Mutations in disordered regions can cause disease by creating dileucine motifs

Meyer K., Kirchner M., Uyar B., Cheng J.Y., Russo G., Hernandez-Miranda L.R., Szymborska A., Zauber H., Rudolph I.M., Willnow T.E., Akalin A., Haucke V., Gerhardt H. Birchmeier C., Kühn R., Krauss M., Diecke S., Pascual J.M., Selbach M.

This is the final version of the accepted manuscript. The original article has been published in final edited form in:

Cell
2018 SEP 20 ; 175(1): 239-253
2018 SEP 06 (first published online: final publication)
doi: 10.1016/j.cell.2018.08.019

Publisher: Cell Press / Elsevier

Mutations in disordered regions can cause disease by creating dileucine motifs

Katrina Meyer[1], Marieluise Kirchner[1], Bora Uyar[1], Jing-Yuan Cheng[1], Giulia Russo[2], Luis R. Hernandez-Miranda[1], Anna Szymborska[1], Henrik Zauber[1], Ina-Maria Rudolph[1], Thomas E. Willnow[1], Altuna Akalin[1], Volker Haucke[2], Holger Gerhardt[1,3,4], Carmen Birchmeier[1], Ralf Kühn[1,4], Michael Krauss[2], Sebastian Diecke[1,4], Juan M. Pascual[5], Matthias Selbach[1,6]

[1] Max Delbrück Center for Molecular Medicine in the Helmholtz Association (MDC), Robert-Rössle-Str. 10, D-13092 Berlin, Germany

[2] Molecular Pharmacology and Cell Biology, Leibniz-Forschungsinstitut für Molekulare Pharmakologie (FMP), Robert-Rössle-Str. 10, D-13092 Berlin, Germany

[3] DZHK (German Centre for Cardiovascular Research), partner site Berlin

[4] Berlin Institute of Health (BIH), 10178 Berlin, Germany

[5] Department of Neurology and Neurotherapeutics, UT Southwestern Medical Center, 5323 Harry Hines Blvd. Dallas, Texas 75390, USA

[6] Charité-Universitätsmedizin Berlin, 10117 Berlin, Germany

Corresponding author:
Matthias Selbach
Max Delbrück Center for Molecular Medicine
Robert-Rössle-Str. 10, D-13092 Berlin, Germany
Tel.: +49 30 9406 3574
Fax.: +49 30 9406 2394
email: matthias.selbach@mdc-berlin.de

Word count: 4,216 (main text)

Abstract

Many disease-causing missense mutations affect intrinsically disordered regions (IDRs) of proteins. Since these mutations do not affect protein structure, the molecular mechanism of their pathogenicity is enigmatic. Here, we employ a peptide-based proteomic screen to investigate the impact of mutations in IDRs on protein-protein interactions. We find that mutations in disordered cytosolic regions of three transmembrane proteins (GLUT1, ITPR1 and CACNA1H) lead to an increased binding of clathrin. In all three cases, the mutation creates a dileucine motif known to mediate clathrin-dependent trafficking. Follow-up experiments on GLUT1 (SLC2A1), the glucose transporter causative of GLUT1 deficiency syndrome, revealed that the mutated protein mislocalizes to intracellular compartments in a model cell line and in patient-derived induced pluripotent stem cells. Mutant GLUT1 interacts with adaptor proteins (APs) *in vitro*, and knocking-down AP-2 reverts the cellular mislocalization and restores glucose transport. A systematic analysis of other known disease-causing variants revealed a significant and specific overrepresentation of gained dileucine motifs in structurally disordered cytosolic domains of transmembrane proteins. Thus, several mutations in disordered regions appear to cause “dileucineopathies”.

Highlights:

- A peptide-based screen detects how mutations affect protein-protein interactions
- Several pathogenic mutations create dileucine motifs and recruit clathrin
- A dileucine motif gain in GLUT1 causes mistrafficking in GLUT1 deficiency syndrome
- Dileucine motif gains recurrently occur in pathogenic mutations

Introduction

Genome sequencing technologies have greatly facilitated the discovery of human protein variants. In many cases it is not known whether such variants cause disease, and even when associations have been established, determining the molecular mechanisms remains a major challenge (Cooper and Shendure, 2011). Most disease-causing missense mutations affect evolutionarily conserved amino acids within structured regions of proteins and destabilize their structure (Subramanian and Kumar, 2006; Yue et al., 2005). However, over 20% of human disease mutations occur in so called intrinsically disordered regions (IDRs) (Vacic et al., 2012). Contrary to the traditional understanding of protein structure and function, it is now clear that IDRs represent a functionally important and abundant part of eukaryotic proteomes (Uversky et al., 2008; Wright and Jane Dyson, 2014). Yet, since IDRs lack a defined tertiary structure and are typically poorly conserved, the classical structure-function paradigm cannot explain how mutations in IDRs cause disease.

We set out to investigate the mechanism of these mutations by analyzing protein-protein interactions (PPIs), which can help to understand how mutations cause disease (Ryan et al., 2013; Wang and Marcotte, 2010). The role of PPIs in disease is highlighted by the enrichment of missense mutations in interaction interfaces of proteins associated with the corresponding disorders (Wang et al., 2012). Moreover, comparing the interaction partners of wild-type proteins and their disease-associated variants can reveal disease mechanisms (Hosp et al., 2015; Zhong et al., 2009). We therefore sought to systematically investigate how mutations in IDRs affect PPIs.

IDRs often harbor short linear motifs (SLiMs) that mediate their function (Fuxreiter et al., 2007; Van Roey et al., 2014). These SLiMs typically fall into two major classes -- motifs that mediate interactions with globular domains and/or motifs which harbor posttranslational modification sites (Tompa et al., 2014). Mutations in IDRs can cause disease by disrupting such motifs or by creating novel ones. A number of examples of such pathogenic changes in motifs have been reported (Cordeddu et al., 2009; Kadaveru et al., 2008; Silvis et al., 2003; Vogt et al., 2005). Additionally, computational studies have revealed that pathogenic mutations often occur in SLiMs (Narayan et al., 2016; Radivojac et al., 2008; Uyar et al., 2014). Despite these insights, however, there has not yet been a systematic experimental analysis of how disease-causing mutations in IDRs affect interactions. One reason for this is that the small binding area between SLiMs and cognate domains results in low binding affinities, which makes it difficult to study these interactions (Neduva and Russell, 2005).

Here, we developed a scalable proteomic screen using synthetic peptides to assess the impact of missense mutations in disordered regions on PPIs. By applying this screen to over 120 known disease-causing mutations we obtained a network of PPIs that are lost or gained as the result of mutations in disordered regions. Within this network, we identified a subnetwork comprising three mutations and five interacting proteins enriched in terms related to clathrin-dependent trafficking. Intriguingly, all three mutations in this subnetwork create novel dileucine motifs in cytosolic tails of transmembrane proteins. Since dileucine motifs mediate clathrin-dependent trafficking,

our findings may provide a mechanistic explanation about how these mutations cause disease. Indeed, experiments on the glucose transporter GLUT1 confirmed that the gained dileucine motif causes protein mislocalization by recruiting adaptor proteins and inducing clathrin-dependent endocytosis. In summary, we show that our scalable proteomic screen can reveal the functional consequences of mutations in disordered regions. Moreover, the data suggest that dileucine motif gains in disordered cytosolic tails of transmembrane proteins are a relatively frequent - and potentially druggable - cause of disease.

Results

A peptide-based interaction screen of disease-causing mutations

We reasoned that quantitative interaction proteomics with immobilized synthetic peptides should enable us to systematically assess the impact of mutations in IDRs. Such peptide pull-downs can maintain specificity even in the setting of low affinity interactions (Schulze and Mann, 2004). We therefore designed a scalable proteomic screen that employs peptides synthesized on cellulose membranes (Fig. 1 A). These membranes carry peptides with 15 amino acids that correspond to IDRs in both the wild-type and mutant form. Membranes are incubated with cell extracts to pull-down interacting proteins. After washing, peptide spots are excised and the proteins associated with them are identified and quantified by shotgun proteomics.

The main challenge in such interaction screens is to distinguish specific interaction partners from non-specific contaminants (Gingras and Raught, 2012; Gstaiger and Aebersold, 2009; Meyer and Selbach, 2015; Smits and Vermeulen, 2016). We addressed this challenge through the use of two levels of quantification. First, two replicates of a pull-down with a specific peptide sequence are compared to all other peptide pull-downs via label-free quantification (LFQ) (Cox et al., 2014). This LFQ-filter selects proteins that bind specifically to a given peptide. Secondly, the screen employs SILAC-based quantification (Mann, 2006) to identify differential interaction partners of the wild-type and disease-causing form of a peptide. This strategy requires incubating the two replicates of the membrane with cell lysates that have been differentially SILAC labeled. Wild-type peptide spots from the heavy pull-down are combined with spots from

the light pull-down that correspond to the mutant forms of the same peptide, and *vice versa*. SILAC ratios give a measure of the degree to which each particular mutations affects a specific interaction.

For the screen we selected 128 mutations in IDRs which are known to cause neurological diseases (Fig. S1, Table S1). We included a peptide from an IDR in the SOS1 protein that contains a proline-rich motif by which it is known to recruit several specific binders via their SH3 domains (Schulze and Mann, 2004). We analysed the 2 x 129 pull-down samples by using high-resolution shotgun proteomics in 45 min runs, resulting in a total measurement time of about eight days. Replicates of the same peptide clustered with a median correlation coefficient (Pearson's R) of 0.87, indicating good reproducibility (Fig. S2A). The LFQ data identified nine specific interactors of the SOS1 peptide, including four of the five that were previously known (Fig. 1 B). In the corresponding SILAC data, seven of the nine LFQ-specific binders show preferential binding to the wild-type compared to the mutant which contains a disrupted proline-rich motif (Fig. 1 C). Importantly, all interactors that are both specific (LFQ) and differential (SILAC) contain SH3 domains. To further assess the relationship between peptide motifs and cognate domains we also analyzed all pull-downs combined. We found that mutations which disrupt a predicted SLiM in the peptide tend to reduce binding of proteins with cognate domains (Fig. S2B). Conversely, the gain of a SLiM in a peptide tends to increase binding to proteins with matching domains. In summary, these data demonstrate that our screen efficiently detects how mutations in IDRs affect interactions mediated by SLiMs.

A quantitative interaction network for disease-associated IDRs

Individual pull-downs typically led to the identification of ~400 proteins. If all of these proteins were considered specific binders, this would correspond to ~400 binary interactions per pulldown and a total of more than 100,000 interactions. However, since many proteins are in fact background binders, we applied our quantitative filters with cut-offs derived from the SOS1 control peptide (Fig. 2 A). About half of the 2 x 128 peptides showed at least one specific binder according to the LFQ-filter (Fig. S2C). Applying the LFQ-filter dramatically reduced the total number of interactions to 618. All of these 618 interactions are specific for the wild-type and/or the mutant form of a peptide as compared to all other peptides in the screen (Table S2). However, not all of these specific interactions are differential, i.e. affected by the mutation. Therefore we next applied the SILAC-filter, which led to a final list of 180 differential interactions (Table S3). 111 of these interactions are lost through mutations in the peptide, while 69 are gained. Of note, since pull-downs can also capture indirect binders, not all of these interactions are necessarily direct.

To provide an overview of the data we displayed all of the differential interactions as a network (Fig. 2 B). This revealed that several wild-type or mutant peptides shared differential interactors, suggesting functional similarities. Moreover, subnetworks were enriched in specific gene ontology terms (supplemental dataset “GOtermClusters.pdf”). Figure 2 B highlights two subnetworks that we find particularly interesting (insets) : One is enriched in proteins connected to clathrin-coated vesicles (see below). The other is

enriched in splicing factors that interact with an IDR corresponding to amino acids 512-526 of fused in sarcoma (FUS). These interactions are disrupted by the R521C mutation. FUS is an RNA-binding protein implicated in amyotrophic lateral sclerosis (ALS) (Deng et al., 2014). R521C and other mutations in the C-terminal region of the protein are thought to be pathogenic because they disrupt a nuclear localization signal (Dormann et al., 2010). Our data suggests that impaired binding of splicing factors could be an additional/alternative explanation for the pathogenicity of this mutation. This observation is interesting because FUS has already been implicated in splicing (Ishigaki et al., 2012; Qiu et al., 2014; Rogelj et al., 2012). In fact, the C-terminal region of the protein was found to interact with serine/arginine-rich splicing factor 10 (SRSF10) even before pathogenic mutations in this region were identified (Yang et al., 1998).

Recruitment of clathrin through gains of dileucine motifs

The finding we considered most interesting is that mutated IDRs from CACNA1H, GLUT1/SLC2A1 and ITPR1 lead to specific interactions with clathrin (Fig. 3 A). The corresponding SILAC data revealed that, in all three cases, clathrin exhibited a strong preference for the mutant form of the peptides over the wild-type (Fig. 3 B). Since clathrin mediates endocytosis and intracellular trafficking of transmembrane proteins, our finding suggests that these mutations might affect protein trafficking. Intriguingly, the three mutations share other features beyond an increased affinity for clathrin: First, all three mutations affect transmembrane proteins -- a calcium channel (CACNA1H) and a glucose transporter (GLUT1) residing in the plasma membrane and an inositol 1,4,5-trisphosphate receptor (ITPR1) located mainly in the endoplasmic reticulum (ER) (Fig. 3

C). Second, all three mutations affect disordered regions exposed to the cytosol, which makes them accessible to cytosolic adaptor proteins that mediate clathrin recruitment. Third, all three mutations involve the change of a proline to a leucine residue and thereby result in the appearance of a novel dileucine motif ("LL") in the IDR (Fig. 3 D). Such motifs are known to recruit clathrin to the plasma membrane or intracellular locations (Pandey, 2009). The classical dileucine motif is [D/E]XXXL[L/I] (Dinkel et al., 2015), but variations of this theme are common (Kozik et al., 2010; Pandey, 2009; Staudt et al., 2016; Traub, 2009).

A dileucine motif gain causes mislocalization of the glucose transporter GLUT1

To assess the functional significance of the dileucine motif gains and clathrin recruitment we selected the P485L mutation in GLUT1/SLC2A1. This mutation causes GLUT1 deficiency syndrome (G1DS), a disorder characterized by seizures and intellectual disability with onset in early infancy (De Vivo et al., 1991; Leen et al., 2010a; Pascual et al., 2008). GLUT1 is mainly expressed in endothelial cells that form the blood-brain-barrier and in astrocytes, facilitating glucose entry into the brain. Pathogenic mutations in GLUT1 impair cerebral glucose flux, leading to permanent encephalopathy.

To determine the impact of the P485L mutation on the subcellular localization, we first generated stable inducible cell lines expressing epitope tagged full-length wild-type or mutant GLUT1. While the wild-type protein mainly localized to the plasma membrane, the P485L mutant displayed an intracellular pattern (Fig. 4 A). Hence, the mutation

indeed appears to cause protein mislocalization. Colocalization experiments with several markers confirm that mutated GLUT1 localizes to endocytic compartments (Fig. 4 B, S3). To more systematically characterize the cellular compartment in which GLUT1_P485L resides, we used BioID as a proximity labeling method (Roux et al., 2012) (Table S4). We performed this experiment in a comparative manner for both wild-type and mutant GLUT1 using SILAC-based quantification (see Methods). Mutated GLUT1 colocalized with proteins involved in membrane trafficking, clathrin-mediated endocytosis, and post-Golgi trafficking (Fig. 4 C). In contrast, wild-type GLUT1 colocalized with plasma membrane-associated proteins.

Adaptor proteins bind to mutant GLUT1 and cause cellular mistrafficking

The BioID experiment identified several subunits of the heterotetrameric vesicular transport adaptor proteins (APs). This finding is particularly relevant because AP-1, AP-2 and AP-3 directly bind to both dileucine motifs and clathrin to mediate cellular transport (Traub and Bonifacino, 2013): Binding of AP complexes to the cargo triggers a conformational change, which opens up the AP complex. It now exposes a 'Clathrin box motif' that leads to recruitment of clathrin, which begins to surround the emerging vesicle bud as a second protein layer.

Intriguingly, all subunits of AP-1, AP-2 and AP-3 that we identified showed increased co-localization with mutated over wild-type GLUT1 in both the forward and reverse (that is, SILAC label swap) BioID experiment (Fig. 5 A). We therefore tested whether the

cytosolic tail of GLUT1 can interact with APs *in vitro* (Fig. 5 B). We found that mutated but not wild-type GLUT1 pulled down both AP-1 and AP-2. Together, these results show that the mutation causes association of GLUT1 with APs, providing a molecular explanation for mistrafficking. Since we did not detect APs as hits in our original screen, we designed targeted assays against peptides from several APs based on their known fragmentation spectra (Zauber et al., 2018). Repeating the peptide pull-downs for GLUT1, CACNA1H and ITPR1 with targeted proteomics as read-out, confirmed that several APs preferentially interact with the mutated peptides (Fig. S4).

APs localize to different intracellular compartments and mediate membrane trafficking in distinct pathways (Park and Guo, 2014a). For example, AP-2 mediates clathrin-dependent endocytosis at the plasma membrane (McMahon and Boucrot, 2011; Pandey, 2009; Staudt et al., 2016; Traub, 2009). The interaction of GLUT1_P485L with AP-2 therefore suggests that internalization of the protein from the plasma membrane contributes to its mislocalization. To test whether mutant GLUT1 is taken up via endocytosis we added fluorescently labeled transferrin to GLUT1 expressing cells. Mutant but not wild-type protein extensively colocalized with endocytosed transferrin (Fig.5 C and D). Finally, if the P485L mutation causes GLUT1 endocytosis via AP-2, inhibiting AP-2 function should restore the correct subcellular localisation. We therefore used siRNAs to knock-down AP-2 (Fig. 5 E). Consistent with our prediction, loss of AP-2 expression rescued the mislocalization of mutated GLUT1 (Fig. 5 F). We also tested the ability of stable cell lines inducibly expressing GLUT1 to take-up radiolabeled glucose. Cells expressing the mutated proteins showed significantly reduced glucose

uptake, which was rescued by knocking down AP-2 (Fig. 5 G). Thus, the P485L mutant can apparently sustain normal glucose uptake when the protein is located in the plasma membrane. Together, these data show that the dileucine motif gain indeed causes clathrin-dependent endocytosis and mistrafficking of GLUT1.

GLUT1 P485L mislocalizes in patient-derived iPSCs

The experiments described so far are based on *in vitro* assays or cell line models expressing tagged variants of GLUT1. We therefore sought to validate our findings by analyzing the behaviour of endogenous GLUT1 in patient cells. To this end, we obtained fibroblasts from a GLUT1 deficient patient harboring the P485L mutation via a skin punch biopsy (see Methods). We then reprogrammed these fibroblasts into induced pluripotent stem cells (iPSCs) by RNA-based transfer of pluripotency factors (Fig. 6 A). The clones obtained showed characteristic expression of pluripotency markers (Fig. S5). We then analyzed the subcellular distribution of GLUT1 in these cells. While GLUT1 was mainly localized in the plasma membrane of control iPSCs, patient-derived cells showed characteristic intracellular accumulations (Fig. 6 B). Co-staining with a post-Golgi SNARE revealed extensive colocalization (Kreykenbohm et al., 2002) (Fig. 6 C). Of note, the patient-derived iPSCs also showed GLUT1 signal at the plasma membrane. This is consistent with the fact that only one GLUT1 allele in the patient is affected (Leen et al., 2010a; Pascual et al., 2008; Slaughter et al., 2009). In summary, these data show that the mislocalization observed in HEK cells can be reproduced in patient cells.

GLUT1 P485L localization at the blood-brain-barrier is perturbed *in vivo*

Next, to study the P485L mutation *in vivo*, we used CRISPR/Cas9 to generate the corresponding mutant in the mouse genome (Fig. 6 E). Heterozygous mice were viable, fertile and did not display any obvious phenotype. However, out of 6 heterozygous crossings, we failed to detect any born homozygous mutant pup. A detailed inspection of maternal delivery illustrated that homozygous mutant pups come to term but die immediately after birth and are removed from the litter by the dams. To histologically analyze GLUT1 distribution in the endothelial cells of the blood-brain-barrier, we dissected embryonic (E) day 14.5-15.5 mice and stained their cerebral cortex with antibodies against GLUT1. We co-stained with anti-ICAM2 to label the luminal plasma membrane and isolectin B4 (IB4) to label the entire endothelial plasma membrane. We observed an overall reduction in GLUT1 staining in heterozygous and homozygous animals with otherwise largely normal vascular morphology (Fig. 6 F). High resolution STED imaging revealed a strong reduction in the density of GLUT1 clusters in endothelial plasma membranes of homozygous mutant mice compared to wild type littermates (Fig. 6 G and H). Hence, the P485L mutation also reduces GLUT1 levels in the plasma membrane *in vivo.*

Gains in dileucine motifs as a general disease mechanism

We next investigated whether dileucine motif gains represent a more general disease mechanism. To test this, we conducted a search of missense mutations known to cause disease and that occur within disordered cytosolic regions of transmembrane proteins.

We found four additional pathogenic dileucine motif gains in the Humsavar database (Fig. 7 A, Table S5). The same search in the ClinVar database returned four additional mutations. For example, two dileucine motif gains affect cytosolic regions of the cystic fibrosis transmembrane conductance regulator (CFTR) and cause cystic fibrosis (Figure S6A, Table S5). In total, these 11 mutations affect 6 different proteins and cause a wide range of diseases.

Since we focused our follow-up experiments on GLUT1, we cannot state with certainty that the other dileucine motif gains also cause protein mistrafficking. Alternatively, the mutations could cause disease by different mechanisms and might just create dileucine motifs as a by-product. If that was the case, dileucine motifs should be homogeneously distributed between disease-causing mutations and non-pathogenic polymorphisms. In contrast, if dileucine motif gains are responsible for pathogenesis, they would be predicted to occur more often in disease than in non-pathogenic variants. Moreover, pathogenic dileucine motif gains should be specific for cytosolic regions of transmembrane proteins since this is where they exert their function. To test these predictions, we compared the frequency of dileucine motif gains that have been found in disease-causing mutations to their appearance in non-pathogenic polymorphisms (from Humsavar). A global survey of all disordered regions of the entire proteome revealed that gains in dileucine motif gains occurred at about the same rate in disease and non-pathogenic variants (OR = 0.81, p-value = 0.319, two-sided Fisher's Exact Test). In the cytosolic tails of transmembrane proteins, however, we observed a 3.7-fold enrichment of dileucine motifs implicated in disease (OR = 3.7, p-value = 0.017, two-sided Fisher's Exact Test, Fig. 7 B). Disordered extracellular regions of transmembrane proteins do

not show this enrichment. Performing the equivalent analysis with the ClinVar database yielded similar results (Figure S6B). We conclude that dileucine motif gains in disordered regions of the cytosolic segments of transmembrane proteins are significantly and specifically enriched in disease. To further assess the significance of this finding, we systematically searched within cytosolic regions of transmembrane proteins for all other annotated SLiMs contained in the ELM database (Dinkel et al., 2015). Intriguingly, of all 263 SLiMs tested, the dileucine motif (LIG_diLeu_1) was the only significantly enriched motif in disease (Fig. 7 C).

Finally, to test if some of the additional dileucine motif gains can cause mistrafficking similar to GLUT1, we performed antibody feeding experiments. To this end, we created chimeric proteins consisting of the IL-2 receptor alpha chain (TAC) fused to mutated and wild-type cytosolic regions of the respective disease protein (see Methods). Cells expressing these constructs are incubated with antibodies against the extracellular region of TAC and allowed to endocytose the chimeric proteins together with bound antibodies. A specific staining protocol is then used to exclusively detect internalized antibodies (Diril et al., 2009). We generated fusion proteins for GLUT1 (as positive control) and seven additional dileucine motif gains. Four of the seven mutations resulted in increased internalization relative to the corresponding wild-type sequences (Fig. 7 D). In addition, we observed mutant-specific interaction of AP-1 and/or AP2 for several selected cytosolic regions in in vitro interaction assays (Fig. S6D). Collectively, these results indicate that several additional pathogenic dileucine motif gains cause protein mistrafficking.

Discussion

Understanding the functional relevance of protein variants is a major challenge in the era of personal genomics -- especially for missense mutations in disordered regions. Our proteomic screen provides a first systematic experimental analysis of how mutations in disordered regions affect protein-protein interactions. Our results show that the method can (i) capture known interactions, (ii) detect how mutations in SLiMs affect binding of cognate domains and (iii) provide novel mechanistic insights into pathogenesis. The peptide-based method is especially useful for mutations in proteins that are otherwise difficult to study, such as large transmembrane proteins. Nevertheless, it is also important to consider the intrinsic limitations of the approach. Most importantly, *in vitro* pull-downs do not necessarily reflect physiological interactions *in vivo*. For example, artifactual binding can occur when combining peptides and proteins that never meet each other in the cell (Gibson et al., 2015). Moreover, taking IDRs out of the context of the full length protein and immobilizing them as short peptides can affect interactions. Finally, amino acids within IDRs often carry posttranslational modifications -- a possibility which we did not consider here. In the future, it will be interesting to include modified peptides, especially since mutations often affect modification sites (Narayan et al., 2016; Radivojac et al., 2008).

Our screen revealed that three mutations in cytosolic tails of transmembrane proteins create dileucine motifs and lead to increased binding of clathrin. Follow-up experiments demonstrated that the dileucine motif gain in GLUT1 causes mislocalization from the plasma membrane to endocytic compartments. We also observed that mutated GLUT1

recruits several adaptor proteins (AP-1, AP-2 and AP-3), thus providing a direct link between the motif gain, clathrin recruitment and GLUT1 mistrafficking. Furthermore, knocking-down AP-2 rescued mislocalization of mutated GLUT1 and restored glucose transport. This finding shows that the aberrant trafficking is -- at least partially -- due to increased endocytosis. However, it should be kept in mind that AP-1 and AP-3 mediate other trafficking events such as transport between endosomes and the trans-Golgi network or transport to lysosomes (Dell'Angelica, 2009; Park and Guo, 2014). The observation that mutated GLUT1 also interacts with these APs therefore suggests that other trafficking events may also be perturbed.

The finding that pathogenic mutations in cytosolic tails of other transmembrane proteins also create dileucine motifs is particularly intriguing. We term diseases caused by such motif gains "dileucineopathies". Whether the other dileucine motif gains cause protein mislocalization similar to GLUT1 remains to be investigated. The observation that these mutations are significantly and specifically enriched in cytosolic domains suggests that at least some of them are functional. Also, we find that four out of seven tested mutations increase internalization of chimeric proteins in antibody feeding experiments. This further supports the view that at least some pathogenic dileucine motif gains cause disease by inducing protein mistrafficking. However, more detailed follow-up experiments are required to test this hypothesis for individual mutations. It is also interesting to note that a pathogenic mutation in the cystic fibrosis transmembrane conductance regulator (CFTR) has been reported to generate a tyrosine-based internalization motif (Silvis et al., 2003).

Why are dileucine-motif gains a recurrent cause of disease? We think this is due to a combination of several factors: First, dileucine motifs are not very complex and can thus easily arise by chance. Second, proline codons can mutate to leucine codons by changing a single nucleotide. Third, proline is overrepresented in IDRs, which are also the sites where the motif needs to be located in order to be functional. Since the genetic cause of many diseases has not yet been identified (Boycott et al., 2013), we expect that more pathogenic dileucine motif gains will soon emerge. Knowing that such gains can be pathogenic will make it easier to classify them as disease-causing among the many variants present in the human population (Cooper and Shendure, 2011). The observation that GLUT1 mislocalization and glucose transport can be rescued suggests that pathogenic dileucine motif gains may be druggable. Whether patients with "dileucineopathies" might benefit from inhibiting specific clathrin-dependent trafficking events remains to be investigated.

Bioinformatic studies have established that pathogenic mutations in disordered regions often affect SLiMs (Narayan et al., 2016; Radivojac et al., 2008; Uyar et al., 2014). However, whether these predicted motif changes really affect protein-protein interactions has not yet been investigated systematically. Moreover, many motifs have not yet been defined and thus escape computational predictions (Tompa et al., 2014). The biochemical approach presented here provides a useful complementary strategy to computational studies. Key advantages of our setup are its scalability (by using synthetic peptides) and specificity (by employing two quantitative filters). While we focused on neurological disorders here, the approach can also be applied to other types

of Mendelian disorders, somatic mutations in cancer and also to non-pathogenic polymorphisms.

Acknowledgements

We are indebted to the patient and her family for providing fibroblasts. We thank Philip Kim (University of Toronto) for helpful discussions and Russ Hodge (MDC) for valuable comments on the manuscript. We also thank Daniel Perez Hernandez (MDC) and Ulf Reimer (JPT) for practical suggestions in setting up the screen. Markus Landthaler (MDC) kindly provided stable cell lines and the Advanced Light Microscopy Technology Platform (MDC) helped with microscopy. We also like to thank Martha Hergeselle and Sven Buchert (both MDC) for excellent technical assistance during cloning, cell culture and mouse experiments. G.R. was funded by the German Research Foundation (DFG, SFB958/A11 to M.Kr.). B.U. acknowledges funding by the German Federal Ministry of Education and Research (BMBF) as part of the RNA Bioinformatics Center of the German Network for Bioinformatics Infrastructure (de.NBI) [031 A538C RBC (de.NBI)]. J.M.P. is supported by National Institutes of Health grants NS077015 and NS094257.

Author contributions

K.M. and M.Ki. established the screen and selected the candidates; K.M. performed most of the wet-lab experiments with contribution from M.Ki. and J.C.; K.M. processed and analyzed the mass spectrometric data, B.U. carried out most of the remaining

bioinformatic analyses (mainly motif based analyses) supervised by A.A.; G.R. carried out the GST-pulldown assays supervised by M.Kr. and V.H.; H.Z. set-up and analyzed targeted proteomic experiments; J.M.P. obtained fibroblasts from the patient; S.D. generated iPSCs from the patient fibroblasts; I.R. performed experiments with patient-derived iPSCs supervised by T.W.. A.S. performed STED imaging and analysis supervised by H.G. ; M.S. conceived and supervised the work; K.M and M.S. wrote the manuscript with input from all authors.

Competing financial interest statement

The authors declare no competing financial interest.

Figure Legends

Fig. 1: Quantitative interaction screen with disease-associated disordered regions
A, Cellulose membranes with synthetic wild-type (circles) and mutated (stars) peptides are incubated with lysate from light (light blue) or heavy (dark blue) SILAC labelled cells to pull-down interacting proteins. Spots are excised, corresponding wild-type/mutant pairs are combined and analyzed by quantitative shotgun proteomics (represented by an Orbitrap). Middle panels: Label free quantification (LFQ) identifies specific interactors by comparing both replicates to all other pull-downs. Volcano plots depict protein enrichment in in the two replicate pull-downs of a given peptide over all other peptide pull-downs (x axis shows mean log2FC, y axis shows -log10 p-values), separately for the wild-type (left) and mutant peptide (right). The threshold (red lines) was derived from the benchmark experiment with the SOS1 peptide (Fig. 1B, for details see Material and Methods). LFQ-specific interactors are depicted in red. SILAC-based quantification identifies differential binders by directly comparing corresponding wild-type and mutant pairs. Differential binders of the wild-type and mutant peptide appear in the upper-right and lower-left quadrants, respectively (for detailed selection criteria see Methods). **B, C** Results for a SOS1-derived peptide with a SH3 domain-binding PxxP motif as a benchmark. **B,** Volcano plot from the LFQ data for wild-type SOS1. Specific binders are shown as red dots. 4 out of 5 known binders (red gene names) are detected. **C,** SILAC log2 fold changes for differential binders of the wild-type and mutant SOS1 peptide. Proteins with SH3 domains are shown with black outlines.

Fig. 2: Differential interactors of wild-type and mutant IDRs
A, Quantitative filters to select specific and differential interactions. Only a minor fraction of all detected interactions is specific (LFQ filter). Moreover, only a fraction of specific interactions are differential (SILAC filter), i.e. show preferential binding to the wild-type or mutant form of a peptide. Mutation-induced interaction losses are more frequent than mutation-induced gains.
B, Network of all differential interactions. Peptides (rectangles) and interacting proteins (ovals) are presented as nodes. The edges indicate preferential binding to the wild-type

(blue) or mutant (red) form of a peptide (edge width indicates SILAC ratios). Highlighted subnetworks are enriched in splicing regulators and clathrin-coated vesicle proteins (see text).

Fig. 3: Recruitment of clathrin by recurrent gains of dileucine motifs

A, Volcano plots for pull-downs with mutated peptides derived from CACNA1H, GLUT1 and ITPR1. Specific binders (relative to all other pull-downs) are highlighted in red. All three peptides specifically interact with clathrin.

B, Corresponding SILAC plots show that clathrin and related proteins preferentially bind to the mutant form of peptides (relative to the wild-type).

C, Graphical representation of the mutation sites. All three mutations affect cytosolic regions of transmembrane proteins. CACNA1H and GLUT1 are located mainly in the plasma membrane and ITPR1 mainly in the ER.

D, Aligning the three peptide sequences reveals a common gain of a dileucine motif.

Fig. 4: A mutation-induced dileucine motif gain causes mislocalization of the glucose transporter GLUT1

A, Confocal images of GLUT1 localization in HEK cells, stably expressing FLAG-GLUT1, reveal that the wild-type is localized mainly at the cell membrane while the P485L mutant is mislocalized to endocytic compartments. (FLAG-GLUT1: green, DAPI: blue). Scale bars: 10 µm. **B**, Colocalization analysis shows extensive colocalization of mutant, but not wild-type GLUT1 with markers of several endocytic compartments. Pearson's thresholded coefficients (as implemented in the Imaris software) was determined for GLUT1 variants with the indicated proteins. Depicted are mean values, error bars show standard deviations of the mean. For exemplary images see Fig. S3.

C, Comparison of proteins colocalizing with wild-type and mutant GLUT1 by proximity labeling (BioID). The upper left panel shows SILAC log2 fold changes from two replicate experiments with swapped isotope labels. Blue and red labeled proteins are enriched by wild-type GLUT1 or mutant GLUT1, respectively. The ten most significant cellular component GO-terms reveal that mutated GLUT1 is involved in clathrin dependent processes and endosomal trafficking. In contrast, wild-type GLUT1 colocalizes with

plasma membrane-associated proteins. The lower panel is colored according to the top three enriched GO-terms and shows the variants typical subcellular compartments. Figure adapted from (Raiborg and Stenmark, 2009).

Fig. 5: Adaptor proteins bind to mutated GLUT1 and cause cellular mistrafficking

A, Adaptor protein complexes AP-1, AP-2 and AP-3 show increased colocalization to GLUT1 due to P485L mutation in replicates of BioID experiment from Fig. 4. Identified subunits of APs are shown in red.

B, P485L mutant but not wild-type Glut1 C-terminal tail interacts with AP-1 and AP-2. Tails were tagged with GST to pull-down interaction partners from mouse brain lysate. Talin is shown as a negative control and is not pulled down from either of the two variants.

C, D, Mutant but not wild-type GLUT1 extensively colocalizes with endocytosed transferrin. HEK cells stably expressing FLAG-GLUT1 are incubated with fluorescently labeled transferrin for 10 min before fixation. Scale bar: 10 µm. **E, F,** Western-blot against AP-2 α and µ subunits shows downregulation after two rounds of siRNA transfection against AP-2 µ. AP-2 knock-down leads to relocalization of GLUT1_P485L to the plasma membrane and hence rescue of the mutation phenotype. Scale bar: 10 µm. **G**, Radioactive glucose uptake after AP-2 knock-down leads to rescue of glucose uptake deficit of GLUT1_P485L. GLUT1 expression was induced by addition of doxycycline and glucose uptake was inhibited by addition of cytochalasin B. % glucose uptake is relative to GLUT1 wild-type, +doxycycline, -cytochalasin B. Depicted are mean values over technical triplicates of three independent experiments. Error bars: SEM. $*$p value <0.05, ** p value <0.01 from a paired, one-sided t-test.

Fig. 6: GLUT1 P485L mislocalizes in patient-derived iPSCs and endothelial cells of the blood-brain barrier in mice

A, A skin sample was taken from a GLUT1 deficiency patient with a heterozygous GLUT1_P485L mutation. Fibroblasts were grown and reprogrammed to iPSCs.

B, Heterozygous GLUT1_P485L mutation leads to partial mislocalization of GLUT1 in patient-derived iPSCs. Scale bars: 10 µm.

C, D, GLUT1_P485L colocalizes with the post-Golgi SNARE VTI1A. Scale bar: 10 µm. **E**, A mouse carrying the GLUT1 P485L mutation was created by CRISPR/Cas9 targeted method. PAM sequence and gRNA are marked in targeted region of SLC2A1 (GLUT1 gene). Sanger sequencing confirmed insertion of mutation (chromatogram: A=green, T=red, C=blue, G=black). **F**, Immunohistological analyses of cortical slices of wild-type, heterozygous and homozygous GLUT1 mutant mice using antibodies against GLUT1 (red) and DAPI (blue) as counterstain (left panels); a higher magnification of a vessel stained by antibodies against GLUT1 (red), IB4 (green) and ICAM2 (blue) is shown in the right panels. **G**, Representative STED images of transverse cross-sections through brain vessels of wild type (+/+), heterozygous (+/P485L) and homozygous (P485L/P485L) mutant mice stained with isolectin B4 and antibodies against GLUT1 and ICAM2. Insets show a fragment of abluminal membrane (IB4 positive, ICAM2 negative) indicated with a black box. Scale bars: 2 µm (main panels) and 0.25 µm (insets). **H**, Quantification of Glut1 signal relative to IB4 signal in vessel membranes (n – number of vessels per genotype, N – number of animals per genotype). Boxplot central line indicates the median, the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively.

Fig. 7: Mutation-induced gains of dileucine motifs are a significant cause of disease

A, A systematic bioinformatic search revealed four additional pathogenic mutations in cytosolic segments of transmembrane proteins that create dileucine motifs.

B, Relative frequency of dileucine motif gains in disease mutations and polymorphisms in different disordered regions (IUPred Score >= 0.4) of the proteome. Dileucine motif gain is significantly enriched only in disordered regions of the cytoplasmic domains of transmembrane proteins (two-sided Fisher's exact test).

C, Comparison of gained motifs in disordered regions of cytoplasmic tails of transmembrane proteins reveals the dileucine motif to have the most significant specific enrichment when compared with polymorphisms.

D, Antibody feeding indicates that four out of seven tested mutations with gain of dileucine motif lead to a gain in endocytosis (candidates from Humsavar and Clinvar for details on selection see Methods). Scale bars: 10µm.

References

Boycott, K.M., Vanstone, M.R., Bulman, D.E., and MacKenzie, A.E. (2013). Rare-disease genetics in the era of next-generation sequencing: discovery to translation. Nat. Rev. Genet. *14*, 681–691.

Cooper, G.M., and Shendure, J. (2011). Needles in stacks of needles: finding disease-causal variants in a wealth of genomic data. Nat. Rev. Genet. *12*, 628–640.

Cordeddu, V., Di Schiavi, E., Pennacchio, L.A., Ma'ayan, A., Sarkozy, A., Fodale, V., Cecchetti, S., Cardinale, A., Martin, J., Schackwitz, W., et al. (2009). Mutation of SHOC2 promotes aberrant protein N-myristoylation and causes Noonan-like syndrome with loose anagen hair. Nat. Genet. *41*, 1022–1026.

Cox, J., Hein, M.Y., Luber, C.A., Paron, I., Nagaraj, N., and Mann, M. (2014). Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. Mol. Cell. Proteomics *13*, 2513–2526.

Dell'Angelica, E.C. (2009). AP-3-dependent trafficking and disease: the first decade. Curr. Opin. Cell Biol. *21*, 552–559.

Deng, H., Gao, K., and Jankovic, J. (2014). The role of FUS gene variants in neurodegenerative diseases. Nat. Rev. Neurol. *10*, 337–348.

De Vivo, D.C., Trifiletti, R.R., Jacobson, R.I., Ronen, G.M., Behmand, R.A., and Harik, S.I. (1991). Defective Glucose Transport across the Blood-Brain Barrier as a Cause of Persistent Hypoglycorrhachia, Seizures, and Developmental Delay. N. Engl. J. Med. *325*, 703–709.

Dinkel, H., Van Roey, K., Michael, S., Kumar, M., Uyar, B., Altenberg, B., Milchevskaya, V., Schneider, M., Kühn, H., Behrendt, A., et al. (2015). ELM 2016—data update and new functionality of the eukaryotic linear motif resource. Nucleic Acids Res. *44*, D294–D300.

Diril, M.K., Schmidt, S., Krauss, M., Gawlik, V., Joost, H.-G., Schürmann, A., Haucke, V., and Augustin, R. (2009). Lysosomal localization of GLUT8 in the testis--the EXXXLL motif of GLUT8 is sufficient for its intracellular sorting via AP1- and AP2-mediated interaction. FEBS J. *276*, 3729–3743.

Dormann, D., Rodde, R., Edbauer, D., Bentmann, E., Fischer, I., Hruscha, A., Than, M.E., Mackenzie, I.R.A., Capell, A., Schmid, B., et al. (2010). ALS-associated fused in sarcoma (FUS) mutations disrupt Transportin-mediated nuclear import. EMBO J. *29*, 2841–2857.

Fuxreiter, M., Tompa, P., and Simon, I. (2007). Local structural disorder imparts plasticity on linear motifs. Bioinformatics *23*, 950–956.

Gibson, T.J., Dinkel, H., Van Roey, K., and Diella, F. (2015). Experimental detection of short regulatory motifs in eukaryotic proteins: tips for good practice as well as for bad. Cell Commun. Signal. *13*.

Gingras, A.-C., and Raught, B. (2012). Beyond hairballs: The use of quantitative mass spectrometry data to understand protein-protein interactions. FEBS Lett. *586*, 2723–2731.

Gstaiger, M., and Aebersold, R. (2009). Applying mass spectrometry-based proteomics to genetics, genomics and network biology. Nat. Rev. Genet. *10*, 617–627.

Hosp, F., Vossfeldt, H., Heinig, M., Vasiljevic, D., Arumughan, A., Wyler, E., Genetic and Environmental Risk for Alzheimer's Disease GERAD1 Consortium, Landthaler, M., Hubner, N., Wanker, E.E., et al. (2015). Quantitative interaction proteomics of neurodegenerative disease proteins. Cell Rep. *11*, 1134–1146.

Ishigaki, S., Masuda, A., Fujioka, Y., Iguchi, Y., Katsuno, M., Shibata, A., Urano, F., Sobue, G., and Ohno, K. (2012). Position-dependent FUS-RNA interactions regulate alternative splicing events and transcriptions. Sci. Rep. *2*, 529.

Kadaveru, K., Vyas, J., and Schiller, M.R. (2008). Viral infection and human disease--insights from minimotifs. Front. Biosci. *13*, 6455–6471.

Kozik, P., Francis, R.W., Seaman, M.N.J., and Robinson, M.S. (2010). A Screen for Endocytic Motifs. Traffic *11*, 843–855.

Kreykenbohm, V., Wenzel, D., Antonin, W., Atlachkine, V., and von Mollard, G.F. (2002). The SNAREs vti1a and vti1b have distinct localization and SNARE complex partners. Eur. J. Cell Biol. *81*, 273–280.

Leen, W.G., Klepper, J., Verbeek, M.M., Leferink, M., Hofste, T., van Engelen, B.G., Wevers, R.A., Arthur, T., Bahi-Buisson, N., Ballhausen, D., et al. (2010a). Glucose transporter-1 deficiency syndrome: the expanding clinical and genetic spectrum of a treatable disorder. Brain *133*, 655–670.

Leen, W.G., Klepper, J., Verbeek, M.M., Leferink, M., Hofste, T., van Engelen, B.G., Wevers, R.A., Arthur, T., Bahi-Buisson, N., Ballhausen, D., et al. (2010b). Glucose transporter-1 deficiency syndrome: the expanding clinical and genetic spectrum of a treatable disorder. Brain *133*, 655–670.

Mann, M. (2006). Functional and quantitative proteomics using SILAC. Nat. Rev. Mol. Cell Biol. *7*, 952–958.

McMahon, H.T., and Boucrot, E. (2011). Molecular mechanism and physiological functions of clathrin-mediated endocytosis. Nat. Rev. Mol. Cell Biol. *12*, 517–533.

Meyer, K., and Selbach, M. (2015). Quantitative affinity purification mass spectrometry: a versatile technology to study protein–protein interactions. Front. Genet. *6*.

Narayan, S., Bader, G.D., and Reimand, J. (2016). Frequent mutations in acetylation and ubiquitination sites suggest novel driver mechanisms of cancer. Genome Med. *8*, 55.

Neduva, V., and Russell, R.B. (2005). Linear motifs: Evolutionary interaction switches. FEBS Lett. *579*, 3342–3345.

Pandey, K.N. (2009). Functional roles of short sequence motifs in the endocytosis of membrane receptors. Front. Biosci. *14*, 5339–5360.

Park, S.Y., and Guo, X. (2014). Adaptor protein complexes and intracellular transport. Biosci. Rep. *34*, 381–390.

Pascual, J.M., Wang, D., Yang, R., Shi, L., Yang, H., and De Vivo, D.C. (2008). Structural signatures and membrane helix 4 in GLUT1: inferences from human blood-brain glucose transport mutants. J. Biol. Chem. *283*, 16732–16742.

Qiu, H., Lee, S., Shang, Y., Wang, W.-Y., Au, K.F., Kamiya, S., Barmada, S.J., Finkbeiner, S., Lui, H., Carlton, C.E., et al. (2014). ALS-associated mutation FUS-R521C causes DNA damage and RNA splicing defects. J. Clin. Invest. *124*, 981–999.

Radivojac, P., Baenziger, P.H., Kann, M.G., Mort, M.E., Hahn, M.W., and Mooney, S.D. (2008). Gain and loss of phosphorylation sites in human cancer. Bioinformatics *24*, i241–i247.

Raiborg, C., and Stenmark, H. (2009). The ESCRT machinery in endosomal sorting of ubiquitylated membrane proteins. Nature *458*, 445–452.

Rogelj, B., Easton, L.E., Bogu, G.K., Stanton, L.W., Rot, G., Curk, T., Zupan, B., Sugimoto, Y., Modic, M., Haberman, N., et al. (2012). Widespread binding of FUS along nascent RNA regulates alternative splicing in the brain. Sci. Rep. *2*, 603.

Roux, K.J., Kim, D.I., Raida, M., and Burke, B. (2012). A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. J. Cell Biol. *196*, 801–810.

Ryan, C.J., Cimermančič, P., Szpiech, Z.A., Sali, A., Hernandez, R.D., and Krogan, N.J. (2013). High-resolution network biology: connecting sequence with function. Nat. Rev. Genet. *14*, 865–879.

Schulze, W.X., and Mann, M. (2004). A novel proteomic screen for peptide-protein interactions. J. Biol. Chem. *279*, 10756–10764.

Silvis, M.R., Picciano, J.A., Bertrand, C., Weixel, K., Bridges, R.J., and Bradbury, N.A. (2003). A mutation in the cystic fibrosis transmembrane conductance regulator generates a novel internalization sequence and enhances endocytic rates. J. Biol. Chem. *278*, 11554–11560.

Slaughter, L., Vartzelis, G., and Arthur, T. (2009). New GLUT-1 mutation in a child with treatment-resistant epilepsy. Epilepsy Res. *84*, 254–256.

Smits, A.H., and Vermeulen, M. (2016). Characterizing Protein-Protein Interactions Using Mass Spectrometry: Challenges and Opportunities. Trends Biotechnol. *34*, 825–834.

Staudt, C., Puissant, E., and Boonen, M. (2016). Subcellular Trafficking of Mammalian Lysosomal Proteins: An Extended View. Int. J. Mol. Sci. *18*.

Subramanian, S., and Kumar, S. (2006). Evolutionary anatomies of positions and types of disease-associated and neutral amino acid mutations in the human genome. BMC Genomics *7*, 306.

Tompa, P., Davey, N.E., Gibson, T.J., and Babu, M.M. (2014). A million peptide motifs for the molecular biologist. Mol. Cell *55*, 161–169.

Traub, L.M. (2009). Tickets to ride: selecting cargo for clathrin-regulated internalization. Nat.

Rev. Mol. Cell Biol. *10*, 583–596.

Traub, L.M., and Bonifacino, J.S. (2013). Cargo recognition in clathrin-mediated endocytosis. Cold Spring Harb. Perspect. Biol. *5*, a016790.

Uversky, V.N., Oldfield, C.J., and Dunker, A.K. (2008). Intrinsically disordered proteins in human diseases: introducing the D2 concept. Annu. Rev. Biophys. *37*, 215–246.

Uyar, B., Weatheritt, R.J., Dinkel, H., Davey, N.E., and Gibson, T.J. (2014). Proteome-wide analysis of human disease mutations in short linear motifs: neglected players in cancer? Mol. Biosyst. *10*, 2626–2642.

Vacic, V., Markwick, P.R.L., Oldfield, C.J., Zhao, X., Haynes, C., Uversky, V.N., and Iakoucheva, L.M. (2012). Disease-associated mutations disrupt functionally important regions of intrinsic protein disorder. PLoS Comput. Biol. *8*, e1002709.

Van Roey, K., Uyar, B., Weatheritt, R.J., Dinkel, H., Seiler, M., Budd, A., Gibson, T.J., and Davey, N.E. (2014). Short linear motifs: ubiquitous and functionally diverse protein interaction modules directing cell regulation. Chem. Rev. *114*, 6733–6778.

Vogt, G., Chapgier, A., Yang, K., Chuzhanova, N., Feinberg, J., Fieschi, C., Boisson-Dupuis, S., Alcais, A., Filipe-Santos, O., Bustamante, J., et al. (2005). Gains of glycosylation comprise an unexpectedly large group of pathogenic mutations. Nat. Genet. *37*, 692–700.

Wang, P.I., and Marcotte, E.M. (2010). It's the machine that matters: Predicting gene function and phenotype from protein networks. J. Proteomics *73*, 2277–2289.

Wang, X., Wei, X., Thijssen, B., Das, J., Lipkin, S.M., and Yu, H. (2012). Three-dimensional reconstruction of protein networks provides insight into human genetic disease. Nat. Biotechnol. *30*, 159–164.

Wright, P.E., and Jane Dyson, H. (2014). Intrinsically disordered proteins in cellular signalling and regulation. Nat. Rev. Mol. Cell Biol. *16*, 18–29.

Yang, L., Embree, L.J., Tsai, S., and Hickstein, D.D. (1998). Oncoprotein TLS Interacts with Serine-Arginine Proteins Involved in RNA Splicing. J. Biol. Chem. *273*, 27761–27764.

Yue, P., Li, Z., and Moult, J. (2005). Loss of protein structure stability as a major causative factor in monogenic disease. J. Mol. Biol. *353*, 459–473.

Zauber, H., Kirchner, M., and Selbach, M. (2018). Picky: a simple online PRM and SRM method designer for targeted proteomics. Nat. Methods *15*, 156–157.

Zhong, Q., Simonis, N., Li, Q.-R., Charloteaux, B., Heuze, F., Klitgord, N., Tam, S., Yu, H., Venkatesan, K., Mou, D., et al. (2009). Edgetic perturbation models of human inherited disorders. Mol. Syst. Biol. *5*, 321.

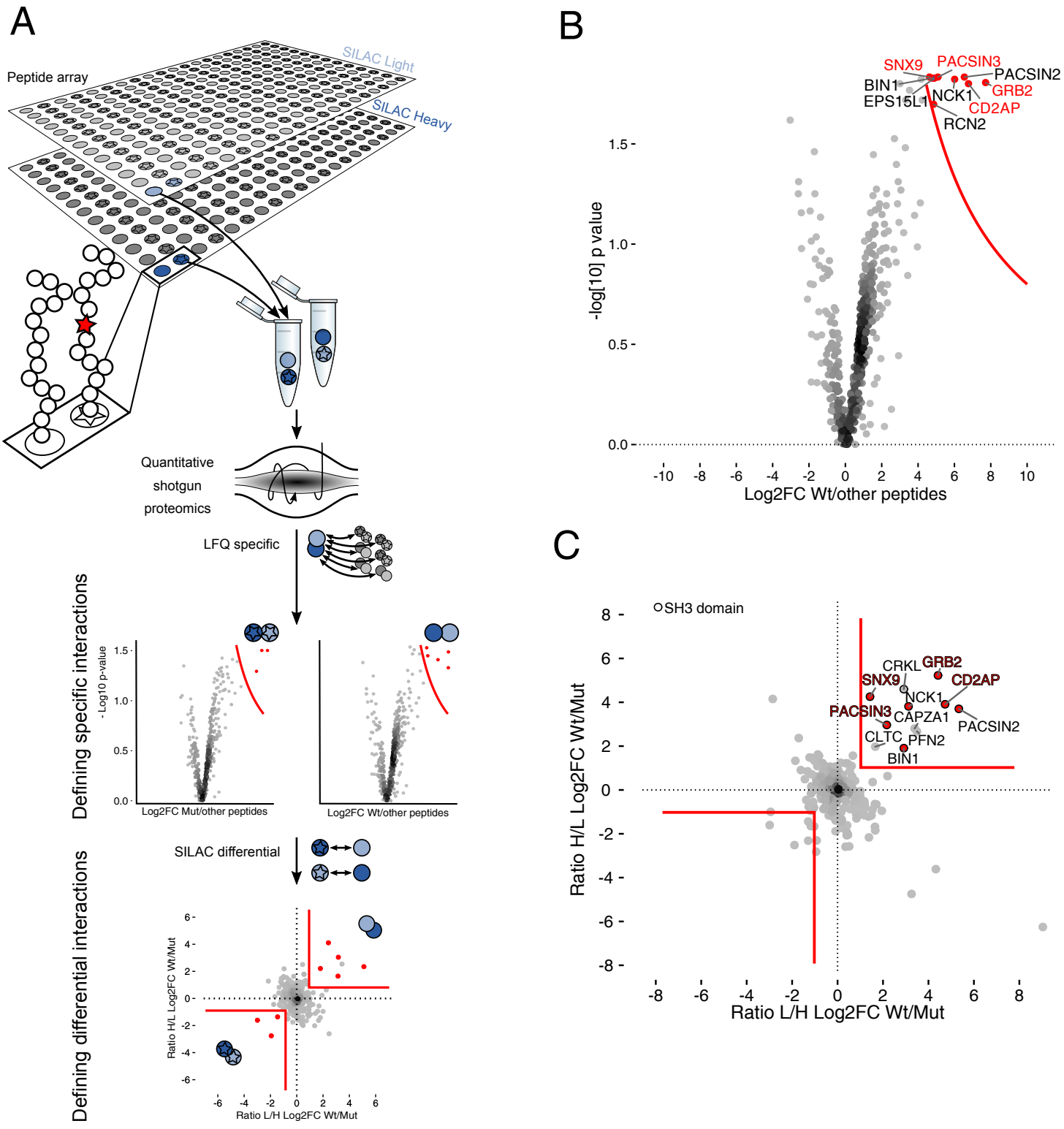
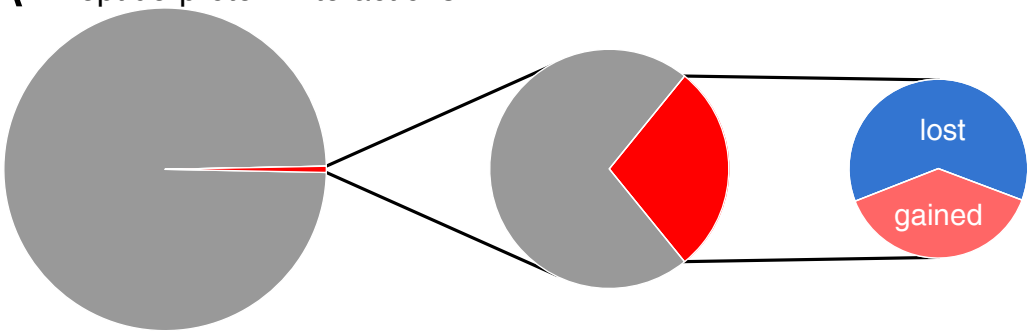

A
Peptide array
SILAC Light
SILAC Heavy
Quantitative
shotgun
proteomics
LFQ specific
Defining specific interactions
- Log10 p-value
Log2FC Mut/other peptides
Log2FC Wt/other peptides
SILAC differential
Defining differential interactions
Ratio H/L Log2FC Wt/Mut
Ratio L/H Log2FC Wt/Mut
B
SNX9
PACSIN3
PACSIN2
BIN1
NCK1
GRB2
EPS15L1
CD2AP
RCN2
-log[10] p value
Log2FC Wt/other peptides
C
SH3 domain
CRKL
GRB2
SNX9
CD2AP
NCK1
PACSIN3
CAPZA1
PACSIN2
CLTC
PFN2
BIN1
Ratio H/L Log2FC Wt/Mut
Ratio L/H Log2FC Wt/Mut

Fig. 1

A Peptide-protein interactions



101,765 detected interactions	618 specific interactions	180 differential interactions

B

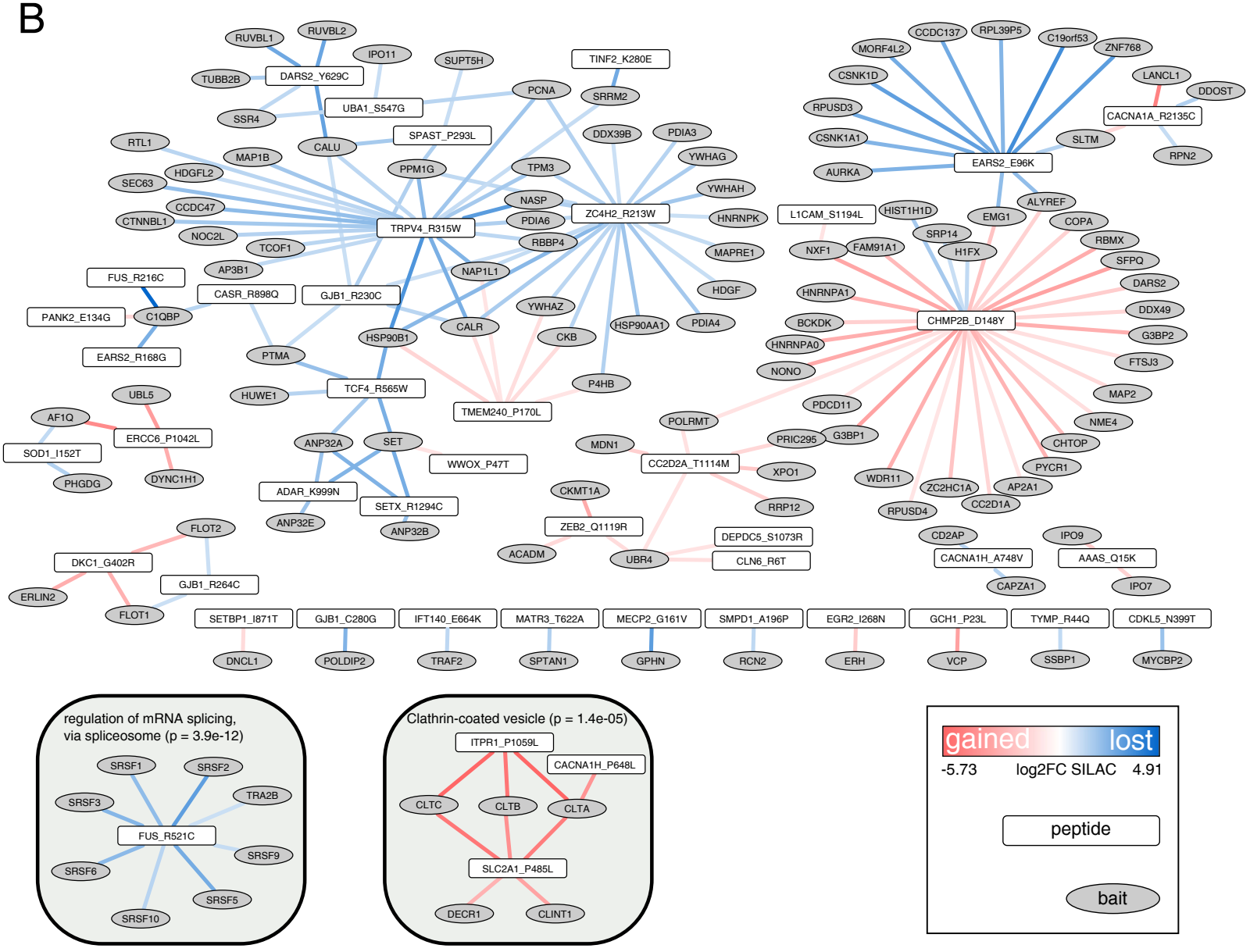


Fig. 2

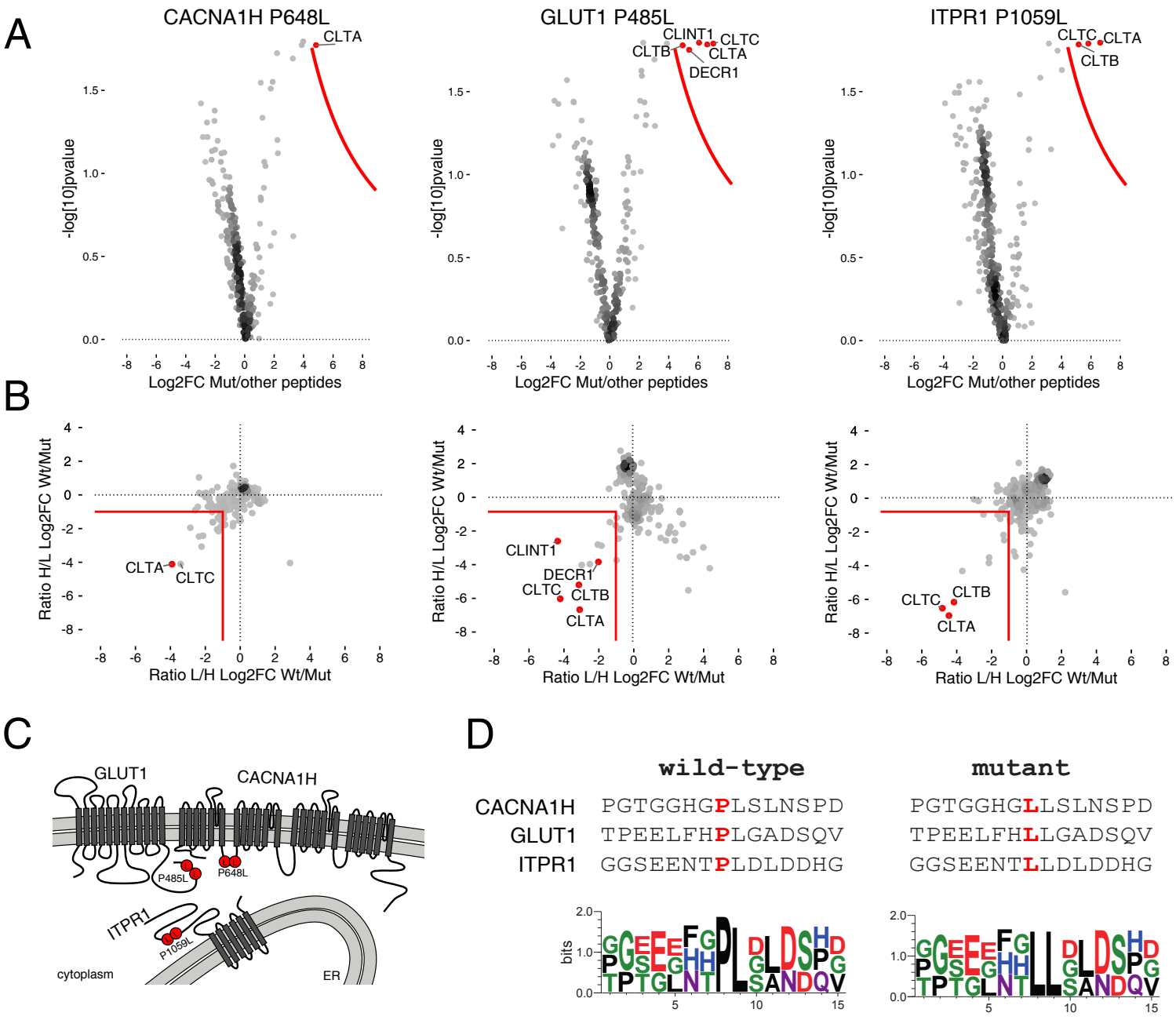
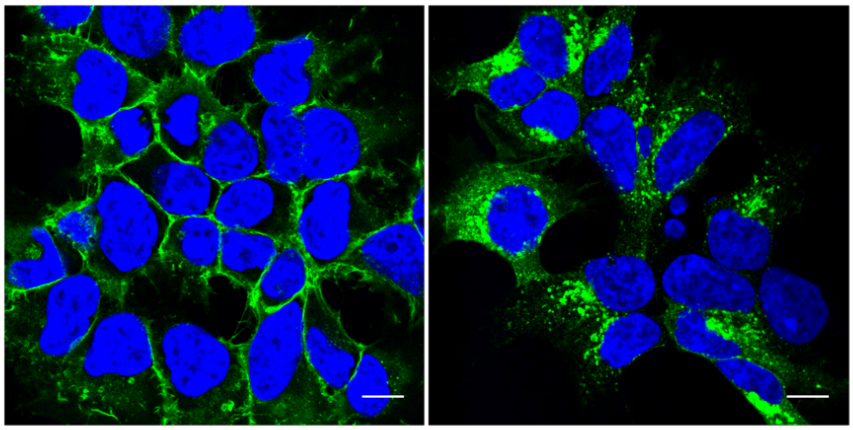
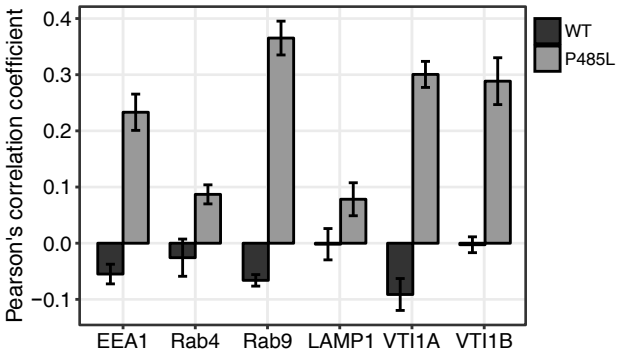

A
CACNA1H P648L
GLUT1 P485L
ITPR1 P1059L
CLTA
CLINT1
CLTC
CLTB
DECR1
-log[10]pvalue
Log2FC Mut/other peptides
B
Ratio H/L Log2FC Wt/Mut
Ratio L/H Log2FC Wt/Mut
C
GLUT1
CACNA1H
ITPR1
P485L
P648L
P1059L
cytoplasm
ER
D
wild-type
mutant
CACNA1H PGTGGHGPLSLNSPD PGTGGHGLLSLNSPD
GLUT1 TPEELFHPLGADSQV TPEELFHLLGADSQV
ITPR1 GGSEENTPLDLDDHG GGSEENTLLDLDDHG
bits

Fig. 3

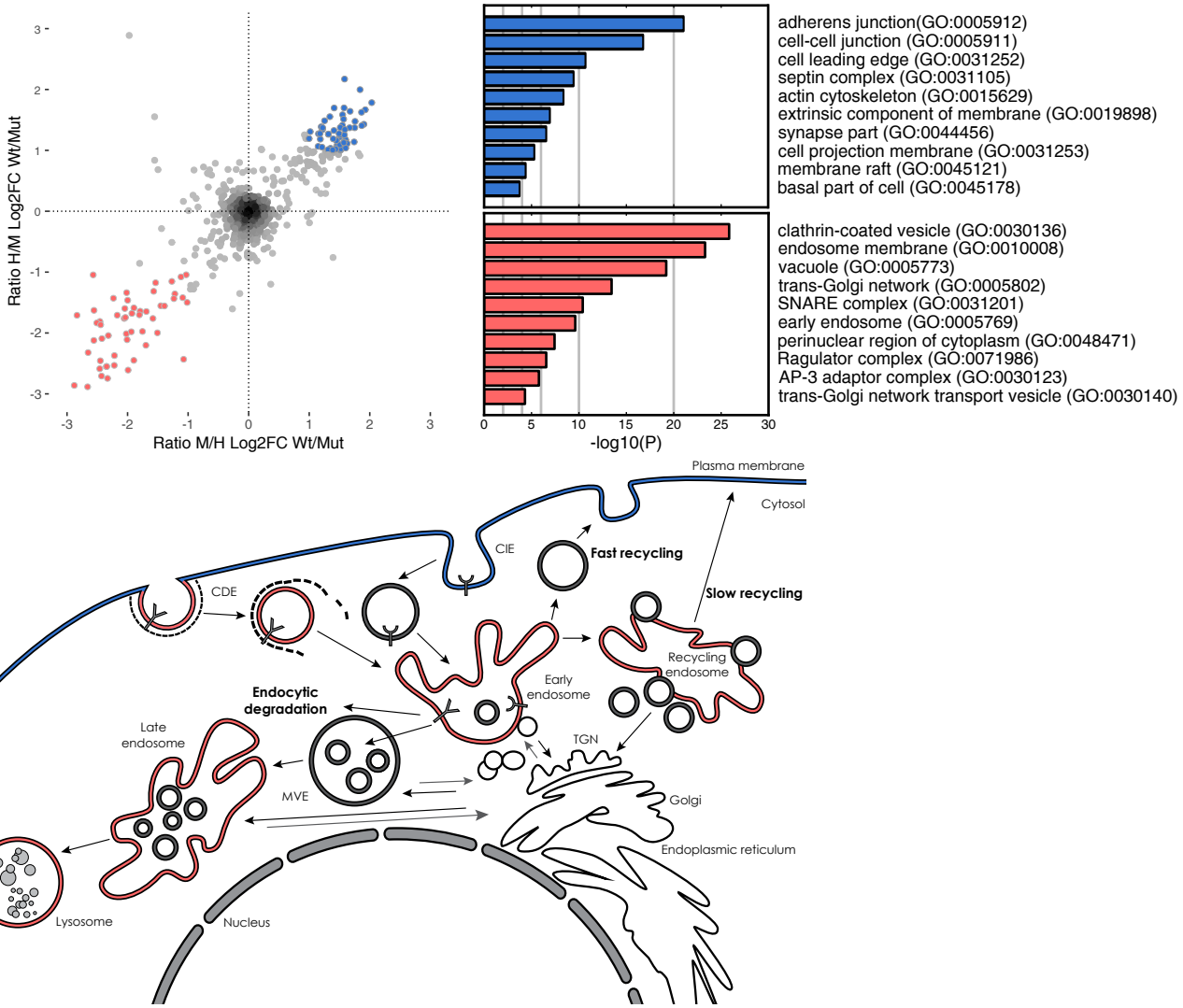
A


FLAG-GLUT1
FLAG-GLUT1_P485L

B


Pearson's correlation coefficient
0.4
0.3
0.2
0.1
0.0
−0.1
EEA1
Rab4
Rab9
LAMP1
VTI1A
VTI1B
WT
P485L

C


Ratio H/M Log2FC Wt/Mut
Ratio M/H Log2FC Wt/Mut
-log10(P)
adherens junction(GO:0005912)
cell-cell junction (GO:0005911)
cell leading edge (GO:0031252)
septin complex (GO:0031105)
actin cytoskeleton (GO:0015629)
extrinsic component of membrane (GO:0019898)
synapse part (GO:0044456)
cell projection membrane (GO:0031253)
membrane raft (GO:0045121)
basal part of cell (GO:0045178)
clathrin-coated vesicle (GO:0030136)
endosome membrane (GO:0010008)
vacuole (GO:0005773)
trans-Golgi network (GO:0005802)
SNARE complex (GO:0031201)
early endosome (GO:0005769)
perinuclear region of cytoplasm (GO:0048471)
Ragulator complex (GO:0071986)
AP-3 adaptor complex (GO:0030123)
trans-Golgi network transport vesicle (GO:0030140)

Plasma membrane
Cytosol
CIE
Fast recycling
Slow recycling
CDE
Recycling endosome
Early endosome
Endocytic degradation
Late endosome
TGN
MVE
Golgi
Endoplasmic reticulum
Lysosome
Nucleus

Fig. 4

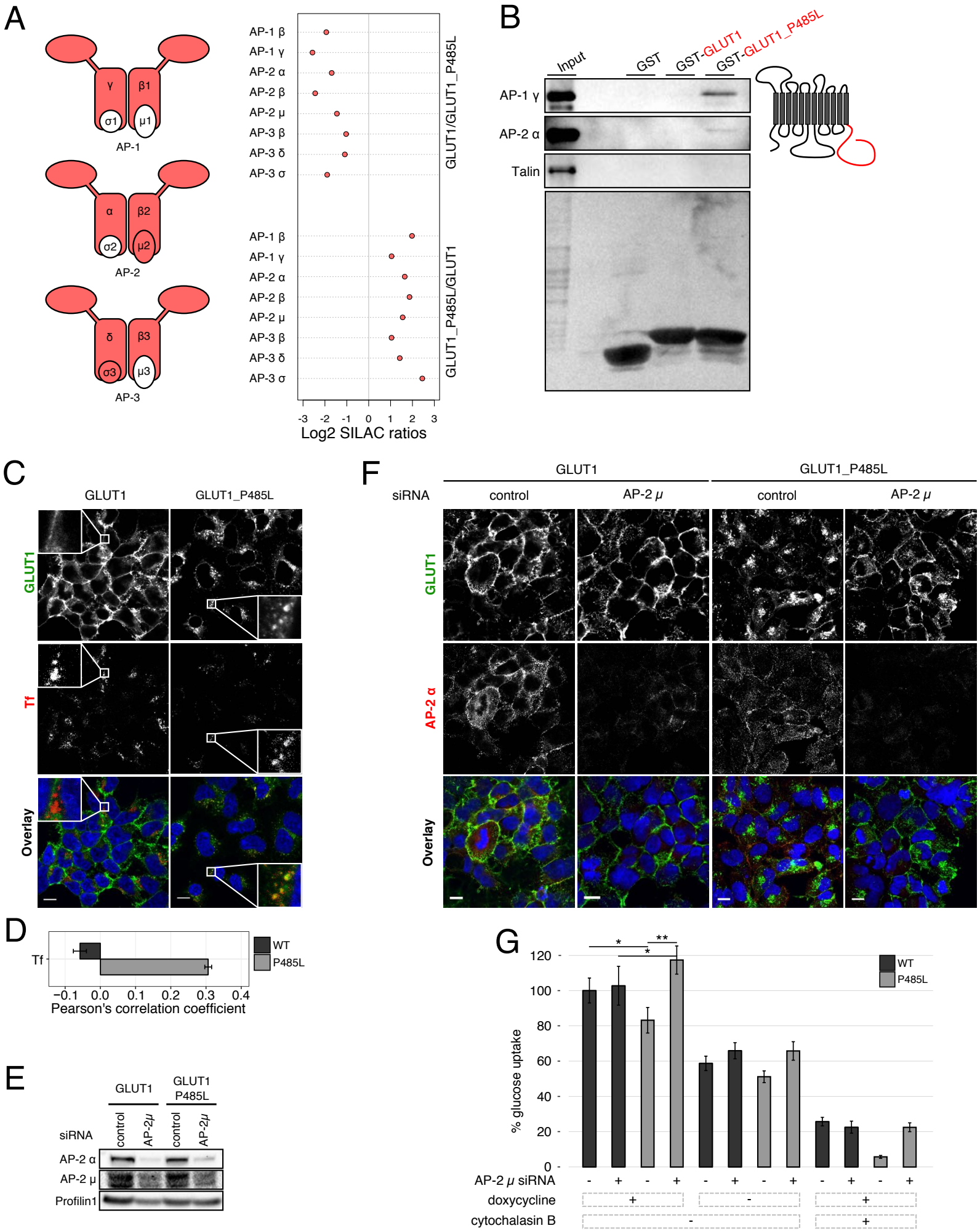

A
γ
β1
σ1
μ1
AP-1
α
β2
σ2
μ2
AP-2
δ
β3
σ3
μ3
AP-3
AP-1 β
AP-1 γ
AP-2 α
AP-2 β
AP-2 μ
AP-3 β
AP-3 δ
AP-3 σ
GLUT1/GLUT1_P485L
GLUT1_P485L/GLUT1
-3 -2 -1 0 1 2 3
Log2 SILAC ratios
B
Input
GST
GST-GLUT1
GST-GLUT1_P485L
AP-1 γ
AP-2 α
Talin
C
GLUT1
GLUT1_P485L
GLUT1
Tf
Overlay
F
siRNA
control
AP-2 μ
GLUT1
AP-2 α
Overlay
D
Tf
WT
P485L
−0.1 0.0 0.1 0.2 0.3 0.4
Pearson's correlation coefficient
E
GLUT1
GLUT1 P485L
siRNA
control
AP-2μ
AP-2 α
AP-2 μ
Profilin1
G
% glucose uptake
0 20 40 60 80 100 120
WT
P485L
AP-2 μ siRNA
doxycycline
cytochalasin B

Fig. 5

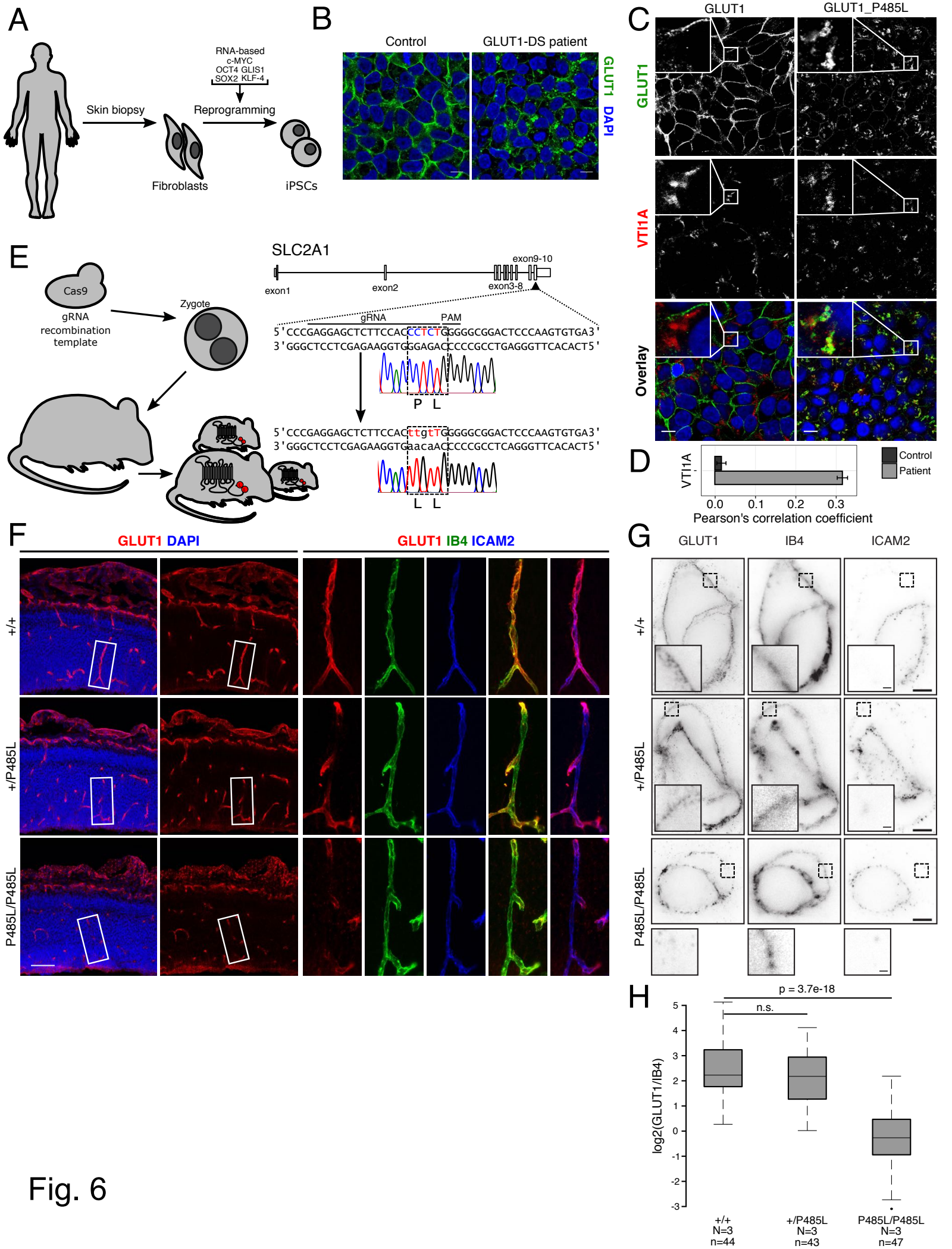

A
RNA-based
c-MYC
OCT4 GLIS1
SOX2 KLF-4
Skin biopsy
Reprogramming
Fibroblasts
iPSCs
B
Control
GLUT1-DS patient
GLUT1 DAPI
C
GLUT1
GLUT1_P485L
GLUT1
VTI1A
Overlay
D
VTI1A
Control
Patient
0.0
0.1
0.2
0.3
Pearson's correlation coefficient
E
Cas9
gRNA
recombination
template
Zygote
SLC2A1
exon9-10
exon1
exon2
exon3-8
gRNA
PAM
5'CCCGAGGAGCTCTTCCACCCTCTGGGGGCGGACTCCCAAGTGTGA3'
3'GGGCTCCTCGAGAAGGTGGGAGACCCCCGCCTGAGGGTTCACACT5'
P L
5'CCCGAGGAGCTCTTCCACttgtTGGGGGCGGACTCCCAAGTGTGA3'
3'GGGCTCCTCGAGAAGGTGaacaACCCCCGCCTCAGGGTTCACACT5'
L L
F
GLUT1 DAPI
GLUT1 IB4 ICAM2
+/+
+/P485L
P485L/P485L
G
GLUT1
IB4
ICAM2
+/+
+/P485L
P485L/P485L
H
p = 3.7e-18
n.s.
log2(GLUT1/IB4)
5
4
3
2
1
0
-1
-2
-3
+/+
N=3
n=44
+/P485L
N=3
n=43
P485L/P485L
N=3
n=47

Fig. 6

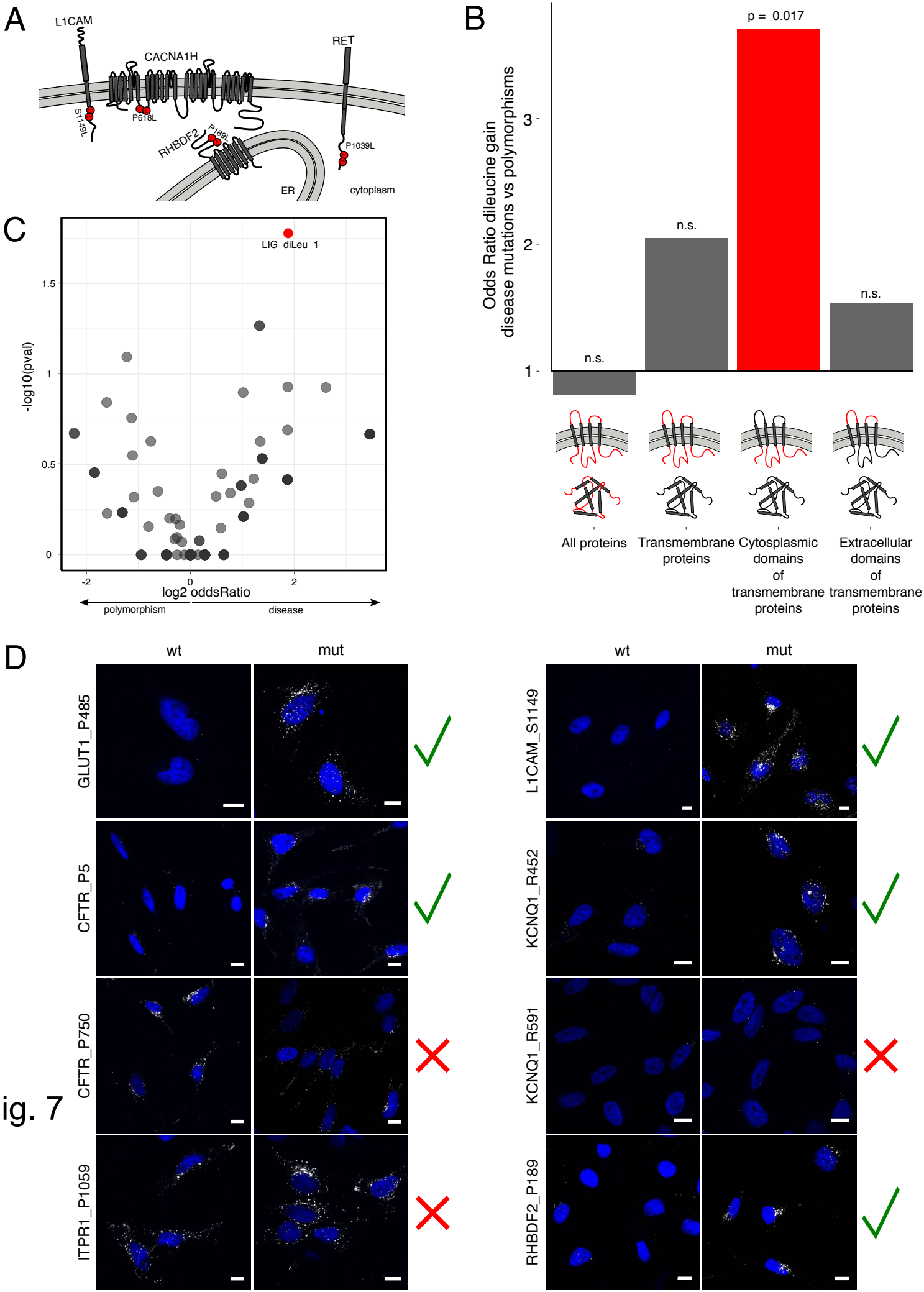

A
L1CAM
CACNA1H
RET
S1149L
P618L
RHBDF2
P189L
P1039L
ER
cytoplasm
B
p = 0.017
n.s.
n.s.
n.s.
Odds Ratio dileucine gain disease mutations vs polymorphisms
All proteins
Transmembrane proteins
Cytosplasmic domains of transmembrane proteins
Extracellular domains of transmembrane proteins
C
LIG_diLeu_1
-log10(pval)
log2 oddsRatio
polymorphism
disease
D
wt
mut
GLUT1_P485
CFTR_P5
CFTR_P750
ITPR1_P1059
L1CAM_S1149
KCNQ1_R452
KCNQ1_R591
RHBDF2_P189

Fig. 7

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Matthias Selbach (matthias.selbach@mdc-berlin.de).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell lines

SH-SY5Y and T-REx™-293 cells were cultured under standard cell culture conditions. In brief, cells were cultured in DMEM (Life Technologies) complemented with 10% fetal calf serum (Pan-Biotech).

Cells used for SILAC based experiments were cultured in SILAC DMEM (Life Technologies) complemented with glutamine (Glutamax, Life Technologies), Pyruvate (Life Technologies), non-essential amino acids (Life Technologies) and 10% dialyzed fetal calf serum (Pan-Biotech). The SILAC DMEM was supplemented with standard L-arginine (Arg0, Sigma-Aldrich) and L-lysine (Lys0, Sigma-Aldrich) (“light”) as in (Schwanhäusser et al., 2011). Alternatively, Arg6 and Lys4 (“medium-heavy”) or Arg10 and Lys8 (“heavy”) were added in place of their light counterparts. Cells were cultured at 37 °C and 5% CO_2.

Flp-In T-Rex GLUT1

We purchased SLC2A1 (GLUT1) from Harvard Plasmid repository (HsCD00378964). A stop codon has been added to the gene with the following primers Fw:TCCCAAGTGTAATTGCCAACTTTCTTGTACAAAGTTG, Rev:ATCAGCCCCCAGGGGATG.

P485L Mutation has been introduced by changing c.1454 C>T (Slaughter et al., 2009) with Q5® Site-Directed Mutagenesis Kit (NEB) Fw:CTGTTCCATCtCCTGGGGGCT, Rev:CTCCTCGGGTGTCTTGTCAC.

SLC2A1 and SLC2A1 mutant have been further cloned into a destination vector with a N-terminal BirA-FLAG Tag (pDEST-pcDNA5-BirA-FLAG N-term (Couzens et al., 2013))

with Gateway cloning strategy (Thermo Fisher Scientific). HEK293 Flp-In T-Rex cells (Invitrogen) that exhibit tetracycline-inducible expression of BirA-FLAG-GLUT1 or BirA-FLAG-GLUT1_P485L were generated using the Flp-In system developed by Life Technologies according to the manufacturer's protocol.

Patient-derived iPSCs

Fibroblasts were obtained from a GLUT1 deficient patient with the P485L mutation. The voluntary informed consent process was documented in writing as approved in advance by the University of Texas Southwestern Medical Center Institutional Review Board. This included information regarding the de-identification of the sample and the adherence to HIPAA regulations.

A 4 mm single-use, sterile skin punch was applied to the lateral surface of the left shoulder after the skin had been cleansed with iodine solution in aseptic fashion followed by injection of 0.5 mL of 1% unbuffered lidocaine with a vasoconstrictor. Prilocaine and lidocaine cream had been previously applied to the area. The punch was advanced by rotation under pressure and the explant was severed from its base and harvested in culture medium containing complete DMEM plus 20% fetal bovine serum and placed on ice until the explant was divided for culture the same day. The explant was divided into 12-15 evenly sized pieces and each piece maintained in a 10 cm dish at 37 °C until fibroblast confluence was reached. The cells were then treated with trypsin and passaged into a T-25 flask for further expansion. Fibroblasts were grown to approximately 50% confluence in the T-25 flask. They were then suspended with trypsin and frozen over dry ice in complete DMEM medium with 10% DMSO at a density of 10^6 cells/ml per vial prior to storage and shipment on dry ice.

The patient fibroblast were reprogrammed using the mRNA reprogramming kit ReproRNA™-OKSGM from Stem Cell Technologies according to the instructions. In brief, 1x 10^5 fibroblast cells were plated on Geltrex coated 6-well plate using regular DMEM media with 10% FBS. The day after the cells were transfected with the ReproRNA- OKSGM construct using the ReproRNA transfection reagents and growth Media with B18R. The next 5 days the growth media was changed every day and

supplemented with B18R and 0.8 µg/ml Puromycin. After 8 days the growth media was exchanged by ReproTeSR and first colonies appeared after 14 days. In total 5 clones were picked and established using mTESR-1 media. As a control for the experiments the following fibroblasts (NHDF-Ad-Der Fibroblasts, C-2511, LONZA) were reprogrammed using the Epi5™ Episomal iPSC Reprogramming Kit from Thermo Fisher Scientific following the vendor's instructions. The established lines and clones were registered and named using the Human pluripotent stem cell registry (https://hpscreg.eu/): BIHi037-(A-E).

The iPSCs used for the experiments were characterized using the PSC 4-Marker immunocytochemistry kit from Thermo Fisher Scientific following the instructions of the protocol. In addition to the 4 markers (OCT4, SOX2, TRA-1-60 and SSEA4) included in the kit, the expression of another pluripotency marker NANOG (Nanog PA1-097, Thermo Fisher Scientific) was analyzed.

Human iPSC cultures were maintained on plates coated with hESC-Qualified Matrigel™ (Corning) in mTESR-1 medium (Stem Cell Technologies) following the manufacturer's instruction. All cells were cultured at 37 °C in humidified atmosphere containing 5% O2 and 5% CO2. Cells were passaged using StemPro Accutase (Thermo Fisher) and replated in mTESR-1 medium w ith the addition of 10 µM ROCK inhibitor Y-27632 (LC Laboratories).

Animal model

Glut P485 mice were produced by microinjection of C57BL/6N zygotes with Cas9 protein (IDT), synthetic guide RNA (IDT) (5'GAGGAGCTCTTCCACCCTCT3') and a mutagenic single stranded deoxyoligonucleotide (IDT) (5'TAGCTGCCTGTGCTCCAGAGAGATCCTTGGGCTGCAGGGAGCAGGCCGGGCTGGGTGTGGGGCTCCTCACACTTGGGAGTCCGCCCCCAacaaGTGGAAGAGCTCCTCGGGTGTCTTGTCACTTTGG3') as recombination template, as described (Wefers et al., 2017) . Reagents were diluted in microinjection buffer (10 mM Tris, 0.1 mM EDTA, pH 7.2), filtrated through a centrifugal filter (Millipore, UFC30LG25) and stored in single use aliquots at -80°C. For microinjections, zygotes were obtained by mating of

C57BL/6N males with superovulated C57BL/6N females (Charles River, Sulzbach, Germany). Zygotes were injected into one pronucleus following standard procedures (Ittner and Götz, 2007). Injected zygotes were transferred into pseudo-pregnant NMRI female mice to obtain live pups. All mice showed normal development and appeared healthy. Mice were handled according to institutional guidelines under experimentation license no. G0162/12 approved by the Landesamt für Gesundheit und Soziales (Berlin, Germany) and housed in standard cages in a specific pathogen-free facility on a 12 h light/dark cycle with *ad libitum* access to food and water.

METHOD DETAILS

Peptide-protein interaction screen

Candidate selection

Disease mutations in humans were taken from UniProt annotations (UniProt Consortium, 2012) of Online Mendelian Inheritance in Man, OMIM®. McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University (Baltimore, MD), https://omim.org/. This dataset consists of experimentally validated missense mutations that contribute to inherited diseases. Inherited disease mutations were downloaded from UniProt (http://www.uniprot.org/docs/humsavar.txt, Release: 2015_07 of 24-Jun-2015, (Famiglietti et al., 2014)). Only mutations that were associated to 'Disease' were kept. 'Unclassified' mutations or 'Polymorphisms' were excluded. The 26,649 disease mutations were further filtered by applying a disorder cut-off. Disorder tendencies of 15 amino acids (AAs) long peptides, with the AA mutated in disease if possible located at position eight, were predicted using IUPred (Dosztányi et al., 2005) using the 'SHORT' profile considering sequential neighbourhood of 25 residues. IUPred disorder scores above 0.5 denote regions of the proteins that have 95% likelihood to be disordered. For filtering, the mean disorder score for all 15 AA as well as the mutation position were required to be >0.5. This resulted in 1,878 disease mutations in disordered regions. Next we assigned disease classes to 3,119 different diseases included in the humsavar

database by combining a manual approach with automatic annotation with the Human Phenotype Ontology database, HPO (Köhler et al., 2017). We selected 305 mutations causing neurological diseases. After manual inspection, we remained with 128 mutations causing 124 distinct neurological diseases that were used for the peptide-protein interaction screen.

Experimental setup

Peptides of 15 AAs, in total 128 wild-type peptide and 128 related peptides containing the disease causing mutation (256 peptides) plus one control peptide pair were synthesized *in situ* on cellulose membrane using PepTrack™ techniques (JPT Peptide Technologies, Berlin, Germany). Two of those peptide filters were moistened in cell lysis buffer [50 mM HEPES pH 7.6 at 4 °C, 150 mM NaCl, 1 mM EGTA, 1 mM MgCl2, 10% Glycerol, 0.5% Nonidet P-40, 0.05% SDS and 0.25% sodium deoxycholate, supplemented with protease inhibitor (Roche) and benzonase (Merck)]. In order to reduce nonspecific binding, the membrane was incubated with 1 mg/ml yeast t-RNA (Invitrogen) for 10 min and then washed twice with cell lysis buffer. The entire peptide libraries were incubated with 15 ml of light or heavy SILAC labeled cell lysate (5 mg/ml) from SH-SY5Y cells for 2 h. Membranes were washed three times and air dried.

Sample preparation for mass spectrometric analysis

Single spots were punched out from cellulose membrane with a 2 mm diameter ear punch (Carl Roth) and SILAC pairs were placed together in a 96-well plate (Thermo Fisher Scientific) prepared with 30 µl of denaturation buffer [6 M urea (Sigma-Aldrich), 2 M thiourea (Sigma-Aldrich), 10 mM HEPES, pH 8]. Samples were reduced by incubating with 10 µl of 3.3 mM DTT (Sigma-Aldrich) for 30 min at RT, followed by an alkylation step using 10 µl of 18.3 mM iodoacetamide (IAA) (Sigma-Aldrich) for 60 min at RT. The samples were first digested using 1 µg endopeptidase LysC (Wako, Osaka, Japan) for 4 h. The samples were diluted by adding 100 µl of 50 mM ammonium bicarbonate (pH = 8.5), and finally digested with 1 µg trypsin (Promega) for 16 h. The digestion was stopped by acidifying each sample to pH < 2.5 by adding 10% trifluoroacetic acid solution. The peptide extracts were purified and stored on stage tips

according to (Rappsilber et al., 2003).

LC-MS/MS analysis

Peptides were eluted using Buffer B (80% Acetonitrile and 0.1% formic acid) and organic solvent was evaporated using a speedvac (Eppendorf). Samples were diluted in Buffer A (5% acetonitrile and 0.1% formic acid). Peptides were separated on a reversed-phase column with 45 min gradient with a 250 nl/min flow rate of increasing Buffer B concentration on a High Performance Liquid Chromatography (HPLC) system (Thermo Fisher Scientific). Peptides were ionized using an electrospray ionization (ESI) source (Thermo Fisher Scientific) and analyzed on a Q-exactive plus Orbitrap instrument (Thermo FisherScientific). Dynamic exclusion for selected precursor ions was 30 s. The mass spectrometer was run in data dependent mode selecting the top 10 most intense ions in the MS full scans, selecting ions from 300 to 1700 m/z (Orbitrap resolution: 70,000; target value: 1,000,000 ions; maximum injection time of 120 ms). The resulting MS/MS spectra from the Orbitrap had a resolution of 17,500 after a maximum ion collection time of 60 ms with a target of reaching 100,000 ions.

Data analysis

The resulting raw files were analyzed using MaxQuant software version 1.5.2.8 (Cox and Mann, 2008). Default settings were kept except that ‘match between runs’ and ‘re-quantify’ was turned on. Lys8 and Arg10 were set as labels and oxidation of methionines and N-terminal acetylation were defined as variable modifications. Carbamidomethyl of cysteines was set as fixed modification. The *in silico* digests of the human Uniprot database (2015-12), a FASTA file containing all peptides used for pull-down and a database containing common contaminants were done with Trypsin/P. The false discovery rate was set to 1% at both the peptide and protein level and was assessed by in parallel searching a database containing the reversed sequences from the Uniprot database. Following statistics and figures were done using R (R version 3.2.1, RStudio Version 1.0.143). The resulting text files were filtered to exclude reverse database hits, potential contaminants, and proteins only identified by site. We imputed

missing LFQ-intensity values with random noise simulating the detection limit of the mass spectrometer (Keilhauer et al., 2015). To this end, imputed values are taken from a log normal distribution with 0.25× the standard deviation of the measured, logarithmized values, down-shifted by 1.8 standard deviations. In this way, we obtained a distribution of quantitative values for each protein across samples. For determination of specific interactions, two replicated pull-downs for the same peptide were tested against all other pull-downs, excluding the corresponding variant peptide, by the nonparametric Mann–Whitney U test. Resulting p-values and fold-changes (log2 space) have been plotted as volcano plots to determine cut-offs. We used an approach that uses a graphical formula to combination a fold-change and p-value cut-off (Keilhauer et al., 2015): $-log10(p) \geq \frac{c}{|x|-x_0}$ with x: enrichment factor of a protein, p: p-value of the Mann–Whitney U test calculated from replicates, x_0: fixed minimum enrichment, c: curvature parameter. The curvature parameter c determines the maximum acceptable p-value for a given enrichment x.

The parameters c and x_0 can be optimized based on prior knowledge of known true and false positives (Keilhauer et al., 2015). Here, cut-offs were chosen according to known interaction partners of the SOS1 control peptide (Keilhauer et al., 2015; Schulze and Mann, 2003). This resulted in x_0=0, c=8.

This cut-off was applied to all other pull-downs to separate specific binders from background. SILAC ratios were normalized by subtracting the median SILAC ratio of every experiment from all SILAC ratios in that experiment. To define interaction partners that bind differentially to wild-type and mutant peptide, a SILAC cut-off has been defined. For wild-type specific interaction partners, the mean log2 SILAC ratio of the two replicates needed to be >1 and none of the two ratios <0 (mutant specific mean log2 SILAC ratio < -1 and none of the two ratios >0). Resulting figures were modified in Inkscape (0.91).

PRM

LC-MS/MS analysis

Peptides were separated by reverse phase chromatography on an effective 150 min gradient (0, 2, 100, 30, 15, 1 and 5 min with 2, 4, 20, 30, 60, 90 and 90 % of buffer B with 90 % acetonitrile) and analyzed on a Q-Exactive HFx (Thermo Fisher Scientific). The PRM settings were: 30,000 resolution; 5e5 AGC target; 1.6 m/z isolation window; 60 ms max ion injection time. The inclusion list for the PRM method was generated using Picky (Zauber et al., 2018) with SILAC option enabled and a retention time window of 30 min. Predicted retention-times were calibrated in Picky with a complex sample of 100 ng Pierce HeLa Protein standard (Thermo Fisher Scientific) immediately before the PRM measurements.

Analysis of PRM data

Traces of all fragments from precursors in the spectral library (as exported from picky) were extracted from all rawfiles using the Thermo MSFileReader and the MSFileReader.py bindings written by François Allen. For each light or heavy scan the normalized spectral contrast angle (SCN) was calculated (Toprak et al., 2014). Peaks were manually selected and required a SCN > 0.4 and Fragment Matches > 4 in the light or heavy channel. Further Peaks needed to be within a similar retention time range across all different measurements. Ratios for each fragment using the maximum intensity of each peak were calculated. The median log2 transformed ratio (log2FC) for each peptide in each raw-file was calculated from selected fragment ratios: The five highest abundant fragments were selected from the peak with the highest detected SCN. Peptide log2FC were plotted as boxplot distributions in a protein centric manner across the different experiments (Fig. S4A).

BioID

Medium-heavy and heavy labelled T-REx™-293 cells have been induced for 24 h with 0.1 µg/ml Doxycycline to induce expression of GLUT1 (wild-type, wt) or GLUT1_P485L (mutant, mut). Light labelled cell lines from both GLUT1 and GLUT1_P485L have been

left uninduced and served as a control for background binding. SILAC labeling allowed for quantitative comparison of proteins that have been proximity labelled by the transiently expressed constructs (Forward experiment: Light - Control, Medium-heavy - wt, Heavy - mut; Label swap experiment: Light - Control, Medium-heavy - mut, Heavy - wt). During the induction period all cell lines have been incubated for 24 h in cell culture medium containing biotin. BioID experiment has been performed essentially as in (Couzens et al., 2013), with minor adaptations.
Mass spec setup and analysis was done similarly as to samples from peptide pull-downs, but on bead digested peptides were separated on a 2,000 mm monolithic column with a 100 µm inner diameter filled with C18 material that was kindly provided by Yasushi Ishihama (Kyoto University) using a 4 h linear gradient with a 300 nl/min flow rate of increasing Buffer B concentration on a High Performance Liquid Chromatography (HPLC) system (Thermo Fisher Scientific). The resulting raw files were analyzed using MaxQuant software version 1.5.2.8 (Cox and Mann, 2008). Default settings were kept except that ‘match between runs’ and ‘re-quantify’ was turned on. Lys4 and Arg6 or Lys8 and Arg10 were set as labels and oxidation of methionines and n-terminal acetylation were defined as variable modifications. Carbamidomethyl of cysteines was set as fixed modification. The *in silico* digests of the human Uniprot database (2015-12), a fasta file containing the sequence of BirA-FLAG-GLUT1 and a database containing common contaminants were done with Trypsin/P. The false discovery rate was set to 1% at both the peptide and protein level and was assessed by in parallel searching a database containing the reversed sequences from the Uniprot database.
Biotinylated proteins with a wild-type to mutant enrichment ratio (log2FC) >1 or <-1 have been considered as significant. These proteins have been analysed for gene ontology enrichment of cellular components with http://metascape.org (Tripathi et al., 2015).

FLAG-GLUT1 localization

HEK293 Flp-In T-Rex cells with BirA-FLAG-GLUT1 or BirA-FLAG-GLUT1_P485L have been seeded on coverslips coated with poly-l-lysine (Sigma-Aldrich). After induction for 24 h in doxycycline (0.1 µg/ml) containing media, cells were fixed with 4% PFA

(paraformaldehyde). Standard procedures were used for immunostaining. Cells have been stained against FLAG 1:200 (F1804, Sigma-Aldrich). Nucleus has been stained with DAPI (Sigma-Aldrich). FLAG staining was accompanied by staining to one of the following endosomal markers and with the following dilutions: anti-EEA1 (Cell Signaling Technology, 1:100); anti-Rab4 (Cell Signaling Technology, 1:100); anti-Rab9 (Cell Signaling Technology, 1:100); anti-LAMP1 (Cell Signaling Technology, 1:100). Mouse anti-FLAG staining was substituted by rabbit anti-GLUT1 (Merck Millipore, 1:500) to costain mouse monoclonal antibodies: anti-VTI1A (BD Biosciences, 1:100); anti-VTI1B (BD Biosciences, 1:100). Secondary antibodies all come from Invitrogen. For colocalization analysis three z-stacks of 5-10 cells each have been quantified for each marker with Imaris v8.4.1 (see “Quantification and statistical analysis” for details).

Transferrin uptake

Essentially as in “FLAG-GLUT1 localization”. Additionally, after 24 h cells were serum-starved for 1 h and used for Transferrin (Tf) uptake. For Tf uptake, cells were treated with 10 µg ml^{-1} Tf-Alexa568 (Life Technologies) for 10 min at 37 °C. For colocalization analysis three z-stacks of more than 15 cells each have been quantified with Imaris v8.4.1 (see “Quantification and statistical analysis” for details).

FLAG-GLUT1 localization under AP-2 µ knock-down

To rescue the GLUT1_P485L phenotype, clathrin mediated endocytosis (CME) has been inhibited by knocking down AP-2 µ and hence the adaptor complex responsible for recognition of cargo for CME.

On day 1, cells were seeded in 6-well plates. On day 2, cells were transfected with 25 nM final siRNA concentration (AP-2 µ: ON-TARGETplus Human AP2M1 (Dharmacon) and non-target: ON-TARGETplus Non-targeting Pool (Dharmacon)) according to DharmaFECT (Dharmacon) transfection protocol. 24 h after the transfection, medium was replaced with complete medium to reduce cytotoxicity and incubated for another 24 h. On day 4, siRNA transfection was repeated as described for day 2. On day 5, cells have been seeded in a 24-well plate onto coverslips coated with

poly-L-lysine (Sigma-Aldrich) for microscopy and into a 6-well plate for western blot analysis. Doxycycline (0.1 µg/ml) has been added to the medium to induce expression of the GLUT1 constructs. After induction for 48 h, cells in 24-well plates were fixed with 4% PFA. Standard procedures were used for immunostaining. Cells were stained with rabbit polyclonal GLUT-1 antibody 1:200 (Merck Millipore) and co-stained with mouse monoclonal anti-alpha adaptin antibody [AP6] 1:200 (Abcam). Secondary antibodies with Alexa fluorophores have all been purchased from Invitrogen. Nucleus has been stained with DAPI (Sigma-Aldrich).
Lysate from cells in 6-well plates has been used for western blotting, α and µ2 subunits of AP-2 were detected using mouse monoclonal antibodies from Thermo Fisher Scientific and BD transduction, respectively. Profilin 1 was stained as a loading control with polyclonal rabbit antibody from CST. Horseradish peroxidase coupled secondary antibodies were purchased from GE Healthcare. Proteins were detected with chemiluminescence substrate (Perkin Elmer) on a ChemiDoc MP Imaging System (Bio-Rad) and quantified with Image Lab 5.2.1.

Antibody feeding assay

An antibody feeding assay was used to study the gain of endocytosis by gain of dileucine motifs. For antibody internalization assay, genes and cytoplasmic regions were chosen according to the following criteria: All 11 disease mutations from Humsavar and Clinvar (‘Pathogenic’ or ‘Conflicting interpretations of pathogenicity’, in case ‘pathogenic’ or likely pathogenic was included in the different interpretations) that lead to a gain of a dileucine motif have been considered. All regions +/-7 AAs of the mutation have been analysed according to Eukaryotic Linear Motif (ELM) database (Dinkel et al., 2016). CACNA1H_P618L and RET_P1039L were not considered for the assay since wild-type variants of the peptides already harbour trafficking motifs. For GLUT1_P485L the whole cytoplasmic C-terminus was amplified via PCR adding EcoRV 5’ and NotI 3’. All other seven constructs were generated by inserting the region surrounding the mutation position with Q5® Site-Directed Mutagenesis Kit (NEB) resulting in a 15 AA insert (we were not able to generate a construct for CACNA1H_P648L). Chimeras consisting of one of the cytoplasmic regions and the

human TAC antigen (interleukin-2 receptor α chain, CD25) were constructed based on a TAC construct (Diril et al., 2009). HeLa cells were transiently transfected with the TAC chimera constructs using jetPRIME (Polyplus-transfection). Two days after transfection, cells were labelled with anti-TAC IgG (Santa Cruz Biotechnology) (1 : 1000 diluted in Opti-MEM; Invitrogen) for 30 min at 4 °C. After one change of medium (to Opti-MEM at 37 °C), plasma membrane antigens were allowed to internalize for 30 min at 37 °C. The cells were then fixed with 4% paraformaldehyde (Sigma-Aldrich) for 10 min on ice, and surface-bound TAC antibody was blocked using goat anti-mouse serum [goat anti-mouse IgG (Thermo Fisher Scientific at a 1 : 5 dilution in goat serum dilution buffer, consisting of 30% normal goat serum (Sigma-Aldrich), 450 mm NaCl in 20mM sodium phosphate buffer pH 7.4] for 2 h at room temperature. Cells were permeabilized and blocked with goat serum dilution buffer containing 0.2% saponin for 10 min. For detection of internalized TAC antibody, a goat anti-mouse Alexa Fluor 488-conjugated IgG (Invitrogen) was added for 1 h. Cells were then washed three times for 10 min each with sodium phosphate buffer containing 0.02% saponin. For total TAC staining, the specimens were incubated for 1 h with TAC antibody diluted 1 : 1000. As secondary antibody, an Alexa Fluor 594-conjugated goat anti-mouse IgG (Invitrogen) was added for 30 min, and nuclei were stained using DAPI (Sigma-Aldrich). Cells were washed, and coverslips were mounted in ProLong® Gold antifade reagent (Invitrogen). For imaging, cells with positive signal in the 594 channel were chosen. All cells in Fig. 7 D are positive for total TAC staining (594). We have seen that the level of internalised TAC chimera (488) does not correlate with the amount of total TAC staining (594), between samples and in the same sample, and hence we have decided to exclude this channel from visualization.

GLUT1 localization in iPSCs

Human iPSCs were seeded on coverslips coated with hESC-Qualified Matrigel™ (Corning). Cells were fixed with 4% PFA, stained with rabbit polyclonal GLUT1 antibody 1:200 (Merck Millipore) and costained with mouse monoclonal VTI1a antibody 1:100 (BD Biosciences). Secondary antibodies with Alexa fluorophores have all been purchased from Invitrogen.

Fluorescence microscopy from cell culture

Images from FLAG-GLUT1 localization were acquired by Leica DMI6600 confocal laser scanning microscope with an HCX PL APO 63.0/1.40 oil objective. Transferrin uptake, GLUT1-localisation under AP-2 µ knock down, antibody feeding assay and GLUT1 in iPSCs were acquired by a Zeiss LSM 700 confocal laser scanning microscope with an EC Plan-Neofluar/NA1.3 40x oil objective or a EC Plan-Apochromat/NA1.4 63x oil objective. Images were further processed with Fiji (Schindelin et al., 2012). For colocalization analysis see “Quantification and statistical analysis”.

Immunofluorescence in mouse tissue

E14-E15.5 embryos were obtained by Caesarian section from pregnant dam on day 14-15.5 post-coitus. Whole-mount embryos were dissected in ice-cold phosphate buffer and fixed for 2 h with a solution of 4% PFA in ice-cold phosphate buffer and cryoprotected overnight in 30% sucrose in phosphate buffer at 4°C. Whole embryo heads were sectioned in a horizontal plane using a cryostat to obtain 12-16µm sections.

Sample preparation for confocal microscopy

Essentially as in (Hernández-Miranda et al., 2011). Brain sections were incubated in blocking buffer 1 (5% horse serum and 0.1% Triton™-X 100 made in phosphate buffer) for 1 h at room temperature. Then, sections were incubated overnight in blocking buffer 1 containing the following antibodies: rabbit anti-Glut1 (1:200; Merck Millipore #07-1401), rat anti-ICAM2 (1:100; BD Biosciences #553326) and Isolectin GS-IB4 Alexa Fluor 488 conjugate (1:100, Thermo Fisher Scientific #I21411) at 4°C. Next, sections were washed three times in ice-cold phosphate buffer and incubated for 3 h in blocking buffer 1 containing Cy3 horse anti-rabbit (1:500; Jackson Lab), Cy5 horse anti-rat (1:500; Jackson Lab) and DAPI at room temperature. Fluorescence was imaged on a Zeiss LSM 700 (Jena, Germany) confocal microscope in a non-blind manner.

Sample preparation for STED microscopy

Sections were washed twice for 5 minutes with PBS to remove the embedding resin and incubated in 0.2% Triton X-100 in blocking buffer 2 (1% bovine serum albumin, 1% fetal calf serum in PBS) for 1 hr at room temperature. Samples were incubated with primary antibodies anti-Glut1 (rabbit anti-human, Merck Millipore #07-1401) and anti-ICAM2 (rat anti-mouse CD102, BD Biosciences #553326) at 1:100 dilution in blocking buffer 2 overnight at 4°C. Next, samples were washed three times for 5 minutes in PBS and incubated with STAR Red goat anti-rabbit antibody (Abberior, #2-0012-011-9), Alexa Fluor 594 donkey anti-rat antibody (Thermo Fisher Scientific #A21209), and Isolectin GS-IB4 Alexa Fluor 488 conjugate (Thermo Fisher Scientific #I21411), all diluted at 1:500 in blocking buffer 2. Subsequently, samples were washed three times for 5 minutes in PBS and mounted in Abberior Mount Solid Antifade mounting reagent (Abberior, #4-0100-007-4) under #1.5 coverslips (22x50 mm, VWR #631-0138) and allowed to cure overnight at room temperature.

STED imaging and image analysis

IB4 and ICAM2 signals were used to assess the positions of luminal (IB4 and ICAM2 positive) and abluminal (IB4 positive, ICAM2 negative) vessel membranes. Cross-sections of vessels (10-20 per animal) were selected for imaging in areas where luminal and abluminal membranes were clearly distinguishable, typically in the vicinity of the endothelial cell nucleus.

STED images were acquired using Abberior STED microscope equipped with 640 nm, 561 nm and 485 nm pulsed excitation lasers, 775 nm and 595 nm pulsed depletion lasers, UPlanSApo 100x/1.40 Oil objective (Olympus), 509/22 (GFP), 605/50 (Cy3) and 685/70 (Cy5) bandpass emission filters and spatial light modulators for STED beam shaping and alignment. Emitted light was collected with avalanche photodiode detectors using 8 ns-wide detection time gates. 120 µs total pixel dwell time per channel was used. Cy3 and Cy5 channels were acquired by line switching followed by the acquisition of the GFP channel.

STED images in 488 nm and 640 nm channels were aligned using reference images of fluorescent beads (Tetraspeck, 100 nm, ThermoFisher Scientific #T7280). Centers of

beads were determined by centroid fit and resulting positions were used as control points to calculate an affine transformation between the 488 nm and 640 nm channels. To quantify the average amount of membrane-localized Glut1 per vessel, a measurement area containing all pixels within 300 nm of manually segmented abluminal membrane was created. Luminal membranes were not included in the analysis due to frequent collapse of vessels during sample preparation. The ratio between mean Glut1 and mean IB4 signal was used as a measure of Glut1 to account for the amount of membrane in the measurement area, imaging depth and antibody penetration differences between samples. The ratio of Glut1 to IB4 was further corrected using images of Tetraspeck beads for relative intensity fluctuations between Cy5 and GFP detection channels between imaging sessions. Statistical significance was assessed using unpaired Student's t-tests of log2 transformed data. The analysis was performed using ImageJ and Matlab 2015 (Mathworks, Inc).

Radioactive glucose uptake under AP-2 µ knock-down

AP-2 knock down has been performed as described before (FLAG-GLUT1 localization under AP-2 µ knock-down). Only that cells have been seeded in triplicates in a 24-well plate without coverslips. Radioactive glucose uptake was performed mainly as in (Shi and Kandror, 2008). Radioactive glucose cocktail was prepared by adding 10 µL of 3H-2-deoxy-D-glucose in ethanol:water solution (specific activity, 5–10 Ci (185–370 GBq) / mmol) (Perkin Elmer) to a 2.0-mL tube and left open for 5 min to evaporate ethanol. 1.6 mL of KRH(−) glucose and 16 µL of cold 2-DOG (100X) stock solution (100 mM 2-deoxy-D-glucose in KRH (−) glucose) (Sigma-Aldrich) were added to the tube. Cells in each well were rinsed with DMEM (without serum, SFM) warmed to 37°C and SFM was added to cells slowly and carefully by the side of the well in order to avoid detachment of cells. Cells were incubated with 0.5 mL of SFM (in case of +dox containing 1µg/µl doxycycline) per well for 2 h at 37°C. Cells in each well were washed twice with 2 mL of KRH(−) glucose buffer (121 mM NaCl, 4.9 mM KCl, 1.2 mM MgSO4, 0.33 mM CaCl2, 12 mM HEPES, pH 7.4) at 37°C. 225µl of KRH(-) containing 25 µM final Cytochalasin B (dissolved in DMSO) or 0.5% DMSO were added to each well. Immediately after, 25 µL of radioactive glucose cocktail was added to all wells. Samples were incubated at 37°C

for 1 h and then transferred on ice. Radioactive glucose cocktail was aspirated, and ice-cold KRH (+) glucose (121 mM NaCl, 4.9 mM KCl, 1.2 mM MgSO4 0.33 mM CaCl2:12 mM HEPES, 25 mM D- (+)- Glucose, pH 7.4.) was added to terminate the reaction. Cells were washed once more with ice-cold KRH(+) glucose. Plate was transferred to room temperature, and 400 µL of 0.1% SDS in KRH (−) glucose have been added to each well, incubated at room temperature for 10 min, and thoroughly resuspend to homogeneity. 100 µl of the lysate was kept to measure protein concentration with DC protein assay kit I (BioRad). 300 µL of lysates have been transferred in scintillation vials containing 4 mL of Rotiszint eco plus scintillation fluid (Carl Roth) and count in a Liquid Scintillation Analyzer (Tri-Carb 2800TR, PerkinElmer) for 1 min per vial. These numbers represent “Counts in the samples”. In parallel, 10 µL of the radioactive glucose cocktail have been mixed with 290 µL of 0.1% SDS in KRH (−) glucose and this mixture was measured under the same conditions. This number represents “Counts in the cocktail”. The amount of intracellular 2-deoxyglucose was calculated using the following formula: $\left[\frac{[Counts\ in\ the\ sample]\times 1000}{[Counts\ in\ the\ cocktail]\times 0.03\times[C]\times t}\right]$pmol/mg×min, where [C] is protein concentration in mg/ml and t is the total time of incubation with radioactive glucose in min. All resulting values have been divided by the overall mean value from all wild-type GLUT1 (+ doxycycline, -cytochalasin B) and multiplied by 100 to receive relative values for glucose uptake (%). For test of statistical significance, the mean values of three technical replicates were calculated from three biological replicates. Values labeled with an asterisk were tested for statistical significance in a paired, one-tailed t-test. Depicted values are mean values over all replicates and error bars show standard error of mean (SEM) over all replicates.

GST pulldown assay

GLUT1 cytoplasmic c-terminal tail has been amplified from pDEST_pcDNA5_FLAG_BirA GLUT1 or GLUT1P485L with Fw: tatatcGAATTCGTTCCTGAGACTAAAGGC, Rev: aacaatGCGGCCGCTTACACTTGGGAATCAGCC. This resulted in C-terminal tail amino acids 451-492 (UniProt P11166). Added EcoRI and NotI restriction sites have been used to insert the PCR product into pGEX6P1.

Other cytoplasmic regions have been ordered as gBlocks Gene Fragments (IDT) from the region +/- 20AA of the mutation position, with an additional 5'- EcoRV restriction site and 3'- Stop codon - NotI restriction site. After restriction, the gene fragments have been inserted into pGEX6P2.

Expression of GST-tagged proteins was induced for 5 h at 22 °C by addition of isopropyl thio-β-d-galactoside (0.5 mM) to E.Coli BL21 in 2X YT medium (0.8 OD). To lyse the cells, bacterial pellets were resuspended in PBS and left on ice for 15 min in presence of PMSF (1mM), cyanase (4U/µL) and lysozyme (1 mg·mL−1). Then, Triton X-100 was added to a 0.5% and cells were sonicated for 2 min. Lysates were centrifuged for 15 min at 50 000 x g. 300 µL of glutathion-coupled beads were added to the supernatant and rotated end-over-end for 2 h at 4 °C. Beads were washed three times with PBS / 0.1% Triton X-100 and once with PBS.

Pulldown experiments were performed using mouse brain extracts. Mouse brains were homogenized in buffer (20 mM HEPES, 320 mM sucrose, pH 7.5) containing protease inhibitor cocktail (Sigma-Aldrich). The homogenate was centrifuged at 1000 x g for 10 min and the supernatant was supplemented with 1% Triton X-100, 50 mM KClMgCl2, and kept on ice for 10 min with occasional vortexing. The lysate was cleared by centrifugation at 17,000 x g for 15 min and at 178,000 x g for 15 min. The supernatant was recovered and used at a concentration of 7.5 mg protein/mL.

The pulldown experiments were performed using 85 µg of GST fusion proteins and 0.6 mL protein extract by end-over-end rotation for 3 h. The samples were washed four times with buffer containing 20 mM Hepes, 50 mM KCl, 2 mM MgCl2, Triton X-100 (1%) and once in the same buffer without detergent. Proteins were eluted from the beads twice with Laemmli buffer and analysed by Western blotting. The following antibodies and dilutions were used: mouse anti-talin 1:1000 (Sigma-Aldrich), mouse anti-γ1 adaptin of AP-1 1:500 (BD Biosciences), mouse anti α-adaptin of AP-2 1:200 (BD Biosciences), horseradish peroxidase-conjugated goat anti-mouse 1:2000 or 1:5000 (Jackson labs).

Analysis of human missense variants and short linear motifs (SLiMs)

SLiM regular expression patterns

262 annotated SLiM class definitions (regular expression patterns) were downloaded from the Eukaryotic Linear Motif (ELM) database (Dinkel et al., 2016). In order to analyse dileucine motifs, an additional motif '.LL.' was added to this compilation and named 'LIG_diLeu_1' in order to conserve the naming convention followed by the ELM database.

Pathogenic and non-pathogenic missense variants

Humsavar dataset: For the analysis of the missense variants that lead to *de novo* SLiM instances in protein sequences **Uniprot Humsavar dataset** (version 12-Apr-2017) (Famiglietti et al., 2014) was downloaded and filtered for missense variants. Variants that are classified as 'Disease' or 'Polymorphism' in this dataset were selected.
ClinVar dataset: Clinically relevant genomic variation data annotated in the ClinVar database (Landrum et al., 2016) was downloaded from the ftp server (ftp.ncbi.nlm.nih.gov/pub/clinvar/tab_delimited/variant_summary.txt.gz) in tab-delimited format (latest update on 25th of March, 2017). The downloaded table was filtered for assembly version GRCh38, and variants of type 'single nucleotide variant' were kept. In order to integrate the ClinVar annotations with other kinds of annotations available from the Uniprot database, these nucleotide variants were translated to the Uniprot protein sequences to obtain single amino-acid substitutions using the Ensembl Variant Effect Predictor (version 82) (VEP) (McLaren et al., 2016). The output of VEP tool was filtered to only keep missense variants such that the translated amino-acid substitution occurs at exactly the same amino-acid at the same position of the Uniprot sequence with the same gene name as those of the annotation in the ClinVar dataset ('Name' field). Thus, 98,219 unique single amino-acid substitutions (missense variants) from 4,298 Uniprot sequences were obtained. Variants primarily annotated with clinical significance levels 'Pathogenic', 'Pathogenic/Likely pathogenic', or 'Likely pathogenic' were grouped as

‘Disease’ variants, while variants annotated with ‘Benign’, ‘Benign/Likely benign’, or ‘Likely benign’ were grouped into the ‘Polymorphism’ variants.

Protein domains

PFAM domain annotations of proteins were downloaded from the PFAM database (ftp://ftp.ebi.ac.uk/pub/databases/Pfam//releases/Pfam30.0/proteomes/9606.tsv.gz) (Finn et al., 2016).

SLiM - PFAM associations

PFAM domains and SLiM classes that are known to interact were downloaded from the ELM database (http://elm.eu.org/interactiondomains).

Analysis of gain of SLiMs via missense variants in disordered regions

For each reviewed human protein from Uniprot (20191 proteins), the disorder scores of each residue were calculated using IUPred (using the ‘short’ setting). Using a IUPred disorder score cut-off of 0.4, the missense variants in disordered regions were selected. The missense variants that overlap PFAM domains were further filtered out based on the PFAM domain annotations found in the protein feature files downloaded from Uniprot in GFF format (e.g. the link to the GFF file for GLUT1 is http://www.uniprot.org/uniprot/P11166.gff). These protein feature files were also used to detect the transmembrane proteins and their cytoplasmic/extracellular regions. The missense variants in disordered regions and not overlapping any PFAM domains were further classified as variants from 1) the whole proteome, 2) the transmembrane proteins (only those that have annotation of at least one cytoplasmic domain or an extracellular domain, in total 3836 proteins), 3) the cytoplasmic domains of transmembrane proteins, and 4) extracellular domains of transmembrane proteins. For each of these classes, the number of disease-causing variants and the number of polymorphisms that lead to a gain of SLiMs was counted and a two-sided Fisher’s Exact Test was applied to see if there is a statistically significant difference for the likelihood of

a given class of SLiMs to be gained via disease-causing variant compared to that of polymorphisms.

Peptide-Protein Interaction Network Analysis

180 peptide-protein interactions that passed the strict LFQ filter and showed significant differential SILAC ratios between wild-type and mutant forms of the peptides were used to compose a peptide-protein interaction network. The network was visualized using Cytoscape 3.5.1 (Shannon et al., 2003). Sub-graphs of the significant interactions were generated using R package igraph (version 1.0.1) (Csardi and Nepusz, 2006) (using *fastgreedy.community* function) and visualized using the R packages ggnetwork (Briatte, 2016) and ggplot2 (Wilkinson, 2011). Enriched GO terms for each sub-graph were calculated using the topGO R package (Alexa and Rahnenfuhrer, 2016).

DATA AND SOFTWARE AVAILABILITY

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (Vizcaíno et al., 2016) partner repository. The accession number for the peptide-protein interaction screen dataset is ProteomeXchange: PXD010027. The accession number for the PRM dataset is ProteomeXchange: PXD010005. The accession number for the BioID dataset on GLUT1 is ProteomeXchange: PXD010061.

.

QUANTIFICATION AND STATISTICAL ANALYSIS

The type of statistical test (e.g., Mann–Whitney U test or t-test) is annotated in the Figure legend and/or in the Methods segment specific to the analysis. In addition, statistical parameters such as the value of n, mean/median, SEM, SD and significance level are reported in the Figures and/or in the Figure Legends. When * are used to signify the significance level the key is reported in the respective Figure legend. Statistical analyses were generally performed using R as described in Methods and Resources for each individual analysis, if not stated differently.

Colocalization analysis

Imaris v8.4.1 was used for the quantitative colocalization analysis. The original z-stack images were adjusted by adding an adequate mask on the respective red channel to subtract background noise (Costes et al., 2004). The threshold for the mask was uniformly adjusted in each staining experiment. Automatic thresholding was used to define the area where a colocalization would be determined and the statistics was calculated for each colocalization channel (Costes et al., 2004). For the images whose observed correlation was not statistically significant in comparison to randomized images, the colocalization channel was built without additional thresholding on the masked dataset. The resulting thresholded Pearson's coefficients were exported. The number of images and cells in the analyses is stated in the respective Method sections.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Mouse monoclonal anti-FLAG	Sigma-Aldrich	Cat# F3165; RRID:AB_259529
Rabbit polyclonal anti-GLUT1	Merck Millipore	Cat# 07-1401; RRID:AB_1587074
Rat monoclonal anti-ICAM2	BD Biosciences	Cat# 553326; RRID:AB_394784
Isolectin GS-IB4 Alexa Fluor 488 conjugate (IB4)	Thermo Fisher Scientific	Cat# I21411; RRID:AB_2314662
Mouse monoclonal anti-Vti1a	BD Biosciences	Cat# 611220; RRID:AB_398752
Mouse monoclonal anti-Vti1b	BD Biosciences	Cat# 611405; RRID:AB_398927
Rabbit polyclonal anti-EEA1	Cell Signaling Technology	Cat# 2411S; RRID:AB_2096814

Rabbit monoclonal anti-Rab4	Abcam	Cat# ab13252; RRID:AB_2269374
Rabbit monoclonal anti-Rab9	Cell Signaling Technology	Cat# 5118S; RRID:AB_10621426
Rabbit monoclonal anti-LAMP1	Cell Signaling Technology	Cat# 9091; RRID:AB_2687579
Mouse monoclonal anti-gamma Adaptin (AP-1 γ)	BD Biosciences	Cat# 610385; RRID:AB_397768
Mouse monoclonal anti-AP50 (AP-2 μ)	BD Biosciences	Cat# 611351; RRID:AB_398873
Mouse monoclonal anti- alpha Adaptin (AP-2 α)	Thermo Fisher Scientific	Cat# MA3-061; RRID:AB_2056321
Mouse monoclonal anti-alpha Adaptin (AP-2 α)	Abcam	Cat# ab2730; RRID:AB_303255
Mouse monoclonal anti - IL-2 R alpha (TAC)	Santa Cruz Biotechnology	Cat# sc-65258, RRID:AB_631112
Chemicals, Peptides, and Recombinant Proteins		
Deoxy-D-glucose, 2-[1,2-3H (N)]-, Specific Activity: 5-10 Ci (185-370 GBq)/mmol, 250 μCi (9.25 MBq)	Perkin Elmer	NET328250UC
L-arginine-HCl (Arg0)	Sigma-Aldrich	A6969; CAS: 1119-34-2
L-arginine-HCl($^{13}C_6$) (Arg6)	Sigma-Aldrich	643440; CAS: 201740-91-2
L-arginine-HCl($^{13}C_6$,$^{15}N_4$) (Arg10)	Sigma-Aldrich	608033; CAS: 202468-25-5
L-lysine-HCl (Lys0)	Sigma-Aldrich	L8662; CAS: 657-27-2
L-lysine-2HCl(4,4,5,5-D4) (Lys4)	Cambridge Isotope Laboratories	DLM-2640-PK; CAS: 657-26-1
L-lysine-HCl($^{13}C_6$,$^{15}N_2$) (Lys8)	Silantes	211604102
Cytochalasin B	Sigma-Aldrich	C2743; CAS: 14930-96-2
Deposited Data		

Peptide-protein interaction screen dataset	This paper	PXD010027
PRM dataset	This paper	PXD010005
BioID	This paper	PXD010061
Humsavar	http://www.uniprot.org/docs/humsavar.txt	Release: 2015_07 of 24-Jun-2015, (Famiglietti et al., 2014)
Clinvar	ftp.ncbi.nlm.nih.gov/pub/clinvar/tab_delimited/variant_summary.txt.gz	latest update on 25th of March, 2017 (Landrum et al., 2016)
Experimental Models: Cell Lines		
Flp-In T-Rex GLUT1	This paper	N/A
Flp-In T-Rex GLUT1_P485L	This paper	N/A
Human: Patient-derived iPSCs	This paper	https://hpscreg.eu/: BIHi037-(A-E)
Experimental Models: Organisms/Strains		
Mouse: C57BL/6N: GLUT1_P485L	This paper	N/A
Oligonucleotides		
ON-TARGETplus Human AP2M1 (1173) siRNA - SMARTpool	Dharmacon	Cat# L-008170-00-0005
ON-TARGETplus Non-targeting Pool	Dharmacon	Cat# D-001810-10-05
Primers TAC chimera constructs	This paper	See TableS6

Recombinant DNA		
SLC2A1 (GLUT1)	Harvard Plasmid Repository	HsCD00378964
Software and Algorithms		
Imaris v8.4.1	Bitplane	N/A
MaxQuant v1.5.2.8	http://www.biochem.mpg.de/5111795/maxquant	(Cox and Mann, 2008)

References

Alexa, A., and Rahnenfuhrer, J. (2016). topGO: Enrichment Analysis for Gene Ontology. R package version 2.24.0. CRAN.

Briatte, F. (2016). ggnetwork: Geometries to Plot Networks with “ggplot2”, R package version 0.5.1. CRAN.

Costes, S.V., Daelemans, D., Cho, E.H., Dobbin, Z., Pavlakis, G., and Lockett, S. (2004). Automatic and quantitative measurement of protein-protein colocalization in live cells. Biophys. J. *86*, 3993–4003.

Couzens, A.L., Knight, J.D.R., Kean, M.J., Teo, G., Weiss, A., Dunham, W.H., Lin, Z.-Y., Bagshaw, R.D., Sicheri, F., Pawson, T., et al. (2013). Protein interaction network of the mammalian Hippo pathway reveals mechanisms of kinase-phosphatase interactions. Sci. Signal. *6*, rs15.

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat. Biotechnol. *26*, 1367–1372.

Csardi, G., and Nepusz, T. (2006). The igraph software package for complex network research. InterJournal, Complex Systems 1695.

Dinkel, H., Van Roey, K., Michael, S., Kumar, M., Uyar, B., Altenberg, B., Milchevskaya, V., Schneider, M., Kühn, H., Behrendt, A., et al. (2016). ELM 2016--data update and new functionality of the eukaryotic linear motif resource. Nucleic Acids Res. *44*, D294–D300.

Diril, M.K., Schmidt, S., Krauss, M., Gawlik, V., Joost, H.-G., Schürmann, A., Haucke, V., and Augustin, R. (2009). Lysosomal localization of GLUT8 in the testis--the EXXXLL motif of GLUT8

is sufficient for its intracellular sorting via AP1- and AP2-mediated interaction. FEBS J. *276*, 3729–3743.

Dosztányi, Z., Csizmok, V., Tompa, P., and Simon, I. (2005). IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. Bioinformatics *21*, 3433–3434.

Famiglietti, M.L., Estreicher, A., Gos, A., Bolleman, J., Géhant, S., Breuza, L., Bridge, A., Poux, S., Redaschi, N., Bougueleret, L., et al. (2014). Genetic variations and diseases in UniProtKB/Swiss-Prot: the ins and outs of expert manual curation. Hum. Mutat. *35*, 927–935.

Finn, R.D., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Mistry, J., Mitchell, A.L., Potter, S.C., Punta, M., Qureshi, M., Sangrador-Vegas, A., et al. (2016). The Pfam protein families database: towards a more sustainable future. Nucleic Acids Res. *44*, D279–D285.

Hernández-Miranda, L.R., Cariboni, A., Faux, C., Ruhrberg, C., Cho, J.H., Cloutier, J.-F., Eickholt, B.J., Parnavelas, J.G., and Andrews, W.D. (2011). Robo1 regulates semaphorin signaling to guide the migration of cortical interneurons through the ventral forebrain. J. Neurosci. *31*, 6174–6187.

Ittner, L.M., and Götz, J. (2007). Pronuclear injection for the production of transgenic mice. Nat. Protoc. *2*, 1206–1215.

Keilhauer, E.C., Hein, M.Y., and Mann, M. (2015). Accurate protein complex retrieval by affinity enrichment mass spectrometry (AE-MS) rather than affinity purification mass spectrometry (AP-MS). Mol. Cell. Proteomics *14*, 120–135.

Köhler, S., Vasilevsky, N.A., Engelstad, M., Foster, E., McMurry, J., Aymé, S., Baynam, G., Bello, S.M., Boerkoel, C.F., Boycott, K.M., et al. (2017). The Human Phenotype Ontology in 2017. Nucleic Acids Res. *45*, D865–D876.

Landrum, M.J., Lee, J.M., Benson, M., Brown, G., Chao, C., Chitipiralla, S., Gu, B., Hart, J., Hoffman, D., Hoover, J., et al. (2016). ClinVar: public archive of interpretations of clinically relevant variants. Nucleic Acids Res. *44*, D862–D868.

McLaren, W., Gil, L., Hunt, S.E., Riat, H.S., Ritchie, G.R.S., Thormann, A., Flicek, P., and Cunningham, F. (2016). The Ensembl Variant Effect Predictor. Genome Biol. *17*, 122.

Rappsilber, J., Ishihama, Y., and Mann, M. (2003). Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. Anal. Chem. *75*, 663–670.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods *9*, 676–682.

Schulze, W.X., and Mann, M. (2003). A Novel Proteomic Screen for Peptide-Protein Interactions. J. Biol. Chem. *279*, 10756–10764.

Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011). Global quantification of mammalian gene expression control. Nature *473*, 337–342.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. *13*, 2498–2504.

Shi, J., and Kandror, K.V. (2008). Study of glucose uptake in adipose cells. Methods Mol. Biol. *456*, 307–315.

Slaughter, L., Vartzelis, G., and Arthur, T. (2009). New GLUT-1 mutation in a child with treatment-resistant epilepsy. Epilepsy Res. *84*, 254–256.

Toprak, U.H., Gillet, L.C., Maiolica, A., Navarro, P., Leitner, A., and Aebersold, R. (2014). Conserved peptide fragmentation as a benchmarking tool for mass spectrometers and a discriminating feature for targeted proteomics. Mol. Cell. Proteomics *13*, 2056–2071.

Tripathi, S., Pohl, M.O., Zhou, Y., Rodriguez-Frandsen, A., Wang, G., Stein, D.A., Moulton, H.M., DeJesus, P., Che, J., Mulder, L.C.F., et al. (2015). Meta- and Orthogonal Integration of Influenza “OMICs” Data Defines a Role for UBR4 in Virus Budding. Cell Host Microbe *18*, 723–735.

UniProt Consortium (2012). Reorganizing the protein space at the Universal Protein Resource (UniProt). Nucleic Acids Res. *40*, D71–D75.

Vizcaíno, J.A., Csordas, A., Del-Toro, N., Dianes, J.A., Griss, J., Lavidas, I., Mayer, G., Perez-Riverol, Y., Reisinger, F., Ternent, T., et al. (2016). 2016 update of the PRIDE database and its related tools. Nucleic Acids Res. *44*, 11033.

Wefers, B., Bashir, S., Rossius, J., Wurst, W., and Kühn, R. (2017). Gene editing in mouse zygotes using the CRISPR/Cas9 system. Methods *121-122*, 55–67.

Wilkinson, L. (2011). ggplot2: Elegant Graphics for Data Analysis by WICKHAM, H. Biometrics *67*, 678–679.

Zauber, H., Kirchner, M., and Selbach, M. (2018). Picky: a simple online PRM and SRM method designer for targeted proteomics. Nat. Methods *15*, 156–157.

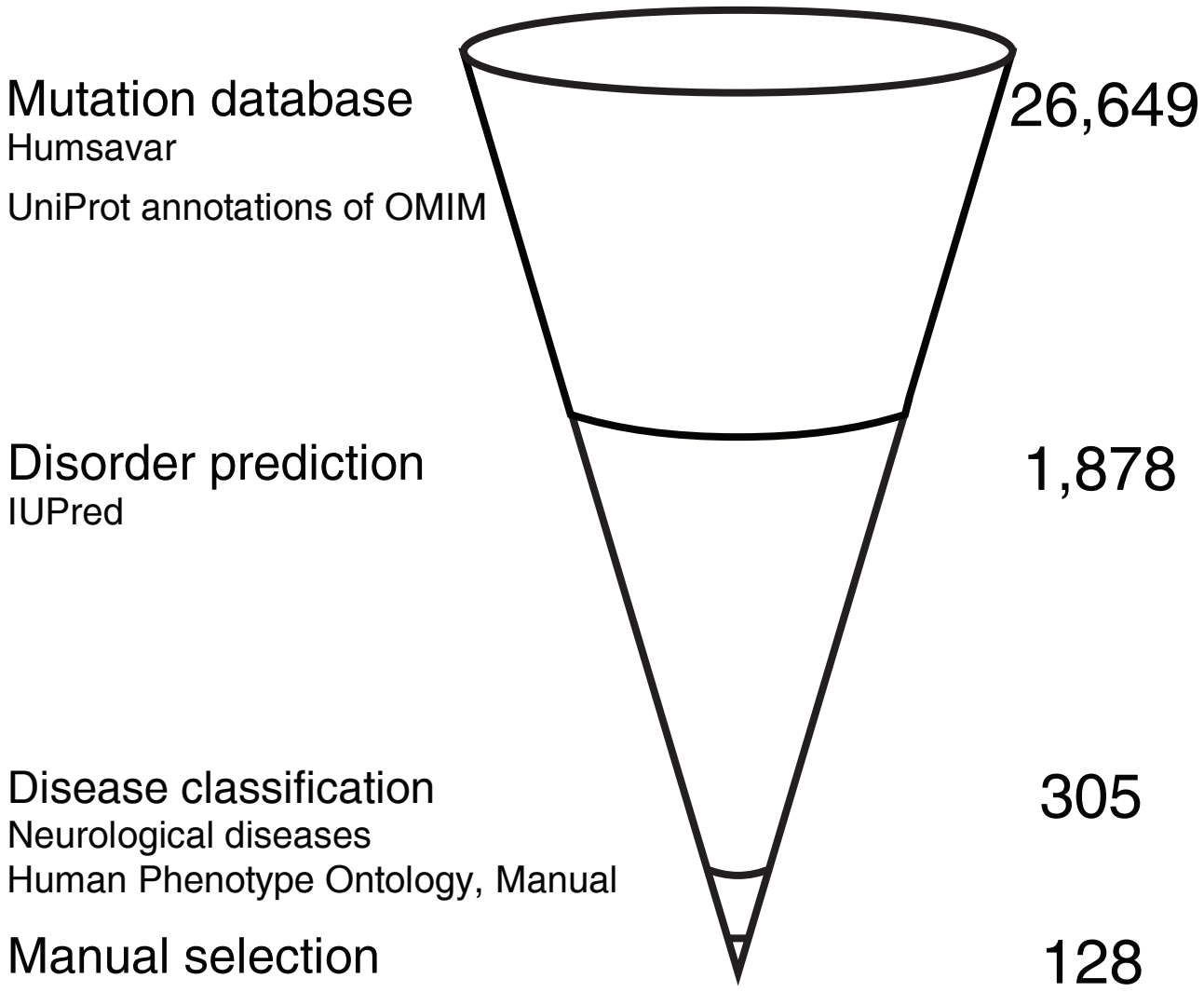
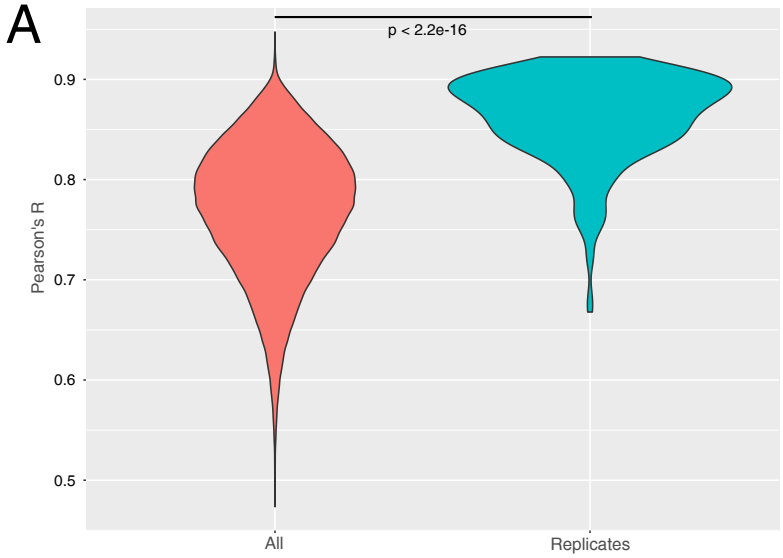

Mutation database
Humsavar
UniProt annotations of OMIM
26,649
Disorder prediction
IUPred
1,878
Disease classification
Neurological diseases
Human Phenotype Ontology, Manual
305
Manual selection
128

Fig. S1


A
p < 2.2e-16
Pearson's R
0.9
0.8
0.7
0.6
0.5
All
Replicates

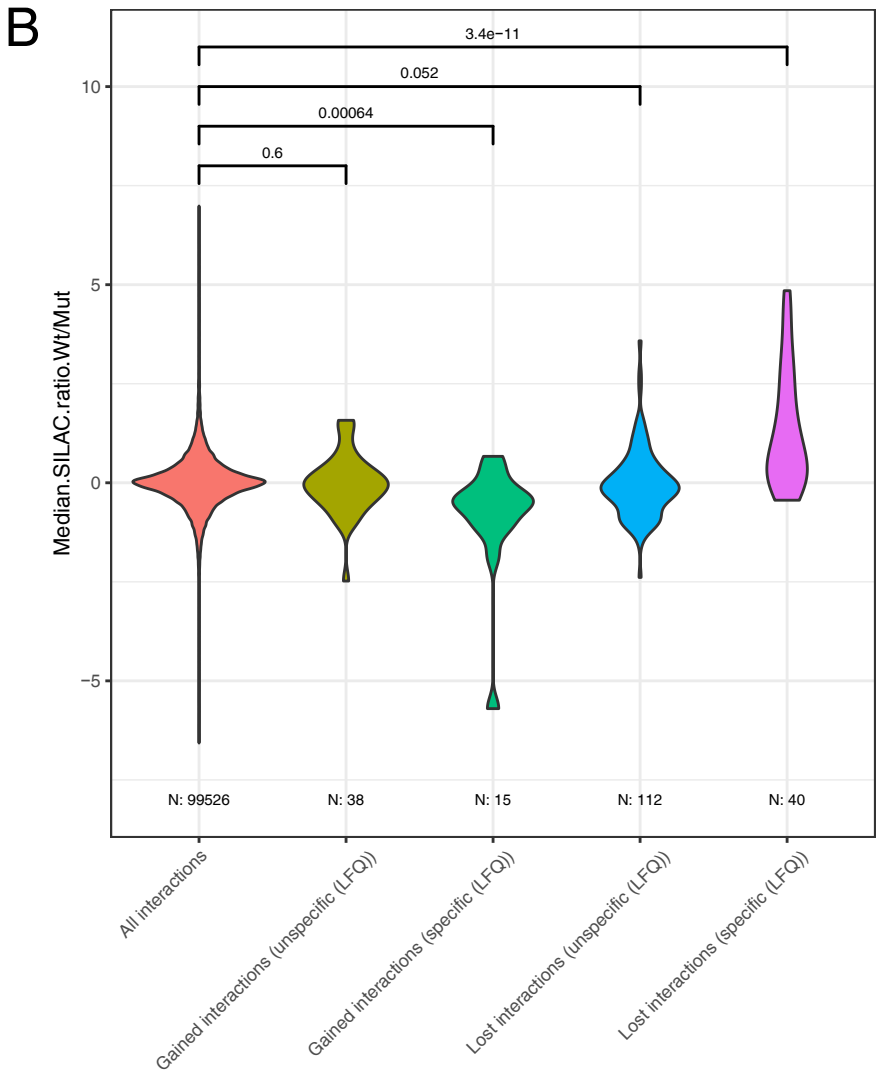
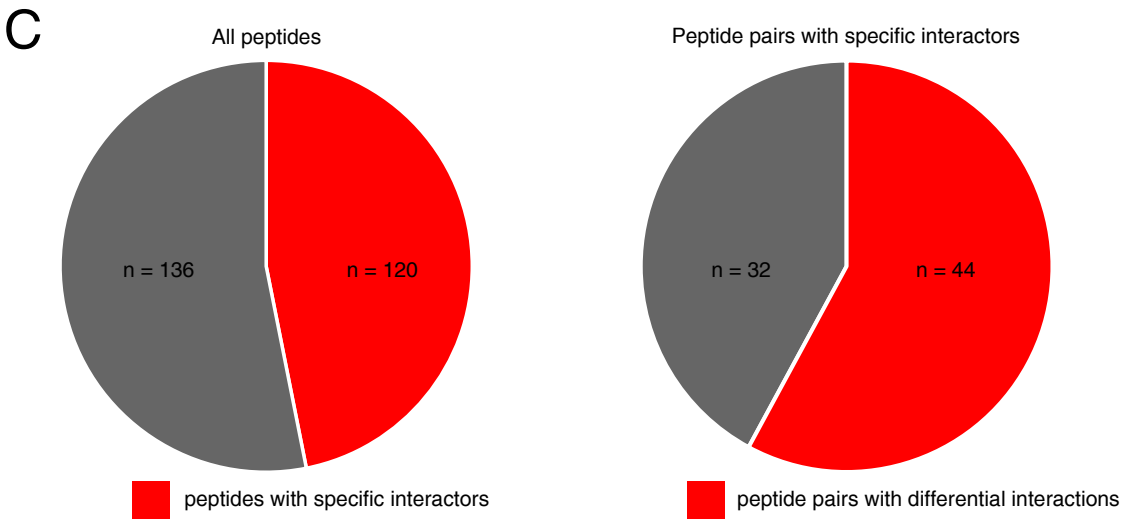

B
3.4e−11
0.052
0.00064
0.6
Median.SILAC.ratio.Wt/Mut
10
5
0
−5
N: 99526
N: 38
N: 15
N: 112
N: 40
All interactions
Gained interactions (unspecific (LFQ))
Gained interactions (specific (LFQ))
Lost interactions (unspecific (LFQ))
Lost interactions (specific (LFQ))

Fig. S2


C
All peptides
n = 136
n = 120
peptides with specific interactors
Peptide pairs with specific interactors
n = 32
n = 44
peptide pairs with differential interactions

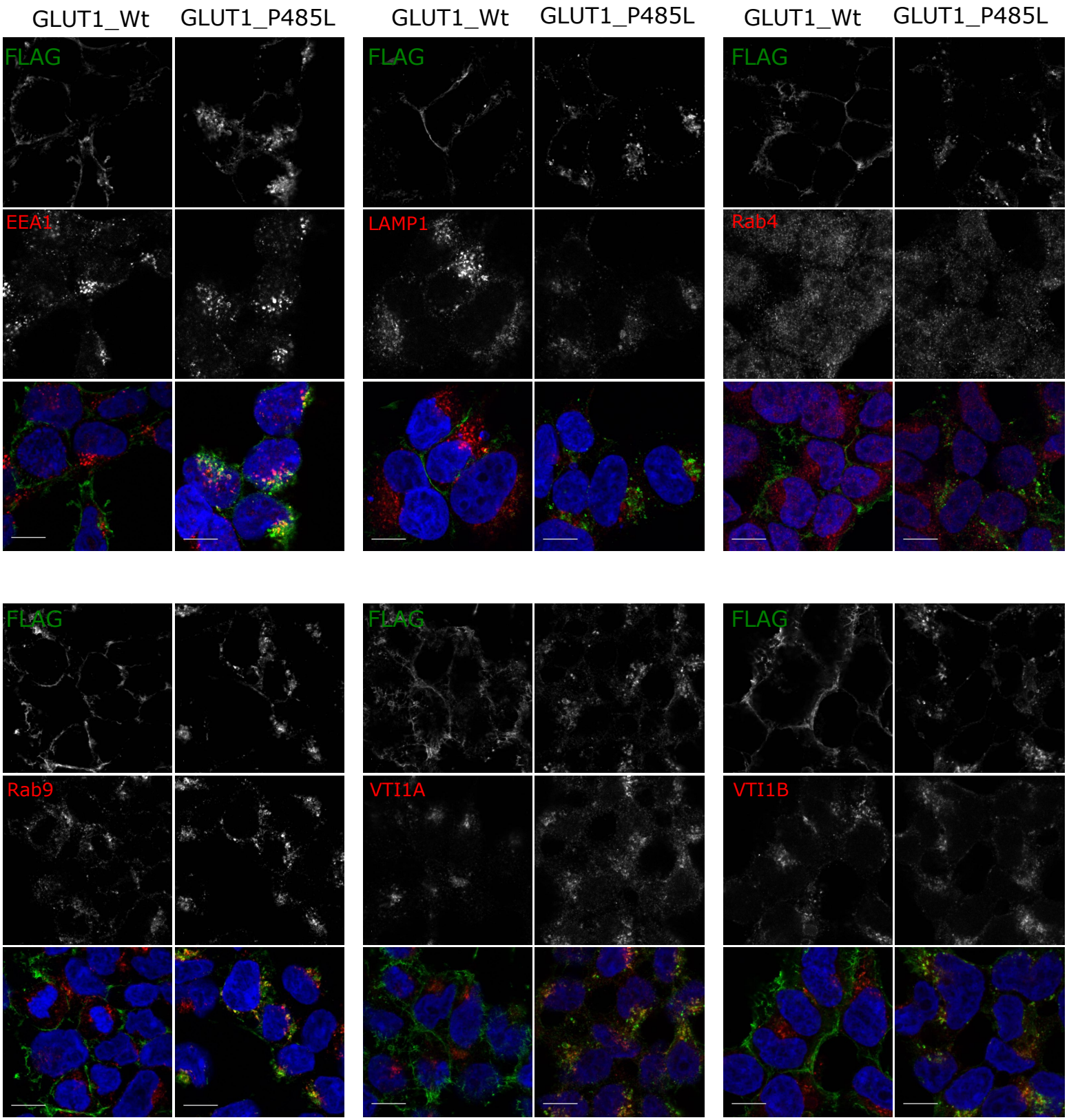
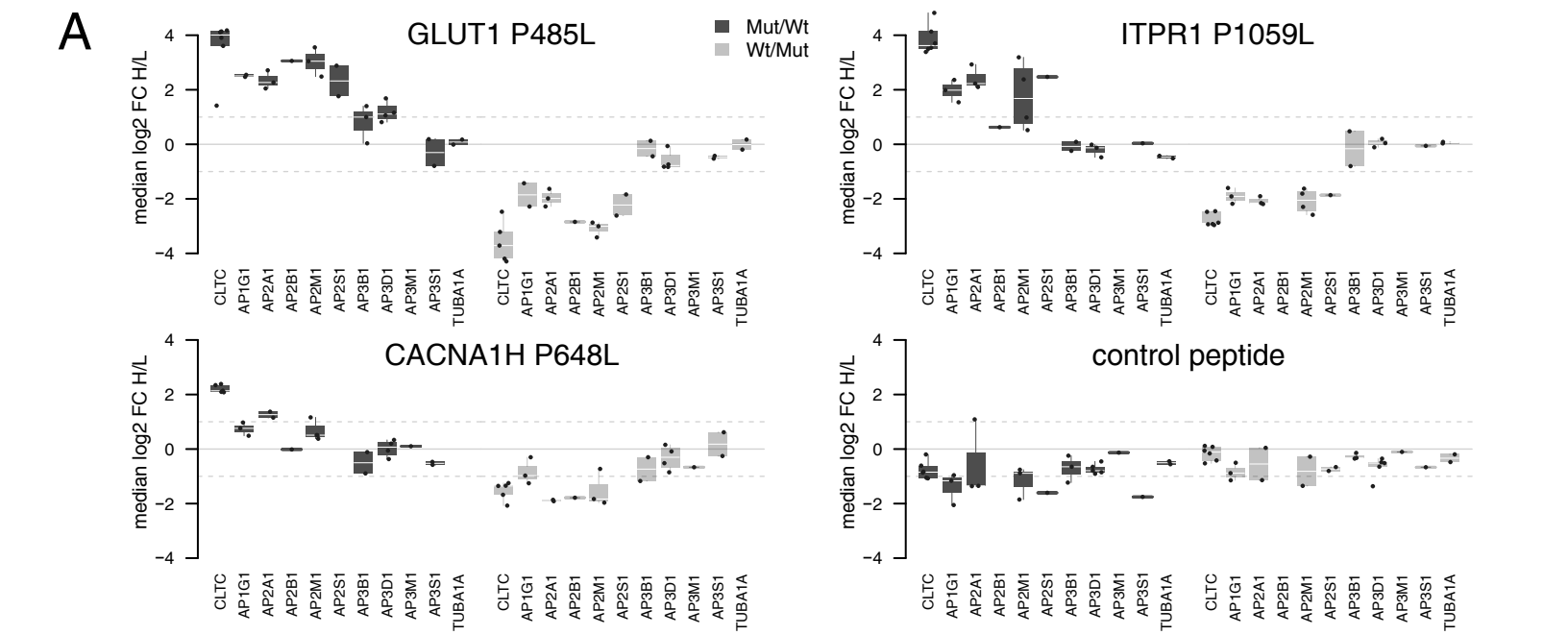

GLUT1_Wt
GLUT1_P485L
GLUT1_Wt
GLUT1_P485L
GLUT1_Wt
GLUT1_P485L
FLAG
EEA1
FLAG
LAMP1
FLAG
Rab4
FLAG
Rab9
FLAG
VTI1A
FLAG
VTI1B

Fig. S3


A
GLUT1 P485L
ITPR1 P1059L
CACNA1H P648L
control peptide
Mut/Wt
Wt/Mut
median log2 FC H/L
CLTC
AP1G1
AP2A1
AP2B1
AP2M1
AP2S1
AP3B1
AP3D1
AP3M1
AP3S1
TUBA1A

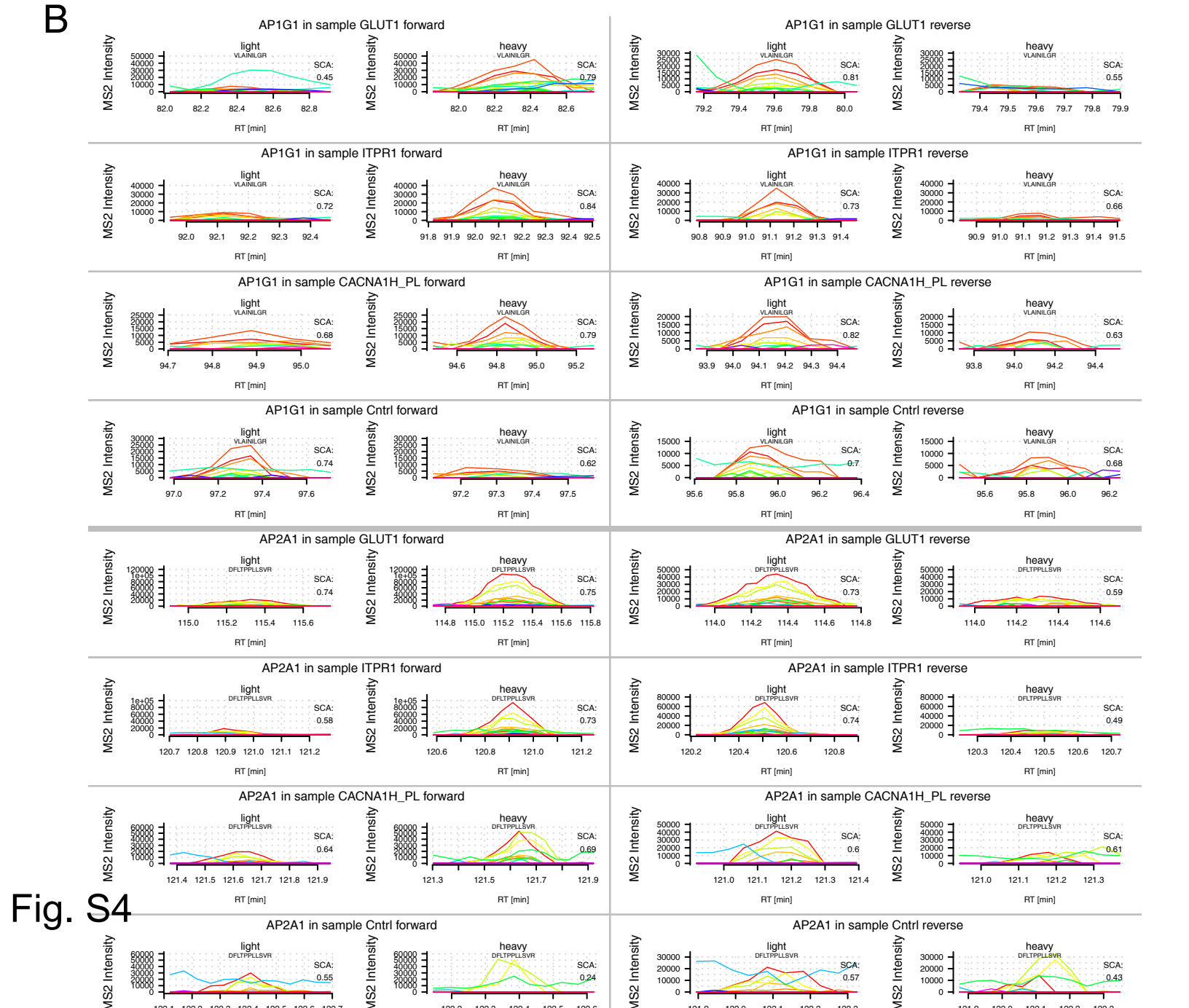

B
AP1G1 in sample GLUT1 forward
AP1G1 in sample GLUT1 reverse
AP1G1 in sample ITPR1 forward
AP1G1 in sample ITPR1 reverse
AP1G1 in sample CACNA1H_PL forward
AP1G1 in sample CACNA1H_PL reverse
AP1G1 in sample Cntrl forward
AP1G1 in sample Cntrl reverse
AP2A1 in sample GLUT1 forward
AP2A1 in sample GLUT1 reverse
AP2A1 in sample ITPR1 forward
AP2A1 in sample ITPR1 reverse
AP2A1 in sample CACNA1H_PL forward
AP2A1 in sample CACNA1H_PL reverse
AP2A1 in sample Cntrl forward
AP2A1 in sample Cntrl reverse
light
heavy
MS2 Intensity
RT [min]
SCA:

Fig. S4

Pluripotency Immunocytochemistry BIHi037-A

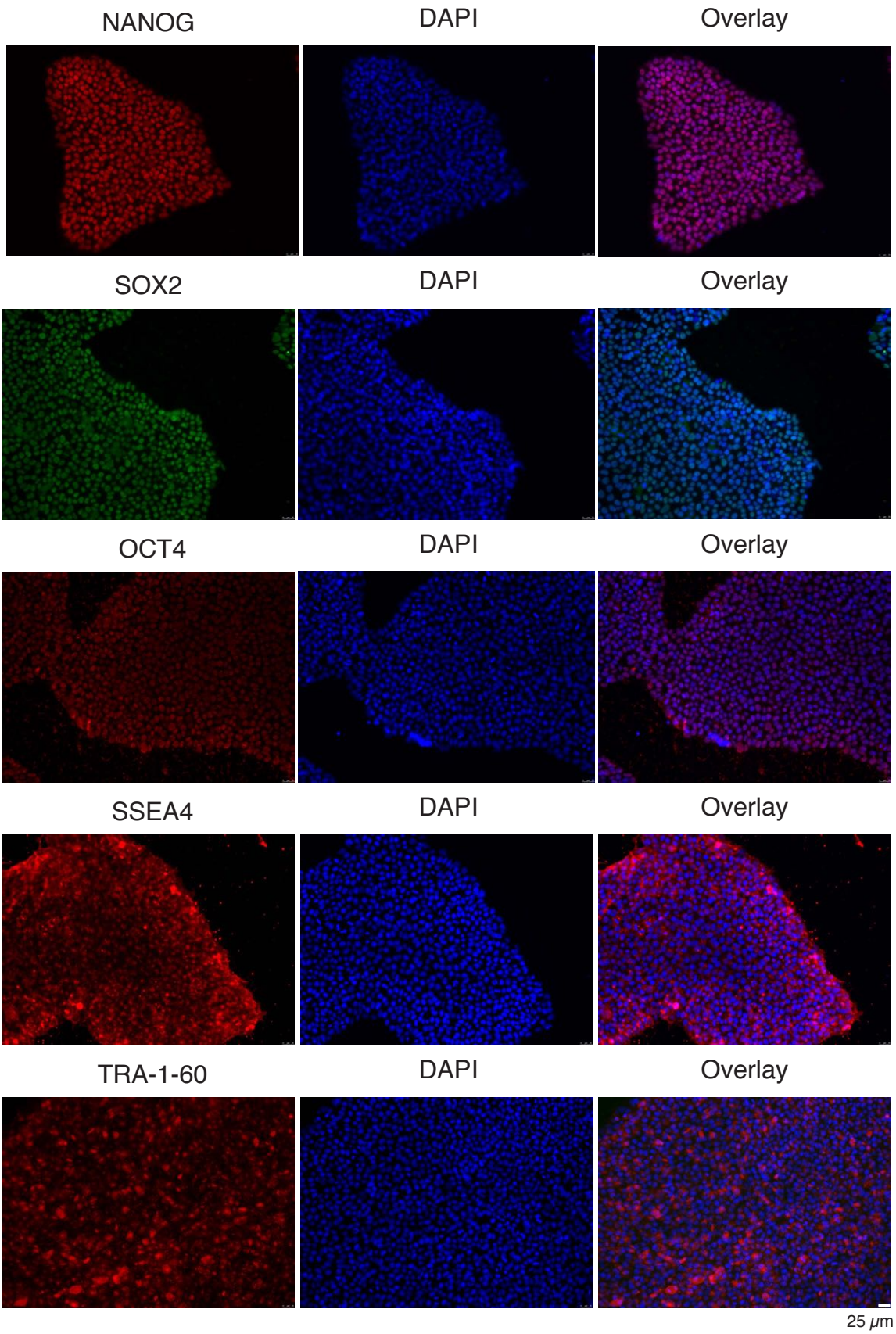
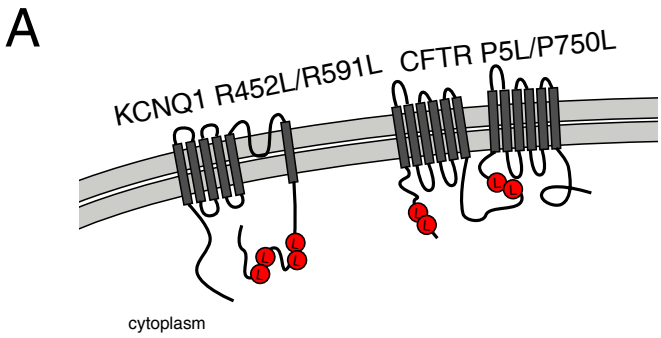
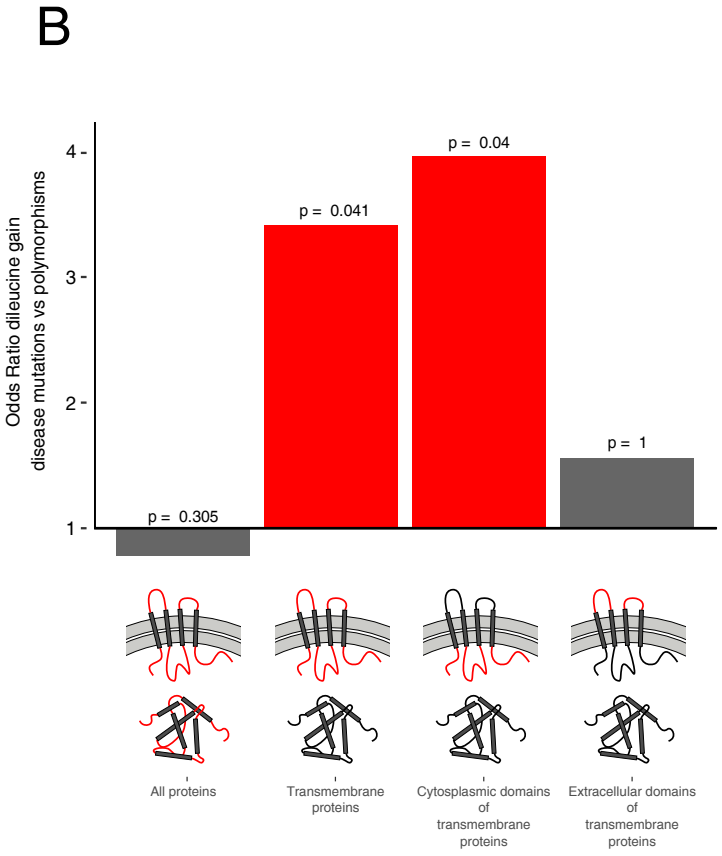
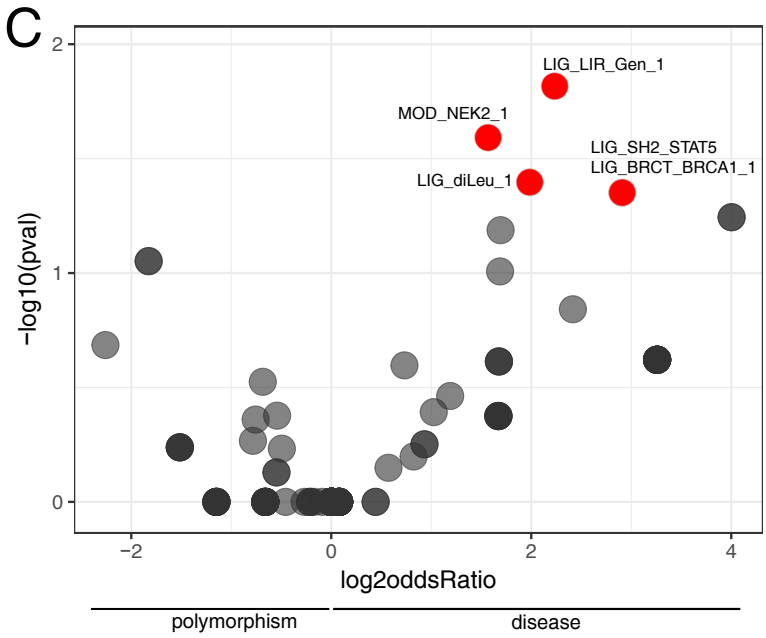


Fig. S5


A
KCNQ1 R452L/R591L
CFTR P5L/P750L
cytoplasm


B
Odds Ratio dileucine gain
disease mutations vs polymorphisms
p = 0.305
p = 0.041
p = 0.04
p = 1
All proteins
Transmembrane proteins
Cytosplasmic domains of transmembrane proteins
Extracellular domains of transmembrane proteins


C
−log10(pval)
log2oddsRatio
LIG_LIR_Gen_1
MOD_NEK2_1
LIG_SH2_STAT5
LIG_BRCT_BRCA1_1
LIG_diLeu_1
polymorphism
disease

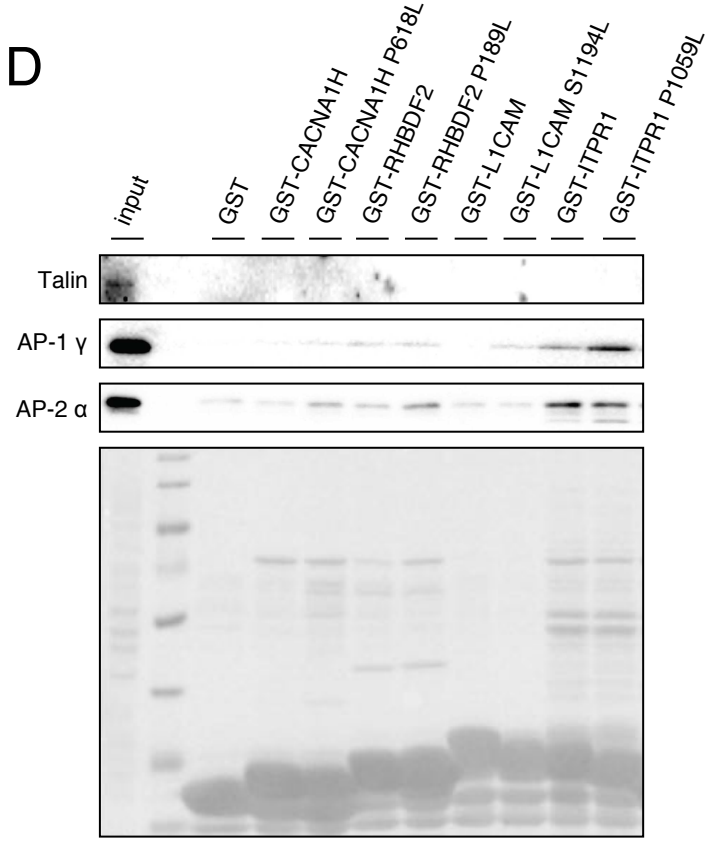

D
input
GST
GST-CACNA1H
GST-CACNA1H P618L
GST-RHBDF2
GST-RHBDF2 P189L
GST-L1CAM
GST-L1CAM S1194L
GST-ITPR1
GST-ITPR1 P1059L
Talin
AP-1 γ
AP-2 α

Fig. S6

Fig. S1: Candidate selection for peptide-protein interaction screen. Related to Fig. 1.

Candidates were selected from missense disease mutations in the Humsavar database (Uniprot) by selecting mutations in disordered regions that cause neurological diseases (see Methods).

Fig. S2: Related to Fig. 2.

A, **Reproducibility of technical replicates** Pearson's R shows significantly higher correlation of technical replicates than correlations between all pull-downs according to Welch Two Sample t-test.

B, Impact of specificity cut-off (LFQ) and differential cut-off (SILAC) on peptide candidates After applying the specificity cut-off (derived from control peptide, see Methods) on all interactions, only about half of the 256 (128 variant pairs) peptides showed at least one specific binder according to the LFQ-filter (left pie chart, red). These 120 peptides relate to 76 peptide pairs with specific interactions of wild-type and/or mutant peptide. More than half of all 76 peptide pairs with specific interactors show differential interaction between the variants after applying the SILAC cut-off (see Methods) (right pie chart, red).

C, **SILAC ratio distributions of detected interactions that can be explained by presence of SLiMs in the peptides and PFAM domains in the interaction partners** Peptide-protein interactions detected in the screen were classified as 'gained' or 'lost' according to the following criteria: An interaction is classified as 'gained' if the mutant peptide sequence matches a SLiM pattern that does not match the wild-type peptide sequence and the mutant peptide has an interaction partner that contains a compatible PFAM domain to bind that SLiM instance. On the other hand, an interaction is classified as 'lost' if the wild-type peptide sequence matches a SLiM pattern that is not matched in the mutant peptide sequence and the wild-type peptide has an interaction partner that contains a compatible PFAM domain to bind that SLiM instance. Gained and lost interactions are further sub-classified as 'LFQ positive' and 'LFQ negative' depending on whether the peptide-protein interaction has an LFQ value that passes a looser version of the LFQ cut-off (i.e. p-value < 0.05 and log2 fold enrichment > 1. The median SILAC ratio distributions (wild-type versus mutant) of each of these four categories of interactions ('Gained interactions - LFQ negative', 'Gained interactions - LFQ positive', 'Lost interactions - LFQ negative', and

‘Lost interactions - LFQ positive’) are compared with the median SILAC ratio distributions of all detected interactions from the array using a Wilcoxon-Mann-Whitney test. Compared to the background distribution of median SILAC ratios (in red), the gained interactions that pass the LFQ filter (in green) show a significant negative skew while the lost interactions that pass the LFQ filter (in purple) show a significant positive skew.

Fig. S3: Related to Fig. 4.
Quantification of Fig.4 B is based on these example images for colocalization study of GLUT1 wild-type and P485L mutant to markers of endocytic compartments. GLUT1 (green), Marker (red), DAPI (blue)

Fig. S4: Adaptor proteins bind preferentially to mutant variant peptides carrying a dileucine. Related to Fig. 5.
A, A highly sensitive, targeted mass spectrometry technique (parallel reaction monitoring, PRM) reveals that adaptor proteins (AP-1, AP-2, AP-3) bind preferentially to peptides carrying a dileucine. **B**, Raw elution profiles of transitions from two example peptides from the proteins AP1G1 and AP2A1 for all forward and label swap (reverse) experiments. Ratios were calculated based on extracted intensity informations from the light and heavy channel in each experiment for each transition. Different transitions are displayed with different colors. The spectrum contrast angle (SCA) indicates similarity to the corresponding spectrum in the spectrum library and is displayed for every peak.

Fig. S5: iPSCs show pluripotency markers. Related to Fig. 6.
iPSCs (BIHi037-A) were reprogrammed from a G1DS patient carrying a heterozygous GLUT1 P485L mutation and were stained for known pluripotency markers (see Methods).

Fig. S6: Mutation-induced gains of dileucine motifs are a significant cause of disease. Related to Fig. 7.
A, A systematic bioinformatic search from the Clinvar Database revealed four additional mutations, with pathogenic indications, in cytosolic segments of transmembrane proteins that create dileucine motifs. (Depicted here are only

mutations from Table S4/Clinvar that are classified ‘Pathogenic’ or ‘Conflicting interpretations of pathogenicity’, in case ‘pathogenic’ or ‘likely pathogenic’ was included in the different interpretations, that had not been found with the Humsavar analysis before.) Overlap between Humsavar and Clinvar is: RET_P1039L, L1CAM_S1194L, ITPR1_P1059L, RHBDF2_P189L).

B, Relative frequency of dileucine motif gains in disease mutations (Clinical significance levels ‘Pathogenic’, ‘Pathogenic/Likely pathogenic’, or ‘Likely pathogenic’) and polymorphisms (Clinical significance levels ‘Benign’, ‘Benign/Likely benign’, or ‘Likely benign’) in different disordered regions (IUPred Score >= 0.4) of the proteome. Dileucine motif gain is significantly enriched in disordered regions of transmembrane proteins. Enrichment becomes more significant when looking only at disordered cytoplasmic domains (two-sided Fisher’s exact test).

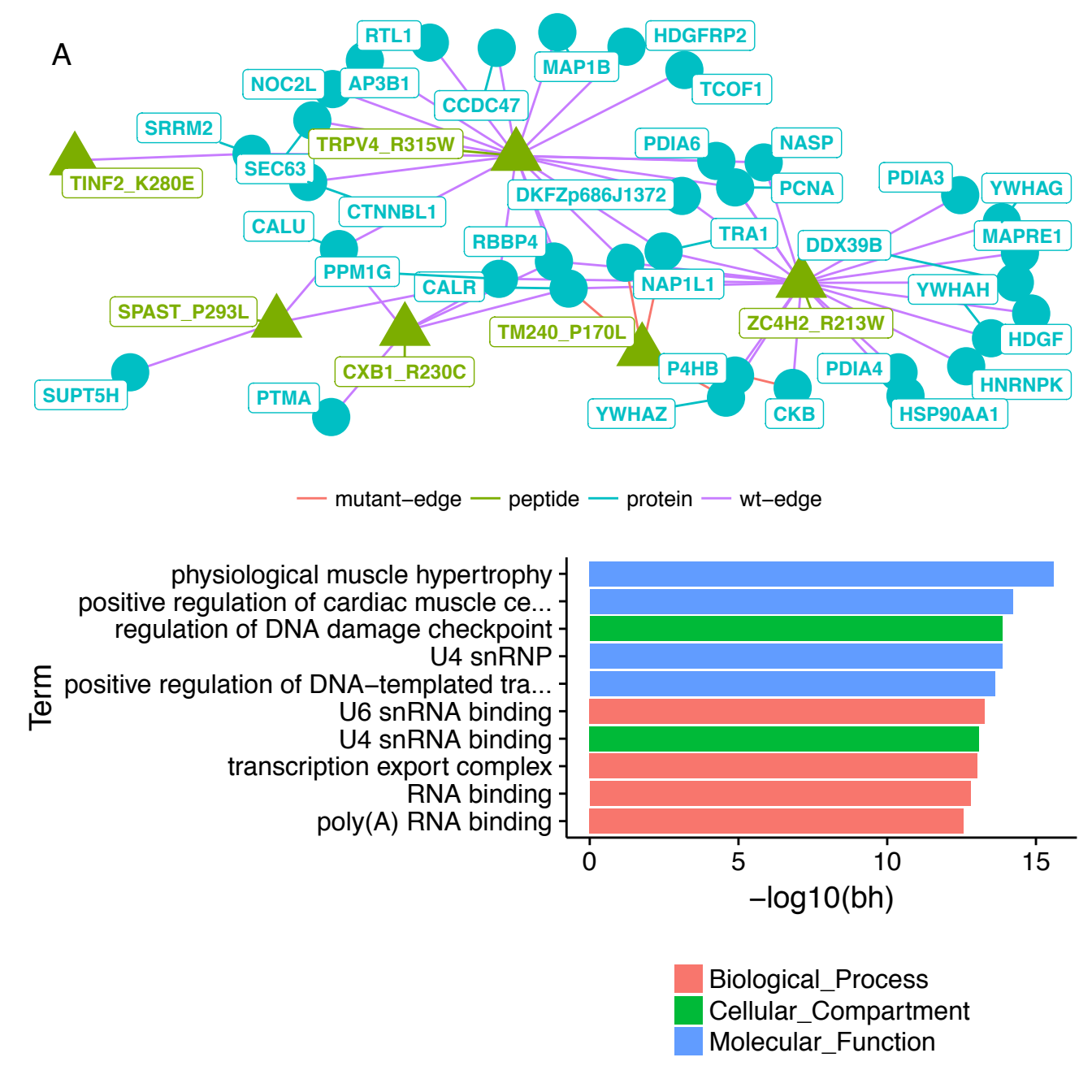
C, Comparison of gained motifs in disordered regions of cytoplasmic tails of transmembrane proteins reveals the dileucine motif is one of the most significant, specific enriched motifs when compared with polymorphisms. (ELM Motif LIG_LIR_Nem_3 (-log10(pval)=2.723722 , log2oddsRatio=2.63761) has been taken out of graphical representation because it is functional only in nematodes and represents a less specific form of LIG_LIR_Gen_1.)

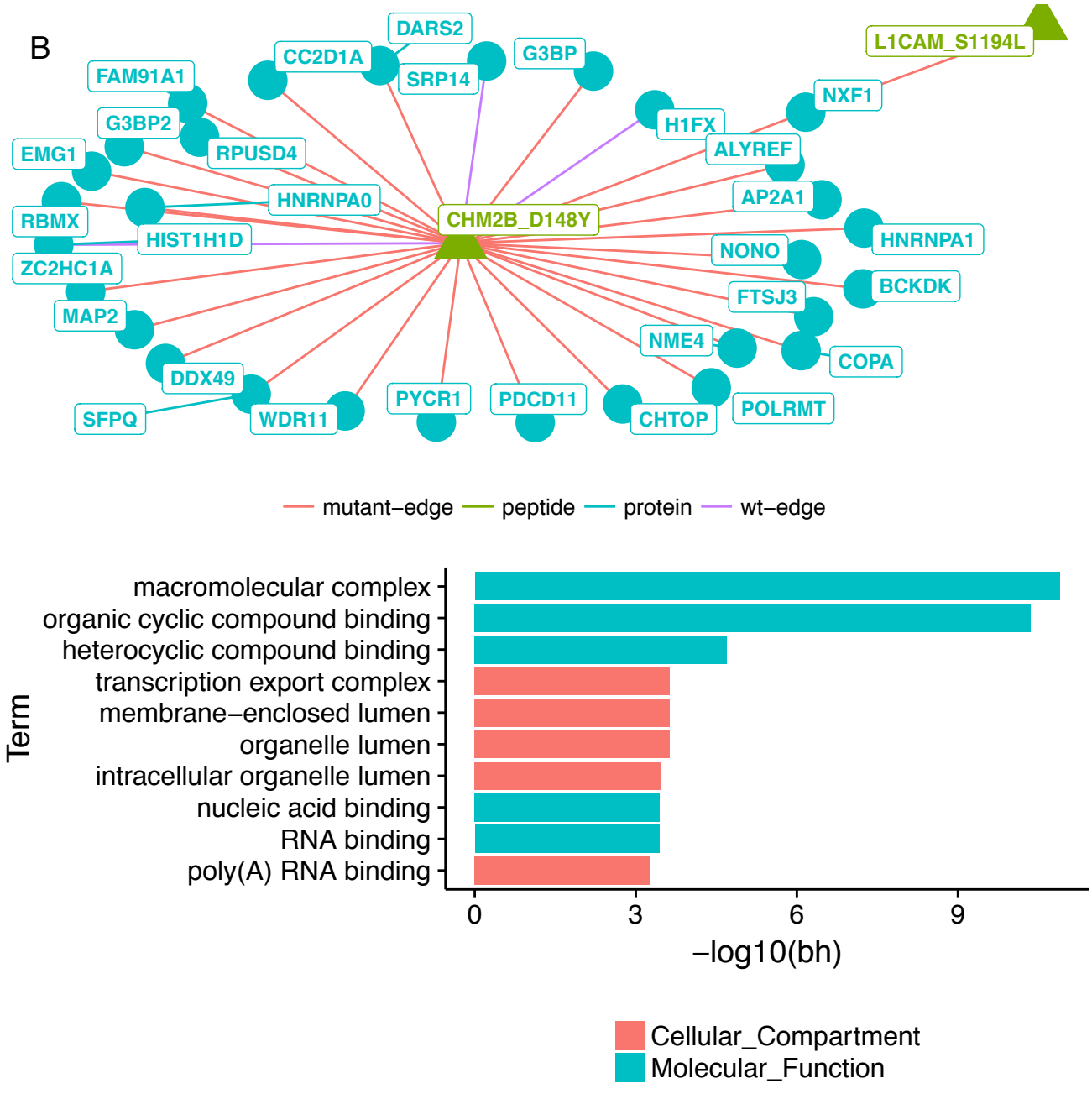
D, Dileucine containing peptides interact with adaptor proteins. Mutant tails of CACNA1H, L1CAM and ITPR1 show increased binding of AP-1 compared to wild-type tails. Mutant tails of CACNA1H and RHBDF2 show increased binding of AP2 compared to wild type tails. Tails were tagged with GST to pull-down interaction partners from mouse brain lysate. Talin is shown as a negative control and is not pulled down from any of the constructs.

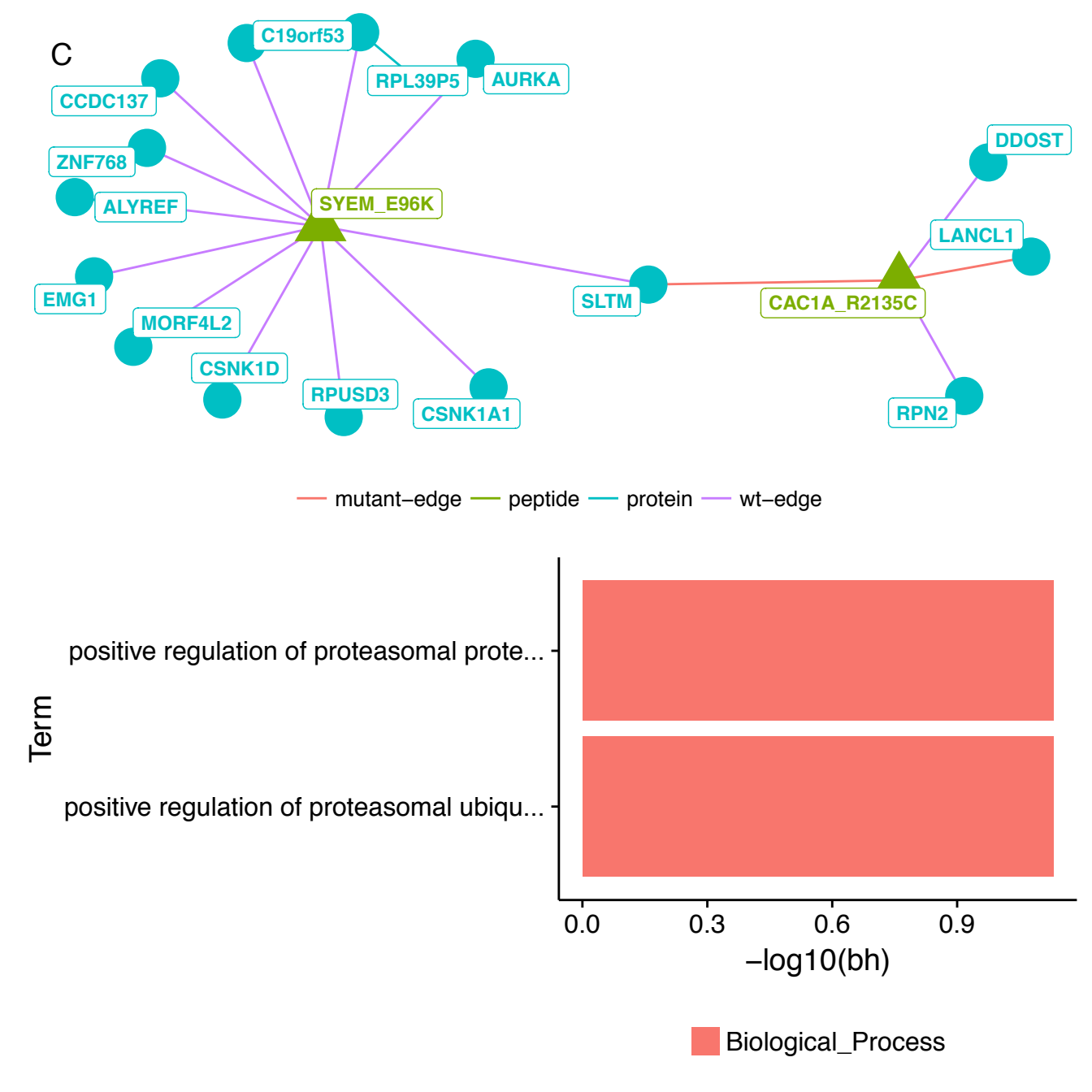
GO Term analysis of subclusters of the peptide-protein interaction network composed of significant differential interactions

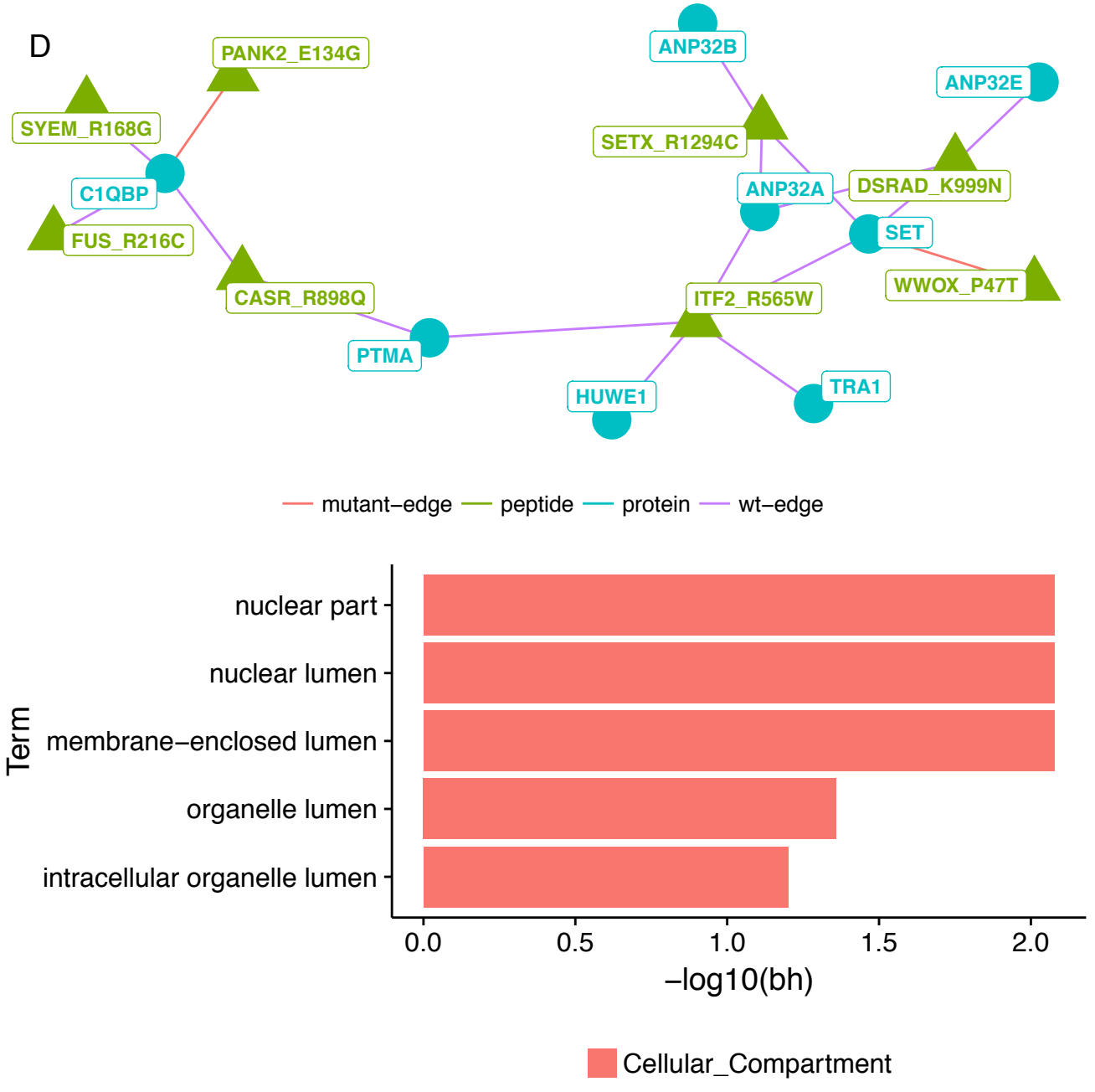
180 peptide-protein interactions that passed the LFQ filter and showed significant differential SILAC ratios between wild-type and mutant forms of the peptides were used to compose a peptide-protein interaction network. Communities (sub-graphs) of this network was extracted using the fastgreedy.community function of the igraph R package (See Methods). GO term enrichment was calculated for the nodes of each subgraph of the network. Within the subgraphs, the edges (interactions) are colored

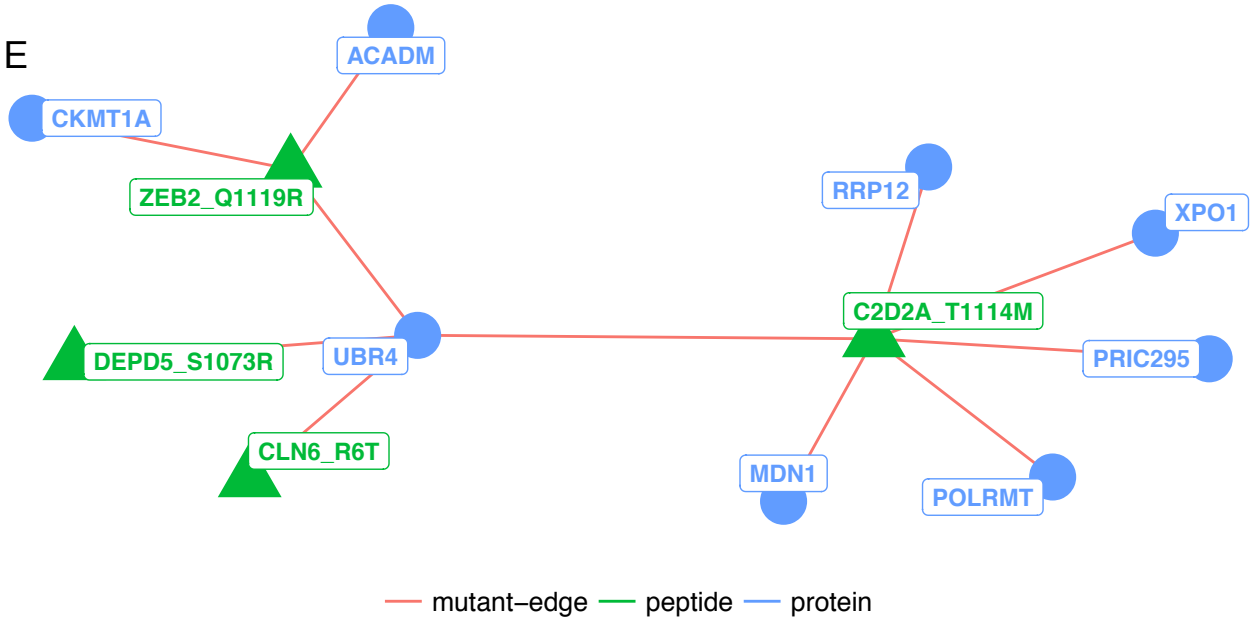
differentially depending on whether the interaction is with a mutant or wild-type peptide. The peptides are depicted as triangles while the proteins are depicted as circles and they are differentially colored. Below the subgraph figure, a bar plot of the top 10 enriched GO terms are displayed (differentially colored by the GO term category as ‘Biological Process’, ‘Molecular Function’ or ‘Cellular Compartment), in which the x-axis shows the log10 p-values corrected for multiple-testing correction using the Benjamini Hochberg method.


A
RTL1
HDGFRP2
MAP1B
TCOF1
NOC2L
AP3B1
CCDC47
SRRM2
TRPV4_R315W
PDIA6
NASP
TINF2_K280E
SEC63
PCNA
PDIA3
YWHAG
DKFZp686J1372
CTNNBL1
CALU
TRA1
MAPRE1
RBBP4
DDX39B
PPM1G
NAP1L1
YWHAH
CALR
SPAST_P293L
ZC4H2_R213W
HDGF
TM240_P170L
P4HB
PDIA4
CXB1_R230C
HNRNPK
SUPT5H
PTMA
YWHAZ
CKB
HSP90AA1
mutant−edge
peptide
protein
wt−edge
physiological muscle hypertrophy
positive regulation of cardiac muscle ce...
regulation of DNA damage checkpoint
U4 snRNP
positive regulation of DNA−templated tra...
U6 snRNA binding
U4 snRNA binding
transcription export complex
RNA binding
poly(A) RNA binding
Term
0
5
10
15
−log10(bh)
Biological_Process
Cellular_Compartment
Molecular_Function


B
DARS2
CC2D1A
SRP14
G3BP
L1CAM_S1194L
FAM91A1
NXF1
G3BP2
H1FX
EMG1
RPUSD4
ALYREF
HNRNPA0
AP2A1
RBMX
CHM2B_D148Y
HIST1H1D
HNRNPA1
ZC2HC1A
NONO
BCKDK
FTSJ3
MAP2
NME4
COPA
DDX49
PYCR1
PDCD11
POLRMT
SFPQ
WDR11
CHTOP
mutant−edge
peptide
protein
wt−edge
Term
macromolecular complex
organic cyclic compound binding
heterocyclic compound binding
transcription export complex
membrane−enclosed lumen
organelle lumen
intracellular organelle lumen
nucleic acid binding
RNA binding
poly(A) RNA binding
0
3
6
9
−log10(bh)
Cellular_Compartment
Molecular_Function

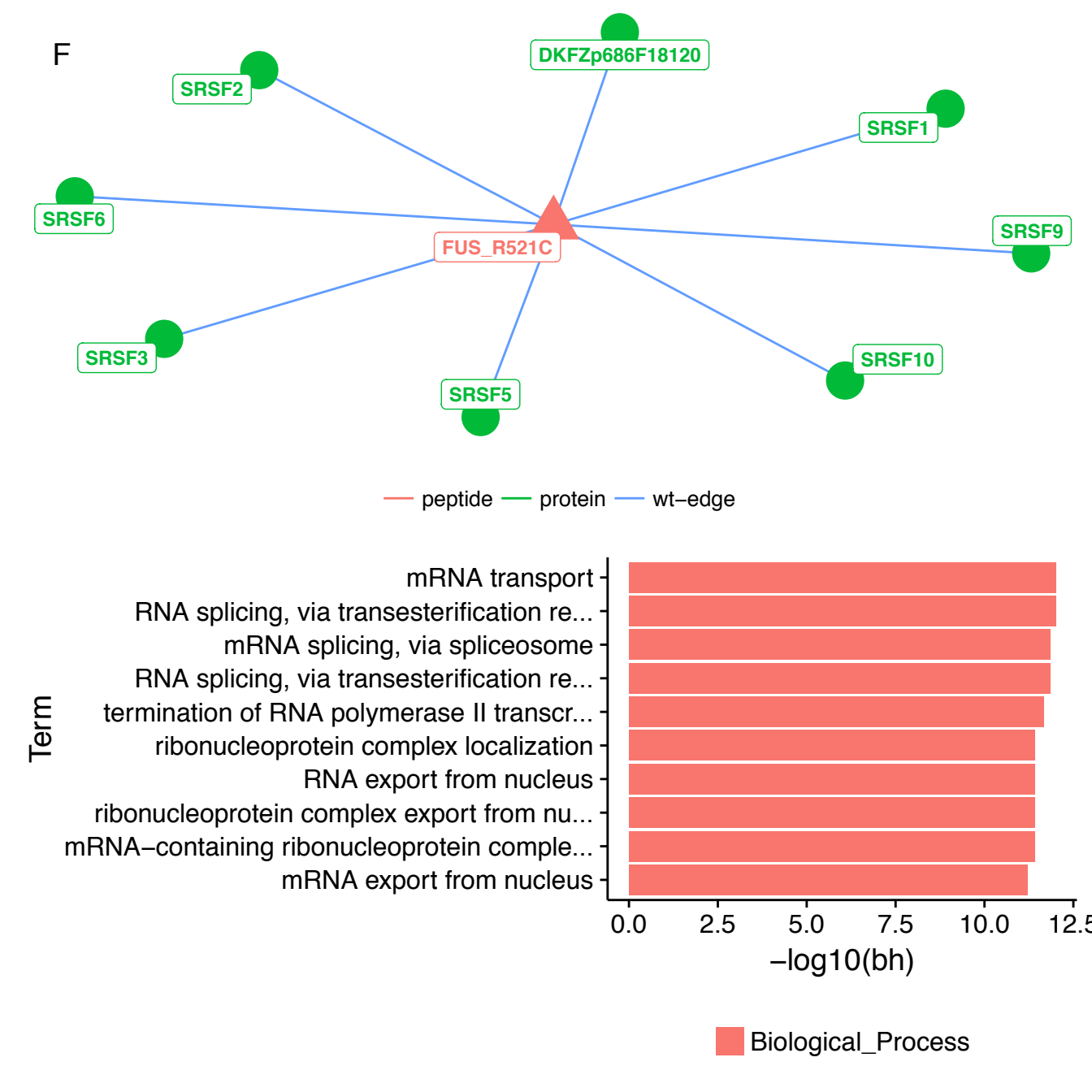

C
C19orf53
CCDC137
RPL39P5
AURKA
ZNF768
ALYREF
SYEM_E96K
EMG1
MORF4L2
CSNK1D
RPUSD3
CSNK1A1
SLTM
DDOST
LANCL1
CAC1A_R2135C
RPN2
mutant−edge
peptide
protein
wt−edge
Term
positive regulation of proteasomal prote...
positive regulation of proteasomal ubiqu...
0.0
0.3
0.6
0.9
−log10(bh)
Biological_Process

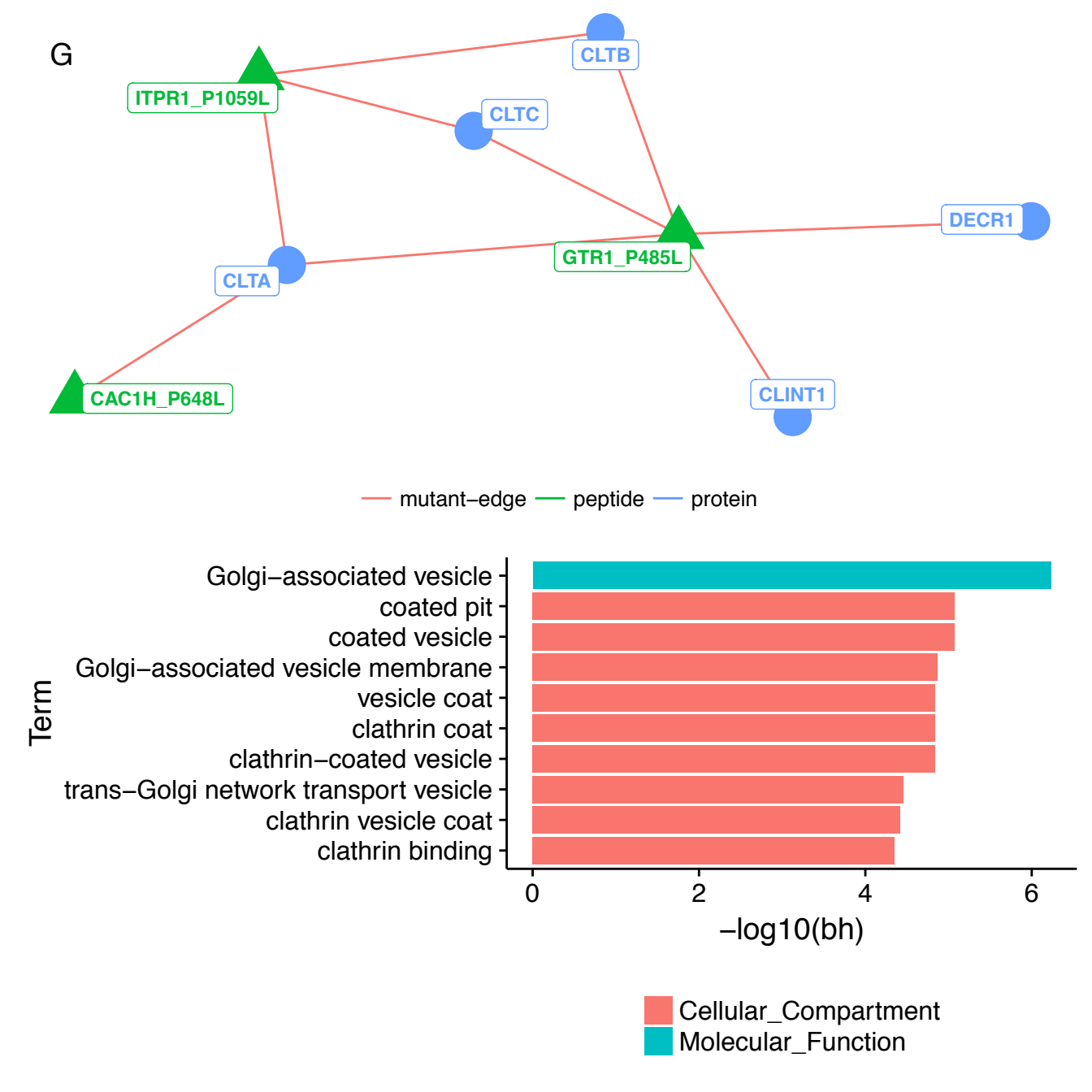

D
PANK2_E134G
SYEM_R168G
C1QBP
FUS_R216C
CASR_R898Q
PTMA
ANP32B
ANP32E
SETX_R1294C
ANP32A
DSRAD_K999N
SET
WWOX_P47T
ITF2_R565W
HUWE1
TRA1
mutant−edge
peptide
protein
wt−edge
Term
nuclear part
nuclear lumen
membrane−enclosed lumen
organelle lumen
intracellular organelle lumen
0.0
0.5
1.0
1.5
2.0
−log10(bh)
Cellular_Compartment

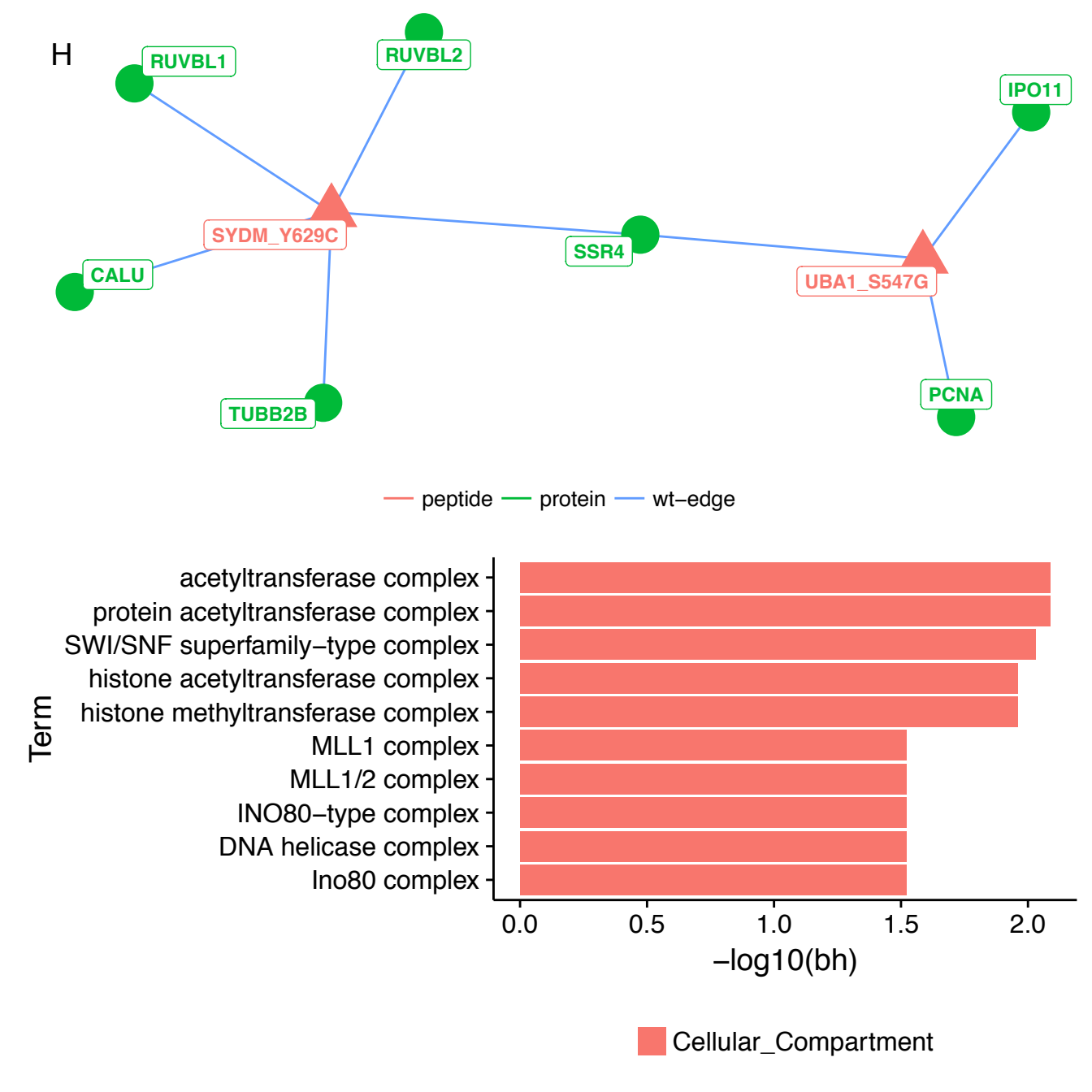


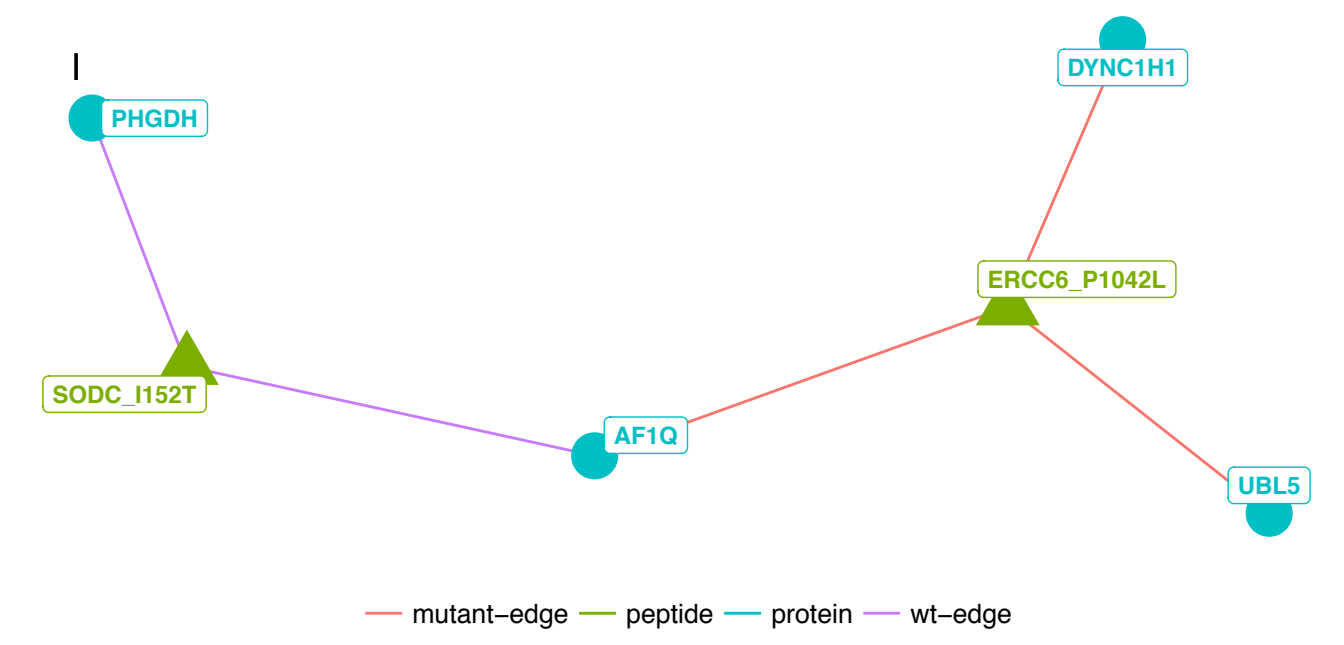
Term

−log10(bh)


F
DKFZp686F18120
SRSF2
SRSF1
SRSF6
SRSF9
FUS_R521C
SRSF3
SRSF10
SRSF5
peptide
protein
wt−edge
Term
mRNA transport
RNA splicing, via transesterification re...
mRNA splicing, via spliceosome
RNA splicing, via transesterification re...
termination of RNA polymerase II transcr...
ribonucleoprotein complex localization
RNA export from nucleus
ribonucleoprotein complex export from nu...
mRNA−containing ribonucleoprotein comple...
mRNA export from nucleus
0.0
2.5
5.0
7.5
10.0
12.5
−log10(bh)
Biological_Process

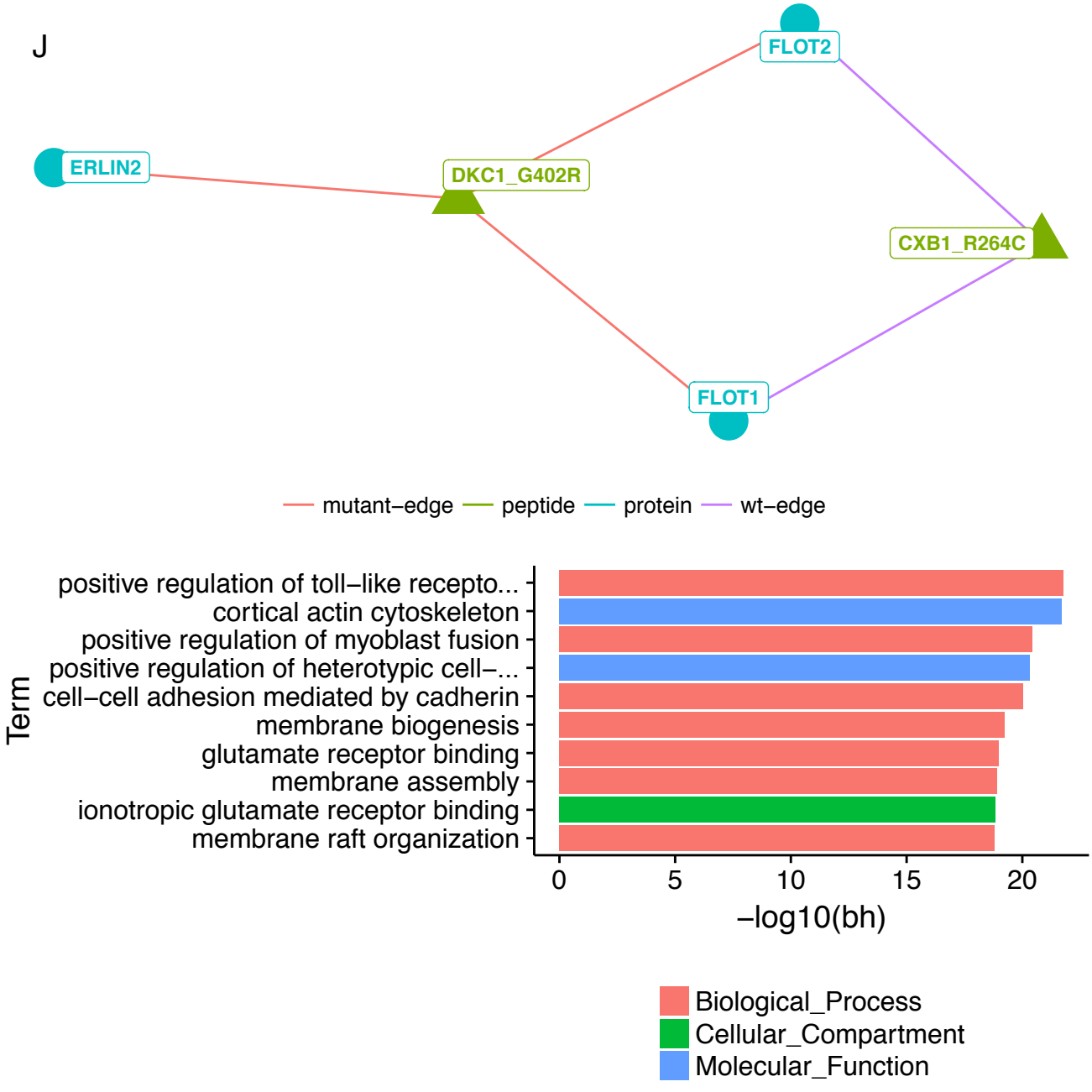

G
CLTB
ITPR1_P1059L
CLTC
DECR1
GTR1_P485L
CLTA
CAC1H_P648L
CLINT1
mutant−edge
peptide
protein
Term
Golgi−associated vesicle
coated pit
coated vesicle
Golgi−associated vesicle membrane
vesicle coat
clathrin coat
clathrin−coated vesicle
trans−Golgi network transport vesicle
clathrin vesicle coat
clathrin binding
0
2
4
6
−log10(bh)
Cellular_Compartment
Molecular_Function

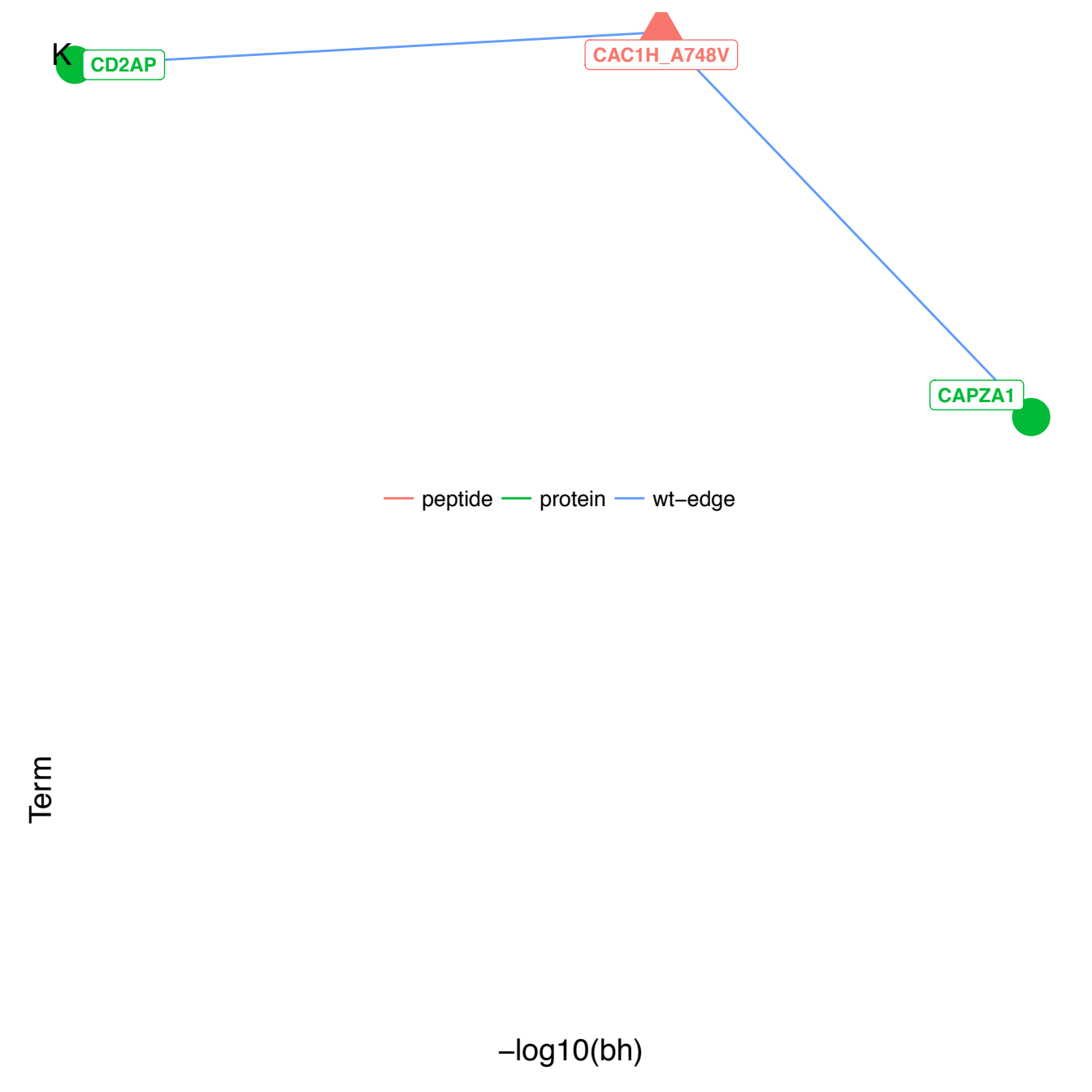

H
RUVBL1
RUVBL2
IPO11
SYDM_Y629C
SSR4
UBA1_S547G
CALU
TUBB2B
PCNA
peptide
protein
wt−edge
Term
acetyltransferase complex
protein acetyltransferase complex
SWI/SNF superfamily−type complex
histone acetyltransferase complex
histone methyltransferase complex
MLL1 complex
MLL1/2 complex
INO80−type complex
DNA helicase complex
Ino80 complex
0.0
0.5
1.0
1.5
2.0
−log10(bh)
Cellular_Compartment

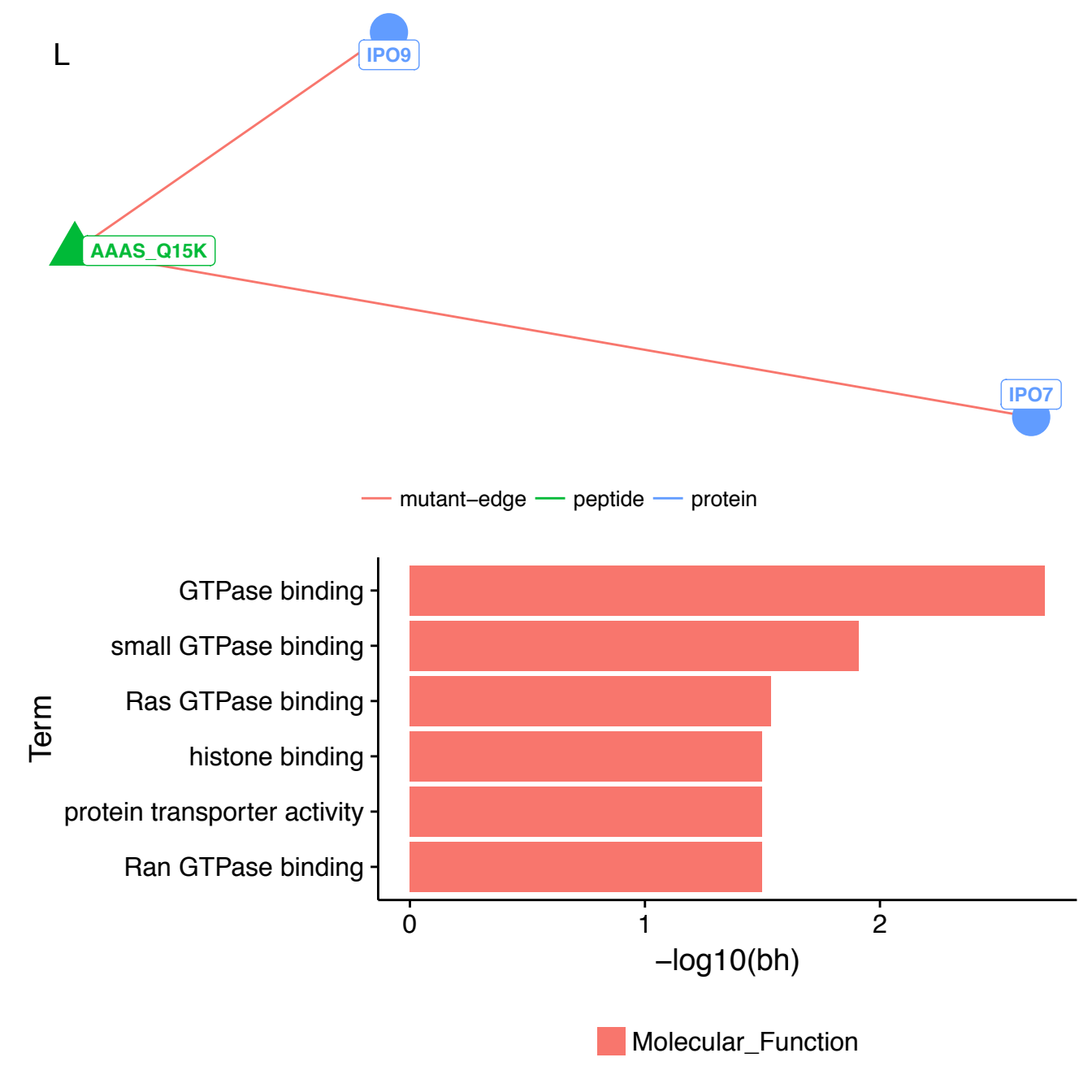


Term

−log10(bh)


J
FLOT2
ERLIN2
DKC1_G402R
CXB1_R264C
FLOT1
mutant−edge
peptide
protein
wt−edge
positive regulation of toll−like recepto...
cortical actin cytoskeleton
positive regulation of myoblast fusion
positive regulation of heterotypic cell−...
cell−cell adhesion mediated by cadherin
membrane biogenesis
glutamate receptor binding
membrane assembly
ionotropic glutamate receptor binding
membrane raft organization
Term
0
5
10
15
20
−log10(bh)
Biological_Process
Cellular_Compartment
Molecular_Function


K
CD2AP
CAC1H_A748V
CAPZA1
peptide
protein
wt−edge
Term
−log10(bh)


L
IPO9
AAAS_Q15K
IPO7
mutant−edge
peptide
protein
Term
GTPase binding
small GTPase binding
Ras GTPase binding
histone binding
protein transporter activity
Ran GTPase binding
0
1
2
−log10(bh)
Molecular_Function