

Supplementary Data file 3

Plots of 29 ALS genes showing all reported ALS-implicated variants including REVEL and CADD scores. Visual summary of ALS-implicated variants identified within each ALS gene between 1993 and March 2021. Central horizontal bars represent the cDNA sequence of the gene, with variants plotted along the x-axis according to their cDNA position. REVEL (missense variants) and CADD phred (all variants) scores are denoted by the height of points on the y-axis, above and below the cDNA bar respectively. Genomic context is indicated by point shape, as exonic (circle) and intronic (diamond), with the latter plotted at the exon bounds as indicated by vertical dashed lines. The 3' untranslated region (UTR) is specifically plotted for *FUS* due to the large number of variants present in this region. A density plot of REVEL scores for missense variants reported in FALS (red) versus SALS exclusively (blue) is provided, with horizontal dashed lines indicating group median.

Visual summary of ALS-implicated variants identified within each ALS gene between 1993 and March 2021. Panels A and B) The horizontal bars on the x-axis represent the cDNA sequence of the gene, with variants plotted along the length of the gene according to their cDNA position. CADD phred (A, all variants) and REVEL (B, missense variants) scores are denoted by the height of points on the y-axis. Genomic context is indicated by point shape, as exonic (circle) and intronic (diamond), with the latter plotted at the exon bounds as indicated by vertical dashed lines. The 3' untranslated region (UTR) is specifically plotted for *FUS* due to the large number of variants that lie in this region. Panels C and D) Density plots of CADD phred (C, all variants) and REVEL (D, missense variants) reported in FALS (red) versus SALS exclusively (blue). Bolded vertical lines indicate group median score, while grey dashed lines denote standard threshold bins.

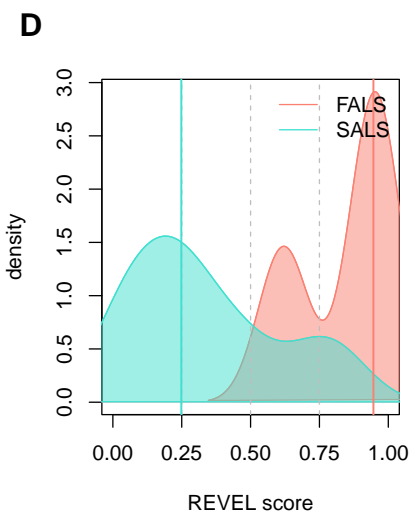
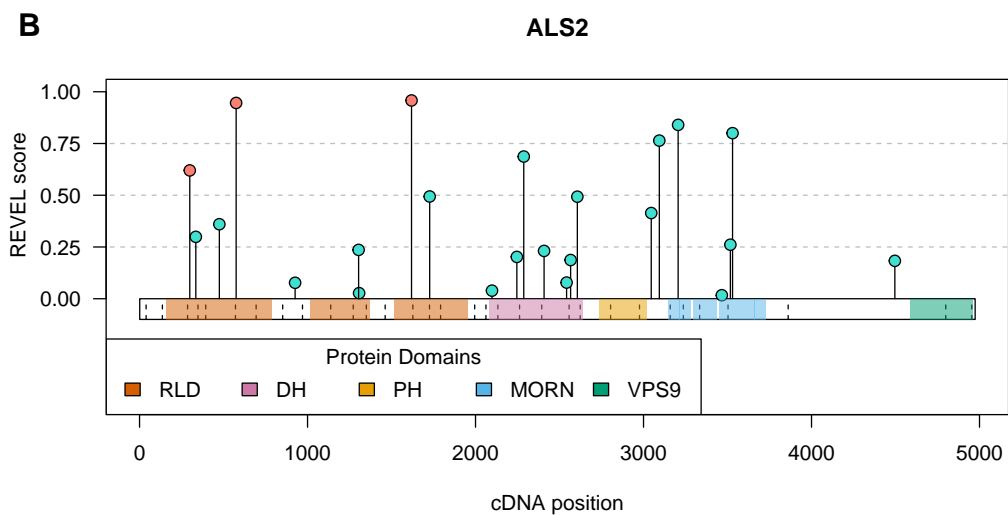
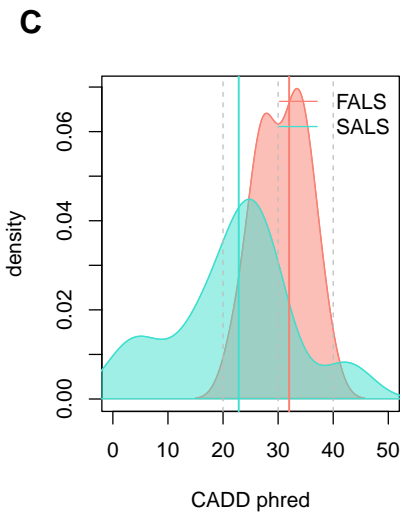
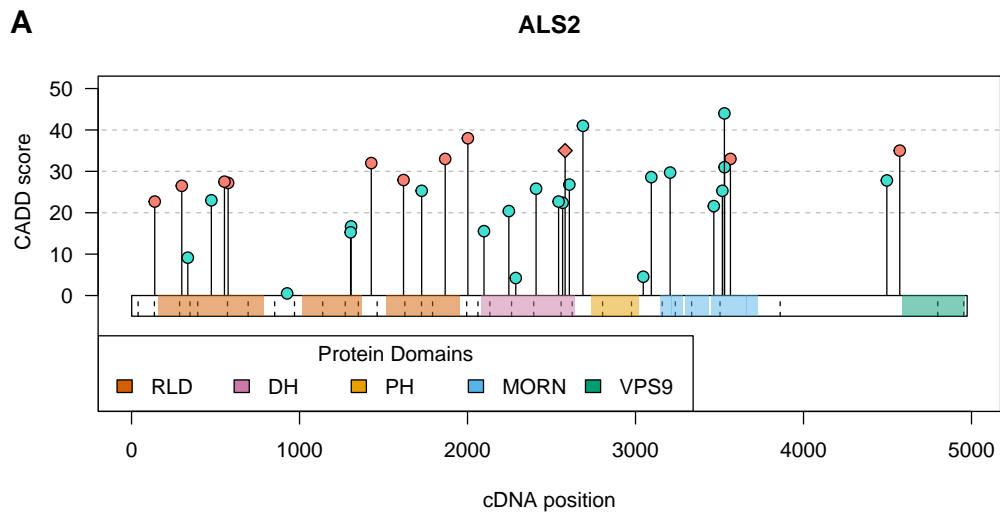
<i>ALS2</i>	4
<i>ANG</i>	5
<i>C21orf2</i>	6
<i>CCNF</i>	7
<i>CHCHD10</i>	8
<i>CHMP2B</i>	9
<i>CYLD</i>	10
<i>DCTN1</i>	11
<i>ELP3</i>	12
<i>FUS</i>	13
<i>HNRNPA1</i>	14
<i>HNRNPA2B1</i>	15
<i>KIF5A</i>	16
<i>MATR3</i>	17
<i>NEFH</i>	18
<i>NEK1</i>	19
<i>OPTN</i>	20
<i>PFN1</i>	21
<i>SETX</i>	22
<i>SOD1</i>	23
<i>SPG11</i>	24
<i>SQSTM1</i>	25
<i>TARDBP</i>	26
<i>TBK1</i>	27
<i>TIA1</i>	28

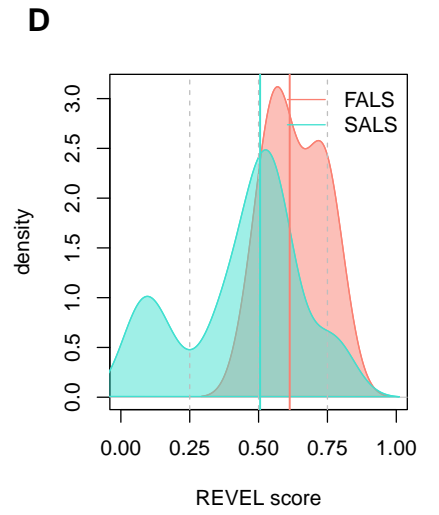
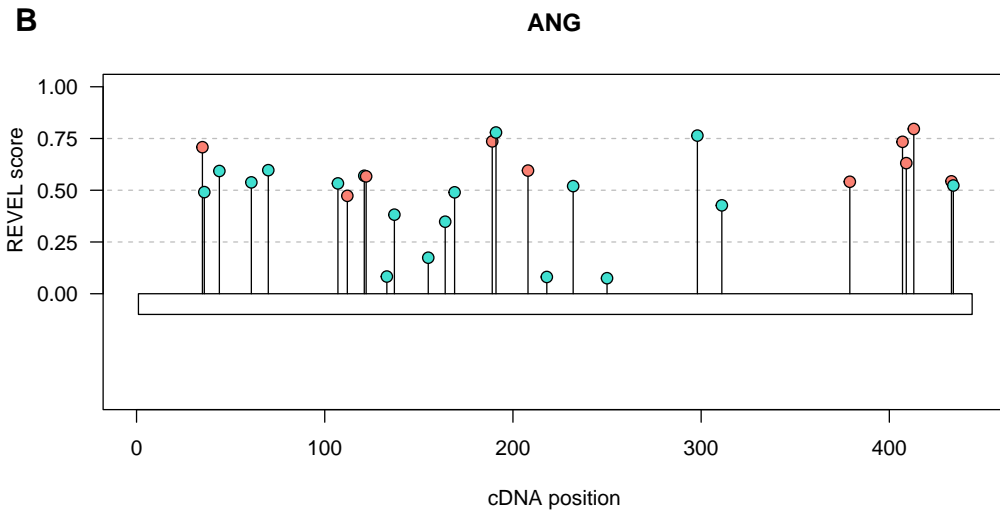
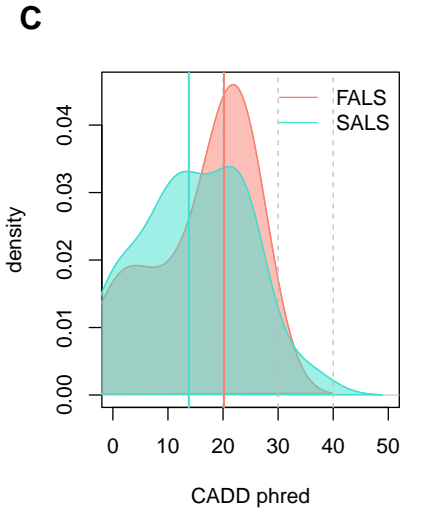
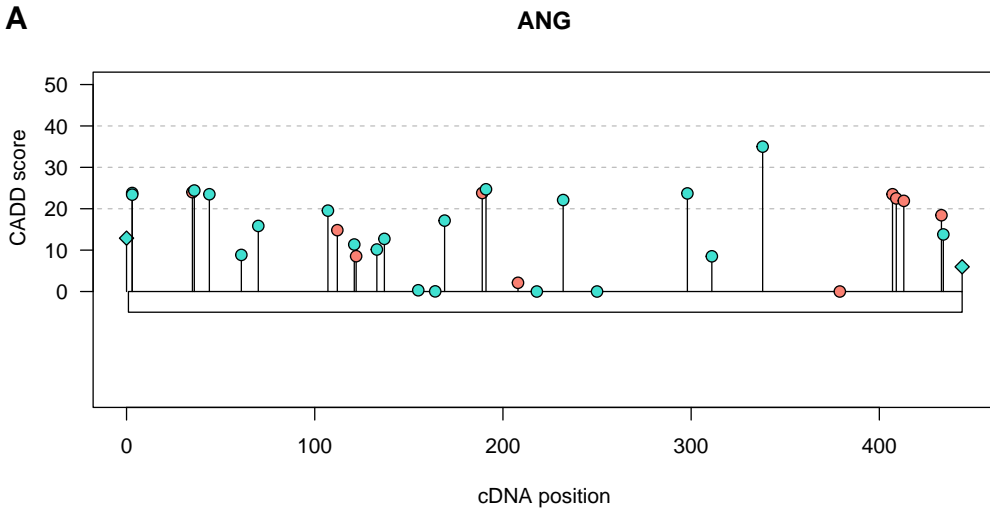
<i>TUBA4A</i>	29
<i>UBQLN2</i>	30
<i>VAPB</i>	31
<i>VCP</i>	32

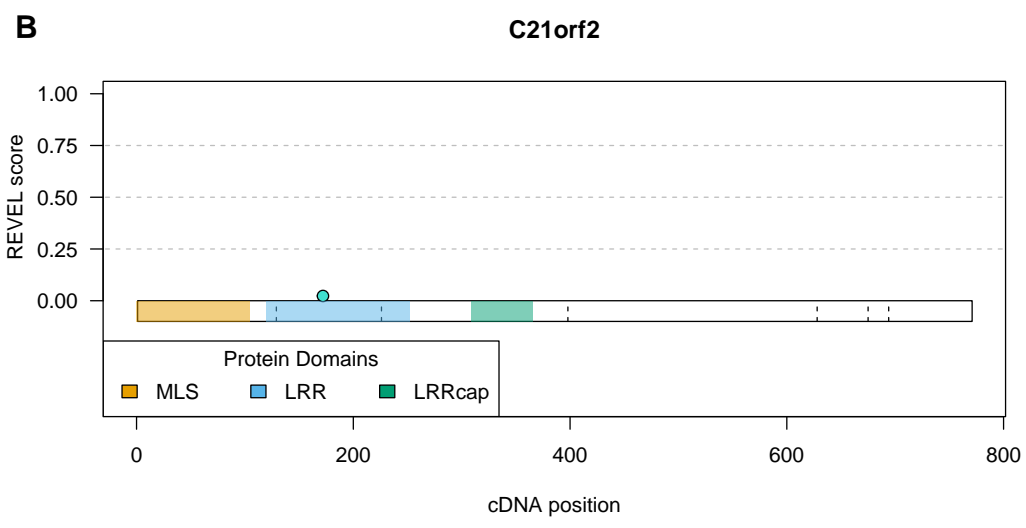
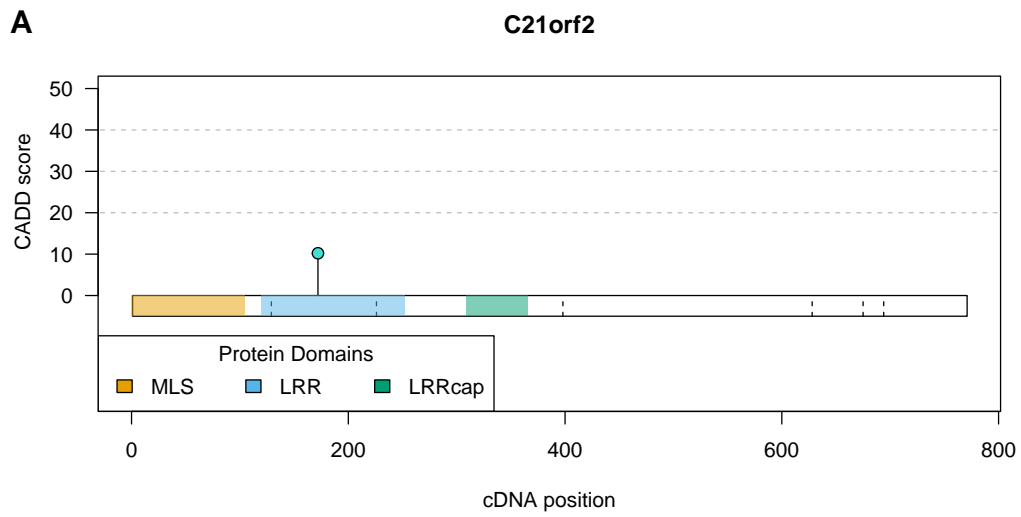
Protein domain abbreviations:

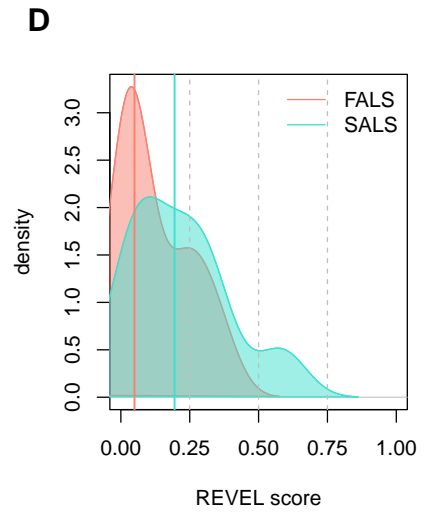
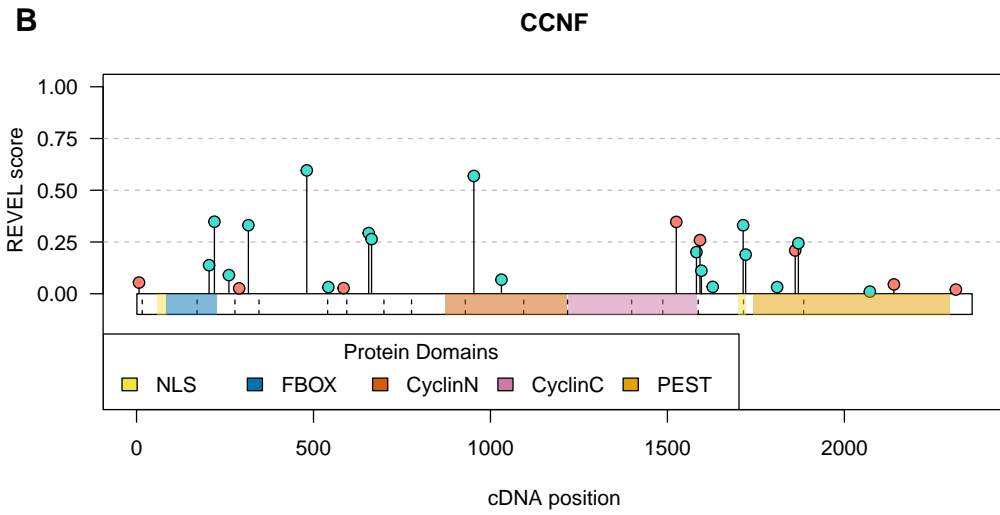
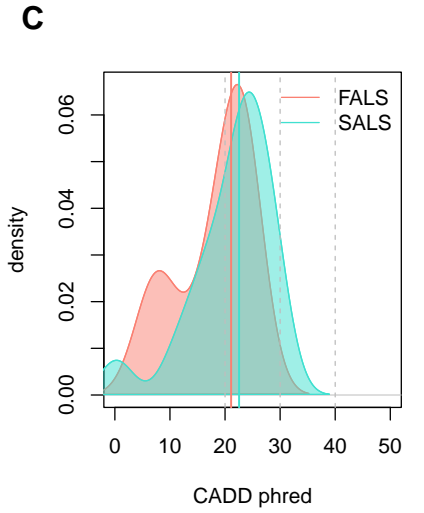
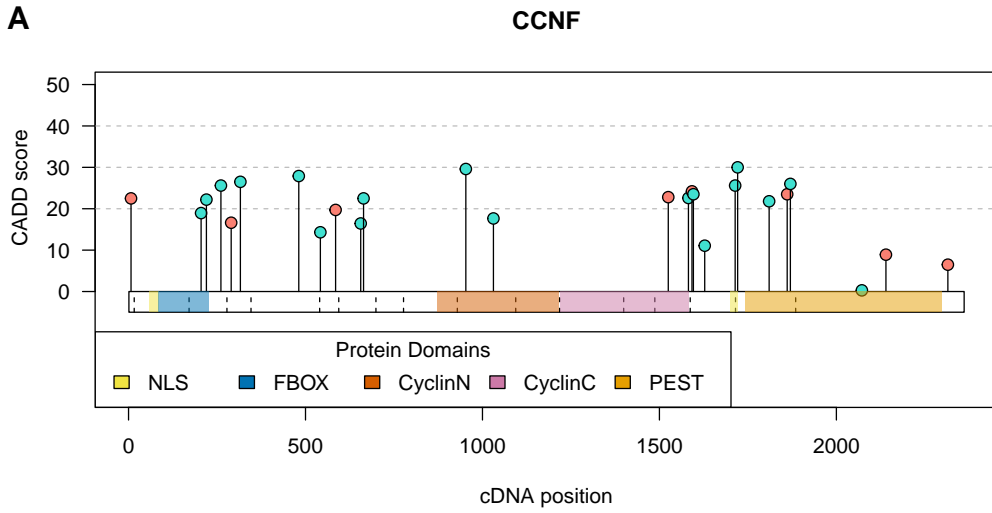
AP-CTD, Aggregation-prone C-terminal domain
ATPase D1, ATPase (oligomerisation) domain 1
ATPase D2, ATPase (activity) domain 2
Basic, Basic domain
CAP-Gly, Cytoskeleton-associated protein (CAP), glycine-rich domain
Catalytic domain, Catalytic domain
CCD-LIR, Coiled-coil and LC3-interacting region domain
CCD, Coiled-coil domain
CHCH, Coiled coil-helix-coiled coil-helix domain
CTD, C-terminal domain
CyclinC, Cyclin C domain
CyclinN, Cyclin N domain
DH, Diffuse B cell lymphoma (dbl) homology domain
Dynactin, Dynactin domain
FBOX, F-box domain
GHF1, Glycosyl hydroxylase F1 signature
GNAT, Gcn5-related N-acetyltransferase domain
GRD, Glycine-rich domain
GTPase, GTPase domain
HC, Helicase core
Head, Head domain
Hydrophobic, Hydrophobic domain
KSP, KSP (lysine-serine-proline) motif
LIR, LC3-interacting region
LRR, Leucine-rich repeat
LRRcap, Leucine-rich repeat cap
LZ, Leucine zipper domain
M9 NLS, M9 nuclear localisation signal
MIM, Microtubule interacting and transport (MIT)-interacting motif
MLS, Mitochondrial localisation signal
MORN, MORN domain
Motor, Globular motor domain
MSP, Major sperm protein domain
Myb, Myb domain
NA, Ubiquitin-related domains
NA, Zinc finger domain
NES, Nuclear export signal
NLS, Nuclear localisation signal
NTD, N-terminal domain
PB1, PB1 domain
PBD, Polyubiquitin binding domain
PEST, PEST (proline-glutamine-serine-threonine)-rich motif
PH, Pleckstrin homology domain
PK, Protein kinase domain

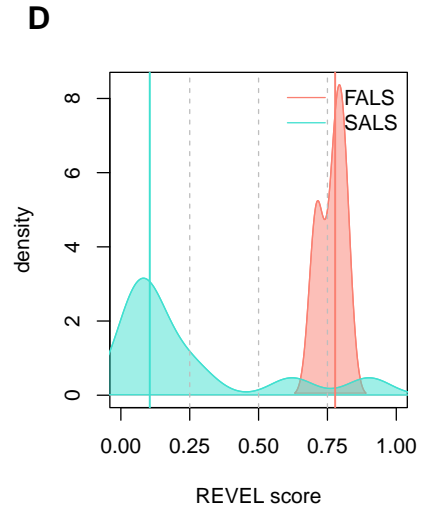
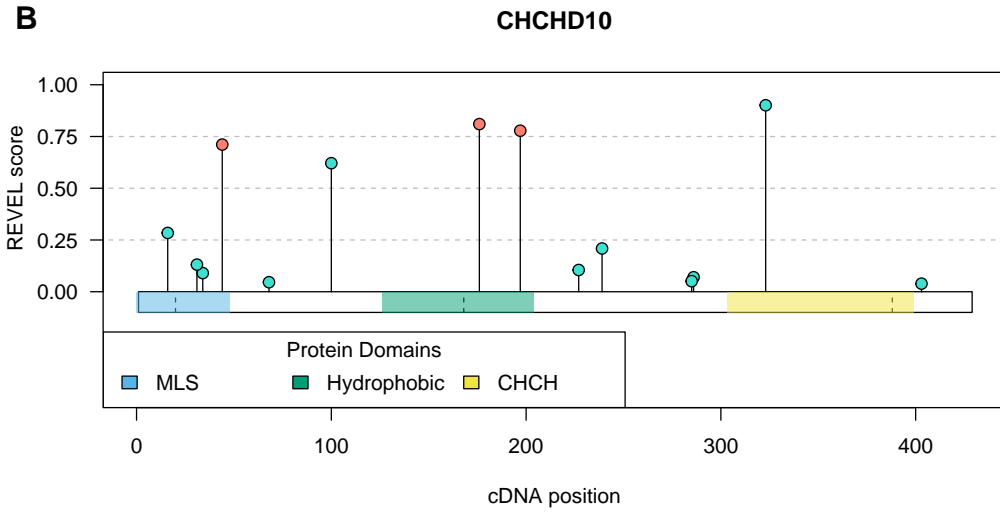
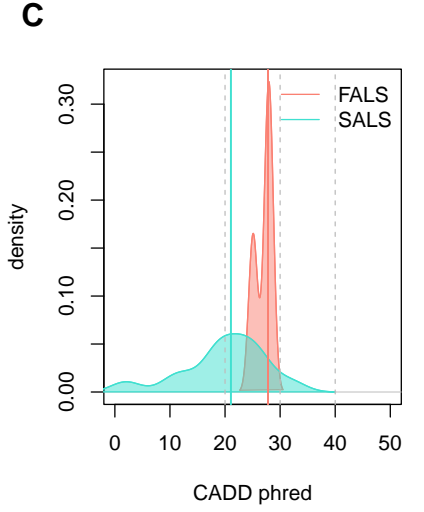
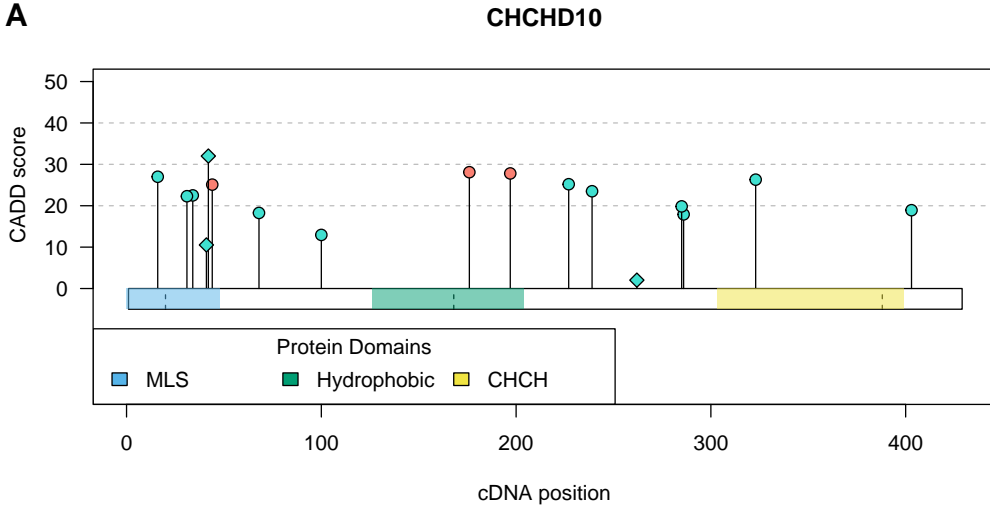
PLD, Prion-like domain
PXX, PXX repeats
QGSY, QGSY (glutamine-glycine-serine-threonine)-rich region
RGG, RGG (arginine-glycine-glycine)-rich motif
RLD, Regulator of chromosome condensation 1 (RCC1) like domain
Rod, Rod domain
RRM, RNA recognition motif
S/T PK, S/T (Serine/Threonine) protein kinase domain
SAM, Radical S-adenosyl-L-methionine (SAM) domain
Stalk, Alpha-helical coiled-coil stalk domain
STI1, Heat shock chaperonin-binding motif
Tail, Globular tail domain
Tail, Tail domain
TMD, Transmembrane domain
TRAF6, TNF receptor-associated factor 6 binding site
UBA, Ubiquitin-associated domain
UBD, Ubiquitin-binding domain
ULD, Ubiquitin-like domain
USP, Ubiquitin specific protease catalytic domain
VPS9, VPS9 domain
ZnF (C2H2), Zinc finger (C2H2-type) domain
ZnF (CCHC NOA), Zinc finger (CCHC NOA-type) domain
ZnF (ZZ), Zinc finger (ZZ-type) domain
ZnF, Zinc finger domain (unspecified)

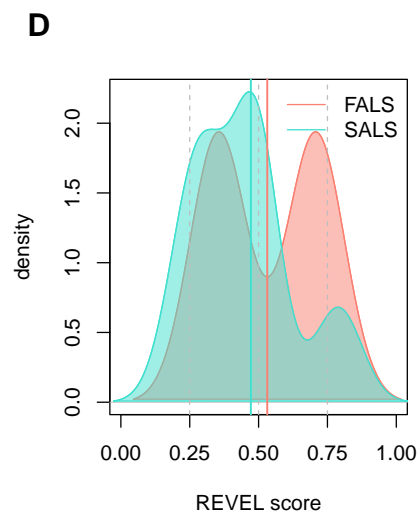
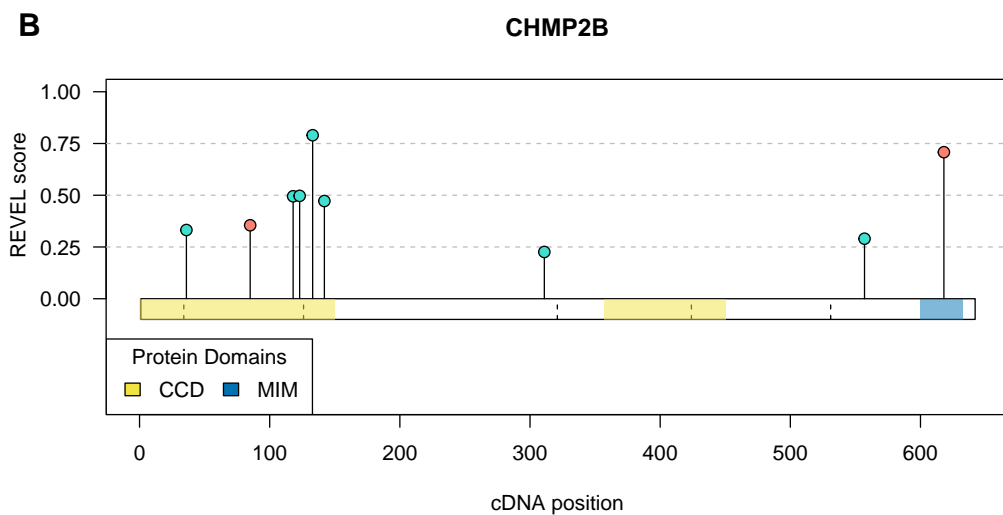
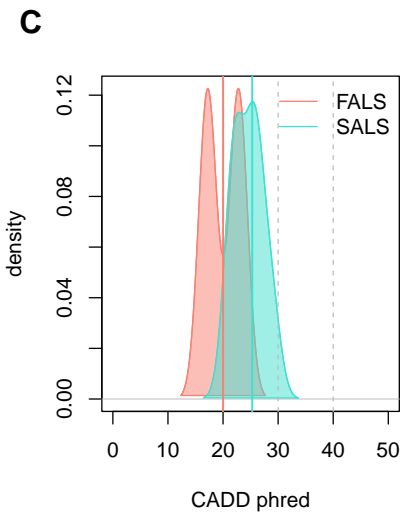
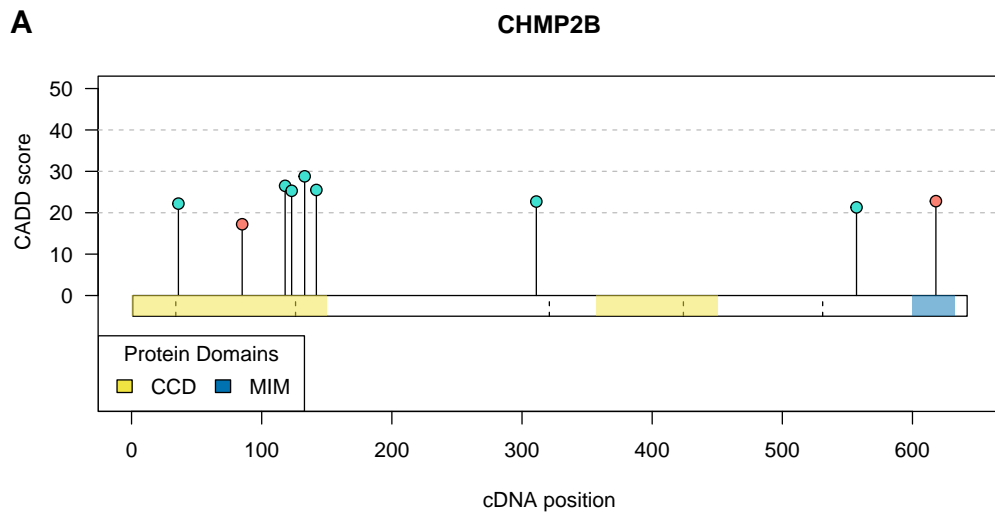


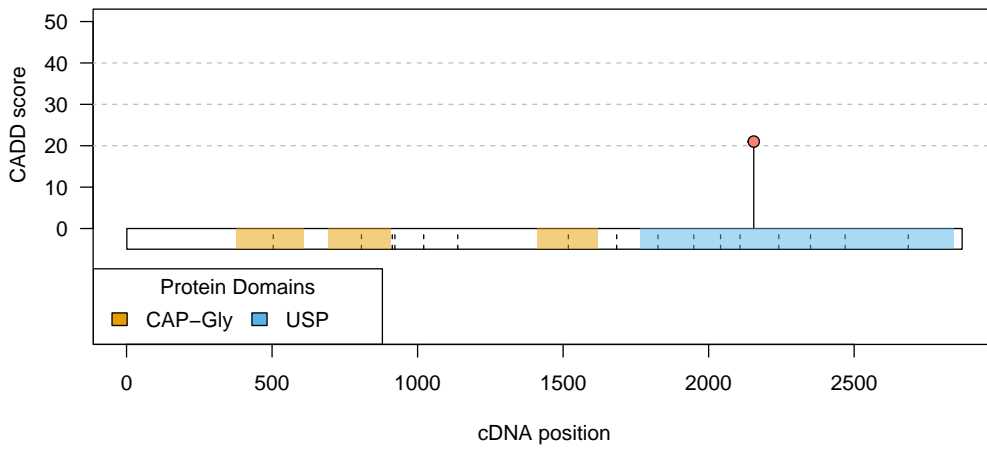










A**CYLD****B****CYLD**