Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Complex regulation of Gephyrin splicing is a determinant of inhibitory postsynaptic diversity

Raphael Dos Reis
Institute for Neurosciences of Montpellier (INM), Univ. Montpellier, CNRS, INSERM
Etienne Kornobis
Biomics, C2RT, Institut Pasteur https://orcid.org/0000-0001-7712-8270
Alyssa Pereira
Institute for Neurosciences of Montpellier (INM), Univ. Montpellier, CNRS, INSERM
Frédéric Tores
Bioinformatics Platform, Paris Descartes University, IMAGINE Institute, Paris
Judit Carrasco
Max Planck Institute of Immunobiology and Epigenetics https://orcid.org/0000-0002-4619-9452
Céline Jahannault-Talignani
Institut de Génomique Fonctionnelle
Patrick Nitschké
Université Paris Descartes https://orcid.org/0000-0002-2094-3298
Christian Muchardt
Epigenetics and RNA metabolism in human diseases – CNRS – UMR8256 – Institut de Biologie Paris-Seine
Andreas Schlosser
Institute of Structural Biology, Rudolf Virchow Center for Experimental Biomedicine, University of Würzburg https://orcid.org/0000-0003-0612-9932
Hans Maric
Rudolf Virchow Center, Center for Integrative and Translational Bioimaging, University of Würzburg
https://orcid.org/0000-0002-2719-4752
Fabrice Ango
IGF https://orcid.org/0000-0002-5548-209X
Eric Allemand (eric.allemand@inserm.fr)
INSERM https://orcid.org/0000-0002-9617-5336

Article

Keywords:

Posted Date: December 16th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-1111489/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Version of Record: A version of this preprint was published at Nature Communications on June 18th, 2022. See the published version at https://doi.org/10.1038/s41467-022-31264-w.

Complex regulation of Gephyrin splicing is a determinant of inhibitory postsynaptic diversity

One Sentence Summary: Alternative splicing generates a myriad of Gephyrin isoforms at iPSD

Abstract:

Gephyrin (*GPHN*) regulates the clustering of postsynaptic components at inhibitory synapses and is involved in pathophysiology of neuropsychiatric disorders. Here, we uncover an extensive diversity of *GPHN* transcripts that are tightly controlled by splicing during mouse and human brain development. Proteomic analysis reveals at least a hundred isoforms of GPHN incorporated at inhibitory Glycine and GABA-$_A$ receptors containing synapses. They exhibit different localization and postsynaptic clustering properties, and altering the expression level of one isoform is sufficient to affect the number, size, and density of inhibitory synapses in cerebellar Purkinje cells. Furthermore, we discovered that splicing defects reported in neuropsychiatric disorders are carried by multiple alternative *GPHN* transcripts, demonstrating the need for a thorough analysis of the *GPHN* transcriptome in patients. Overall, we show that alternative splicing of *GPHN* is an important genetic variation to consider in neurological diseases and a determinant of the diversity of postsynaptic inhibitory synapses.

Introduction:

As many as 60 different types of inhibitory interneurons co-exist in the brain [1,2] with distinct morphologies, physiological properties, connectivity patterns, and functions. Each of these interneurons established various heterogeneous inhibitory synapses that are arguably the quintessence of synaptic diversity. As such, it is surprising that our understanding of interneuron development and function is not supported by a better characterization of the molecules settling the diversity of their synapses. Recent studies using single-cell transcriptomic identified cell-type-specific repertoires of cell-surface and synaptic proteins expressed by various cell types, but how these factors shape the synaptic heterogeneity remain largely unknown [1,3].

GPHN is the central organizer of the inhibitory postsynaptic density (iPSD). *Gphn*-deficient mice lose the postsynaptic clustering of receptors [4-7], and therefore its activity is essential for inhibitory synaptic transmission [5,8]. GPHN harbors two protein domains, an N-terminal G-domain and a C-terminal E-domain, which are connected by an unstructured linker region [9]. Structural analysis suggests that GPHN postsynaptic clustering is formed by a network of G-domain trimers connected by E-domain dimers[10]. The linker region and the G-domain bind to multiple inhibitory synaptic molecules including GABA-$_A$ receptor subunits [11-15]. Besides the function of GPHN at inhibitory synapses, the G- and E-domains catalyze the last step of molybdenum cofactor biosynthesis [16]. Genetic alteration of this activity is associated with extensive neurological defects and early lethality [17,18]. In addition, dysfunction of GPHN-mediated neurotransmission has been implicated in severe disorders, such as Alzheimer's disease, autism, schizophrenia, epilepsy, and also in hyperekplexia [19-23].

The *Gphn* gene exhibits a complex intron-exon structure that consists of 29 exons, of which 9 are subjected to alternative splicing in tissue-specific manners [24]. Although alternative exons inclusion in *Gphn* transcripts were proposed to change the properties and functions of GPHN protein, only 2 and 16 expressed sequence tags (EST) are currently annotated for mice and humans respectively. Most significantly, patients with temporal lobe epilepsy without any genetic mutation express four abnormal *GPHN* splice variants[25]. Expressions of these irregular splice variants curtailed inhibitory synapse formation and were proposed as the main culprits for inhibitory circuit dysfunction. Yet, the precise repertoire of *GPHN* transcripts remains largely unknown and precludes our understanding of the function of GPHN at inhibitory synapses.

Here, using a targeted gene approach strategy combined with long-read sequencing, we have completely reframed the *GPHN* expression landscape and unveiled its extensive regulation by alternative splicing in mouse brain and human tissues. Such a large number of transcripts contributes significantly to the broad proteome of GPHN isoforms that interact with inhibitory Glycine and GABA-$_A$ receptors. Each GPHN transcript is regulated during brain development, and individual isoforms have different synaptic clustering properties. We show that inhibitory synapses harbor distinct combinations of GPHN protein isoforms, either at synapses present at specific subcellular locations, or containing distinct GABA-A receptors. Changing the level of a single splice variant in the pool of other *GPHN* transcripts is sufficient to affect the number, size, and density of inhibitory synapses in the cerebellum. In addition, aberrant GPHN splice variants identified in neuropsychiatric diseases are expressed at low levels in the healthy brain, indicating that their fine tuning is critical for the normal function of inhibitory synapses. Altogether, our data show that the extensive splicing regulation of *GPHN* expression is a determinant for inhibitory synapse diversity under physiological and pathological conditions

Results:

Extensive splicing regulation of *Gphn* during brain development in mice.

To explore the repertoire of *Gphn* transcripts during mouse brain development, we performed a targeted *Gphn* RNA analysis using Third Generation Sequencing technology developed by Pacific Bioscience (PACBIO). This approach provides high throughput long-read sequencing of full-length transcripts, and therefore the complete analysis of exon combinations incorporated in distinct splice variants. *Gphn* cDNAs were amplified from the mouse cortex and cerebellum at four postnatal stages (Fig. 1A). Among the 114458 sequences, our custom-made bioinformatic pipeline retained 42033 high-quality circular consensus sequences (CCS; and Fig.S1A). They were further grouped by clusters of identical sequences, each defining a *Gphn* alternative mRNA transcript (Fig. S1B and 1B). In total, we identified 277 unique transcripts (*Gphn*-1 to *Gphn*-277, ordered by descending abundance; Fig. S1B and Table S1), demonstrating that *Gphn* is subjected to extensive alternative splicing regulation of 40 exons (Fig.S1C and S1D). Among the 9 splicing cassettes described in the literature, splice cassettes G1, C4b, and E1

remained undetectable in our dataset, suggesting that they are expressed outside the brain and the E2 cassette was only detected in two splice variants (*Gphn*-197 and *Gphn*-181) (Fig. S1D). Our data further unveiled five entirely novel *Gphn* exons, as well as multiple exons annotated to alternative transcript(s) but, so far, without experimental validation. We also detected hundreds of new *Gphn* splice junctions, as well as several unannotated alternative 3' splice sites (ss) (Fig. S1E to S1I). Each exon is part of a myriad of different transcripts, emphasizing how difficult it is to detect a unique transcript by conventional method. in particular to amplify a specific PCR product and assess its expression level (Fig. S1C).

Using available public database, we built a library of 2.9 billion short reads sequenced to analyze the global gene expression, and found that a significant part of the exon-exon junctions identified in this study were not detected previously, unveiling the importance of targeted long read sequencing to capture the complexity of *Gphn* expression (Fig. 1C and Table S2). Our data completely reframes our understanding of *Gphn* expression. In contrast to the current view that counts 20 *Gphn* exons as constitutive we found that all internal exons were alternatively spliced. To the best of our knowledge, *Gphn* is the first gene (>10 exons) for which all internal exons are subjected to alternative splicing regulation (SM1 and Fig. S1B and S1C).

Next, we analyzed the expression of specific transcripts in space and time comparing SMRT sequencing in cerebellum and cortex at the four post-birth developmental stages (Fig. 2D). The most abundant transcript *Gphn*-1 was stably expressed in both tissues (less than 15% variation between P6 to P39). In contrast, many other abundant transcripts (*Gphn*-2 to -60) displayed distinct expression profiles across developmental stages, and several of them got different regulation in cortical and cerebellar cortices. By limiting RT-PCR assays only to transcripts containing unannotated alternative 3'ss, we validated their expression including those detected with only one CCS (*Gphn*-192, -203 and -244, Fig. 1E, S1E-I). Several variants displayed a differential expression in brain versus muscle or heart, supporting a fine-tuning tissue-specific regulation of *Gphn* transcripts diversity. To further explore this possibility, we used publicly available RNA-seq data to expand our analysis over eight different cell types of the central nervous system, including neurons and astrocytes (Fig. 1F, Table S3 and [26]). We detected several of the unannotated *Gphn* exons identified in our PACBIO sequencing (E9, E16, E22),

and acquired evidence for cell-type specific regulation of *Gphn* exons through alternative splicing in the brain. Altogether, our data redefines the *Gphn* locus with forty, rather than thirty exons as described previously (Fig. S1D) and reveals a previously unforeseen number of *Gphn* transcripts regulated during brain development.

***Gphn* splicing remodels functional domains of GPHN through a complex isoform proteome.**

Whether the *Gphn* transcriptome is translated to protein isoforms or rather retained in the nucleus as "transcriptional noise" remains a critical question [27]. To tackle this question, we first used purified polysomes from the mouse brain, we performed RT-PCR analysis and confirmed that novel *Gphn* transcripts are associated with the translation machinery, independently on their expression levels (Fig. S2A). Second, we generated the theoretical proteome of GPHN *in silico* by tracking open reading frames (ORFs) spanning at least three hundred nucleotides (Fig. 2A and Table S4) and filtered Translation Initiation Starts (TISs) with ribosome profiling data as biological evidence of their activation [28]. Our approach retained 9 out of 31 potential TISs and 253 out of 402 ORFs. Interestingly, we found cases in which several alternative transcripts encoded the same ORF suggesting that multiple distinct spliced messenger RNAs could regulate a unique protein isoform. Removal of duplicated ORFs and those subjected to degradation by nonsense-mediated mRNA decay due to existence of premature stop codon (PTC), led to list the theoretical GPHN proteome to 154 potential protein isoforms. The predicted molecular weights of the distinct GPHN isoforms are surprisingly close, suggesting that the separation of most abundant isoforms by conventional electrophoresis cannot discriminate them, thus leading to the detection of essentially a single band in western blots using anti-GPHN antibodies (Fig. S2B and S2C). Third, we performed in-depth proteomic analysis of the GPHN proteins isolated from the mouse brain to interrogate existing GPHN isoforms expression from our theoretical prediction. To this end, GPHN isoforms were isolated using a functional interaction assay that pulled-down all protein isoforms capable of interacting with GABAergic and Glycinergic receptors (Fig. 2B and [29,30]). Because GPHN is known to form multimeric assembly, we expect that this approach will also isolate associated isoforms with no receptor binding site. Tandem Mass Spectrometry (MS) analysis was performed with the theoretical proteome list of peptides resulting from the proteolytic cleavage of the novel GPHN ORFs and identified 2428 unique peptides that encompassed over 95% of the predicted GPHN proteome (Fig. 2B and Table S5). In particular,

the peptides corresponding to 22 undocumented exon-exon junctions were validated as well as those matching 3 new exons, and 1 uncanonical TIS (Fig. S2E and S2F). Altogether, MS provided experimental evidence supporting the expression of more than 140 new GPHN isoforms in mouse brain demonstrating that the complexity of the *Gphn* transcriptome is indeed transferred to its proteome. Thus, our data reveal a myriad of GPHN isoforms expressed at the postsynaptic densities (iPSD) of inhibitory synapses.

Distinct inhibitory synapses display a specific pattern of GPHN isoforms in cerebellum.

To study the complexity of GPHN isoforms at iPSD, a differential immunofluorescent assay was developed. Our approach used four antibodies, each recognizing different epitopes in the central region and E domain of mouse GPHN (Fig. S3A). The protein isoform GPHN-1 that harbors all of the epitopes was detected by all antibodies, while other isoforms lacking one of the epitopes were differentially recognized by a specific combination of antibodies (Fig. S3B and S3C).

Using these antibodies on cerebellar slices, we found that clusters of GPHN are heterogeneously marked, indicating a diversity of incorporated epitopes which was consistent with specific combinations of GPHN isoforms (Fig. 3B, 3C and S3D, S3E). To explore whether heterogeneous GPHN clusters were synaptic. We stained synaptic sites using specific presynaptic proteins (GAD-65 or VGAT) together with GPHN antibodies in cerebellar slices of mouse brain (Fig. S3D). Although GPHN is the ubiquitous marker of inhibitory postsynaptic sites, we found that none of the GPHN antibodies singly labelled all inhibitory synapses (Fig. S3E). We repeated this assay with combinations of anti-GPHN antibodies and found varying degrees of overlap in GPHN labeling at synapses. Our quantification further revealed the variety of GPHN epitopes combination at individual inhibitory synapses (Fig. 3B and 3C). Thus, *Gphn* expression is highly heterogeneous at inhibitory synapses.

One remarkable feature of inhibitory interneurons is the specific and highly structured axonal arbors with subcellular synapse specificity [e.g., dendrites, somata, or axon initial segments (AISs)] to control the input, integration, and output of their target cells. Therefore, we characterize the profiles of GPHN epitopes associated with inhibitory synapse at specific subcellular domains. For this purpose, we analyzed GPHN expression at Axon Initial Segments (AIS), soma and dendrites of cerebellar Purkinje cells (Fig. 3D). By combining one, two or three

GPHN antibodies simultaneously, we found that subcellular domain-specific synapses harbor distinct combinations of GPHN epitopes (Fig. 3D, S3F and S3G), supporting a specific subcellular distribution of GPHN isoforms.

At the molecular level, GPHN interacts with several GABA-A and Gly receptors using a universal binding domain[30], although the molecular mechanisms that control the specific enrichment of one receptor subunit over the other remained unknown, we characterized GPHN expression pattern at selected synapses including α1, α3 or α6 GABA-AR subunits in cerebellar cortex [31-34] (Fig. 3E, S3H to S3J). We discovered a distinct pattern of GPHN epitopes at each synapse. Altogether, our data show that GPHN is differentially distributed in inhibitory synapses depending on their subcellular localization and the expression of subtype specific inhibitory receptors .

GPHN isoforms have different properties at inhibitory synapses.

The diversity of GPHN proteins unveil several isoforms for which alternative splicing deeply altered the canonic functional domains required for GPHN synaptic activities [24]. In particular, we identified isoforms with unknown function that arbor important deletion in the G- (GPHN-5, 7 and 14) or the E-domains (Gphn-10, 28, 32, 42 and 49), as well as new insertion in the central domain (GPHN-6 and Gphn-8) (Fig. S4A). Scarlet-tagged GPHN isoforms constructs were assayed for their aggregation properties and inhibitory synapse localization in hippocampal primary neuronal culture (Fig. S4B). Mature primary hippocampal neurons expressing isoforms were analyzed 14 days in culture after lentiviral transduction. Most isoforms that harbor the complete E domain (GPHN-1, GPHN-5, GPHN-6 and GPHN-8) formed clusters at inhibitory synapses in contrast to GPHN-14 for which expression appeared both diffused and clustered (Fig. 4A). Interestingly, skipping of exon 10 is the only difference in the exon architecture of *Gphn*-1 and -14 showing how alternative splicing can influence the expression pattern of GPHN isoforms at the synapse (Fig. S4C). In contrast, proteins isoforms with altered E-domain (GPHN-7, GPHN-10, GPHN-28, GPHN-32, GPHN-42 and GPHN-49) were severely affected in their clustering capacity and displayed a diffuse distribution in neurons (Fig. 4A). Interestingly, we noticed that these isoforms have a similar localization to GPHN fragments previously characterized for their dominant negative activity that impair the anchoring of GABA-A receptors at the synapse [20,35].

We further analyzed the cluster properties (size, number and synaptic localization) of each GPHN isoforms able to multimerize and found that the number and densities of these clusters at the dendrites were different (Fig. 4C and 4D). GPHN-1 induced the highest density of clusters while the presence of short additional splicing cassettes in *Gphn*-6 and *Gphn*-8 reduced cluster numbers (Fig. S4D). Removal of exons affecting the integrity of G domain in GPHN-5 and GPHN-14 led to a critical decrease of cluster density (Fig. 4C and S4C, S4E). Interestingly, the size of clusters containing GPHN-6 and GPHN-8 were also significantly larger (~30%) suggesting that both isoforms have better multimerization properties (Fig. 4D). Because the central unstructured region has many post-translational modifications that regulate the size of GPHN clusters, it is also possible that GPHN-6 and GPHN-8 are differently regulated by post-translational modifications[36].

GPHN-5, GPHN-8 and GPHN-14 were more often localized at inhibitory synapses than GPHN-1 and GPHN-6 (Fig. 4E). We also showed that the density of inhibitory synapses in dendrites was influenced by the expression of specific GPHN isoforms *in vitro* (Fig. 4F), suggesting that a change in the level of a specific GPHN isoform could modulate the density of synapses.

To explore how change in the level of GPHN isoforms affect synapse formation *in vivo*, we injected lentiviral constructs in the mouse brain, and quantified the number of inhibitory synapses in cerebellum (Fig. 5A and 5B). Our experimental conditions led to an increase in the expression of GPHN isoform 12 days post-injection, mimicking therefore an up regulation of a specific splice variant in the pool of all alternative *Gphn* transcripts expressed endogenously. Exogenous expression of GPHN-1, GPHN-6, GPHN-10, GPHN-32 and GPHN-49 led to a similar synapses density than the control Scarlett (Fig. 5C). In contrast, GPHN-14, GPHN-28 and GPHN-42 exhibited a decreased number of synapses, while GPHN-8 enhanced the density of synapses, unveiling negative and positive dominant activities of several isoforms. These data demonstrated how the balance of *Gphn* alternative transcripts can direct the inhibitory synapse formation and/or maintenance *in vivo*. In addition, the remodeling of functional domains by alternative splicing is a mechanism to enrich selectively distinct GPHN isoforms to $GABA_A$ circuitry and control their clustering properties. These results sustain our initial observation about the heterogeneity of endogenous GPHN epitopes and further document that distinct types of inhibitory synapses are assembled with an array of different GPHN isoforms.

To further test that specific GPHN isoforms are heterogeneously distributed, and are not present at all synapses of an individual cell. We analyzed the distribution of GPHN-1 and GPHN-8 on the dendrite of Purkinje cells *in vivo*. We found that the GPHN isoforms were recruited to post-synaptic sites with distinct efficiency: GPHN-1 and -8 were detected at 50% and 65% of GAD65 positive synapses respectively (Fig. 6A and 6B). These results confirmed our initial observations in primary cultures of hippocampal neurons although the differences between isoforms were less pronounced (Fig. 4E). In addition, we also found that both isoforms were detected at the axon initial segment of PCs, i.e. at the site where the action potential is initiated following integration of excitatory and inhibitory inputs (Fig. 6C). These results show that GPHN isoforms regulate inhibitory synapse properties at distinct subcellular domains. Our *in vivo* results corroborated our *in vitro* experiments (Fig. 4E) and demonstrated that specific inhibitory synapses are settled with different ratios of GPHN isoforms.

Analysis of *GPHN* splice variant diversity is required to uncover pathological splicing regulation.

Expression of aberrant GPHN splice variants is associated with several neurological disorders such as epilepsy, autism, schizophrenia, Alzheimer's disease, and hyperekplexia (Fig. 8F), prompting a deeper understanding of the regulation of GPHN splice variants in humans. Here, we characterized GPHN expression in the fetal and adult brains, cerebellum, and 18 other human tissues using targeted long-read sequencing (Fig. 7A, S5A). We found an extensive diversity of *GPHN* transcripts in the human tissues, with 1040 unique mRNA produced from the alternative inclusion of 42 exons, some of which have tissue-specific expression profiles (Fig. 7A, S5B, S5C and Table 6). As observed in Mouse, all internal exons are alternatively included in messengers, and *GPHN* expresses a major transcript encompassing ~67% of total RNA level (Fig. 1B, 7B). Distribution analysis of splice variants highlighted 226 transcripts as a common core to all tissues, while some displayed a tissue specific expression, suggesting different cellular functions associated with distinct GPHN isoforms (Fig. 7C, 7D). Interestingly, regardless of *GPHN* expression level, all tissues had a comparable number of *GPHN* splice variants (Fig. 7D and 7E), except in the brain which had the largest number of GPHN alternative transcripts. Among the 3 brain samples, the pool of mRNAs was very similar suggesting a dedicated

regulation of *GPHN* splicing to this tissue (Fig. S5D). Moreover, the diversity of *GPHN* expression displayed a high regulation between the adult and fetal stages, but also between the cerebellum and whole brain, demonstrating an important modulation during development and across different brain areas (Fig. 7F and Table S6). Overall, our data revealed a much more diverse and regulated expression of human *GPHN* than the 16 alternative transcripts currently annotated in databases.

Among the *GPHN* transcriptome, we sought transcripts containing exon-exon junctions mimicking the genetic variations of *GPHN* previously characterized in patients with neurological disorders (Fig. 8A and 8B). Surprisingly, almost all pathogenic exon-exon junctions were detected in healthy human brain samples and we noticed that they were all included in transcripts expressed at low level, from 0.02 to 2.7% of the total *GPHN* expression (Fig. 8C). Although many of these exon-exon junctions have been classified as irregular splicing events uncovered in patients, our work demonstrates that they are actually canonic splicing events expressed in healthy individuals. Furthermore, these exon-exon junctions are present in multiple transcripts, in particular the junction between exon 5 and 12 that was detected in 82 different transcripts (Fig. 8C). We noted that the abundance and diversity of these pathogenesis-related transcripts are differently controlled through brain areas. Overall, our study shows that a detailed appreciation of the GPHN splice variant landscape is required to evaluate aberrant expression of GPHN isoforms in neurological diseases.

Discussion

Despite the progress made in understanding the formation of inhibitory synapses, little is known about the molecular diversity of inhibitory synapses. Here, we demonstrate that the major scaffolding protein of inhibitory synapses, namely GPHN, is a marker of synapse diversity. *GPHN* expresses a complex landscape of transcripts tightly regulated by alternative splicing to encode a myriad of protein isoforms. We show that they are differentially assembled to specific subclasses of inhibitory synapses including those with distinct GABA-$_A$ receptors or connecting at different subcellular localization. Diversity of *GPHN* expression is highly regulated during developmental stages, between different tissues or brain areas. Among the diversity of gephyrin isoforms, the change in the expression level of a single variant is sufficient to affect the formation/maintenance, size and density of inhibitory synapses *in vivo*. Furthermore, we demonstrate that aberrant *GPHN* splicing, reported in neuropsychiatric pathologies (schizophrenia, Autism Spectrum Disorders (ASDs), and epileptogenesis), are encompassed in a myriad of alternative transcripts that are expressed at low levels in brain of healthy individuals. This finding argues for further in-depth analysis of the *GPHN* transcriptome in patients to decipher what part of its expression might be affected and which isoform is clinically relevant.

Splicing regulation of *Gphn*

The regulation of mouse and human *GPHN* splicing must be controlled by a complex network of factors, especially because we found that all internal exons are alternatively spliced (Fig. S1B, 1F and S5A). Alternative splicing is catalysed by the spliceosome, an enzyme including more than 200 factors, as well as the influence of other gene expression machinery such as chromatin and RNA polymerase II[37]. Analysis of global gene expression in knockout mice or high-throughput screening of RNA-binding proteins have highlighted the splicing factors Nova, Sam68, PTBP2, and Rbfox as potential regulator of *GPHN* splicing[38-40]. However, it is likely that *GPHN* splicing regulators are not limited to this small set of factors. Furthermore, each human tissue is expected to have distinct regulation of *GPHN* splicing, especially in the brain where its expression is extremely diverse.

Although single-cell analysis has delineated cardinal classes of inhibitory interneurons with a specific transcriptional design to encode their synaptic communication, our study and others show an underestimation of splicing regulation[1,41-45]. Analysis of global gene expression

using high-throughput RNA-seq technologies (NGS or TGS) can only provide an overall assessment of RNAs transcribed by each gene. Indeed, the most expressed genes and the most abundant splice variants saturate by competition the sequencing capacities, which hinders the analysis of less represented splice variants. In contrast, targeted RNA-seq analysis leads to a more qualitative analysis of gene expression reducing the intrinsic complexity of samples. Applying this strategy to study *GPHN* allowed us to unveil its complex, fine and dynamic regulation by alternative splicing. The procedure used to analyze gene expression is therefore a critical point to evaluate the portfolios of alternative mRNA isoforms. Recent studies have reached similar conclusions and demonstrated that overlooked isoforms may carry novel disease-relevant gene functions[41,46].

Expression of GPHN isoforms shape inhibitory synapse diversity

How inhibitory synapses diversity is regulated remains entirely unclear, in contrast to excitatory synapses, for which specificity and diversity result primarily from the combination of several postsynaptic scaffolding proteins encoded by different genes [47-52]. Here, we found that the gene encoding for the central organizer of iPSD expresses more than hundreds of isoforms, and each has a unique structural organization, and therefore distinct partner binding properties at iPSD[24]. GPHN isoforms do not localize at all inhibitory synapses and specific GABA-$_A$R containing synapses displayed different patterns of GPHN epitopes, suggesting that each synapse contains a specific ratio of GPHN isoforms. The specific GPHN isoforms distribution within a single cell could represent local activity-dependent regulation at specific GABA-AR containing synapses. Indeed, four aberrant splicing isoforms of *GPHN* were observed following cellular stress[25]. In addition, NOVA has been shown to regulate the alternative splicing of Gephyrin[39] and NOVA function is modulated by electrical activity[53], thus, highly suggesting that *GPHN* splicing variant could also be regulated during change in the excitation/inhibition balance in neurons. Conversely, exons microdeletion in *GPHN* induced neuronal dysfunction leading to neuropsychiatric diseases. Several exon skipping in the G-domain is sufficient to disrupt gephyrin clusters. Disruption of gephyrin clusters was shown to decrease the amplitude and frequency of spontaneous GABAergic synaptic currents[54]. Since transient receptor-scaffold interactions govern the “diffusion trapping” of the receptors at postsynaptic sites, the distinct synaptic scaffold properties, distribution and dynamics of GPHN isoforms are expected to impact

the synaptic strength[55,56]. Altogether, our study identifies GPHN isoforms as a key factor to better understand changes in inhibitory synapse function and diversity in health and diseases.

***GPHN* splicing regulation in human neuropathologies.**

Characterization of pathological genetic variations in *GPHN* gene locus has so far been limited to the search for coding sequence mutations using data of exome sequencing and, more recently, whole genome sequencing. However, ~90% of SNPs associated with traits/diseases are annotated in intronic (45%) or intergenic (43%) regions[57]. Therefore, it is likely that among these non-coding genetic variations, some dysregulate *GPHN* splicing. Either by affecting cis-splicing regulatory sequences located in introns, or by changing the expression of splicing factor that modulate *GPHN* splice variants. Targeted RNA-seq analysis of full-length transcripts should be extended in the future to analyze the expression of factors involved in the synapse plasticity, especially in the context of patients affected with neuronal disorders. This approach should unveil the genetic mechanism behind the dysregulation of highly conserved genes and open new avenues for treating these diseases.

References

1 Paul, A., Crow, M., Raudales, R., He, M., Gillis, J. & Huang, Z. J. Transcriptional Architecture of Synaptic Communication Delineates GABAergic Neuron Identity. *Cell* **171**, 522-539 e520, doi:10.1016/j.cell.2017.08.032 (2017).

2 Wamsley, B. & Fishell, G. Genetic and activity-dependent mechanisms underlying interneuron diversity. *Nat Rev Neurosci* **18**, 299-309, doi:10.1038/nrn.2017.30 (2017).

3 Huang, Z. J. & Paul, A. The diversity of GABAergic neurons and neural communication elements. *Nat Rev Neurosci* **20**, 563-572, doi:10.1038/s41583-019-0195-4 (2019).

4 Essrich, C., Lorez, M., Benson, J. A., Fritschy, J. M. & Luscher, B. Postsynaptic clustering of major GABAA receptor subtypes requires the gamma 2 subunit and gephyrin. *Nat Neurosci* **1**, 563-571, doi:10.1038/2798 (1998).

5 Feng, G., Tintrup, H., Kirsch, J., Nichol, M. C., Kuhse, J., Betz, H. & Sanes, J. R. Dual requirement for gephyrin in glycine receptor clustering and molybdoenzyme activity. *Science* **282**, 1321-1324, doi:10.1126/science.282.5392.1321 (1998).

6 Kim, E. Y., Schrader, N., Smolinsky, B., Bedet, C., Vannier, C., Schwarz, G. & Schindelin, H. Deciphering the structural framework of glycine receptor anchoring by gephyrin. *EMBO J* **25**, 1385-1395, doi:10.1038/sj.emboj.7601029 (2006).

7 Tretter, V., Jacob, T. C., Mukherjee, J., Fritschy, J. M., Pangalos, M. N. & Moss, S. J. The clustering of GABA(A) receptor subtypes at inhibitory synapses is facilitated via the direct binding of receptor alpha 2 subunits to gephyrin. *J Neurosci* **28**, 1356-1365, doi:10.1523/JNEUROSCI.5050-07.2008 (2008).

8 Kneussel, M., Brandstatter, J. H., Laube, B., Stahl, S., Muller, U. & Betz, H. Loss of postsynaptic GABA(A) receptor clustering in gephyrin-deficient mice. *J Neurosci* **19**, 9289-9297 (1999).

9 Tyagarajan, S. K. & Fritschy, J. M. Gephyrin: a master regulator of neuronal function? *Nat Rev Neurosci* **15**, 141-156, doi:10.1038/nrn3670 (2014).

10 Sander, B., Tria, G., Shkumatov, A. V., Kim, E. Y., Grossmann, J. G., Tessmer, I., Svergun, D. I. & Schindelin, H. Structural characterization of gephyrin by AFM and SAXS reveals a mixture of compact and extended states. *Acta Crystallogr D Biol Crystallogr* **69**, 2050-2060, doi:10.1107/S0907444913018714 (2013).

11 Favuzzi, E., Deogracias, R., Marques-Smith, A., Maeso, P., Jezequel, J., Exposito-Alonso, D., Balia, M., Kroon, T., Hinojosa, A. J., E, F. M. & Rico, B. Distinct molecular programs regulate synapse specificity in cortical inhibitory circuits. *Science* **363**, 413-417, doi:10.1126/science.aau8977 (2019).

12 Fuhrmann, J. C., Kins, S., Rostaing, P., El Far, O., Kirsch, J., Sheng, M., Triller, A., Betz, H. & Kneussel, M. Gephyrin interacts with Dynein light chains 1 and 2, components of motor protein complexes. *J Neurosci* **22**, 5393-5402, doi:20026552 (2002).

13 Giesemann, T., Schwarz, G., Nawrotzki, R., Berhorster, K., Rothkegel, M., Schluter, K., Schrader, N., Schindelin, H., Mendel, R. R., Kirsch, J. & Jockusch, B. M. Complex formation between the postsynaptic scaffolding protein gephyrin,

profilin, and Mena: a possible link to the microfilament system. *J Neurosci* **23**, 8330-8339 (2003).

14 Grosskreutz, Y., Hermann, A., Kins, S., Fuhrmann, J. C., Betz, H. & Kneussel, M. Identification of a gephyrin-binding motif in the GDP/GTP exchange factor collybistin. *Biol Chem* **382**, 1455-1462, doi:10.1515/BC.2001.179 (2001).

15 Sabatini, D. M., Barrow, R. K., Blackshaw, S., Burnett, P. E., Lai, M. M., Field, M. E., Bahr, B. A., Kirsch, J., Betz, H. & Snyder, S. H. Interaction of RAFT1 with gephyrin required for rapamycin-sensitive signaling. *Science* **284**, 1161-1164, doi:10.1126/science.284.5417.1161 (1999).

16 Mendel, R. R. The molybdenum cofactor. *J Biol Chem* **288**, 13165-13172, doi:10.1074/jbc.R113.455311 (2013).

17 Reiss, J. & Hahnewald, R. Molybdenum cofactor deficiency: Mutations in GPHN, MOCS1, and MOCS2. *Hum Mutat* **32**, 10-18, doi:10.1002/humu.21390 (2011).

18 Reiss, J., Lenz, U., Aquaviva-Bourdain, C., Joriot-Chekaf, S., Mention-Mulliez, K. & Holder-Espinasse, M. A GPHN point mutation leading to molybdenum cofactor deficiency. *Clin Genet* **80**, 598-599, doi:10.1111/j.1399-0004.2011.01709.x (2011).

19 Agarwal, S., Tannenberg, R. K. & Dodd, P. R. Reduced expression of the inhibitory synapse scaffolding protein gephyrin in Alzheimer's disease. *J Alzheimers Dis* **14**, 313-321, doi:10.3233/jad-2008-14305 (2008).

20 Dejanovic, B., Djemie, T., Grunewald, N., Suls, A., Kress, V., Hetsch, F., Craiu, D., Zemel, M., Gormley, P., Lal, D., Euro, E. D. w. g., Myers, C. T., Mefford, H. C., Palotie, A., Helbig, I., Meier, J. C., De Jonghe, P., Weckhuysen, S. & Schwarz, G. Simultaneous impairment of neuronal and metabolic function of mutated gephyrin in a patient with epileptic encephalopathy. *EMBO Mol Med* **7**, 1580-1594, doi:10.15252/emmm.201505323 (2015).

21 Dejanovic, B., Lal, D., Catarino, C. B., Arjune, S., Belaidi, A. A., Trucks, H., Vollmar, C., Surges, R., Kunz, W. S., Motameny, S., Altmuller, J., Kohler, A., Neubauer, B. A., Epicure, C., Nurnberg, P., Noachtar, S., Schwarz, G. & Sander, T. Exonic microdeletions of the gephyrin gene impair GABAergic synaptic inhibition in patients with idiopathic generalized epilepsy. *Neurobiol Dis* **67**, 88-96, doi:10.1016/j.nbd.2014.02.001 (2014).

22 Fang, M., Shen, L., Yin, H., Pan, Y. M., Wang, L., Chen, D., Xi, Z. Q., Xiao, Z., Wang, X. F. & Zhou, S. N. Downregulation of gephyrin in temporal lobe epilepsy neurons in humans and a rat model. *Synapse* **65**, 1006-1014, doi:10.1002/syn.20928 (2011).

23 Hales, C. M., Rees, H., Seyfried, N. T., Dammer, E. B., Duong, D. M., Gearing, M., Montine, T. J., Troncoso, J. C., Thambisetty, M., Levey, A. I., Lah, J. J. & Wingo, T. S. Abnormal gephyrin immunoreactivity associated with Alzheimer disease pathologic changes. *J Neuropathol Exp Neurol* **72**, 1009-1015, doi:10.1097/01.jnen.0000435847.59828.db (2013).

24 Fritschy, J. M., Harvey, R. J. & Schwarz, G. Gephyrin: where do we stand, where do we go? *Trends Neurosci* **31**, 257-264, doi:10.1016/j.tins.2008.02.006 (2008).

25 Forstera, B., Belaidi, A. A., Juttner, R., Bernert, C., Tsokos, M., Lehmann, T. N., Horn, P., Dehnicke, C., Schwarz, G. & Meier, J. C. Irregular RNA splicing curtails

postsynaptic gephyrin in the cornu ammonis of patients with epilepsy. *Brain* **133**, 3778-3794, doi:10.1093/brain/awq298 (2010).

26 Zhang, Y., Chen, K., Sloan, S. A., Bennett, M. L., Scholze, A. R., O'Keeffe, S., Phatnani, H. P., Guarnieri, P., Caneda, C., Ruderisch, N., Deng, S., Liddelow, S. A., Zhang, C., Daneman, R., Maniatis, T., Barres, B. A. & Wu, J. Q. An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. *J Neurosci* **34**, 11929-11947, doi:10.1523/JNEUROSCI.1860-14.2014 (2014).

27 Pertea, M., Shumate, A., Pertea, G., Varabyou, A., Breitwieser, F. P., Chang, Y. C., Madugundu, A. K., Pandey, A. & Salzberg, S. L. CHESS: a new human gene catalog curated from thousands of large-scale RNA sequencing experiments reveals extensive transcriptional noise. *Genome Biol* **19**, 208, doi:10.1186/s13059-018-1590-2 (2018).

28 Fields, A. P., Rodriguez, E. H., Jovanovic, M., Stern-Ginossar, N., Haas, B. J., Mertins, P., Raychowdhury, R., Hacohen, N., Carr, S. A., Ingolia, N. T., Regev, A. & Weissman, J. S. A Regression-Based Analysis of Ribosome-Profiling Data Reveals a Conserved Complexity to Mammalian Translation. *Mol Cell* **60**, 816-827, doi:10.1016/j.molcel.2015.11.013 (2015).

29 Maric, H. M., Hausrat, T. J., Neubert, F., Dalby, N. O., Doose, S., Sauer, M., Kneussel, M. & Stromgaard, K. Gephyrin-binding peptides visualize postsynaptic sites and modulate neurotransmission. *Nat Chem Biol* **13**, 153-160, doi:10.1038/nchembio.2246 (2017).

30 Maric, H. M., Kasaragod, V. B., Hausrat, T. J., Kneussel, M., Tretter, V., Stromgaard, K. & Schindelin, H. Molecular basis of the alternative recruitment of GABA(A) versus glycine receptors through gephyrin. *Nat Commun* **5**, 5767, doi:10.1038/ncomms6767 (2014).

31 Laurie, D. J., Seeburg, P. H. & Wisden, W. The distribution of 13 GABAA receptor subunit mRNAs in the rat brain. II. Olfactory bulb and cerebellum. *J Neurosci* **12**, 1063-1076 (1992).

32 Fritschy, J. M., Panzanelli, P., Kralic, J. E., Vogt, K. E. & Sassoe-Pognetto, M. Differential dependence of axo-dendritic and axo-somatic GABAergic synapses on GABAA receptors containing the alpha1 subunit in Purkinje cells. *J Neurosci* **26**, 3245-3255, doi:10.1523/JNEUROSCI.5118-05.2006 (2006).

33 Eyre, M. D. & Nusser, Z. Only a Minority of the Inhibitory Inputs to Cerebellar Golgi Cells Originates from Local GABAergic Cells. *eNeuro* **3**, doi:10.1523/ENEURO.0055-16.2016 (2016).

34 Nusser, Z., Sieghart, W., Stephenson, F. A. & Somogyi, P. The alpha 6 subunit of the GABAA receptor is concentrated in both inhibitory and excitatory synapses on cerebellar granule cells. *J Neurosci* **16**, 103-114 (1996).

35 Costa, J. T., Mele, M., Baptista, M. S., Gomes, J. R., Ruscher, K., Nobre, R. J., de Almeida, L. P., Wieloch, T. & Duarte, C. B. Gephyrin Cleavage in In Vitro Brain Ischemia Decreases GABAA Receptor Clustering and Contributes to Neuronal Death. *Mol Neurobiol* **53**, 3513-3527, doi:10.1007/s12035-015-9283-2 (2016).

36 Tyagarajan, S. K., Ghosh, H., Yevenes, G. E., Imanishi, S. Y., Zeilhofer, H. U., Gerrits, B. & Fritschy, J. M. Extracellular signal-regulated kinase and glycogen synthase kinase 3beta regulate gephyrin postsynaptic aggregation and GABAergic synaptic function in a calpain-dependent mechanism. *J Biol Chem* **288**, 9634-9647, doi:10.1074/jbc.M112.442616 (2013).

37 Allemand, E., Myers, M. P., Garcia-Bernardo, J., Harel-Bellan, A., Krainer, A. R. & Muchardt, C. A Broad Set of Chromatin Factors Influences Splicing. *PLoS Genet* **12**, e1006318, doi:10.1371/journal.pgen.1006318 (2016).

38 Licatalosi, D. D., Mele, A., Fak, J. J., Ule, J., Kayikci, M., Chi, S. W., Clark, T. A., Schweitzer, A. C., Blume, J. E., Wang, X., Darnell, J. C. & Darnell, R. B. HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature* **456**, 464-469, doi:10.1038/nature07488 (2008).

39 Ule, J., Jensen, K. B., Ruggiu, M., Mele, A., Ule, A. & Darnell, R. B. CLIP identifies Nova-regulated RNA networks in the brain. *Science* **302**, 1212-1215, doi:10.1126/science.1090095 (2003).

40 Weyn-Vanhentenryck, S. M., Mele, A., Yan, Q., Sun, S., Farny, N., Zhang, Z., Xue, C., Herre, M., Silver, P. A., Zhang, M. Q., Krainer, A. R., Darnell, R. B. & Zhang, C. HITS-CLIP and integrative modeling define the Rbfox splicing-regulatory network linked to brain development and autism. *Cell Rep* **6**, 1139-1152, doi:10.1016/j.celrep.2014.02.005 (2014).

41 Ray, T. A., Cochran, K., Kozlowski, C., Wang, J., Alexander, G., Cady, M. A., Spencer, W. J., Ruzycki, P. A., Clark, B. S., Laeremans, A., He, M. X., Wang, X., Park, E., Hao, Y., Iannaccone, A., Hu, G., Fedrigo, O., Skiba, N. P., Arshavsky, V. Y. & Kay, J. N. Comprehensive identification of mRNA isoforms reveals the diversity of neural cell-surface molecules with roles in retinal development and disease. *Nat Commun* **11**, 3328, doi:10.1038/s41467-020-17009-7 (2020).

42 Clark, M. B., Wrzesinski, T., Garcia, A. B., Hall, N. A. L., Kleinman, J. E., Hyde, T., Weinberger, D. R., Harrison, P. J., Haerty, W. & Tunbridge, E. M. Long-read sequencing reveals the complex splicing profile of the psychiatric risk gene CACNA1C in human brain. *Mol Psychiatry* **25**, 37-47, doi:10.1038/s41380-019-0583-1 (2020).

43 Treutlein, B., Gokce, O., Quake, S. R. & Sudhof, T. C. Cartography of neurexin alternative splicing mapped by single-molecule long-read mRNA sequencing. *Proc Natl Acad Sci U S A* **111**, E1291-1299, doi:10.1073/pnas.1403244111 (2014).

44 Schreiner, D., Nguyen, T. M., Russo, G., Heber, S., Patrignani, A., Ahrne, E. & Scheiffele, P. Targeted combinatorial alternative splicing generates brain region-specific repertoires of neurexins. *Neuron* **84**, 386-398, doi:10.1016/j.neuron.2014.09.011 (2014).

45 Sun, W., You, X., Gogol-Doring, A., He, H., Kise, Y., Sohn, M., Chen, T., Klebes, A., Schmucker, D. & Chen, W. Ultra-deep profiling of alternatively spliced Drosophila Dscam isoforms by circularization-assisted multi-segment sequencing. *EMBO J* **32**, 2029-2038, doi:10.1038/emboj.2013.144 (2013).

46 Chau, K. K., Zhang, P., Urresti, J., Amar, M., Pramod, A. B., Chen, J., Thomas, A., Corominas, R., Lin, G. N. & Iakoucheva, L. M. Full-length isoform transcriptome of the developing human brain provides further insights into autism. *Cell Rep* **36**, 109631, doi:10.1016/j.celrep.2021.109631 (2021).

47 Grant, S. G. Toward a molecular catalogue of synapses. *Brain Res Rev* **55**, 445-449, doi:10.1016/j.brainresrev.2007.05.003 (2007).

48 Grant, S. G. N. The Synaptomic Theory of Behavior and Brain Disease. *Cold Spring Harb Symp Quant Biol* **83**, 45-56, doi:10.1101/sqb.2018.83.037887 (2018).

49 Grant, S. G. N. Synapse molecular complexity and the plasticity behaviour problem. *Brain Neurosci Adv* **2**, 2398212818810685, doi:10.1177/2398212818810685 (2018).

50 Micheva, K. D., Busse, B., Weiler, N. C., O'Rourke, N. & Smith, S. J. Single-synapse analysis of a diverse synapse population: proteomic imaging methods and markers. *Neuron* **68**, 639-653, doi:10.1016/j.neuron.2010.09.024 (2010).

51 O'Rourke, N. A., Weiler, N. C., Micheva, K. D. & Smith, S. J. Deep molecular diversity of mammalian synapses: why it matters and how to measure it. *Nat Rev Neurosci* **13**, 365-379, doi:10.1038/nrn3170 (2012).

52 Zhu, F., Cizeron, M., Qiu, Z., Benavides-Piccione, R., Kopanitsa, M. V., Skene, N. G., Koniaris, B., DeFelipe, J., Fransen, E., Komiyama, N. H. & Grant, S. G. N. Architecture of the Mouse Brain Synaptome. *Neuron* **99**, 781-799 e710, doi:10.1016/j.neuron.2018.07.007 (2018).

53 Eom, T., Zhang, C., Wang, H., Lay, K., Fak, J., Noebels, J. L. & Darnell, R. B. NOVA-dependent regulation of cryptic NMD exons controls synaptic protein levels after seizure. *Elife* **2**, e00178, doi:10.7554/eLife.00178 (2013).

54 Yu, W., Jiang, M., Miralles, C. P., Li, R. W., Chen, G. & de Blas, A. L. Gephyrin clustering is required for the stability of GABAergic synapses. *Mol Cell Neurosci* **36**, 484-500, doi:10.1016/j.mcn.2007.08.008 (2007).

55 Choquet, D. & Triller, A. The dynamic synapse. *Neuron* **80**, 691-703, doi:10.1016/j.neuron.2013.10.013 (2013).

56 Petrini, E. M. & Barberis, A. Diffusion dynamics of synaptic molecules during inhibitory postsynaptic plasticity. *Front Cell Neurosci* **8**, 300, doi:10.3389/fncel.2014.00300 (2014).

57 Hindorff, L. A., Sethupathy, P., Junkins, H. A., Ramos, E. M., Mehta, J. P., Collins, F. S. & Manolio, T. A. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci U S A* **106**, 9362-9367, doi:10.1073/pnas.0903103106 (2009).

58 Mauger, O., Klinck, R., Chabot, B., Muchardt, C., Allemand, E. & Batsche, E. Alternative splicing regulates the expression of G9A and SUV39H2 methyltransferases, and dramatically changes SUV39H2 functions. *Nucleic Acids Res* **43**, 1869-1882, doi:10.1093/nar/gkv013 (2015).

59 Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R. & Genome Project Data Processing, S. The Sequence

Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079, doi:10.1093/bioinformatics/btp352 (2009).

60 Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K. & Madden, T. L. BLAST+: architecture and applications. *BMC Bioinformatics* **10**, 421, doi:10.1186/1471-2105-10-421 (2009).

61 Slater, G. S. & Birney, E. Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics* **6**, 31, doi:10.1186/1471-2105-6-31 (2005).

62 Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930, doi:10.1093/bioinformatics/btt656 (2014).

63 Lou, W. P., Baser, A., Klussmann, S. & Martin-Villalba, A. In vivo interrogation of central nervous system translatome by polyribosome fractionation. *J Vis Exp*, doi:10.3791/51255 (2014).

64 Kielbasa, S. M., Wan, R., Sato, K., Horton, P. & Frith, M. C. Adaptive seeds tame genomic sequence comparison. *Genome Res* **21**, 487-493, doi:10.1101/gr.113985.110 (2011).

65 Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841-842, doi:10.1093/bioinformatics/btq033 (2010).

66 Mendoza, L., Deutsch, E. W., Sun, Z., Campbell, D. S., Shteynberg, D. D. & Moritz, R. L. Flexible and Fast Mapping of Peptides to a Proteome with ProteoMapper. *J Proteome Res* **17**, 4337-4344, doi:10.1021/acs.jproteome.8b00544 (2018).

67 Liu, M. T., Wuebbens, M. M., Rajagopalan, K. V. & Schindelin, H. Crystal structure of the gephyrin-related molybdenum cofactor biosynthesis protein MogA from Escherichia coli. *J Biol Chem* **275**, 1814-1822, doi:10.1074/jbc.275.3.1814 (2000).

68 Xiang, S., Nichols, J., Rajagopalan, K. V. & Schindelin, H. The crystal structure of Escherichia coli MoeA and its relationship to the multifunctional protein gephyrin. *Structure* **9**, 299-310, doi:10.1016/s0969-2126(01)00588-3 (2001).

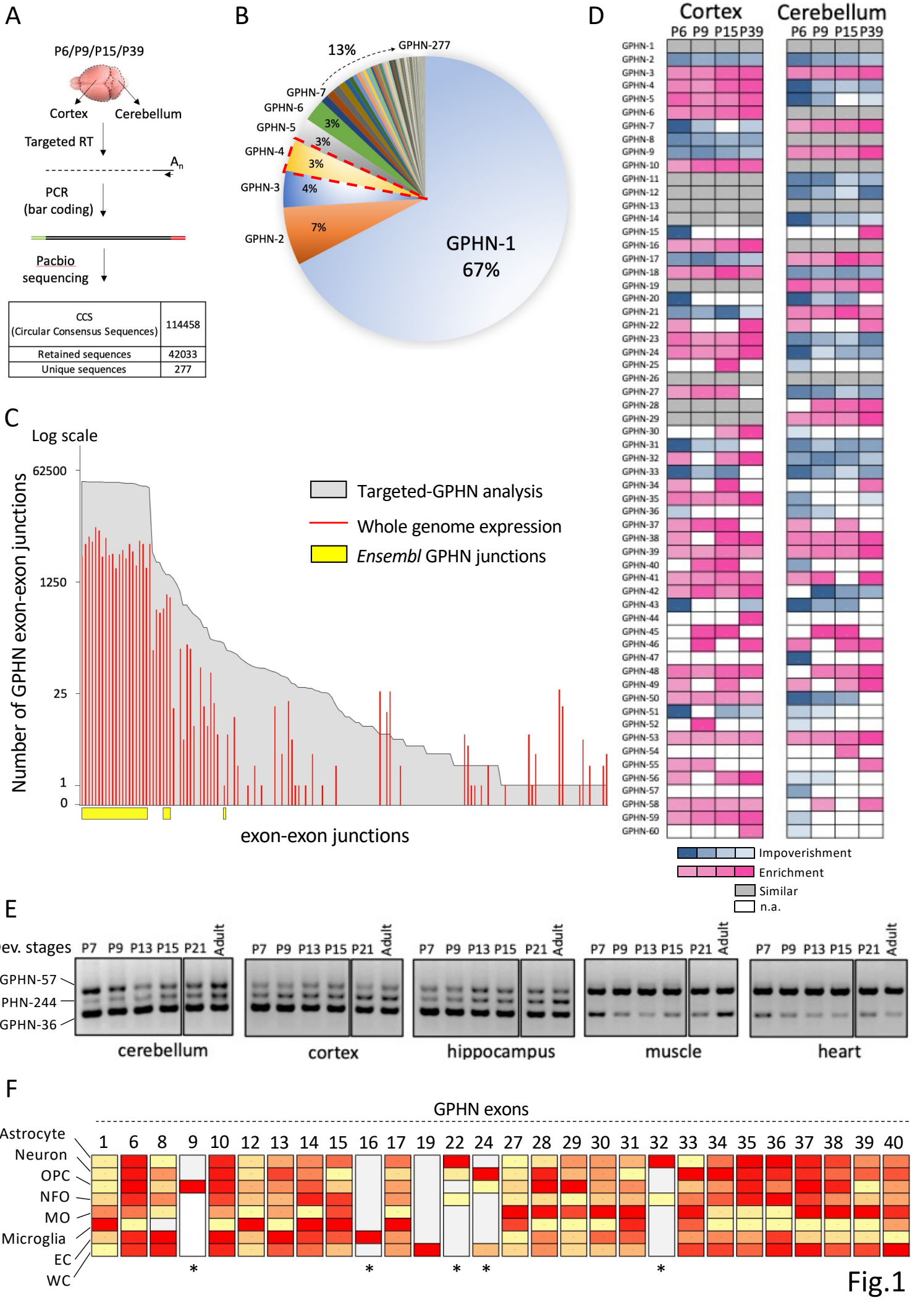
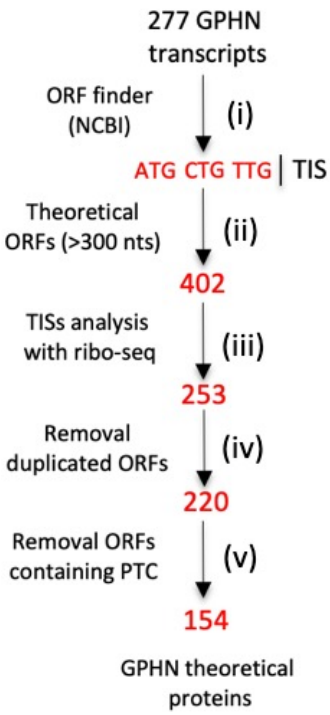
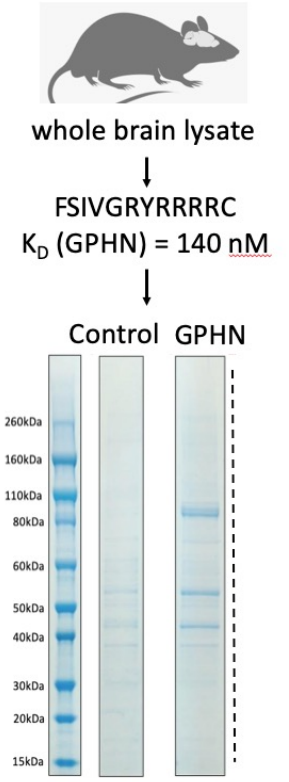

A
P6/P9/P15/P39
Cortex
Cerebellum
Targeted RT
An
PCR
(bar coding)
Pacbio
sequencing
CCS (Circular Consensus Sequences) 114458
Retained sequences 42033
Unique sequences 277
B
13%
GPHN-277
GPHN-7
GPHN-6
GPHN-5
3%
GPHN-4
3%
3%
GPHN-3
4%
GPHN-2
7%
GPHN-1
67%
C
Log scale
Number of GPHN exon-exon junctions
62500
1250
25
1
0
Targeted-GPHN analysis
Whole genome expression
Ensembl GPHN junctions
exon-exon junctions
D
Cortex
Cerebellum
P6 P9 P15 P39
GPHN-1
GPHN-2
GPHN-3
GPHN-4
GPHN-5
GPHN-6
GPHN-7
GPHN-8
GPHN-9
GPHN-10
GPHN-11
GPHN-12
GPHN-13
GPHN-14
GPHN-15
GPHN-16
GPHN-17
GPHN-18
GPHN-19
GPHN-20
GPHN-21
GPHN-22
GPHN-23
GPHN-24
GPHN-25
GPHN-26
GPHN-27
GPHN-28
GPHN-29
GPHN-30
GPHN-31
GPHN-32
GPHN-33
GPHN-34
GPHN-35
GPHN-36
GPHN-37
GPHN-38
GPHN-39
GPHN-40
GPHN-41
GPHN-42
GPHN-43
GPHN-44
GPHN-45
GPHN-46
GPHN-47
GPHN-48
GPHN-49
GPHN-50
GPHN-51
GPHN-52
GPHN-53
GPHN-54
GPHN-55
GPHN-56
GPHN-57
GPHN-58
GPHN-59
GPHN-60
Impoverishment
Enrichment
Similar
n.a.
E
Dev. stages
P7 P9 P13 P15 P21 Adult
GPHN-57
GPHN-244
GPHN-36
cerebellum
cortex
hippocampus
muscle
heart
F
GPHN exons
1 6 8 9 10 12 13 14 15 16 17 19 22 24 27 28 29 30 31 32 33 34 35 36 37 38 39 40
Astrocyte
Neuron
OPC
NFO
MO
Microglia
EC
WC
*
*
*
*
*

Fig.1

A


277 GPHN transcripts
ORF finder (NCBI)
(i)
ATG CTG TTG | TIS
Theoretical ORFs (>300 nts)
(ii)
402
TISs analysis with ribo-seq
(iii)
253
Removal duplicated ORFs
(iv)
220
Removal ORFs containing PTC
(v)
154
GPHN theoretical proteins

B


whole brain lysate
FSIVGRYRRRRC
K_D (GPHN) = 140 nM
Control
GPHN
260kDa
160kDa
110kDa
80kDa
60kDa
50kDa
40kDa
30kDa
20kDa
15kDa

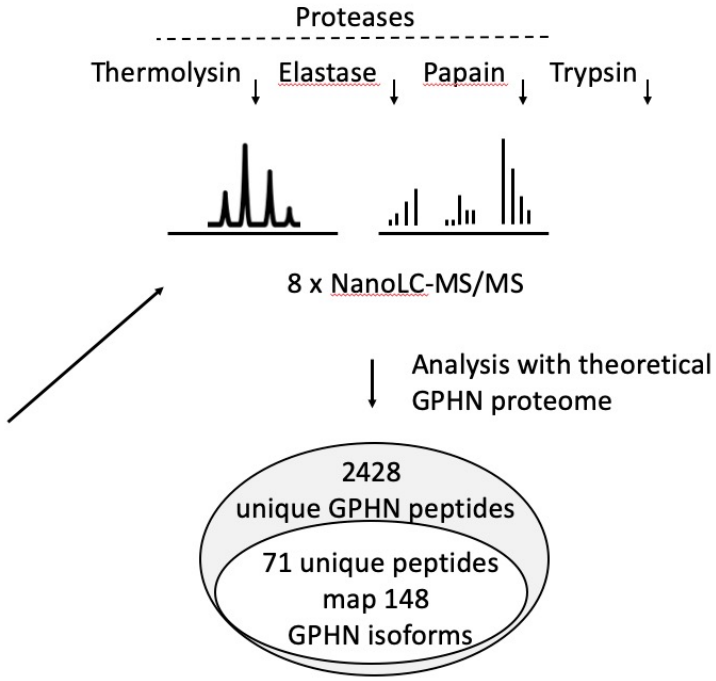
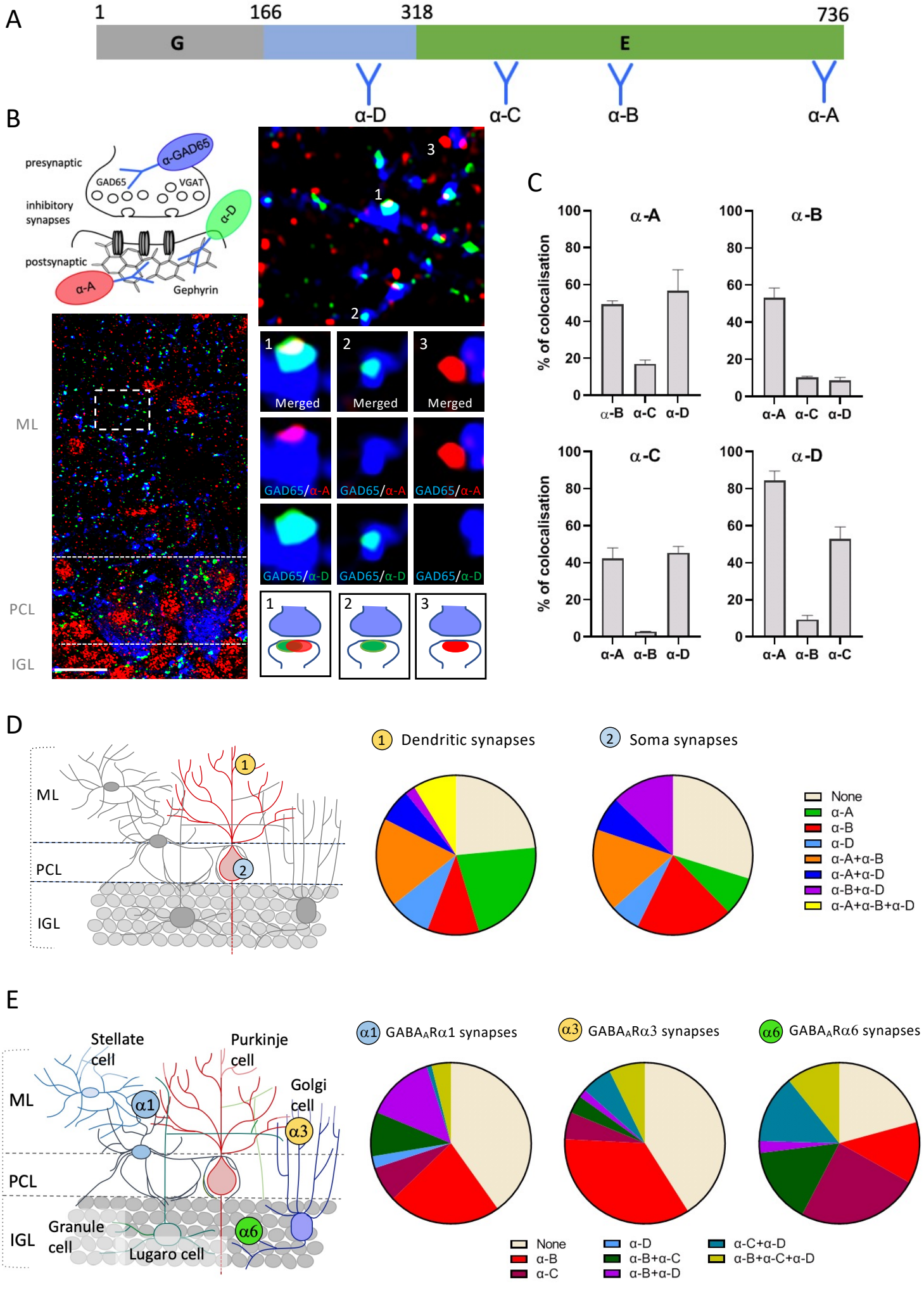
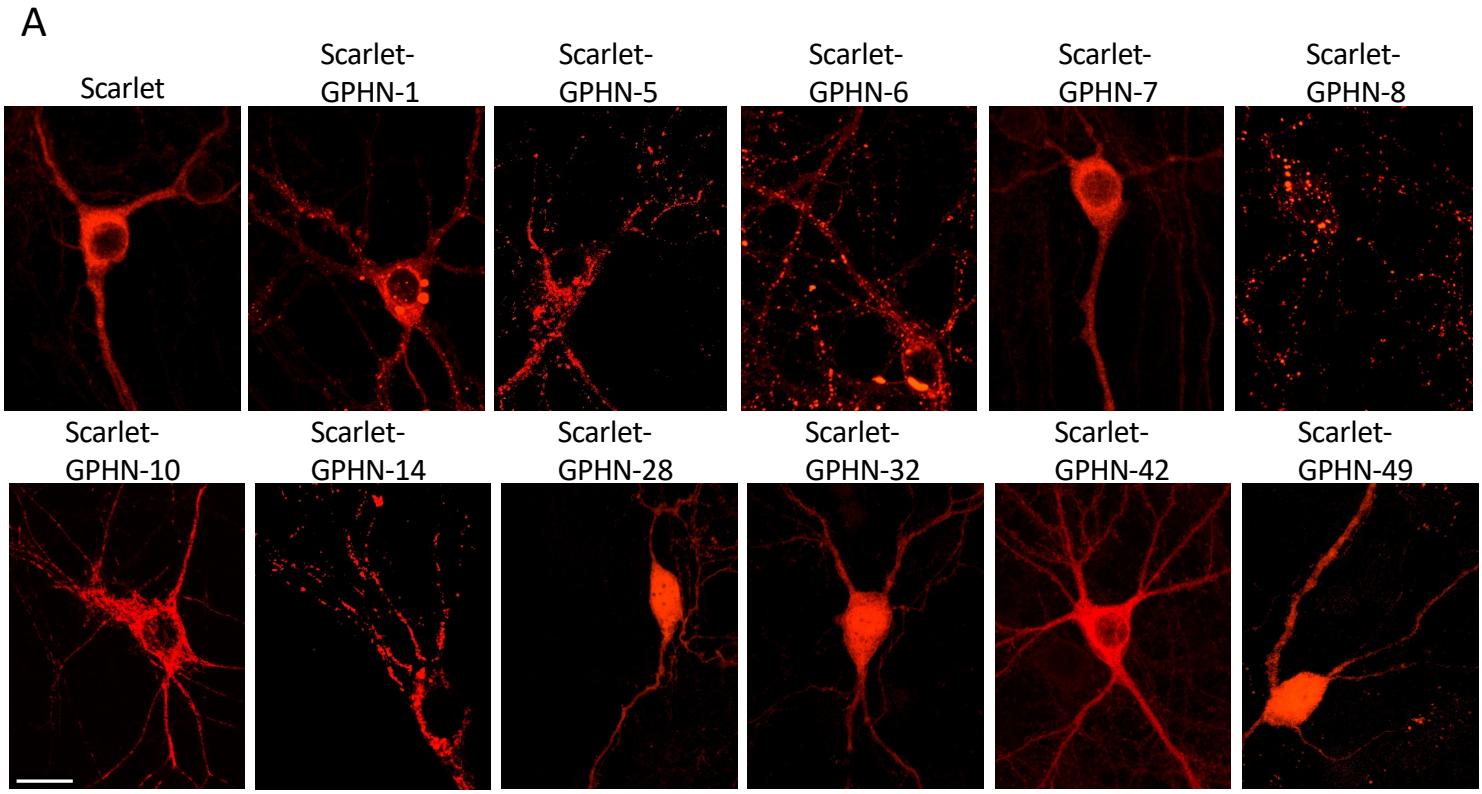
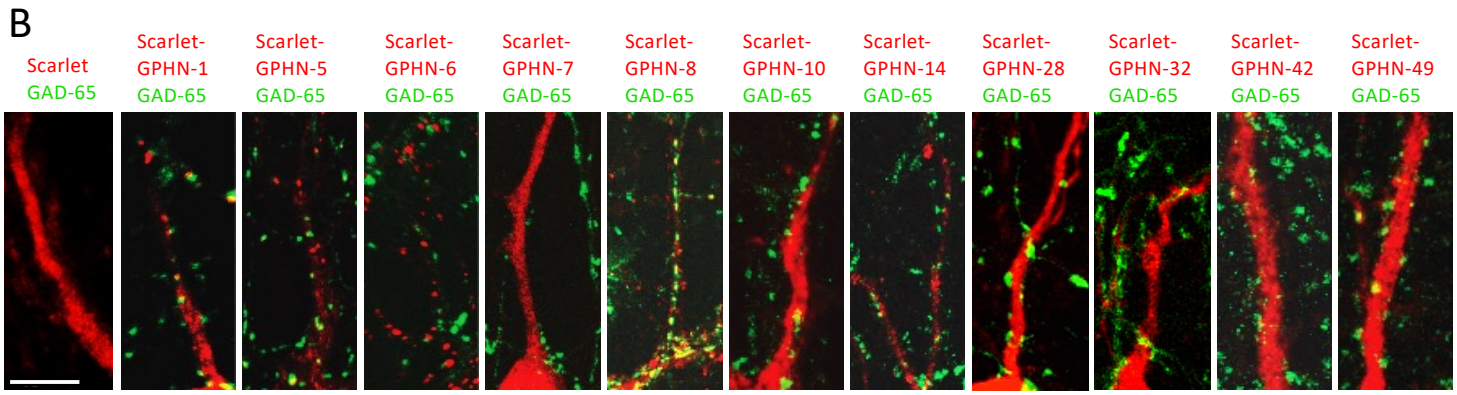

Proteases
Thermolysin
Elastase
Papain
Trypsin
8 x NanoLC-MS/MS
Analysis with theoretical GPHN proteome
2428 unique GPHN peptides
71 unique peptides map 148 GPHN isoforms

Fig.2


A
1
166
318
736
G
E
α-D
α-C
α-B
α-A
B
presynaptic
α-GAD65
GAD65
VGAT
inhibitory synapses
α-D
postsynaptic
α-A
Gephyrin
ML
PCL
IGL
Merged
GAD65/α-A
GAD65/α-D
C
% of colocalisation
α-A
α-B
α-C
α-D
D
ML
PCL
IGL
1 Dendritic synapses
2 Soma synapses
None
α-A
α-B
α-D
α-A+α-B
α-A+α-D
α-B+α-D
α-A+α-B+α-D
E
Stellate cell
Purkinje cell
Golgi cell
Granule cell
Lugaro cell
α1
α3
α6
GABAAα1 synapses
GABAAα3 synapses
GABAAα6 synapses
None
α-B
α-C
α-D
α-B+α-C
α-B+α-D
α-C+α-D
α-B+α-C+α-D


A
Scarlet
Scarlet-GPHN-1
Scarlet-GPHN-5
Scarlet-GPHN-6
Scarlet-GPHN-7
Scarlet-GPHN-8
Scarlet-GPHN-10
Scarlet-GPHN-14
Scarlet-GPHN-28
Scarlet-GPHN-32
Scarlet-GPHN-42
Scarlet-GPHN-49


B
Scarlet GAD-65
Scarlet-GPHN-1 GAD-65
Scarlet-GPHN-5 GAD-65
Scarlet-GPHN-6 GAD-65
Scarlet-GPHN-7 GAD-65
Scarlet-GPHN-8 GAD-65
Scarlet-GPHN-10 GAD-65
Scarlet-GPHN-14 GAD-65
Scarlet-GPHN-28 GAD-65
Scarlet-GPHN-32 GAD-65
Scarlet-GPHN-42 GAD-65
Scarlet-GPHN-49 GAD-65

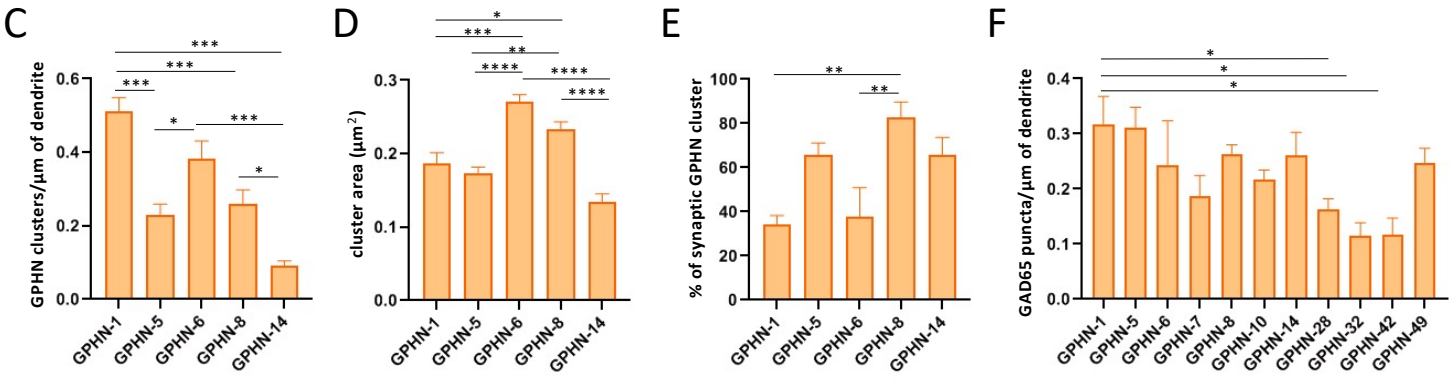
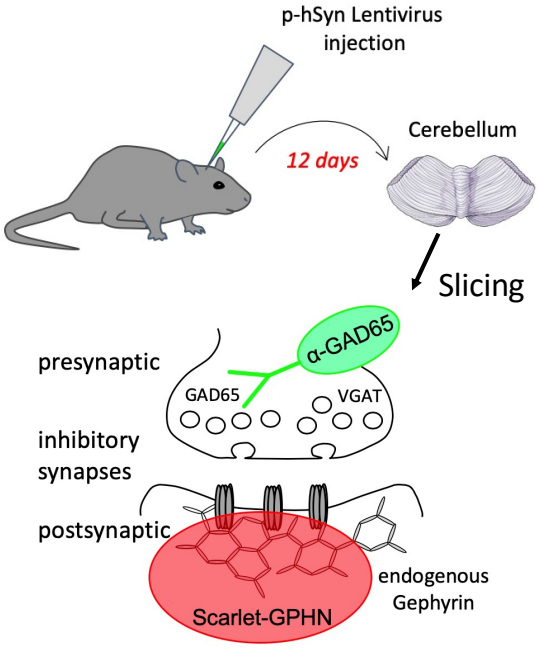
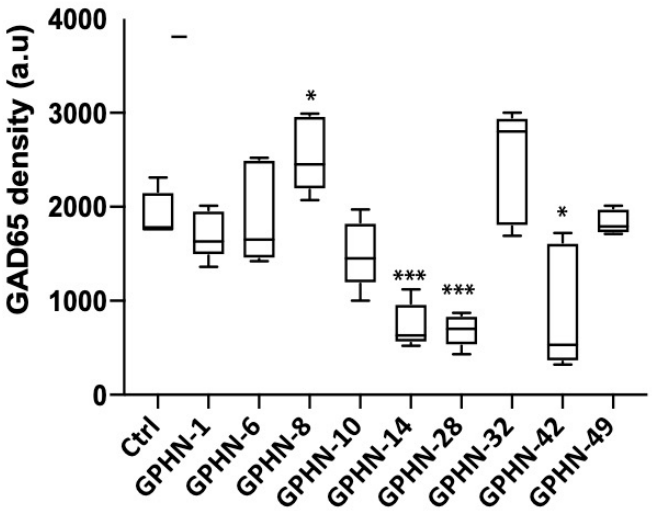

C
GPHN clusters/µm of dendrite
0.0
0.2
0.4
0.6
GPHN-1
GPHN-5
GPHN-6
GPHN-8
GPHN-14
D
cluster area (µm²)
0.0
0.1
0.2
0.3
E
% of synaptic GPHN cluster
0
20
40
60
80
100
F
GAD65 puncta/µm of dendrite
0.0
0.1
0.2
0.3
0.4
GPHN-7
GPHN-10
GPHN-28
GPHN-32
GPHN-42
GPHN-49

Fig.4

A


p-hSyn Lentivirus
injection
Cerebellum
12 days
Slicing
presynaptic
α-GAD65
GAD65
VGAT
inhibitory
synapses
postsynaptic
Scarlet-GPHN
endogenous
Gephyrin

C


GAD65 density (a.u)
4000
3000
2000
1000
0
*
***
***
*
Ctrl
GPHN-1
GPHN-6
GPHN-8
GPHN-10
GPHN-14
GPHN-28
GPHN-32
GPHN-42
GPHN-49

B

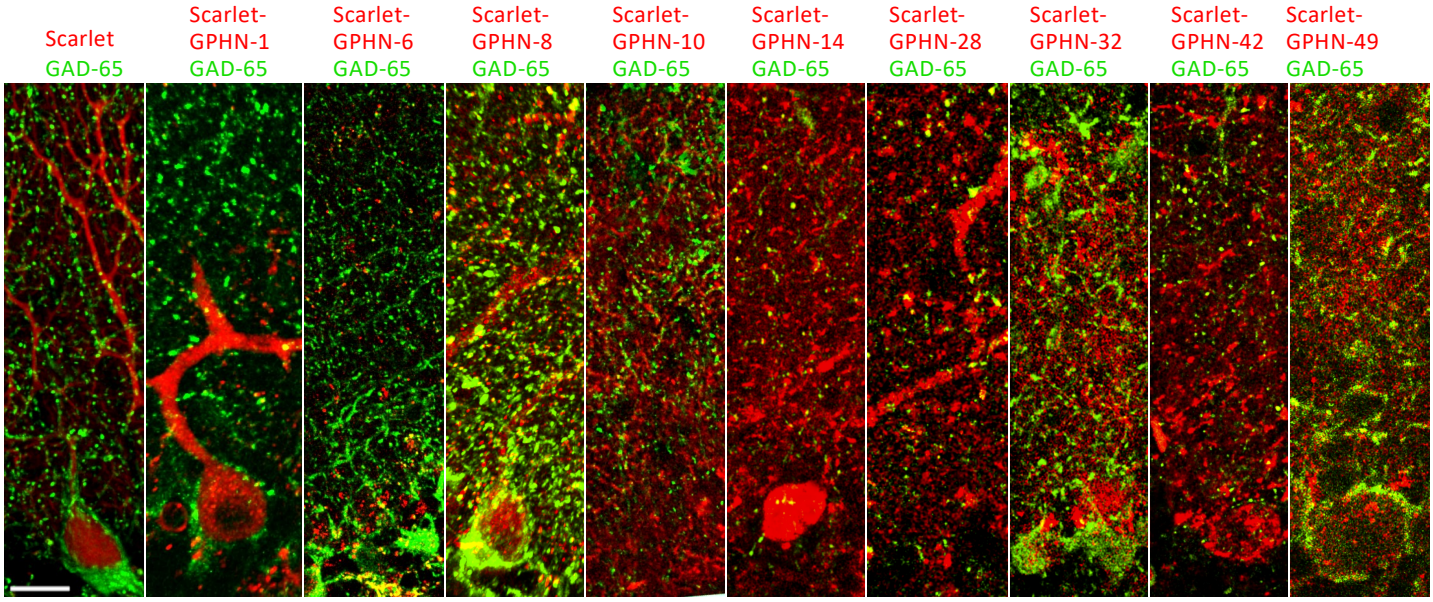

Scarlet
GAD-65
Scarlet-
GPHN-1
GAD-65
Scarlet-
GPHN-6
GAD-65
Scarlet-
GPHN-8
GAD-65
Scarlet-
GPHN-10
GAD-65
Scarlet-
GPHN-14
GAD-65
Scarlet-
GPHN-28
GAD-65
Scarlet-
GPHN-32
GAD-65
Scarlet-
GPHN-42
GAD-65
Scarlet-
GPHN-49
GAD-65

Fig.5

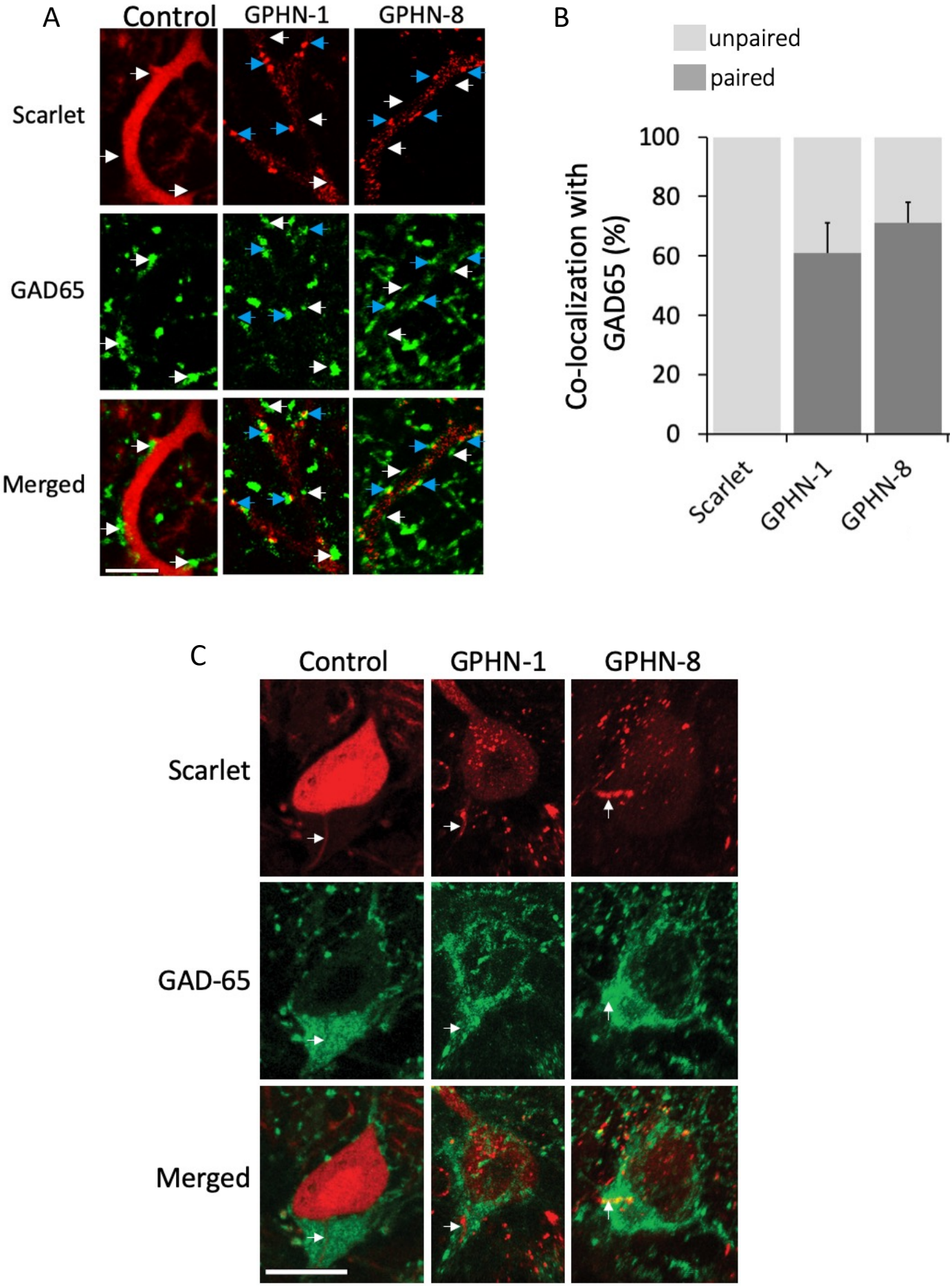
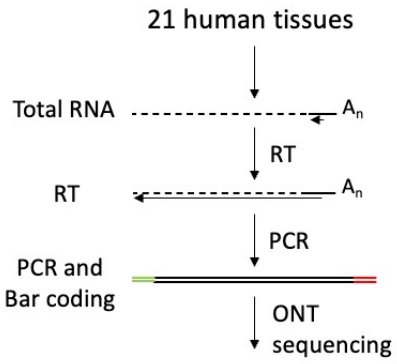

A
Control
GPHN-1
GPHN-8
Scarlet
GAD65
Merged
B
unpaired
paired
Co-localization with GAD65 (%)
100
80
60
40
20
0
Scarlet
GPHN-1
GPHN-8
C
Control
GPHN-1
GPHN-8
Scarlet
GAD-65
Merged

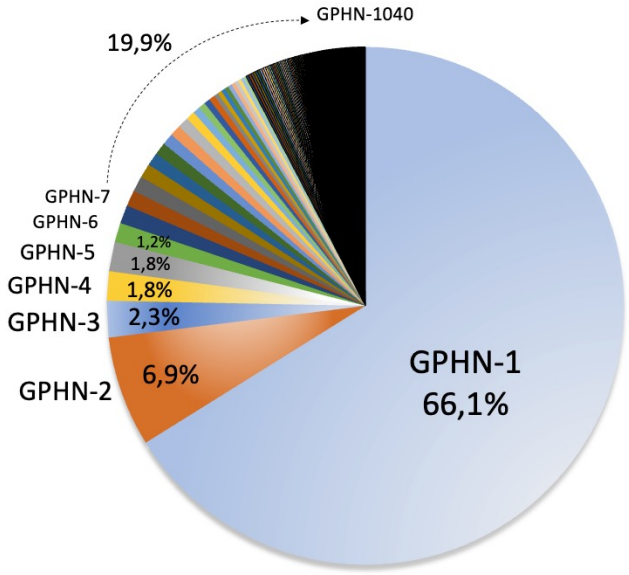
Fig.6

A



Raw sequences	$8{,}37.10^6$
Retained sequences	$5{,}23.10^6$
Unique sequences	1040

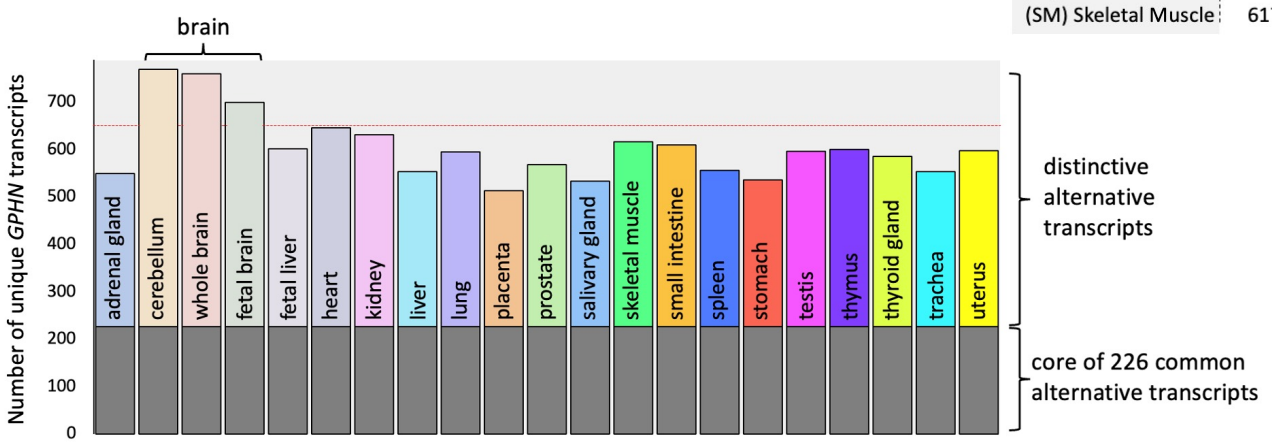
B



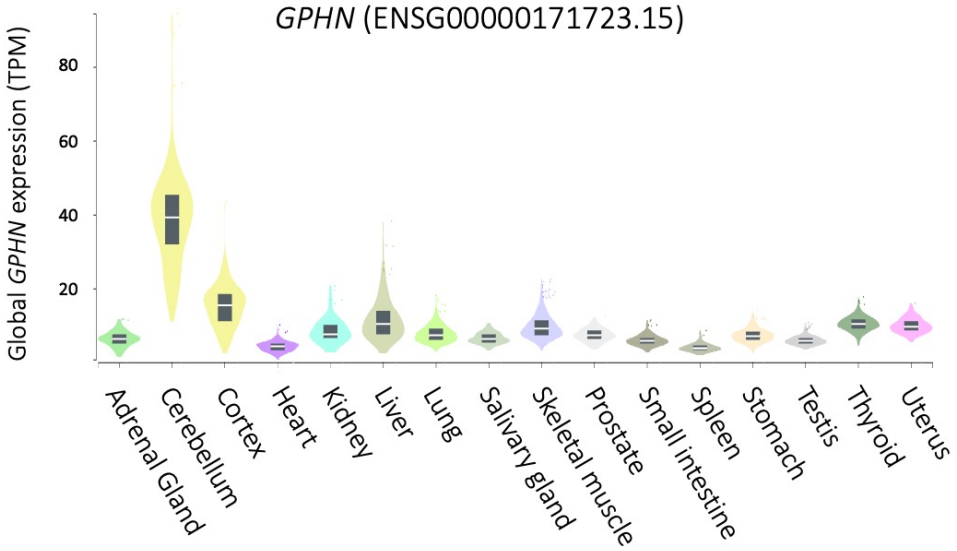
C

	Alternative transcripts
(P) Placenta	513
(FB) Fetal Brain	700
(FL) Fetal Liver	602
(BW) Brain (whole)	760
(C) Cerebellum	770
(SG) Salivary Gland	534
(Tr) Trachea	554
(TG) Thyroid Gland	586
(T) Thymus	601
(Lg) Lung	595
(H) Heart	646
(L) Liver	553
(Sp) Spleen	556
(St) Stomach	536
(AG) Adrenal Gland	550
(K) Kidney	632
(SI) Small Intestine	610
(Pr) Prostate	568
(U) Uterus	598
(Te) Testis	596
(SM) Skeletal Muscle	617

D



E



F

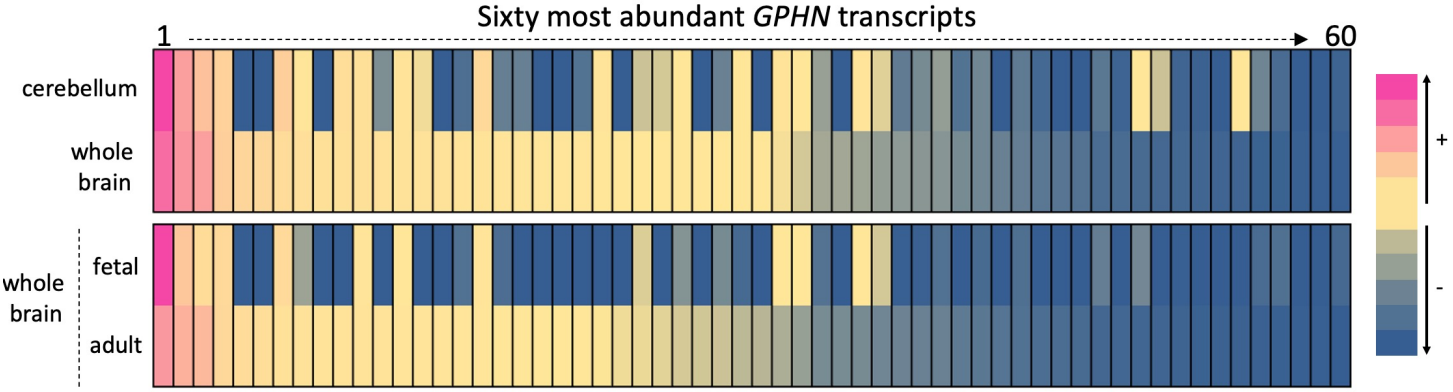


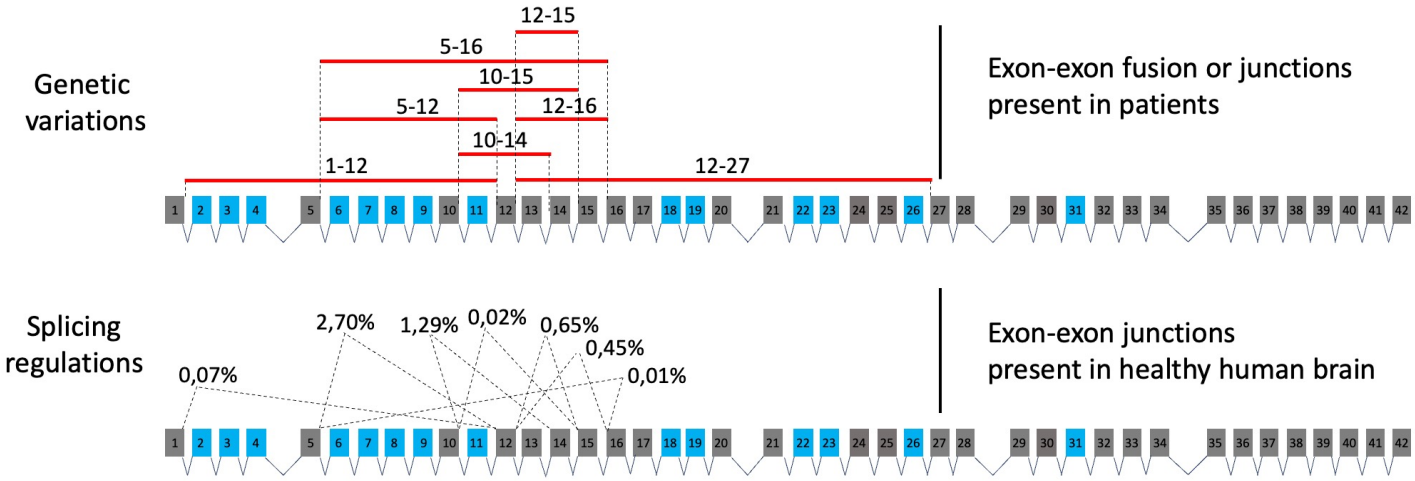
Fig.7

A

Diseases	Genetic variations	previous exon nomenclature	references	novel exon nomenclature (this study)
Seizure, ASD, SCZD	Deletion	fusion exon 3-5	Lionel et al., 2013	fusion exon 10-14
IGE	Deletion	fusion exon 1-4	Dejanovic et al. 2014	fusion exon 1-12
IGE	Deletion	fusion exon 4-10	Dejanovic et al. 2014	fusion exon 12-27
TLE	irregular splicing	EEJ 3-9	Forstera et al., 2010	EEJ 5-16
TLE	irregular splicing	EEJ 5-9	Forstera et al., 2010	EEJ 12-16
TLE	irregular splicing	EEJ 4-8	Forstera et al., 2010	EEJ 10-15
TLE	irregular splicing	EEJ 3-5	Forstera et al., 2010	EEJ 5-12
TLE	irregular splicing	EEJ 5-8	Forstera et al., 2010	EEJ 12-15
EE	missense mutation G375D	exon17 G>A	Dejanovic et al., 2015	exon 29 G>A

TLE (Temporal Lobe Epilepsy)
IGE (Idiopathic Generalized Epilepsy)
EE (Epileptic Encephalopathy)
ASD (Autism syndrom disorder)
SCZD (schizophrenia)

B



C

	whole brain		fetal brain		adult cerebellum	
novel exon nomenclature (this study)	% of total *GPHN* expression	unique transcripts	% of total *GPHN* expression	unique transcripts	% of total *GPHN* expression	unique transcripts
fusion exon 10-14	3,276	59	0,313	17	0,294	14
fusion exon 1-12	0,057	5	0,089	9	0,055	4
fusion exon 12-27	ND	ND	ND	ND	ND	ND
EEJ 5-16	0,011	2	0,013	1	0,016	2
EEJ 12-16	0,281	12	0,804	21	0,271	11
EEJ 10-15	0,024	2	0,021	1	0,012	1
EEJ 5-12	2,927	82	2,042	73	3,127	73
EEJ 12-15	0,543	29	0,752	33	0,651	28
exon 29 G>A	ND	ND	ND	ND	ND	ND

Fig.8

Legend of figures:

Fig. 1: Splicing regulation of *Gphn* expression in mouse brain.

(A) Procedure developed to analyze *Gphn* expression by long read sequencing. *Gphn* transcriptomes expressed at each developmental stage were prepared separately and combined in a multiplexed library sequenced with the PacBio technology. Number of retained Circular Consensus Sequences (CCS) and unique sequences are indicated at the bottom of the draw. **(B)** Distribution and percentages of alternative transcripts expressed by *Gphn* in Mouse brain. **(C)** Graph displaying the *Gphn* exon-exon junctions (EEJ) found in this study (grey), in $2{,}872.10^9$ short read sequences obtained by global gene expression analysis (red), in the *Ensembl* database (yellow). **(D)** Heat map graphical representation of the 60 most expressed *Gphn* transcripts at P6, P9, P15 and P39 in cortex and cerebellum. **(E)** Example of new transcripts including the selection of an alternative 3' splice site in exons 14 which are detected by RT-PCR. Amplicons corresponding to the *Gphn*-36, *Gphn*-57, *Gphn*-244 transcripts are analysed in 6 mouse developmental stages, 3 brain areas, skeletal muscles and heart. **(F)** Quantification of *Gphn* exons expressed in various neuronal cell types of mouse brain using data from (19). Heat map graphical representation displays independent levels of each exon in different cell types (yellow for lowly expressed, and a linear gradient of red for highly expressed). Stars indicated the unannotated and validated exons (see supplemental Fig. S1D).

Fig. 2: *Gphn* expresses a myriad of GPHN protein isoforms.

(A) Filtering pipeline used in this study to generate GPHN theoretical proteome, (i) Translation Initiation Start (TIS) identification, (ii) retained ORFs including at least 300 nucleotides, (iii) ORFs initiated with a TIS confirmed by biological evidences, (iv) removal of duplicated ORFs, (v) withdrawal of ORFs including PTC. **(B)** Workflow to isolate GPHN protein isoforms present at GABA-$_A$ and Glycine inhibitory synapses and MS processing to analyze them. GPHN was isolated from whole brain lysate via affinity purification using a neurotransmitter receptor peptide (FSIVGRYRRRC; K_D (GPHN) = 140 nM). Nano-LC MS/MS identified 2428 unique GPHN peptides, in which 71 corresponded to 148 novel GPHN protein isoforms.

Fig. 3: Heterogeneous distribution of endogenous GPHN reveals a diversity in the assembly of inhibitory synapses.

(A) Schematic of the GPHN-1 protein sequence in which are mapped the epitopes detected by four GPHN antibodies (α-A, α-B, α-C and α-D). **(B)** GAD-65/VGAT colocalization with GPHN epitopes α-A and α-D in cerebellar slices. Enlarged views of the white dashed-box show the heterogeneous staining of synaptic GPHN epitopes. Scale bar: 15 µm. **(C)** Quantification of GAD-65/VGAT colocalization with multiple combinations of the GPHN antibodies. **(D and E)** At the left, schematic of cerebellar cortex cellular organization in which are displayed specific inhibitory synapses (numbered-circle) present at molecular layer (ML), Purkinje cell layer (PCL), internal granular layer (IGL). At the right, quantification of one or multiple GPHN

epitopes at specific inhibitory synapses, for distinct neuronal sub-localisation (**D**; n=3 mice), and specific subunits of $GABA_A$-R (**E**; n=3 mice).

Fig.4: GPHN isoforms have distinct synaptic properties.
Eleven Scarlet-tagged GPHN isoforms were analysed in hippocampal primary neuronal culture. Exogenous GPHN isoforms were selected from the pool of most expressed *Gphn* transcripts in cortex and cerebellum (see Fig.1D). (**A**) Global localisation of GPHN isoforms in neuronal cells. (**B**) Confocal analysis of proximal dendrites in neuronal cells expressing Scarlet or Scarlet-tagged GPHN isoforms, presynaptic side was stained with anti-GAD-65. **(C)** Density of GPHN clusters along the dendrite (n=12 cells from 3 cultures). **(D)** Size of the GPHN clusters. **(E)** Percentage of GPHN clusters associated with a GAD-65 punctum. **(F)** Density of GAD-65 puncta along the dendrite. ****P<0,0001, ****P*<0,0010, **P<0,01, **P*<0,05.**)**. One-way analysis of variance (ANOVA) for (C), (D) and (E). **P*<0,05. Kruskal-Wallis test for (F). Scale bar: 25 µm (A) and 15 µm (B).

Fig.5: Increase of GPHN isoform levels modulates the number of inhibitory synapses in mouse cerebellum.
Analysis of inhibitory synapses connected to Purkinje cells in mouse cerebellar slices after transduction of lentivirus expressing Scarlet-tagged GPHN isoforms. **(A)** Schematic displaying the experimental procedure. **(B)** Cerebellar cortex slices in which inhibitory synapses are stained at pre- and postsynaptic sides using anti-GAD65 antibody (green) and Scarlet fluorescence (red) respectively (n=3 mice for each isoform). **(C)** Quantification of GAD-65 density detected in panel B (n=3 mice) ****P*<0,0010, **P*<0,05. Student t-test. Scale bar: 15 µm.

Fig.6: Different recruitment efficiency of GPHN isoforms to inhibitory synapses in mouse cerebellum.
Comparative analysis of two GPHN isoforms having different clustering properties in mouse cerebellar Purkinje cells. Experimental procedure was performed like in Fig.5A using three different viruses expressing respectively exogenous Scarlet and Scarlet-tagged GPHN-1 and -8. (**A**) Representative confocal images that display the exogenous proteins in PC dendrites of cerebellar slices. White arrows point to GAD65 synapses connecting PC dendrites lacking exogenous proteins, while blue arrows point to GAD65 synapses containing postsynaptic exogenous proteins. (**B**) Quantification of inhibitory synapse co-localization of GAD65 and exogenous proteins. (**C**) Different localization of Scarlet or Scarlet-GPHN isoforms at axon initial segment (AIS) of Purkinje cells. Note that only GPHN-1 and -8 are detected at the AIS, while other isoforms such as GPHN-28 are not.

Fig.7: Splicing regulation of *GPHN* expression in 21 human tissus.
(A) Procedure and number of sequences obtained by long-read sequencing with the Oxford Nanopore technology. Amplification of *GPHN* transcriptome from each 21 human tissues was

processed independently and mixed in a multiplexed library. **(B)** Distribution and percentages of alternative transcripts expressed by *GPHN* in the human body. **(C)** Table displaying the diversity of splice variants identified in each tissue. **(D)** Graphical representation showing the proportion of splice variants that were detected in all tissues (core of common 226 alternative transcripts) and those distinctly expressed in one or more tissue (s). **(E)** Graphical representation of global expression level of *GPHN* in 16 distinct human tissues using the data provided by the Genotype-Tissue Expression (GTEx) project (https://gtexportal.org/home/). **(F)** Heat map graphical representation of the 60 most expressed *GPHN* transcripts in adult and fetal human brain, as well as the whole brain versus the cerebellum. The yellow colour indicates an identical expression level between samples, while the blue scale shows a decrease and the pink an increase.

Fig.8: Expression of pathological *GPHN* transcripts is repressed in the healthy human brain.

(A) Table summarizing genetic variations and irregular splicing of *GPHN* identified in patients with neurological disorders. Diseases are referenced below the table, while previous and new exon annotations are indicated in two separated columns. **(B)** Schematic representation of genetic variations found in patients (upper panel) and the corresponding spliced exon-exon junctions identified by this study in the healthy human brain samples (lower panel). Percentages of each splicing event indicated the average detection in adult whole brain, fetal whole brain, adult cerebellum. **(C)** Table summarizing the detection of each exon-exon junctions in brain samples, and the number of unique transcripts containing them.

Materials and Methods

Mice:

G42 (GAD67-GFP) and GlyT2 (GlyT2-GFP) maintained in C57BL/6 background (Charles River Laboratories) were used for immunohistochemistry experiments. Wild-type C57BL/6 were used for stereotaxic injection and immunohistochemistry experiments. Swiss mice (Janvier Laboratories) were used for primary neuron cultures. Animals were kept under standard conditions with controlled temperature and lighting, and received food and water ad libitum. We followed the European and national regulations for the care and use of animals in order to protect vertebrate animals for experimental and other scientific purposes (Directive 86/609).

Tissues preparation:

C57BL/6, G42 or GlyT2 mice aged 40 days were used. Animals were lightly anesthetized with isoflurane and then deeply anesthetized with sodium pentobarbital by intraperitoneal injection. Artificial CerebroSpinal Fluid (ACSF; NaCl 126 mM, KCl 3 mM, NaH_2PO_4 1,25 mM, $NaHCO_3$ 20 mM, $MgSO_4$ 2 mM, Dextrose 20mM, $CaCl_2$ 2mM) was perfused transcardially for 2 minutes. The brains once extracted were placed 2 hours in 4% paraformaldehyde (PFA) in PBS. The brains were then sectioned using a microtome (Leica) to obtain 50 µm slices kept in PBS at 4°C for 2 weeks maximum before use.

PacBio sequencing:

RNA was extracted from mouse cortex and cerebellum tissues as described (*29*), and a primer specific Reverse Transcription (RT) was performed with the primer GPHN-49: GTACTGTGCCTGAGGCTGC. Gphn expression was next amplified using two rounds of PCRs, first 20 cycles with the primers GCAGTCGAACATGTAGCTGACTCAGGTCACCCACGACCATCAAATCCGTGTC and TGGATCACTTGTGCAAGCATCACATCGTAGCAGGATACAGTCAATGATATGTGGACAT GC, and after purification of PCR fragment a second PCR of 23 cycles was applied to each sample with Bar Coded (BC) primers: BC-GCAGTCGAACATGTAGCTGACTCAGGTCAC and BC-TGGATCACTTGTGCAAGCATCACATCGTAG. PCRs were purified with NucleoMag NGS Clean-up (Macherey-Nagel), quantified using Qubit, and the stoichiometric amount of each sample was finally mixed to build the multiplexed library. Library was processed as suggested by PacBio and sequenced, 114458 reads were obtained post filtering with a mean size of 16,508 bases. CCS showed a mean size of 2,3kb.

Oxford Nanopore sequencing:

Total RNAs purified from 21 human tissues were subjected to reverse transcription using the specific primer GPHN-48: CAGGATACAGTCAATGATATGTGGACATGC. Amplification of the *GPHN* transcriptome was processed independently in each sample by two PCR steps[58]. First, a pre-amplification of 20-cycles with specific primers corresponding to the first and last exons,

and then a second amplification of 18-cycles to add the barcodes. The multiplexed library was constructed using the Oxford Nanopore SQK-LSK109 kit and sequenced on MinION device.

PacBio, ONT and short reads analysis:

PacBio analysis: PacBio data were processed using informatic and manual approaches because the Isoseq 3.1.1software provided by PacBio did not provide homogeneous clusters of identical Circular Consensus Sequences (CCS). CCS were obtained from the Cold Spring Harbor sequencing platform and aligned to the Mouse genome GRCm38.p6 using STARlong 2.5.3a[59]. Genomic intervals showing alignments with the CCS (corresponding to potential exons) were extracted using bedtools 2.27.1merge[60] from merged bam files generated with samtools 1.6[61]. Each potential exon was extracted using bedtools 2.27.1 getfasta and checked by eyes to keep only those framed with canonical GT-AG dinucleotides at their splice sites. Retained exons were then aligned to CCS with blastn-short algorithm of blastn 2.2.31[62] (using an evalue threshold of 0.001 and a penalty for nucleotide mismatch of -2) to annotate each CSS with a binary code 1 and 0 for the presence or absence of distinct exon. Theoretical transcripts were constructed by concatenating the exon sequences and according to their annotation, CCS were next aligned against each theoretical transcript sequence (using a gapped global alignment, similar to the Needleman-Wunsch algorithm with Exonerate 2.4.0)[28] to provide an alignment score used as quality control. Similar sequences were clustered together and poorly aligned CCS were removed. Homogeneity of each cluster was checked visually using Snapgene software. Number of CCS present in each cluster was quantified in order to establish an overall assessment of alternative *Gphn* transcript expression. R 3.3.2 and ggplot were used to produce the overall isoform composition plot displayed Fig.1A.

Analysis of *Gphn* expression in brain cell types: *Gphn* exons expressed in distinct neuronal cell types were analyzed using Illumina data from [26] corresponding to total gene expression. An artificial sequence containing all *Gphn* exon junctions was used to align short reads using Hisat 2.1.0 (with the –no-spliced-alignment option). Using FeatureCounts 1.6.4[63], the aligned reads were quantified to assess the expression of each exon. To compare exon expression between cell subtypes, we normalized their quantification by the overall total *Gphn* expression in each cell type.

Analysis to validate exon-exon junction of *Gphn* expression in mouse brain: *Gphn* exon junctions detected using short reads reported in 82 publicly available data sets (listed in Table S1). An artificial sequence containing all exon junctions found in Fig.1A was used to align short reads using Hisat 2.1.0 (with the –no-spliced-alignment option). Using FeatureCounts 1.6.4[63], the aligned reads were quantified to assess junction detection. Finally, we compared them with read counts supporting each junction obtained by PacBio sequencing.

ONT analysis:

Raw data were processed for base calling using GUPPY and long read sequences were analysed with a homemade pipeline. Demultiplexing of samples was processed similarly to the PacBio

analysis and 149K up to 397K raw sequences were obtained per human tissue sample. Using LAST version 1205[64], the sequences were aligned on Human Genome GRCh38 and the corresponding BAM files were reviewed manually to annotate all exons. Exons were used to produce a universal reference and all reads were aligned on this particular reference with LAST. Using standard genomic tools, samtools[59], bedtools[65], a specific transcript barcode was generated to each sequence depending on its exon architecture, then pooled in clusters defining each alternative *GPHN* transcripts. Given the library preparation, we only retained the sequence containing at least the first and last exons in their barcode. Finally, we only retained transcripts seen 10 times or more in tissues.

The mouse and human GPHN exons architecture (Fig.S5C) were compared by aligning mouse features onto the human features using LAST.

DNA constructs:

The cDNAs corresponding to selected *Gphn* alternative transcripts were built using QuikChange II Site-Directed Mutagenesis Kit and cloned in a modified version of pLVX-puro vector for which CMV promoter was replaced by human Syn promoter. Each cDNA is cloned in 5' fusion with Scarlet and Flag-V5 tags. Lentiviral particles were produced using HEK-293 cells co-transfected with pLVX-FV5-Scarlet-GPHNs, p8.7 and pCMV-VSV-G, then concentrated by gradient centrifugation. All constructs used in this study are listed in Fig.S4B and available on request. Expression of exogenous proteins were probed by western blot using anti-V5 (Invitrogen).

***In silico* GPHN theoretical proteome:**

The following description is summarized in Fig.3A and reported in Table S. Each cDNA was processed with ORFfinder (https://www.ncbi.nlm.nih.gov/orffinder/) using a minimal ORF length of 300 nucleotides, "ATG" and alternative initiation codons parameters. The 31 translation initiation sites were next filtered with data published in [66] to retain only those supporting biological evidences. ORFs found more than once were removed as well as those containing a stop codon in exons upstream of exon 40 (such stop codon was considered as a PTC). ORFs obtained from this pipeline were named as the theoretical GPHN proteome. using Protein extracts from cerebellum, cortex, hippocampus, muscle, heart and spleen were prepared and separated in PAGE-SDS to probe endogenous GPHNs by western blots using mab3B11.

RT-PCR analysis:

RNA extracted from cortex and cerebellum tissues at P7, P9, P13, P15, P21 and P40 were retro-transcribed with GPHN-49 (see PacBio library) and amplified by PCR using specific set of primers:

- Fig.S1E: GGAGTCCTCACAGCCCACATAAAC and CTAGCCACCTTGGTGATATCTACAGC

- Fig.S1F: CTCTTGCTGCAAAGTTGACCAACTTTAG and GTTCCTTTGGCCAAAACACACTC
- Fig.S1G: CAAGAAAGGATCTCAGTAGTGCAAGTTG and CTGATACCCTCATTCAAGGCATTGAG
- Fig.S1H: GTCCTCACAGTGGTTGCCG and CTGCATCTTTCTCAGTGCAGGTACAAC
- Fig.S1I: GAGTCCTCACAGGAAAGATTCGGG and GGTGATGCCAAGTCAGTATACACC

Finally, PCRs were analyzed in agarose gel and stained with ethidium bromide.

Polysome preparation:
Cytoplasmic extracts containing translation machinery were prepared from adult mouse brain as described [67] and RNAs were purified after dilution of sucrose fractions with H2O (50% v/v). Ribosomal RNAs (28S, 18S and 5S) were separated in agarose gel stained with ethidium bromide, while *Gphn* isoforms were amplified and analyzed as Fig.S1E.

Preparation of whole-brain lysates:
After cervical dislocation, whole brains from adult male/female C57BL6 wildtype mice were removed from the skull and rapidly cooled to 4°C and homogenized in 1 ml lysate buffer (20 mM HEPES, 100mM KCH3COOH, 40mM KCl, 5mM EGTA, 5mM MgCl2, 5mM DTT, 1mM PMSF, 1% Triton X, cOmpleteTM EDTA-free protease inhibitor cocktail (Roche, Mannheim, Germany), pH 7.2) per 200 mg using a pistol homogenizer (8 strokes at 900 rpm). The homogenate was centrifuged at 10,000 × g and 4°C for 15 min. The supernatant was aliquoted and flash frozen in liquid nitrogen and stored at −80 °C.

Gephyrin isolation from brain lysate:
The resin for gephyrin pulldown was prepared in 400 µl micro-spin columns (Thermo Scientific, Karlsruhe, Germany). The Peptide FSIVGRYPRRRRC (K_D (Gephyrin) = 140 pM) (10.1038/nCHeMBIO.2246) was dissolved in coupling buffer (50 mM Tris, 5mM EDTA, pH 8.5) at a concentration of 1 mM and incubated for 2 h at RT with washed and equilibrated Ultra Link iodoacetyl resin (Thermo Scientific, Karlsruhe, Germany). The resulting resin was washed and residual iodoacetyl groups were quenched by incubation with 1 mM Cysteine for 2 h and stored at 4 °C after extensive washing with a coupling buffer.
Whole-brain lysate was incubated on the prepared FSIVGRYPRRRRC-resin for 30 min at 4°C and washed three times with lysate buffer. Bound proteins were eluted with NuPAGE LDS sample buffer (Life Technologies) at 90°C for 3 min. Tris(2-carboxyethyl)phosphin (TCEP) was added to a final concentration of 20 mM and incubated at 90°C to achieve complete reduction. Subsequently proteins were alkylated using iodoacetamide (120 mM) for 20 min at room

temperature in the dark, 30 µl of the resulting denatured, reduced and alkylated proteins were separated by mass by NuPAGE Novex 4-12 % Bis-Tris gels (Life Technologies) with MOPS buffer. Gels were washed three times for 5 min with water, stained for 45 min with Simply Blue™ Safe Stain (Life Technologies) and washed with water for 1 h.
Gel bands were excised and destained with 30 % acetonitrile in 0.1 M NH_4HCO_3 (pH 8), shrunk with 100 % acetonitrile, and dried in a vacuum concentrator (Concentrator 5301, Eppendorf, Germany). Digests were performed with 0.1 µg protease (trypsin, elastase, thermolysin or papain) per gel band overnight at 37 °C in 0.1 M NH_4HCO_3 (pH 8). After removing the supernatant, peptides were extracted from the gel slices with 5 % formic acid, and extracted peptides were pooled with the supernatant.

NanoLC-MS/MS Analysis
NanoLC-MS/MS analyses were performed on an Orbitrap Fusion (Thermo Scientific) equipped with an EASY-Spray Ion Source and coupled to an EASY-nLC 1000 (Thermo Scientific). Peptides were loaded on a trapping column (2 cm x 75 µm ID, PepMap C18, 3 µm particles, 100 Å pore size) and separated on an EASY-Spray column (25 cm x 75 µm ID, PepMap C18, 2 µm particles, 100 Å pore size) with a 30-minute linear gradient from 3 % to 40 % acetonitrile and 0.1 % formic acid.
Both MS and MS/MS scans were acquired in the Orbitrap analyzer with a resolution of 60,000 for MS scans and 15,000 for MS/MS scans. HCD fragmentation with 35 % normalized collision energy was applied. A Top Speed data-dependent MS/MS method with a fixed cycle time of 3 seconds was used. Dynamic exclusion was applied with a repeat count of 1 and an exclusion duration of 30 seconds; singly charged precursors were excluded from selection. Minimum signal threshold for precursor selection was set to 50,000. Predictive AGC was used with AGC a target value of 2e5 for MS scans and 5e4 for MS/MS scans. EASY-IC was used for internal calibration.

Data Analysis
MS data were analyzed with PEAKS Studio X+ (Bioinformatics Solutions Inc., Canada). Raw data refinement was performed with the following settings: Merge Options: no merge, Precursor Options: corrected, Charge Options: 1-6, Filter Options: no filter, Process: true, Default: true, Associate Chimera: yes. De novo sequencing and database searching were performed with a Parent Mass Error Tolerance of 10 ppm. Fragment Mass Error Tolerance was set to 0.02 Da, and Enzyme was set to none. The following variable modifications have been used: Oxidation (M), pyro-Glu from Q (N-term Q), phosphorylation (STY), acetylation (protein N-terminal) and carbamidomethylation (C). A maximum of 3 variable PTMs were allowed per peptide.
Data were searched against a fasta database concatenated from UniProt_mouse (UP000000589, reference proteome, 55408 proteins, 5-Nov-2019, all variants), predicted (XP_) and validated (NP_) gephyrin sequences from NCBI and a custom ORF database constructed as described

above. The list of identified peptides was filtered to 1% PSM-FDR. All gephyrin-derived peptides were matched to the entries of the ORF database using ProteoMapper [68].

Immunohistochemistry:

Free-floating immunochemistry: Slices were incubated in a blocking solution (0,1% Triton, 5% Horse Serum in Tris Buffer Saline (TBS) for 2 hours. The solution was then replaced by the primary antibody solution (0,1% Triton, 5% Horse Serum in TBS) with antibodies added at the indicated dilution, for an overnight incubation. The slices were washed three times for 10 minutes with Phosphate Buffer Saline (PBS), before the incubation in a solution (0,02% Triton, 5% Horse Serum in TBS) containing the secondary antibodies for 1 hour. The slices were again washed three times for 10 minutes with PBS. The slices were then mounted between slide and coverslip using hardening mounting medium (Vectashield Hard Set).

Hippocampus primary cell culture:

Swiss mice at 18 days of gestation were deeply anesthetized with isoflurane and then euthanized by cervical dislocation. The pups were decapitated and the heads were placed in ice cold PBS supplemented with glucose and antibiotics (Pencillin-Streptomycin). The brains were extracted and the hippocampi dissected and placed in 37°C Neurobasal medium (Gibco). The hippocampi were then dissociated with three glass-pasteur pipettes of decreasing tip widths. The suspension was then centrifuged at 1000 rpm for 8 minutes, the supernatant discarded and the pellet re-suspended in 37°C Neurobasal supplemented with B27, L-glutamine and antibiotics. The neurons were plated onto poly-L-ornithine coated coverslips and cultured for 11 days in an incubator at 37°C and 5% CO_2.

Cell culture infection:

One day after the plating, the neuron cultures were infected using 1µL of a suspension of lentivirus containing the relevant plasmid. The neurons were cultured for 10 more days in the conditions described previously.

Cerebellar stereotaxic injection:

40-50 day old C57BL/6 male mice were used. Mice were deeply anesthetized with isoflurane before being placed on a stereotaxic frame fitted with a custom-made isoflurane mask. Ophthalmic gel (Ocry-gel) was placed onto the mice's eyes and subcutaneous injection of lidocaine (Xylocaine) in the area was made prior to the incision of the skin. Craniotomy was performed with a fine drill bilaterally. Glass capillaries were filled with lentivirus suspensions and lowered into the cerebellar cortex. Injections were performed using a NanoLiter Injector and UMP3000 micropump and controller (World Precision Instruments), 1 µL per hole was dispensed over a three minutes injection. The capillary was left in place for three minutes and then slowly pulled out over a two minutes period. Tissues were then glued using surgical skin

glue (Vetbond). Mice were given buprenorphine (Vetergesic) for analgesia. Animals were kept 10 days before brains were collected for immunohistochemistry assay.

Antibody table:

Antigen	Host	Dilution	Supplier	Catalog N	RRID
GAD-65	Mouse	1:500	Chemicon	MAB351R	RRID:AB_94905
Gphn (A)	Rabbit	1:1000	Abcam	ab32206	RRID:AB_2112628
Gphn (B)	Chicken	1:500	Abcam	ab136343	RRID: none
Gphn (C)	Mouse	1:500	Synaptic System	147111	RRID:AB_887719
Gphn (D)	Guinea Pig	1:250	Synaptic System	147318	RRID:AB_2661777
VGAT	Guinea Pig	1:500	Synaptic System	131004	RRID:AB_887873
α-dystroglycan	Mouse	1:500	Millipore	05-298	RRID:AB_309674
$GABA_AR\alpha 1$	Rabbit	1:1000	Alomone Labs	AGA-001	RRID:AB_2039862
$GABA_AR\alpha 3$	Rabbit	1:1000	Alomone Labs	AGA-003	RRID:AB_2039866

$GABA_AR\alpha 6$	Rabbit	1:1000	Alomone Labs	AGA-004	RRID:AB_2039868
Calbindin	Rabbit	1:1000	Swant	CB38	RRID:AB_2721225
GFP	Chicken	1:1000	Aves Lab	GFP-1020	RRID:AB_10000240

Secondary	**Host**	**Dilution**	**Supplier**	**Catalog N**	**RRID**
Mouse Alexa 488	Donkey	1:500	Molecular probes	A-21202	RRID:AB_141607
Mouse Alexa 405	Goat	1:500	Molecular probes	A-31553	RRID:AB_221604
Mouse Alexa 546	Goat	1:500	Molecular probes	A-11003	RRID:AB_2534071
Rabbit Alexa 546	Goat	1:500	Molecular probes	A-11010	RRID:AB_2534077
Chiken Alexa 488	Goat	1:500	Molecular probes	A-11039	RRID:AB_142924

Guinea pig Alexa 647	Donkey	1:500	Jackson ImmunoResearch Labs	706-605	RRID:AB_2340476

Image acquisition and quantification:

Tissue and culture samples were imaged on a Zeiss LSM780 confocal microscope to obtain a stack; pinhole, detection filter settings, dwell time and step size were kept identical (40X oil immersion objective, 1.3 NA, 1024x1024 resolution, 16 bits, 0,4 µm step size). Photomicrographs were obtained with the following band-pass and long-pass filter setting: alexafluor 405 (Band pass filter: 460/50), alexafluor 488/Cy2 (band pass filter: 505–530), Cy3 (band pass filter: 560–615) and Cy5 (long-pass filter 650). All parameters were held constant for all sections from the same experiment. At least three slices per injection site were used in all immunofluorescence analyses (n = 3 mice / staining).

Hippocampus neuron cultures were repeated at least three times, and infected neurons were analyzed. For each neuron, proximal dendrites (50-100µm) were used as ROI. Analyses were made with Fiji (ImageJ) software. Images were thresholded using MaxEntropy Auto-Threshold command. GPHN clusters or GAD-65 puncta were annotated and counted. Measure Particle function was used to measure GPHN cluster area.

For immunofluorescence analyses at least three mice were used and at least three slices per mice were imaged. For every image, ROI of identical size was selected within the region of interest. For cerebellar slices, images were taken so as to contain all three layers. All analyses were made using the molecular layer with the exceptions of Fig. 4E (α-dg cell body) and supplementary Fig. 4F and 4L where ROI contained only Purkinje cell bodies; and Fig. 4F ($GABA_AR\alpha6$) and supplementary Fig. 4I, where ROI were chosen within the granular cell layer.

For HEK cell cultures, stainings were repeated three times, one image containing several cells were taken for each experiment.

Colocalization between clusters was made with Icy software using the SODA plugin (Statistical Object Distance Analysis) when the number of studied channels was below three. ROI of the relevant areas were delineated onto the images, and the SODA 2 colors or SODA 3 colors protocol was applied. Results were checked once by manual counting using the method explained below to ensure threshold and distance parameters were adequate.

For images with four channels and for the verification of SODA analyses, Fiji (ImageJ) software was used. Images were thresholded using MaxEntropy Auto-Threshold command. Within a ROI, all clusters from each channel were counted and annotated, and clusters were deemed colocalized

when distance separating them was below 210 nm which is the optical resolution of the confocal with the 40x objective. The analysis was done independently by two experimenters.

Statistical analysis

All statistical analyses were performed using GraphPad Prism 8 (GraphPad Softwares). Normality was assessed with the Shapiro-Wilk normality test. Homoscedasticity was assessed with the Barlett's test. Parametric data with the same standard deviation were analyzed by t-test, one-way ANOVA, or two-way ANOVA followed by comparison of multiple samples with Tukey post hoc analysis. Parametric data with significantly different standard deviation were analyzed by Welch's ANOVA followed by comparison of multiple samples with Dunett's T3 post hoc analysis. Non-parametric data were analyzed by the Kruskal-Wallis one-way analysis of variance on ranks followed by comparison of multiple samples with Dunn post hoc analysis. P values <0.05 were considered statistically significant. *$p<0.05$, **$p<0.01$, ***$p<0.001$, ****$p<0.0001$. ns not significant. Data are presented as mean ± SEM.

Supplementary Materials for

This file includes:

Figures S1-S5

Fig.S1: *Gphn* transcripts expressed in cortex and cerebellum of mouse

(A) Schematic boxes describing a mixed pipeline (informatic and manual) to analyze Circular Consensus Sequences obtained by PacBio sequencing. **(B)** The 277 alternative *Gphn* transcripts identified from 42033 SMRT sequences, each line represents a distinct transcript and each column symbolizes one exon named at the bottom. Transcripts are ordered and named *Gphn*-1 to *Gphn*-277 for the most to the least detected respectively. Exons are displayed as a rectangle, green for included, white for skipped and red for exon included with an alternative 3' splice site. The protein domains (G and E) and the central unstructured region (blue) of GPHN are represented at the top. The abundance of each transcript is reported on a bar graph established with a log scale. **(C)** Table displaying how many times an exon participates in the 277 alternative transcripts. **(D)** Illustrations comparing *GPHN* exon architecture reported from Ensembl and literature data (top) versus the exon annotation determined in this study (Bottom). **(E-I)** Analysis by RT-PCR of transcripts carrying exons selected with an alternative 3'ss (exon A, B, C, D, E that correspond to alternative 3'ss of exon, 14, 16, 30, 33 and 34 respectively). At the top of each panel, PCR fragment(s) amplified in 5 different tissues, at 6 distinct developmental stages, were separated in agarose gel and stained with ethidium bromide. At the bottom right, we show a schematic view of expected PCR fragment(s) and their sizes based on selected primers. At the bottom left, we have listed all *Gphn* alternative transcripts carrying each alternative exon.

Fig.S2: *Gphn* transcripts engaged in the translation machinery, and GPHN protein enrichment.

(A) Illustration displaying the procedure used to separate RNA associated with the polyribosome. Brain extracts isolated from mouse perfused with cycloheximide were prepared and separated onto a sucrose gradient to distinguish fractions containing RNA engaged in translation (polyribosome). The treatment of the extract with EDTA dissociates ribosomes and was used as a control. Ribosomal RNA (28S, 18S, 5S) were analyzed in agarose gel for each

fraction, while *Gphn* transcripts showed in Fig.S1E detected using RT-PCR amplification (*Gphn*-57, *Gphn*-244 and *Gphn*-36). **(B)** *In silico* PAGE-SDS graph displaying the molecular weight separation associated with the 154 GPHN theoretical isoforms. On the left side, a molecular weight ladder is shown for reference. On the right side, a density graph shows isoform distribution that doesn't consider expression levels of each isoform. **(C)** Western blot of GPHN isoforms in different mouse tissue using the mab3B11 antibody. **(D)** Cartoon displaying exon-exon junctions (EEJ) or exons included in new ORFs for which mass spectrometry analysis identified corresponding peptides. **(E)** Table displaying the EEJs and the number of corresponding peptides (number of matched peptides) identified by mass spectrometry. Number of GPHN isoforms containing the corresponding peptides are also indicated.

Fig.S3: Heterogeneous distribution of endogenous GPHN in inhibitory synapses.
(A) Schematic representation of the mouse and human protein domain of GPHN with the mapping of epitopes recognized by antibodies used in this study. On the top of each drawing, GPHN ORFs translated from Ensembl ESTs are displayed with the positioning of exon inclusion and exon skipping. **(B)** GPHN-1 is ubiquitously labeled by all four antibody combinations. Confocal images of HEK 293 cells expressing the exogenous Scarlet-GPHN-1 protein and stained with all the possible combinations of anti-GPHN antibodies used in this study. (Scale bar: 15 μm). At the bottom, the epitope of each antibody is displayed on a schematic view of the GPHN primary sequence. **(C)** Specificity of antibodies to detect different GPHN isoforms. Expression of GPHN isoforms (-7, -10, -28 and -32) is detected by only one anti-GPHN antibody as expected by the presence of a single epitope on their primary sequence as schematized below. **(D)** Confocal images of cerebellar cortex slices stained with the presynaptic markers (GAD-65 or VGAT) and anti-GPHN antibodies (a-A, a-B, a-C and a-D). Presynaptic markers are labeled in blue and anti-GPHN in red. Yellow boxes indicate the region magnified below each panel. **(E)** Graph showing the percentage of colocalization between anti-GPHNs and the pre-synaptic marker (n≥4 mice) **P=0,0019; *P<0,05. One-way ANOVA. **(F and G)** *In vivo* labeling of GPHN epitopes present at inhibitory synapses localized on PC dendrites **(F)** and soma **(G)**. Schematic of synapses made by stellate cell axon on PC dendrite (**F**; middle panel) and Basket cell on PC soma (**G**; middle panel). At the left panel, pre- and post-synaptic sites are detected respectively using anti-VGAT (blue) and anti-calbindin (green). Stellate cell and Basket cell inhibitory synapse made on PC are positive to a-dystroglycan (a-dg; red). At the right panel, combined GPHN epitopes labeling at a-dystroglycan positive synapse. **(G-I)** Heterogeneous labeling of GPHN epitopes at inhibitory synapses containing specific $GABA_A$ Rα subunits. Schematics in middle panels show each specific inhibitory synapse. Synapses between GABAergic interneurons and PC dendrites are positive for $GABA_A$ Rα1 (**G**). Synapses made by deep cerebellar nuclei axons on Golgi cell dendrites are positive for $GABA_A$ Rα3 (**H**). Synapses made by Golgi cell axons on Granule cell dendrites are positive for $GABA_A$ Rα6 (**I**). GFP positive inhibitory interneuron dendrites within

the molecular layer are identified using the Gad67-GFP mice. Golgi Cells axon and dendrite are identified using the GlyT2-GFP mice.

Fig.S4: Exogenous expression of several GPHN protein isoforms.

(A) On the right, schematics of GPHN-5, -6, -7, -8, -10, -14, -28, -32, -42, -49 isoforms; additional domains are colored in dark blue, missing domains are shown in dashed lines and altered domains are filled with a striped motif. On the left, the table shows the prediction for GPHN isoforms to form clusters (through G-trimerization (*39*), E-dimerization (*40*)), crucial residues palmitoylation (*27*) or collybistin binding (*13*)), to interact with molecular transport (*11*), to interact with GABAARs or GlyRs (*24*), and to be anchored to the cytoskeleton (*12*). A plus sign means the sites are present in the sequence of the variant, a minus sign means the whole site is absent from the sequence. +/- means that the site is partially present in the sequence. Note that these estimations do not account for the potential conformation changes induced by the addition or removal of a peptide sequence. **(B)** At the left, a table list of the theoretical molecular weights of each GPHN isoform fused to Scarlet with the name of corresponding lentiviral constructs. At the right, exogenous expression of fusion GPHN isoforms is shown using western blotting after transfection of lentiviral constructs in HEK-293 cell line. Exogenous factors are detected by probing of the V5 Tag. **(C-F)** Schematic representation of different *Gphn* transcripts by comparing their homology with *Gphn*-1.

Fig.S5: Analyse of human *GPHN* transcriptome by long read sequencing.

(A) Table reporting the origin of human tissue sample (s) that have been used to extract total RNA and amplify human *GPHN* expression. **(B)** Graph displaying the number of sequences in which the *GPHN* exons were detected by long read sequencing in all human tissues. Values are reported using a log scale. **(C)** Schematic representation of human and mouse *GPHN* gene in which both exon architectures are compared using the percentage of nucleotide homology. **(D)** Table showing the percentage of alternative GPHN transcripts that are common between all analyzed human tissues and brain samples. Heatmap staining is used to indicate the percentage of similarity, with red meaning high overlapping (100%) and white means poor overlapping (0%).

Table S1: Sequences of *Gphn* alternative transcripts.

Table S2: List of short **read sequencing data used in this study and quantification of exon-exon junctions present in *Gphn* transcriptome.**

Table S3: List of RNA-seq data used to analyse *Gphn* expression in various neuronal cells.

Table S4: Analysis of theoretical GPHN proteome.

Table S5: List of GPHN peptides identified by mass spectrometry analysis.

Table S6: Frequency of exon inclusion associated with *GPHN* expression in human tissues.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SuplementalFigures.pdf
- TableS1.xlsx
- TableS2.xlsx
- TableS3.xlsx
- TableS4.xlsx
- TableS5.xlsx
- TableS6.xlsx