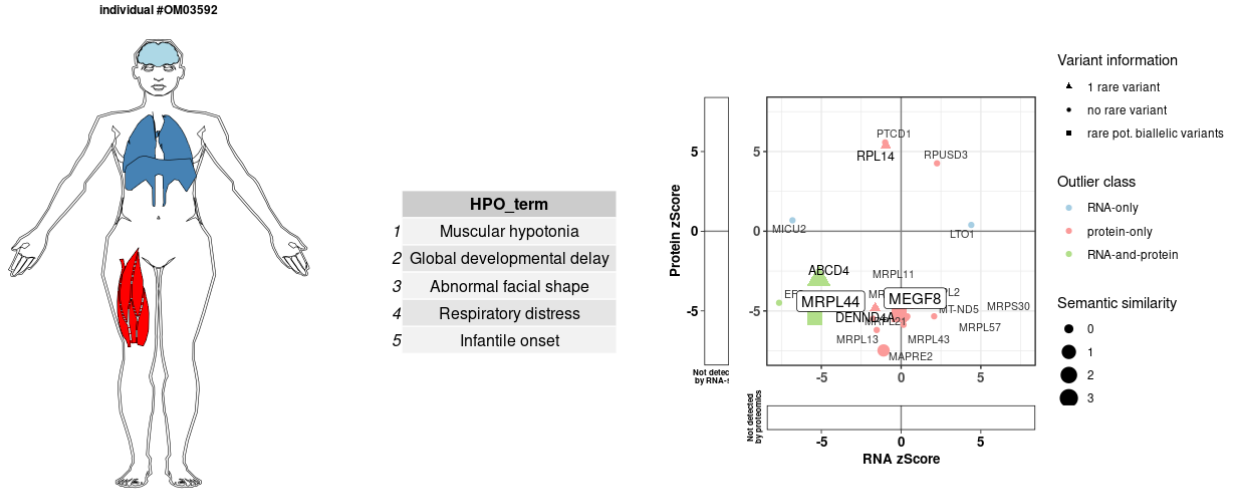
Paediatric male diagnosed with homozygous 5 bp duplication (c.[1016insAAGGC], p.[Tyr340Argfs\*13]) in *GMPPB*. Phenotype of the patient matched the published phenotype of the causal gene (Semantic similarity score = 4.7). *GMPPB* was a significant protein-only underexpression outlier (Protein: FC: 0.40; p-value:  $3.74 \times 10^{-12}$ ; z-score: -6.95). Indel is intronic in the muscle-specific transcript isoform, resulting in a mild splice defect and exonic in brain-specific isoform leading to a frameshift in keeping with the patient presenting with neurological symptoms. Furthermore *GMPPA* protein, direct interaction partner of *GMPPB*, was also found as a significant protein-only underexpression outlier (Protein: FC: 0.48; p-value:  $6.97 \times 10^{-11}$ ; z-score: -6.52).

**Case 9, Sample ID: OM27390, Causal gene: *MORC2***



Infant female diagnosed with de-novo heterozygous missense variant (c.[77C>T], p.[Ala26Val]) in *MORC2*. Phenotype of the patient matched the published phenotype of the causal gene (Semantic similarity score = 2.91). *MORC2* was a significant RNA-and-protein underexpression outlier (RNA: FC: 0.73; p-value:  $3.90 \times 10^{-6}$ ; z-score: -4.86 / Protein: FC: 0.69; p-value:  $5.04 \times 10^{-6}$ ; z-score: -4.56).

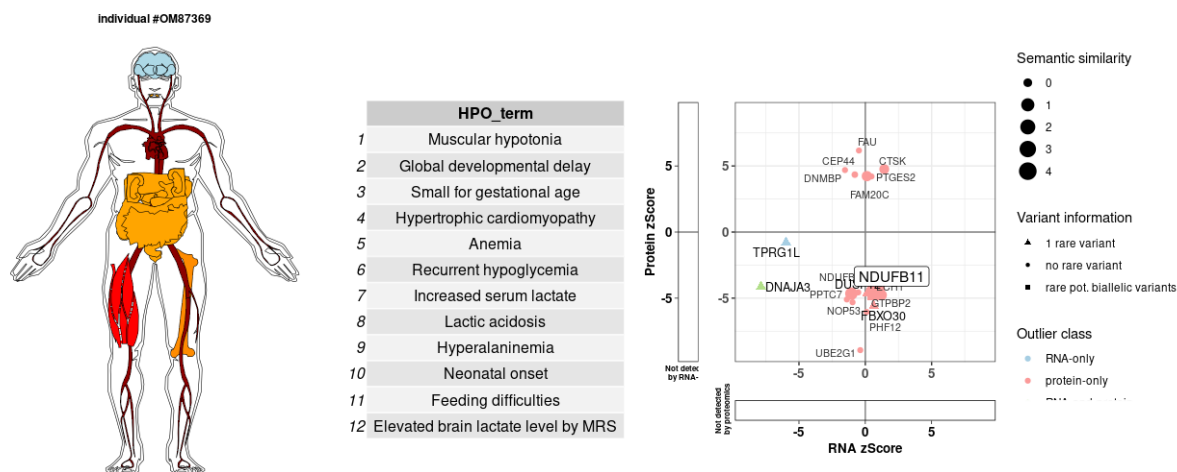
**Case 10, Sample ID: OM03592, Causal gene: *MRPL44***



Infant female diagnosed with homozygous nearsplice variant (c.[179+3A>G], p.[?]) in *MRPL44*. *MRPL44* was a significant RNA-and-protein underexpression outlier (RNA: FC: 0.74; p-value:  $6.68 \times 10^{-8}$ ; z-score: -5.42 / Protein: FC: 0.50; p-value:  $6.44 \times 10^{-8}$ ; z-score: -5.41). Splice defect resulting in several aberrant species: truncation and elongation of exon 1 and retention of intron 1. Furthermore reduction of large mitochondrial ribosomal subunit was observed. Although phenotypic presentation was atypical (Semantic similarity score = 1.54), sufficient evidence of pathogenicity allowed to provide the diagnosis and extend phenotypic spectra of the associated disorder.

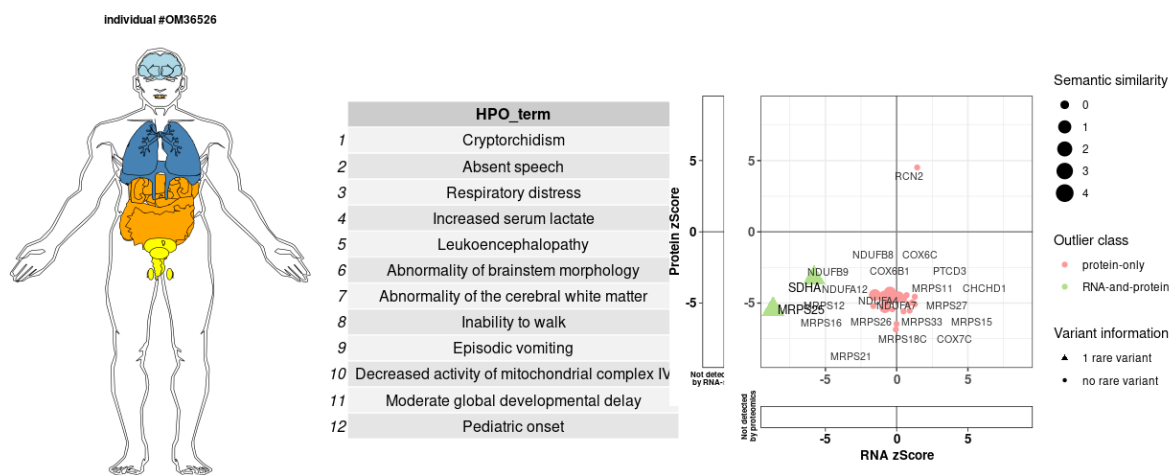


**Case 11, Sample ID: OM87369, Causal gene: *NDUFB11***



Infant male diagnosed with hemizygous missense variant (c.[440T>C], p.[Met147Thr]) in *NDUFB11*. Phenotype of the patient matched the published phenotype of the causal gene (Semantic similarity score = 3.96). *NDUFB11* was a significant protein-only underexpression outlier (Protein: FC: 0.46; p-value:  $6.12 \times 10^{-6}$ ; z-score: -4.52). One further significantly reduced protein outlier *NDUFB10* (Protein: FC: 0.45; p-value:  $3.20 \times 10^{-6}$ ; z-score: -4.66) without potentially biallelic variants as well as a global reduction of respiratory chain complex I was identified.

## Case 12, Sample ID: OM36526, Causal gene: *MRPS25*



Paediatric male diagnosed with homozygous variant (c.[329+75G>A], p.[Glu111\*]) in *MRPS25*. Phenotype of the patient matched the published phenotype of the causal gene (Semantic similarity score = 4.17). Homozygous intronic variants caused exon extension and premature termination, resulting in significant RNA-and-protein underexpression outlier (RNA: FC: 0.43; p-value:  $7.68 \times 10^{-16}$ ; z-score: -8.68 / Protein: FC: 0.42; p-value:  $5.66 \times 10^{-8}$ ; z-score: -5.43). Furthermore reduction of small mitochondrial ribosomal subunit, as well as respiratory chain complexes, was observed.

### **Supplementary References**

Bonora, E., Chakrabarty, S., Kellaris, G., Tsutsumi, M., Bianco, F., Bergamini, C., ... & De Giorgio, R. (2021). Biallelic variants in LIG3 cause a novel mitochondrial neurogastrointestinal encephalomyopathy. *Brain*.

Frésard, L., Smail, C., Ferraro, N. M., Teran, N. A., Li, X., Smith, K. S., ... & Balliu, B. (2019). Identification of rare-disease genes using blood transcriptome sequencing and large control cohorts. *Nature medicine*, 25(6), 911-919.

Jiang, L., Wang, M., Lin, S., Jian, R., Li, X., Chan, J., ... & Doherty, J. A. (2020). A quantitative proteome map of the human body. *Cell*, 183(1), 269-283.