



## Genomic mapping and cellular expression of human *CPG2* transcripts in the *SYNE1* gene



Sven Loebrich<sup>a,b,c,1</sup>, Mette Rathje<sup>a,b,c,1</sup>, Emily Hager<sup>a,b,c</sup>, Bulent Ataman<sup>d</sup>, David A. Harmin<sup>d</sup>, Michael E. Greenberg<sup>d</sup>, Elly Nedivi<sup>a,b,c,\*</sup>

<sup>a</sup> The Picower Institute for Learning and Memory, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

<sup>b</sup> Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

<sup>c</sup> Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

<sup>d</sup> Department of Neurobiology, Harvard Medical School, Boston, MA 02114, USA

### ARTICLE INFO

#### Article history:

Received 18 September 2015

Revised 11 December 2015

Accepted 14 December 2015

Available online 15 December 2015

#### Keywords:

Bipolar disorder

*SYNE1*

*CPG2*

Excitatory synapses

Glutamate receptor internalization

Activity-regulated genes

### ABSTRACT

Bipolar disorder (BD) is a prevalent and severe mood disorder characterized by recurrent episodes of mania and depression. Both genetic and environmental factors have been implicated in BD etiology, but the biological underpinnings remain elusive. Recent genome-wide association studies (GWAS) for identifying genes conferring risk for schizophrenia, BD, and major depression, identified an association between single-nucleotide polymorphisms (SNPs) in the *SYNE1* gene and increased risk of BD. *SYNE1* has also been identified as a risk locus for multiple other neurological or neuromuscular genetic disorders. The BD associated SNPs map within the gene region homologous to part of rat *Syne1* encompassing the brain specific transcripts encoding *CPG2*, a postsynaptic neuronal protein localized to excitatory synapses and an important regulator of glutamate receptor internalization. Here, we use RNA-seq, ChIP-seq and RACE to map the human *SYNE1* transcriptome, focusing on the *CPG2* locus. We validate several *CPG2* transcripts, including ones not previously annotated in public databases, and identify and clone a full-length *CPG2* cDNA expressed in human neocortex, hippocampus and striatum. Using lenti-viral gene knock down/replacement and surface receptor internalization assays, we demonstrate that human *CPG2* protein localizes to dendritic spines in rat hippocampal neurons and is functionally equivalent to rat *CPG2* in regulating glutamate receptor internalization. This study provides a valuable gene-mapping framework for relating multiple genetic disease loci in *SYNE1* with their transcripts, and for evaluating the effects of missense SNPs identified by patient genome sequencing on neuronal function.

© 2015 Elsevier Inc. All rights reserved.

### 1. Introduction

Robust evidence suggests that genetic factors influence an individual's susceptibility to BD (Craddock and Sklar, 2013). Yet, no single, high penetrance gene has been identified with direct causality (Craddock and Sklar, 2009). The largest GWAS to date for identifying genes conferring risk for the five major psychiatric disorders: schizophrenia, BD, major depression, attention deficit-hyperactivity disorder (ADHD) and autism spectrum disorder (ASD), identified the *SYNE1* gene among the top ten risk loci (Cross-Disorder Group of the Psychiatric Genomics, 2013). Recent meta-analyses of the GWAS data identified SNPs in *SYNE1* with genome-wide statistically significant association to BD ( $p < 4 \times 10^{-8}$ ) (Cross-Disorder Group of the Psychiatric Genomics, 2013; Psychiatric, 2011; Xu et al., 2014). This finding was validated in independent samples and the same risk allele was also found to be associated with recurrent unipolar depression (UD) (Green et al., 2013a,b).

*SYNE1* showed the strongest association with BD of any locus in the genome other than *ANK3* (Cross-Disorder Group of the Psychiatric Genomics, 2013), which encodes ankyrin-G, a newly identified regulator of glutamatergic synapses implicated in psychiatric-related behaviors (Leussis et al., 2013; Smith et al., 2014).

The BD associated SNPs in human *SYNE1* are in a region of the gene homologous to brain specific transcripts of rat *Syne1* encoding *CPG2* (Cottrell et al., 2004). *Cpg2* transcripts were identified and isolated from rat brain in a screen for neural activity-regulated genes (Loebrich and Nedivi, 2009; Nedivi et al., 1993), where their expression is responsive to seizure as well as to physiological levels of stimulation (Nedivi et al., 1993, 1996). Two *Cpg2* transcripts have been reported and both are brain-specific splice variants of the *Syne1* gene that are expressed predominantly in the hippocampus, neocortex, striatum, and cerebellum, brain regions strongly associated with electrophysiological paradigms of synaptic plasticity (Cottrell et al., 2004). The *CPG2* protein localizes to excitatory synapses of excitatory neurons, where it facilitates glutamate receptor cycling, consistent with a role in synaptic plasticity (Cottrell et al., 2004). While the BD associated SNPs identified by GWAS are not necessarily themselves disease-causing mutations, their

\* Corresponding author at: 43 Vassar St., Room 46-3239, MA 02139, Cambridge, USA.  
E-mail address: [nedivi@mit.edu](mailto:nedivi@mit.edu) (E. Nedivi).

<sup>1</sup> Co-first authors.

localization to the *CPG2* locus is intriguing given *CPG2*'s role in glutamate receptor internalization and prior implication of the glutamatergic synapse in BD (Cottrell et al., 2004; Sanacora et al., 2008).

In addition to its association with BD and UD, the *SYNE1* locus has been identified as a risk locus for multiple other genetic disorders. Variations in *SYNE1* are known to cause autosomal recessive cerebellar ataxia type 1 (ARCA1) (Gros-Louis et al., 2007), Emery–Dreifuss type muscular dystrophy (Zhang et al., 2007a), and myogenic arthrogryposis, an autosomal recessive form of congenital muscular dystrophy (Attali et al., 2009). A severe *de novo* mutation identified by exome sequencing in sporadic cases of ASD has been mapped to *SYNE1* (O'Roak et al., 2011), in addition to an autosomal recessive autism mutation identified by homozygosity mapping in pedigrees with shared ancestry (Yu et al., 2013). As yet, the relationship between the different mutations and specific protein products from the *SYNE1* locus is unclear. There is no comprehensive transcriptional map of this important gene and little is known regarding expression of the various transcripts in the human brain, or the cellular localization of their products in human neurons.

Here, we use RNA sequencing (RNA-seq), chromatin immunoprecipitation sequencing (ChIP-seq) and rapid amplification of cDNA ends (RACE) to map the human *SYNE1* transcriptome, providing a framework for localizing the multiple disease-associated mutations identified in this complex gene in relation to specific protein coding transcripts. Several *CPG2* transcripts from the human *SYNE1* gene were identified, including ones not previously annotated in public databases. The first full-length human *CPG2* cDNA was cloned, and the corresponding transcript shown to be expressed in the human neocortex. Further, we raised specific polyclonal antibodies against human *CPG2* and used them to confirm that the protein is expressed in several human brain regions. Using a lenti-viral gene knock down/replacement strategy and a surface receptor internalization assay, we demonstrate that human *CPG2* localizes to dendritic spines in rat hippocampal neurons and is functionally equivalent to rat *CPG2* in regulating glutamate receptor

internalization. The conserved function of *CPG2* between rat and human provides a platform for testing the effect of missense SNPs identified by BD patient exome sequencing on neuronal function.

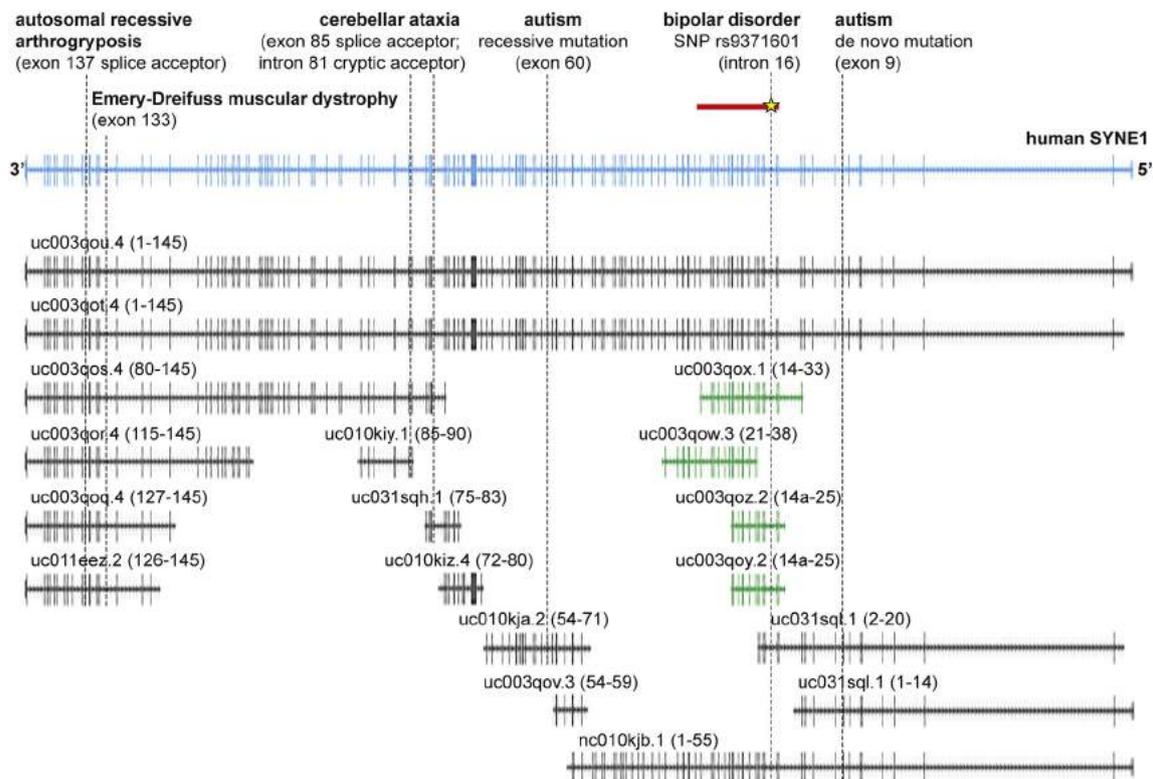
## 2. Materials and methods

### 2.1. Database analysis

Transcript annotations from the human *SYNE1* gene locus (6q25.1–q25.2, NM\_182961) were assembled and mapped using the UCSC Genome browser. All GENCODE basic annotations and transcription support level (tsl) 1–2 annotations from the GENCODE Comprehensive Gene Annotation set are included in the overview (Fig. 1). Exons were mapped based on human spectrin repeat-containing nuclear envelope protein 1 (*SYNE1*, OMIM: 608441).

### 2.2. RNA sequencing and ChIP sequencing of human neurons

RNA-seq and ChIP-seq experiments were performed as previously described (Yu et al., 2013). Primary human neuronal cells were purchased from ScienCell (Carlsbad, CA) and cultured 14 DIV as recommended by the manufacturer. For RNA-seq, total RNA was isolated, ribosomal RNAs were depleted and strand-specific and paired-end cDNA libraries were generated using the Paired-End RNA-seq library kit (Illumina). RNA-seq was performed using HiSeq 2000 (Illumina) and reads were aligned to the human GRCh37/hg19 assembly using Burrows–Wheeler Aligner (BWA) software. For ChIP-seq, cells were cross-linked, lysed and the fragmented chromatin was immunoprecipitated with H3K27Ac (Abcam Cat# ab4729) and H3K4Me3 (Millipore Cat# 07-473) antibodies. The ChIP DNA libraries were constructed using the ChIP-seq DNA Sample Prep Kit (Illumina) and subsequently sequenced using HiSeq 2000. ChIP-seq reads were aligned to the human genome (GRCh37/hg19 assembly) using BWA with default parameters.



**Fig. 1.** The human *SYNE1* locus. The reverse strand of human chromosome 6 spanning the *SYNE1* locus is depicted in blue, note the inverted orientation. Small vertical lines represent exons. Long vertical dashed lines mark sites of disease-associated genetic variations (Attali et al., 2009; Green et al., 2013a; Gros-Louis et al., 2007; O'Roak et al., 2011; Psychiatric, 2011; Yu et al., 2013; Zhang et al., 2007a). Representative transcripts of the nesprins and other *CPG2* transcripts annotated from the UCSC Genome database are shown in black, with exon numbers in parenthesis. Four transcripts (green) span in part the region homologous to rat *Cpg2* (red bar). The yellow star marks the position of the lead SNP identifying *SYNE1* as a risk locus for BD.

### 2.3. Rapid amplification of cDNA ends (RACE)

Total human brain RNA was purchased from Ambion, and cDNA was generated following standard procedures using SuperScript II reverse transcriptase (Life Technologies). For 3' RACE, cDNA was generated using oligo-dT primers. For amplification of cDNA ends, two nested gene specific primers (GSPs) were used for successive PCRs in combination with reverse primers that anneal to a sequence-specific annex in the oligo-dT primer (see Supplementary Table A1 for primer combinations).

For 5' ends, RNA-ligase-mediated (RLM)-RACE was performed (Applied Biosystems). In brief, 5' ends of incomplete, fragmented or partially degraded transcripts were dephosphorylated, rendering them incompetent for ligation to an RNA adapter sequence. Subsequent removal of the cap by tobacco acid pyrophosphatase (TAP) enzyme exposed the intact 5' phosphate to RNA ligase, selecting true 5' ends for ligation to the RNA adapter. cDNA was generated using random decamers. GSPs were designed for nested PCRs, and used in combination with nested primers that annealed to specific sequences in the RNA adapter (see Supplementary Table A1 for primer combinations).

For both 3' RACE and 5' RLM RACE, GSPs were designed to anneal in exon sequences between 300 and 500 bp away from the closest intron junction. Since rat *Cpg2* spans *Syne1* exons 15–34, we tested for 5' ends upstream of exons 12, 14, 16, 17/18, 21, 23 and 27; as well as for 3' ends downstream of exons 24, 27, 28/29, 32 and 37.

Products obtained after nested PCRs were separated on an agarose gel, excised, purified and captured by TOPO cloning. Cloned inserts were categorized by size after EcoRI digestion and verified by Sanger sequencing. Novel cDNA ends were named based on the position of homologous exons in full-length rat *Syne1* cDNA. The 5' ends were named based on the first incorporated *SYNE1* exon after the intronic sequence (i) of the 5'-RLM-RACE end, e.g. i15A. Similarly, 3' ends were named based on the last integrated *SYNE1* exon before the intronic sequence (i) and poly-A-tail, e.g. 34iA.

### 2.4. cDNA cloning

cDNA was derived from whole brain RNA (Ambion) treated with 10 mM methyl mercury hydroxide to suppress secondary structures. Full-length human *CPG2* transcripts were amplified by PCR combining 5'-specific sense primers and 3'-specific antisense primers, where each potential combination of identified 5' and 3' ends was tested individually (see Supplementary Table A2). Long cDNAs were amplified using Taq polymerase and proof reading Pfx DNA polymerase at a 16:1 M ratio. PCR products were separated on agarose gels, excised, purified and captured by TOPO cloning (pTOPO-*hCPG2*-i15A/25iA, pTOPO-*hCPG2*-i15B/25iA and pTOPO-*hCPG2*-i15B/34iA).

The novel cDNA sequences for full-length human *CPG2* (*hCPG2*-i15B/34iA) and short human *CPG2* (*hCPG2*-i15A/25iA) have been submitted to GenBank with accession numbers KU310547 and KU310548, respectively.

The full-length human *CPG2* was subsequently cloned into pcDNA3 with a triple HA-tag using EcoRI and XhoI restriction sites, and into a lenti-viral transfer vector for rat *Cpg2* KD (pFUGW-*Cpg2*-shRNA) (Cottrell et al., 2004), in which the stop codon after *GFP* was removed and *hCPG2* inserted using EcoRI and XhoI restriction sites yielding a GFP-*hCPG2* fusion protein (pFUGW-*GFP-hCPG2* with *Cpg2*-shRNA, Fig. 5).

### 2.5. Northern blotting

Total RNA was isolated from fresh adult Sprague–Dawley rat brain regions or from postmortem, freshly frozen human brain tissue (medial frontal cortex or cerebellum) from three individual subjects with postmortem index of 3–10 h (Massachusetts Alzheimer's Disease Research Center). Rat and human DNA probes were derived from *CPG2* coding

exons in *syne1/SYNE1*. Rat probe: rat brain cDNA was synthesized from RNA using Superscript III first-strand cDNA synthesis (Invitrogen), the probe sequence was amplified by PCR using GSPs (sense: CTGTGACGAGCTAAC, antisense: CTTGACAAGGTTTTTCAG), captured by pTOPO cloning, sequenced and excised by EcoRI digestion. Human probes: P1 and P2 sequences were excised from pTOPO-*hCPG2*-i15B/34iA by BglII and PstI digestion or HindIII digestion, respectively (endogenous restriction sites within *CPG2*). All probes were gel purified and <sup>32</sup>P-dCTP-labeled using random hexamer priming of exo-Klenow fragments. Samples of 10 µg cortical or cerebellar total RNA were separated by chilled electrophoresis on 1% agarose gels containing 5% formaldehyde, and blotted onto nylon membrane. Blots were cross-linked by UV light and blocked with Ultrahyb solution (Ambion). Probes were applied at 1 million counts/mL in a 5 mL volume overnight at 42 °C, then vigorously washed using standard procedures. Kodak Biomax film was exposed through an intensifying screen for 1–2 days at –80 °C.

### 2.6. Western blotting

HEK293 cells were transfected with pcDNA3-*HA-hCPG2* or pFUGW (*GFP* control). After 24 h, cells were lysed in 150 mM NaCl, 25 mM Tris-HCl (pH 7.4), 1% Triton-X-100, 1% sodium deoxycholate, and 0.1% SDS, with protease inhibitors (Roche). Samples of 10 µg total protein were diluted in loading buffer and separated by SDS-PAGE.

Tissue samples of 200–300 mg were collected from the neocortex, hippocampus, striatum or cerebellum of 3 week old Sprague–Dawley rats (n = 3) or healthy human subjects (n = 3, age: 34–83 years, PMI: 3–11 h). Samples were homogenized in 0.32 M sucrose, 40 mM HEPES-KOH (pH 7.4), 1 mM DTT, 1 mM EDTA, and 1 mM EGTA, with protease inhibitors (Roche). Homogenates were centrifuged for 10 min at 1100 ×g and supernatants were collected and centrifuged for 10 min at 9200 ×g. Pellets were dissolved in homogenization buffer with 2% Triton-X-100 and centrifuged for 20 min at 15,000 ×g. Pellets were dissolved in 200 µl of homogenization buffer with 2% SDS and centrifuged at RT for 10 min at 15,000 ×g. Samples of 20 µg total protein were diluted in loading buffer and separated by SDS-PAGE. Proteins were transferred onto nitrocellulose blots, which were subsequently blocked with 2% nonfat milk and 2% BSA in PBS with 0.1% Tween-20.

Guinea pig polyclonal anti-CPG2 antibodies (A002396; np913) were generated against the peptide LEQTKFSKRTESIATQAENLVKE and purified by antigen affinity purification (NEO Peptide). Blots were incubated with anti-CPG2 (1:1000), mouse anti-HA, rabbit anti-GFP (1:5000; Abcam) or mouse anti-PSD95 (1:10,000; UC Davis) antibodies and visualized using an Odyssey infrared imaging system (LI-COR).

### 2.7. Neuronal cultures, lentiviral infection and immunocytochemistry

Rat hippocampal cultures were prepared as described (Rathje et al., 2013). In short, hippocampi from E18 rat embryos were isolated, digested with papain (Worthington) and titrated in plating media (Neurobasal with GlutaMAX (Invitrogen) and 10% FBS), using fire-polished Pasteur pipettes. Cells were then plated on poly-L-lysine coated glass coverslips in 12-well plates at a density of  $1.3 \times 10^4$  cells/cm<sup>2</sup>. After 3 h, the media was replaced with culture media (Neurobasal with GlutaMAX and 2% B27 supplement (Invitrogen)). Ara-C (2 µM) was added to the media at 7 DIV, to block glial proliferation. At 8 DIV, neurons were infected with lentivirus for KD of rat *Cpg2* (pFUGW-*Cpg2*-shRNA) (Cottrell et al., 2004), or at 19 DIV with lentivirus for replacement of the endogenous rat *CPG2* with the GFP-*hCPG2* fusion protein (pFUGW-*GFP-hCPG2* with *Cpg2*-shRNA). Transduction rates were estimated as ~70% for GFP expressing KD neurons and ~20% for GFP-*hCPG2* replacement expressing neurons.

At 24 DIV cells were fixed in 4% formaldehyde for 15 min and then permeabilized for 25 min with 0.2% saponin in PBS with 10% goat serum. Cells were then incubated for 3 h with primary antibodies; guinea pig anti-CPG2 (1:1000; NEO Peptide), mouse anti-PSD95

(1:1000; ABR) and chicken anti-MAP2 (1:5000; Novus Biologicals) or rabbit anti-GFP (1:3000; Abcam), followed by 1 h incubation with secondary antibodies; goat anti-guinea pig IgG-Alexa Fluor 647 (1:500), goat anti-mouse IgG-Alexa Fluor 555 and goat anti-chicken IgG-Alexa Fluor 488 (1:1000) or goat anti-rabbit IgG-Alexa Fluor 488 (1:1000; Molecular Probes). Coverslips were mounted using Fluoromount-G (Southern Biotech) and imaged using a Nikon Eclipse E600 upright microscope with a 40×/1.40 Plan Apo oil immersion objective (Nikon).

### 2.8. Biotin internalization assay

The biotin internalization assay was performed essentially as described (Cottrell et al., 2004) with the following modifications: cortical neurons were isolated from E18 rat brains, digested with papain, titrated in plating media and seeded onto 10 cm poly-L-lysine coated cell culture dishes at a density of  $6.4 \times 10^4$  cells/cm<sup>2</sup>. At 8 DIV, the cultures were infected with lenti-virus for CPG2 KD (pFUGW-Cpg2-shRNA), or at 10 DIV with the pFUGW-GFP-hCPG2 replacement virus. At 15 DIV, cultures were incubated with 100 µg/ml of leupeptin (Sigma) for 1 h at 37 °C and then surface labeled with 1.5 mg/ml sulfo-NHS-SS-biotin for 20 min at 4 °C. Cultures were returned to their original culture media and incubated at 37 °C for 30 min to allow surface receptor internalization. Control cultures were kept at 4 °C. Remaining extracellular biotin was cleaved with 37 mM TCEP in 200 mM Tris-HCl (pH 7.5), 75 mM NaCl, 10 mM EDTA, and 1% BSA followed by 50 mM glutathione in PBS (pH 8.7), 75 mM NaCl, 10 mM EDTA, and 1% BSA, each for 25 min at 4 °C. Cells were lysed as described (Cottrell et al., 2004) and biotinylated proteins were separated by SDS-PAGE. Western blots were probed with rabbit anti-GluA2 (1:1000; Abcam), mouse anti-GluN1 (1:2000; Temicula) or mouse anti-TfR (1:1000; Invitrogen) primary antibodies and developed using the Odyssey infrared system (LI-COR). Receptor internalization was quantified as described (Cottrell et al., 2004) and groups were compared using one-way ANOVA and Tukey's *post hoc* test for multiple comparisons.

## 3. Results

### 3.1. Known SYNE1 transcripts and their relationship to risk loci for various muscle and brain related diseases

Human SYNE1 is a large gene comprised of 145 exons that give rise to multiple transcripts annotated in the UCSC Genome database, only some of which have been verified (Fig. 1). The largest transcript, which spans exons 1–145, encodes a giant isoform of the protein nesprin-1. Shorter putative nesprin isoforms, derived from exons 80–145, contain the C-terminal region of the giant isoform, and range in size from 40 to 1000 kDa. These include specific isoforms expressed in heart and muscle but not in brain (Rajgor et al., 2012; Rajgor and Shanahan, 2013) with cellular localization to the nuclear envelope (Zhang et al., 2001). Mutations related to Emery–Dreifuss type muscular dystrophy (Zhang et al., 2007a) and autosomal recessive arthrogyriposis (Attali et al., 2009), fall within the 3' terminal region of SYNE1 encompassing the short nesprin isoforms.

Three transcripts are reported in the 5' terminal region of SYNE1 encoding N-terminal nesprin products, which include the isoform spanning exons 1–55, a region harboring a *de novo* mutation identified in sporadic cases of ASD (O'Roak et al., 2011). Additional SYNE1 annotations span short central stretches, mostly encoding partial sequences and hypothetical proteins of unknown function. These include a cluster of four sequences that span the rat Cpg2 homologous region approximately from exons 14 to 38. An intronic SNP (rs9371601), situated in the Cpg2 homologous region was identified in a GWAS for BD risk genes (Green et al., 2013a; Psychiatric, 2011). A mutation in SYNE1 identified by whole exome sequencing in familial ASD, maps to one additional internal transcript potentially spanning exons 54 to 71 (Yu

et al., 2013), and splice-altering mutations associated with cerebellar ataxia (Gros-Louis et al., 2007) map to other downstream transcripts overlapping exons 81–85.

### 3.2. Mapping transcriptional activity and transcription start sites in SYNE1

Given the complexity of the human SYNE1 locus, and that many of the annotated transcripts had not been validated, including the internal 'partial' transcripts in the Cpg2 homologous region, we sought to verify their existence and potential transcriptional start sites. We first performed strand-specific RNA-seq on total RNA preparations from cultured embryonic human cortical neurons (Fig. 2A, top row signals).

The origin and the orientation of the RNAs were analyzed genome-wide by mapping the reads to the reference human genome. The RNA-seq showed above background transcriptional activity mainly in exonic regions, with a substantial increase in levels starting around exons 8–10. This suggests modest cortical transcription of the full SYNE1 gene and more pronounced transcription of partial gene products.

We also performed ChIP with antibodies recognizing tri-methylated human histone H3 (H3K4me3) or antibodies recognizing the acetylated H3 protein (H3K27Ac) followed by DNA sequencing, with H3K4me3 marking active promoter regions and H3K27Ac marking both active promoter and enhancer regions (Fig. 2A, green and purple tracks). The ChIP-seq data showed four distinct active promoter regions within SYNE1, suggesting that this gene gives rise to multiple transcript families.

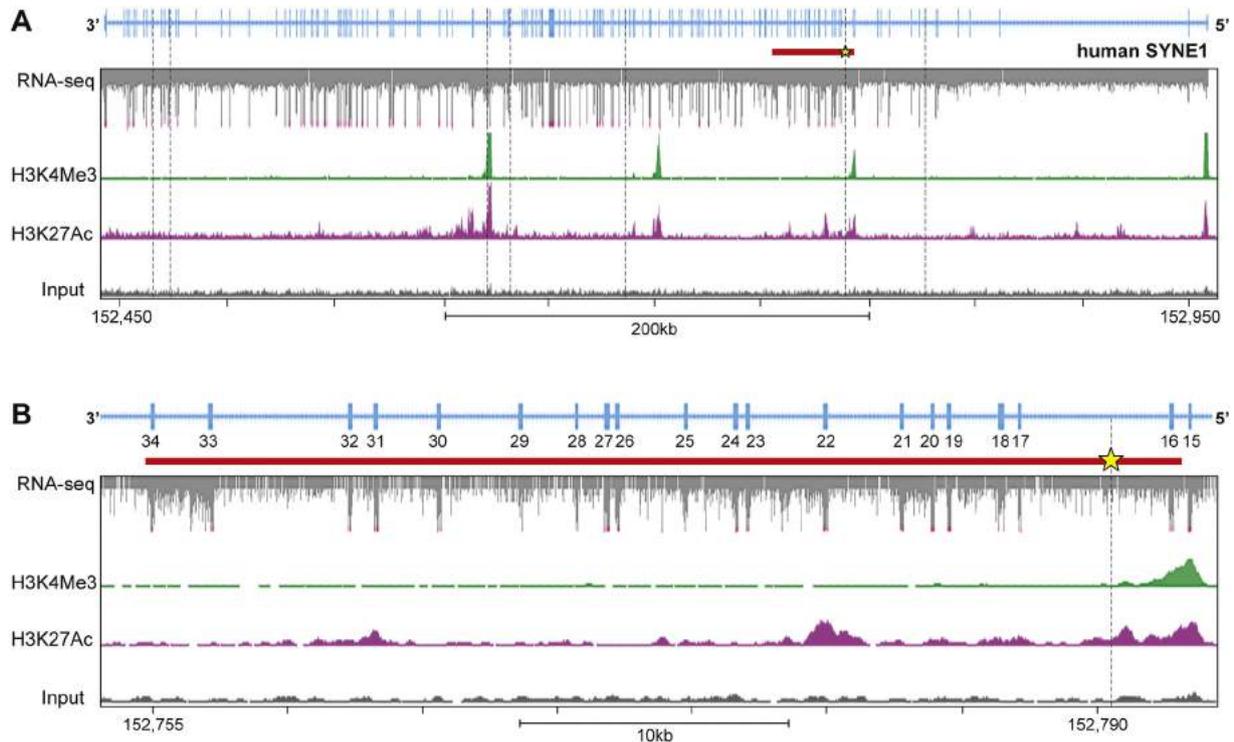
We found a promoter signal in the exon 15/16 region (Fig. 2A and B), near the start site of the rat Cpg2 homologous region of SYNE1 (red bar), suggesting that it may be the promoter for a human CPG2 variant. RNA-seq data showed strong transcriptional activity through potential CPG2 exons 15–34, as well as the intron region between exons 33 and 34 (Fig. 2B). Read-through of the 33–34 intron is a hallmark of the rat Cpg2 mRNA (Cottrell et al., 2004). The human SYNE1 33–34 intron is similar in length to the homologous rat *Syne1* intron, which contains the Cpg2 mRNA stop codon and untranslated region (UTR), and has a sequence identity of ~70%. The predicted human SYNE1 transcript would therefore be approximately 6 kb, similar in size to rat Cpg2.

### 3.3. Identification and cloning of human CPG2 homologs

In rat there are two brain-specific Cpg2 transcripts, derived from alternative splicing of the *Syne1* gene (Cottrell et al., 2004). The ~6 kb full-length rat Cpg2 transcript contains a 2.8 kb open reading frame (ORF), which encodes a 941 amino acid protein (Fig. 3A) spanning exons 16–34. The ORF is flanked by a short 5' UTR and a 2.3 kb 3' UTR that mainly consists of the unspliced intron between exons 33 and 34, with the translation termination site residing a few codons past exon 33 (Cottrell et al., 2004). A second transcript, Cpg2b, contains an additional 5' exon (exon 15) encoding 24 extra N-terminal amino acid residues (Cottrell et al., 2004). The start codon for this second ORF is not present in human SYNE1. The SYNE1 region homologous to rat Cpg2 spans exons 16–34 (Fig. 3A), and four transcripts are annotated in this region (uc003.qox.1, uc003.qow.3, uc003.qoz.2, and uc003.qoy.2, Figs. 1 and 3A).

While these transcripts could potentially represent human CPG2 variants, none correspond to a full-length version, suggesting they may be incomplete, or that other unreported transcripts exist. Further, 5' ends of unverified database sequences can be a result of premature RT-PCR termination, while priming from genomic poly-A stretches can give rise to false 3' ends.

To validate the ends of the potential CPG2 transcripts annotated in the database and to identify novel ends, we performed RACE on cDNA derived from adult human brain total RNA. For identifying 5' ends we used RLM (RNA-ligase mediated)-RACE, which relies on an intact 5' cap structure for amplification. Hence, identified ends are unlikely to result from fragmented or incomplete transcription. Our data confirmed the 5' end of transcript uc003.qoy.2 (Fig. 3A), and identified several



**Fig. 2.** Transcription activity in the human *SYNE1* locus. RNA-seq and ChIP-seq were performed on cultured human embryonic neurons. RNA-seq results are pooled from five individual cultures. A) The exon structure of human *SYNE1* is depicted in blue (note reversed orientation). Black dashed lines indicate the locations of the disease-associated variations; the BD-associated SNP is marked by a yellow star. RNA-seq reads from the full *SYNE1* gene locus are shown in gray with pink tips indicating cut-off of stronger signals. Active promoter regions are visualized by H3K4me3 ChIP-seq (green tracks). Both active promoter and enhancer regions are visualized by H3K27Ac ChIP-seq (purple tracks). ChIP-seq input control is shown as gray tracks on the bottom row. Note the increased signal starting from the region homologous to rat *cpg2* (red bar). B) RNA-seq and ChIP-seq reads in the *Cpg2* homologous region are presented at higher magnification. Note the promoter in the exon 15 region and transcriptional activity indicating read-through between exons 33 and 34.

novel 5' transcript start sites at or around exon 15, near the ChIP-seq mapped promoter (Figs. 2 and 3B).

The 3' ends were considered valid when (i) the identified sequence of the corresponding last exon was followed by the immediately adjacent intronic sequence, (ii) the intronic sequence was terminated by a poly-A tail, and (iii) there were no oligo-A or A-rich stretches in the genomic sequence that could have primed the oligo-dT-based reverse transcription from an unspliced primary transcript (see Supplementary Table A1 for primer combinations). Our data confirmed the 3' ends of all 4 annotated transcripts (Fig. 3A). We also identified three novel 3' transcript ends adjacent to exons 27 and 34, one of which included the unspliced 33–34 introns as also seen in the rat full-length transcript (Fig. 3B). These findings suggest that there may be several alternative *CPG2*-like transcripts in human brain, including one encoding a full-length *CPG2* protein.

To isolate and clone human *CPG2* transcripts we performed PCR on human brain cDNA using end-specific primers intronic to known *SYNE1* transcripts. We tested all combinations of primers that could amplify transcripts corresponding to the 5' and 3' ends identified by RACE (Supplementary Table A2). All PCR products were captured by TOPO cloning and sequenced. Using either the i15A or i15B 5' end primer in combination with the 25iA 3' end primer we amplified two 2.0 kb transcripts (Fig. 3C). Sequencing confirmed the i15B/25iA transcript to be identical to the uc003.qoy.2 annotation. The i15A/25iA transcript is novel, yet both transcripts potentially encode a 562 amino acid protein, homologous to the N-terminal half of rat *CPG2*. Using a combination of i15B/34iA end primers, we also amplified a 6 kb transcript, which represents the first identified human full-length *CPG2* mRNA (Fig. 3D).

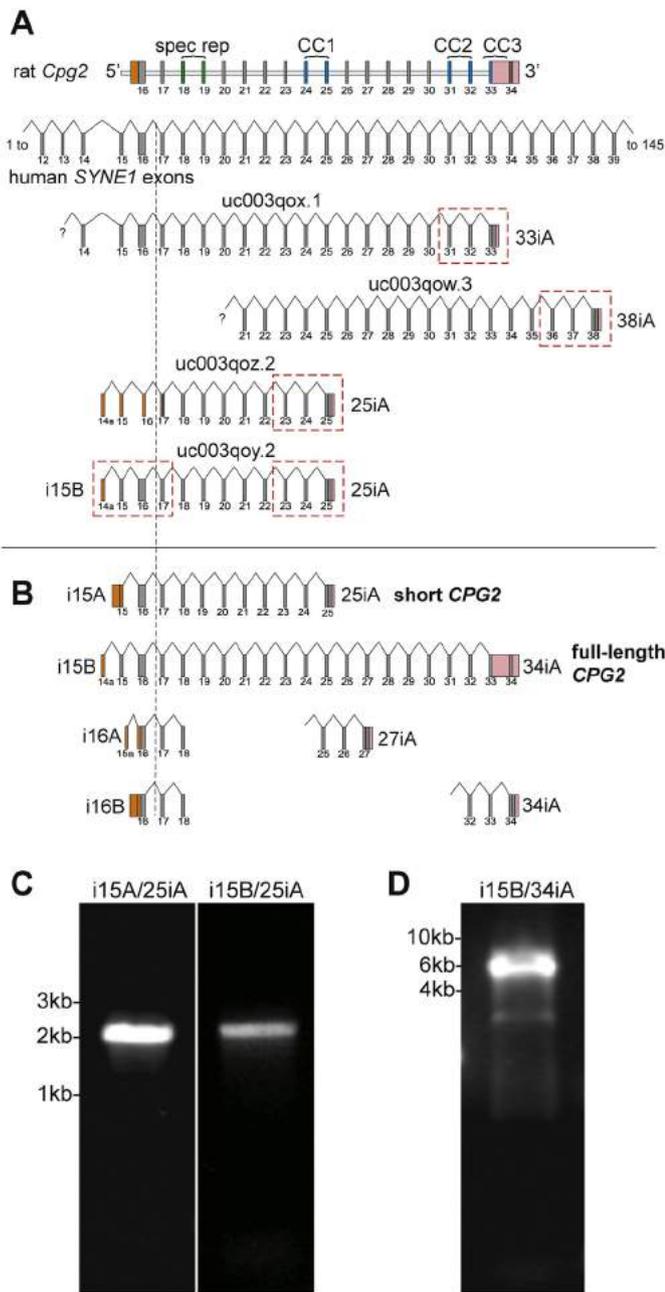
### 3.4. *CPG2* expression in rat and human brain

To compare expression of *CPG2* transcripts in human brain tissue with known rat transcripts (Cottrell et al., 2004), we performed a

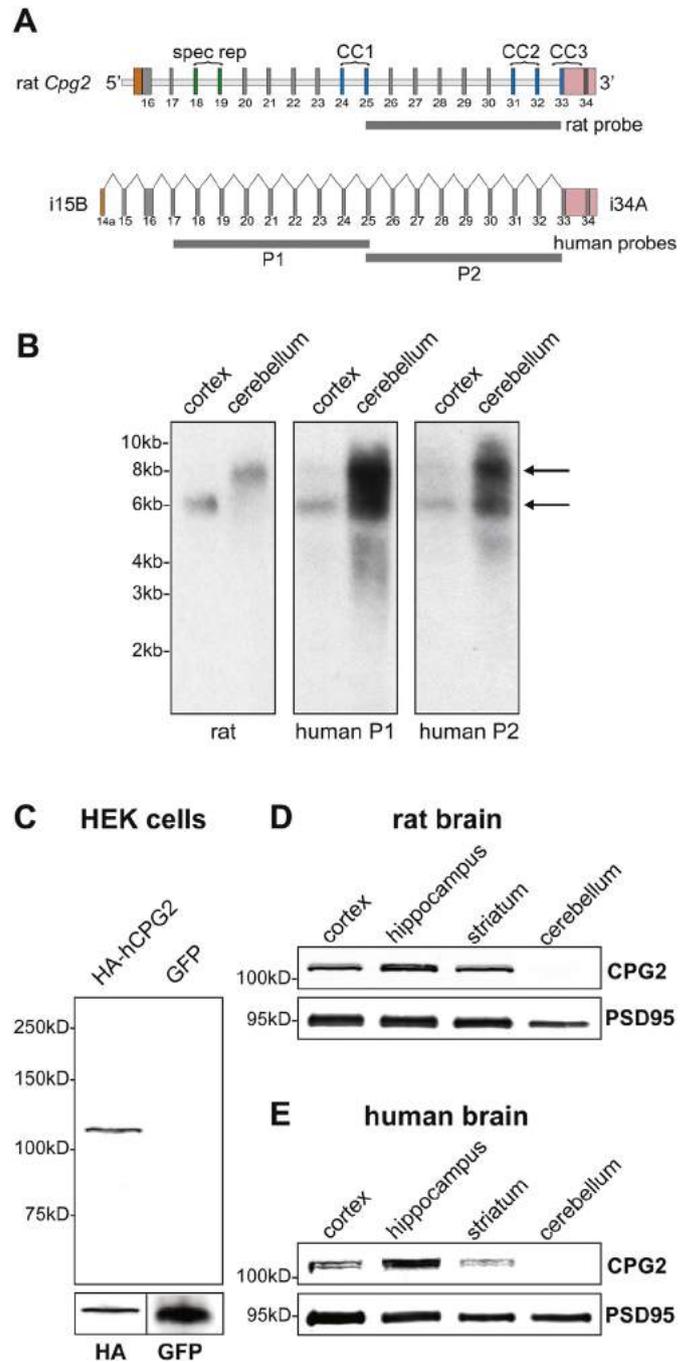
Northern blot analysis on total RNA from the cortex or cerebellum using  $^{32}\text{P}$ -radiolabeled DNA probes designed to detect mRNAs in the *CPG2* encoding region (Fig. 4). A probe derived from the 3' protein coding region of rat *Cpg2* (Fig. 4A) recognized a band at ~6 kb in the rat cortex (Fig. 4B), consistent with previous data (Cottrell et al., 2004). A transcript of similar size was also identified in the human cortex using a probe derived from the human 3' exonic region, verifying the presence of the newly cloned i15B/34iA transcript. In the rat cerebellum, we detected a band at ~8 kb, larger than the full-length rat *Cpg2* mRNA identified in the cortex (Fig. 4B). This may represent a cerebellum specific *Cpg2* transcript, or a non-*Cpg2* cerebellum specific *Syne1* transcript with shared exons. In the human cerebellum, we detected both the ~6 kb and ~8 kb bands, suggesting a more complex *SYNE1/CPG2* transcriptome in the human brain.

To test whether the identified human *CPG2* transcripts were translated into *CPG2* protein in different human brain regions, we performed Western blotting analysis. Since previously developed antibodies against rat *CPG2* did not recognize the human protein, we generated a new antibody against a C-terminal epitope conserved between the rat and human proteins. To verify that this antibody recognizes *CPG2*, Western blots of protein extracts from HEK293 cells transfected with HA-tagged human full length *CPG2* (i15B/34iA) were probed with *CPG2* and HA antibodies. The blots showed a single band at ~115 kDa, consistent with the predicted size of HA-*CPG2* (Fig. 4C). The antibody did not detect a band in control extracts from GFP transfected HEK293 cells. We then probed Western blots containing synaptosome protein extracts from rat (Fig. 4D) or human brain (Fig. 4E) with *CPG2* and PSD95 antibodies. Blots of protein extracts from the rat or human neo-cortex, hippocampus or striatum showed robust bands at ~110 kDa, consistent with the expected size of *CPG2*. However, we did not detect *CPG2* protein expression in synaptosome fractions from rat or human cerebellum. We speculate that the protein product of the ~8 kb

cerebellar transcript detected in the Northern blots analysis, is either not expressed in synaptosome protein fractions or that the derived protein does not contain the CPG2 antibody epitope.



**Fig. 3.** Human *SYNE1* transcripts homologous to rat *Cpg2*. A) Schematic depiction of the rat *Syne1* exon structure in the *Cpg2* coding region (top row) and *Cpg2*-homologous transcripts in humans according to the UCSC Genome database. 5' and 3' UTRs are shown in orange and pink, respectively. Four annotated transcripts overlap with the *Cpg2*-homologous region (uc003.qox.1, uc003.qow.3, uc003.qoz.2 and uc003.qoy.2). Boxed transcript ends were verified by 5' RLM-RACE or 3' RACE. Transcript uc003.qoy.2 was verified and encodes a shorter CPG2 variant not known in rat. B) A full-length human CPG2 cDNA has a 5' end identical to the uc003.qoy.2 and extends to a newly identified 3' end (34iA). The region between exons 33 and 34 is not spliced out and gives rise to a 3 kb 3' UTR, homologous to rat *Cpg2*. Three novel 5' ends around exons 15 and 16 (i15A, i16B, i16A), and two novel 3' ends around exons 27 and 34 (27iA, 34iA) are shown. C) PCR amplification testing 5' RLM-RACE and 3' RACE end-specific primers in all combinations on human brain cDNA yielded two transcripts encoding shorter forms of CPG2. D) A full-length CPG2 transcript was obtained by pairing i15B and 34iA primers. The exon structure of the resulting 6 kb transcript is homologous to rat *Cpg2*, and the derived human protein is 85% identical to rat CPG2.



**Fig. 4.** CPG2 expression in rat and human brain. Northern blotting was performed on total RNA from rat or human cortex or cerebellum probing for *Cpg2*/CPG2 transcripts, using random-hexamer-primed,  $^{32}$ P-radiolabeled DNA probes. A) Schematic of the probes used to detect rat and human *Cpg2*/CPG2 transcripts. Two non-overlapping human probe templates (P1 and P2) were excised by restriction digest, taking advantage of endogenous restriction sites within the ORF. B) Representative northern blots showing a ~6kb-band in rat and human cortices corresponding to the length of the i15B/34iA transcript isolated from human brain cDNA. Note that both rat and human cortices show a longer, not yet characterized 8 kb transcript mainly expressed in the cerebellum. C) Western blotting was performed using antibodies recognizing an epitope conserved in rat and human CPG2. Representative Western blot showing HA-tagged i15B/34iA transcript (*hCPG2*) expressed in HEK293 cells. Lysate from GFP-transfected cells was used as negative control. Representative Western blots showing endogenous CPG2 and PSD95 protein expression in synaptosome fractions from rat (D) or human (E) neocortex, hippocampus or striatum. CPG2 protein expression was not detected in synaptosome fractions from rat nor human cerebellum.

Immunostaining of cultured rat hippocampal neurons with the new CPG2 antibody showed a punctate distribution along the dendrites partially overlapping the staining pattern of the postsynaptic density protein PSD95, consistent with the localization pattern of CPG2 (Fig. 5A and B) (Cottrell et al., 2004; Loebrich et al., 2013).

### 3.5. Human CPG2 expressed in cultured rat neurons is localized to dendritic spines

We next tested whether human CPG2 is synaptically localized, similar to rat CPG2. For this purpose, we first knocked down (KD) endogenous rat CPG2 in cultured neurons using a previously validated lentivirus plasmid expressing a rat *Cpg2*-specific small hairpin RNA (shRNA) driven by the mouse U6 promoter together with a GFP reporter driven by the human ubiquitin-C promoter (see (Cottrell et al., 2004) for details). Neurons infected with the shRNA virus showed robust KD of CPG2 protein in spines (Fig. 5B). It should be noted that we still observed immunostaining at perinuclear sites, which was not decreased with expression of shRNA for KD of CPG2 (data not shown) and could potentially originate from other *Syne1*-derived proteins not expressed in spines.

We then replaced the GFP reporter with a shRNA-resistant GFP-hCPG2 fusion construct that also expressed the *Cpg2*-specific small hairpin RNA (shRNA) (see vector diagram, Fig. 5C). GFP-hCPG2 fusion protein showed a punctate anti-CPG2 staining pattern similar to the spine localization of the endogenous protein (Fig. 5D) overlapping with the anti-GFP staining pattern. These results demonstrate that human CPG2 protein localizes to dendritic spines in hippocampal neurons similar to rat CPG2.

### 3.6. Human CPG2 expressed in cultured rat neurons can functionally replace rat CPG2

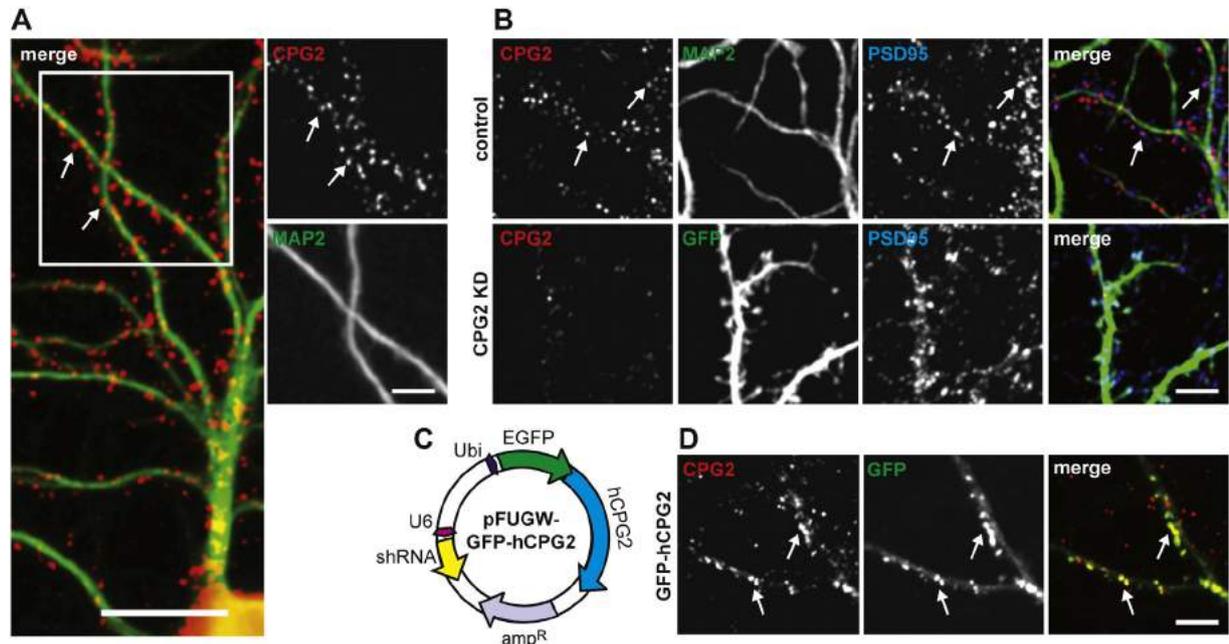
Rat CPG2 has previously been shown to localize to the endocytic zone of dendritic spines, where it regulates endocytosis of synaptic glutamate receptors (Cottrell et al., 2004; Loebrich et al., 2013). To test if

human CPG2 is equally able to facilitate glutamate receptor internalization, we used a surface receptor biotin internalization assay (Cottrell et al., 2004; Loebrich et al., 2013). High-density cortical neuron cultures were either untreated (control), infected with shRNA for KD of endogenous CPG2 or infected with shRNA for CPG2 KD together with GFP-hCPG2 replacement. Surface proteins were biotinylated with cleavable sulfo-NHS-biotin and cultures were incubated at 37 °C for 30 min to allow constitutive internalization of surface receptors. Remaining extracellular biotin was removed by chemical reduction. Cultures kept at 4 °C without internalization (time 0') were used as a control for cleavage efficiency. Cells were lysed and biotinylated protein was isolated using agarose-immobilized neutravidin. Western blots of biotinylated receptor proteins were probed with antibodies recognizing glutamate receptors GluA2, GluN1 or transferrin receptor (TfR) (Fig. 6A). Internalized receptor fractions were calibrated by probing blots of known amounts of uncleaved biotinylated receptor protein.

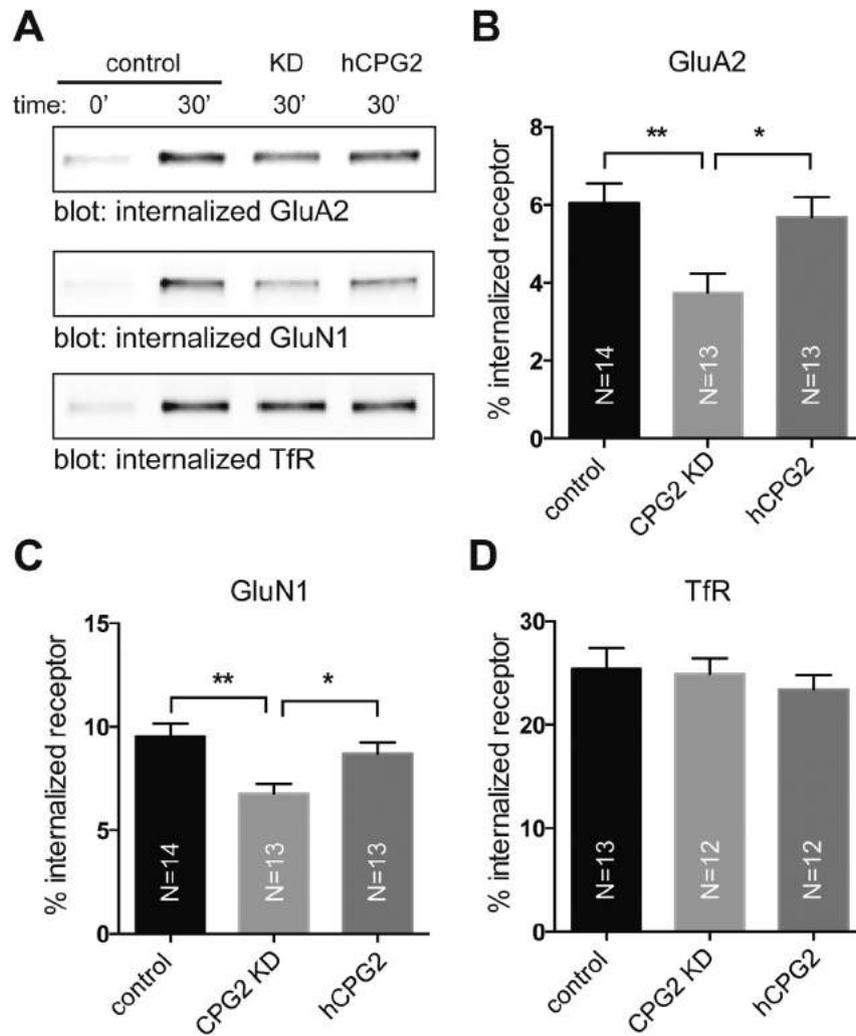
We found that  $6.0 \pm 0.5\%$  of GluA2 containing AMPARs are constitutively internalized after 30 min (Fig. 6B). Consistent with previous findings (Cottrell et al., 2004; Loebrich et al., 2013), KD of CPG2 significantly decreased the level of GluA2 internalization ( $3.7 \pm 0.5\%$ ,  $P < 0.01$ , one-way ANOVA). Viral replacement of endogenous CPG2 with GFP-hCPG2 rescued GluA2 internalization to normal values ( $5.7 \pm 0.5\%$ ). Similarly, CPG2 KD significantly decreased constitutive internalization of GluN1 containing NMDARs by ~30% compared to control levels ( $6.8 \pm 0.4\%$  vs.  $9.5 \pm 0.6\%$ ,  $P < 0.01$ , One-way ANOVA). Viral replacement with human CPG2 also rescued the internalization of GluN1 to control values ( $8.9 \pm 0.5\%$ ). Neither CPG2 KD nor replacement with human CPG2 significantly affected the constitutive internalization of TfR (Fig. 6C). Our results show that human CPG2 is functionally equivalent to rat CPG2 in facilitating glutamate receptor internalization in neurons, suggesting a conserved function for CPG2 in the rat and human brains.

## 4. Discussion

In the past decade, GWAS have attempted to identify genetic variants that confer risk for many human diseases, whose inherited



**Fig. 5.** Spine localization of human CPG2 protein in cultured rat neurons. A) Representative images of cultured hippocampal neurons (24 DIV) stained with antibodies against CPG2 (red) and dendritic protein MAP2 (green). Boxed region is channel-separated in the second column. B) Representative images of control neurons (top row) or neurons expressing shRNA for KD of endogenous CPG2 together with a GFP reporter (second row) stained with antibodies against CPG2 (red), MAP2 (green) and synaptic density protein PSD95 (blue). C) Schematic depiction of the human CPG2 molecular replacement plasmid with shRNA for KD of endogenous CPG2 and a GFP-hCPG2 fusion protein. D) Representative images showing lenti-virus infected hippocampal neurons expressing GFP-hCPG2 fusion protein. Cells were stained with anti-CPG2 (red) and anti-GFP (green). Scale bars: 15  $\mu$ m (A, left panel) and 5  $\mu$ m (A, right panels, B and D). Arrows indicate signal co-localization.



**Fig. 6.** Human and rat CPG2 are functionally equivalent in regulating glutamate receptor internalization. Cultured cortical neurons were either uninfected (control), infected with shRNA for KD of endogenous CPG2 or with shRNA for CPG2 KD together with the hCPG2 replacement. Surface receptors were biotinylated and allowed to internalize at 37 °C for 30 min, after which remaining surface biotin was cleaved. Cultures kept at 4 °C without internalization (time 0') were used as a control for cleavage efficiency. A) Representative Western blots showing the internalized fraction of biotinylated surface receptors. Blots were probed with anti-GluA2, anti-GluN1 or anti-TfR. B–D) Quantification of internalized GluA2 (B), GluN1 (C) and TfR (D) after 30 min is presented as percent internalization calibrated using known loads of an uncleaved biotinylated receptor. KD of endogenous CPG2 decreased the internalization of GluA2 and GluN1 but not the internalization of TfR. Human CPG2 replacement showed receptor internalization rates comparable to control. N = 12–14 individual experiments. Statistical significance was determined by one-way ANOVA and Tukey's *post hoc* test, \* $p < 0.05$ , \*\* $p < 0.01$ .

components remain unexplained (Manolio et al., 2009). In a few cases risk variants identified by GWAS have paved the way for a molecular understanding of disease causes. Crohn's disease, ulcerative colitis, and type I diabetes are all examples where deep sequencing follow-up on GWAS hits have revealed specific disease causing mutations (Lee et al., 2013; Nejentsev et al., 2009). However, experience from these studies and others indicates that when individual gene effects are relatively small, as would be the case for neuropsychiatric diseases such as BD, data from tens of thousands of patients is required in order for GWAS results to be meaningful (Craddock and Sklar, 2009). Only in recent years have patient databases become large and diagnostically detailed enough to confidently identify individual risk genes as well as complex genetic pathways (Cross-Disorder Group of the Psychiatric Genomics, 2013; Psychiatric, 2011).

An additional difficulty, making many GWAS hits challenging to follow up functionally, is the absence of comprehensive transcriptional, and protein coding maps for genes in the region of the identified risk loci. Without knowing the gene products potentially altered by GWAS identified variations, mechanistic testing of disease processes is impossible. Further, knowledge of the human gene products is essential for evaluating the relevance of studies using animal models to human

disease. While annotated human sequence databases, such as the UCSC Genome database, are a good starting point, they suffer from several limitations. Most annotated database sequences are unverified, in particular in relation to their transcript start and stop sites. Due to the way most database sequences are generated from poly-A-primed cDNA templates, transcript ends are frequently a result of early termination of cDNA synthesis and lack the true 5' terminus. This is particularly true for the long transcripts typical of brain tissue (Zylka et al., 2015). In addition, transcript priming can arise from genomic poly-A stretches, thus denoting false 3' termini. Further, current databases are by no means comprehensive and it is likely that additional unreported transcripts exist, especially in the case of genes with differential tissue specific expression patterns. Given these concerns, the first step in evaluating the potential significance of various mutations and polymorphisms that fall near GWAS risk loci is mapping of the transcriptional and protein coding region of interest.

Here, we map the *SYNE1* gene, a large gene with a complex transcriptional profile, which has been implicated in a variety of diseases. The first group includes muscular dystrophies; Emery–Dreifuss type muscular dystrophy and autosomal recessive arthrogyrosis, which are conditions that severely affect skeletal and cardiac muscle function

and might be associated with muscle specific transcripts of *SYNE1* (Attali et al., 2009; Puckelwartz et al., 2009; Zhang et al., 2001, 2007a). The mutations for these degenerative neuromuscular diseases have been mapped to the *SYNE1* 3' end (Rajgor and Shanahan, 2013). Mice with deletion of the *Syne1* 3' end encoding the Klarsicht, ANC-1, *Syne1* Homology (KASH) domain, a consensus sequence for nuclear envelope binding, display disrupted anchoring of nuclei to the cytoskeleton in skeletal muscle, which significantly affects motor nerve innervation (Zhang et al., 2010). Both *SYNE1* and the homologous gene *SYNE2* are expressed in skeletal muscle cells and have overlapping cellular functions (Rajgor et al., 2012; Zhang et al., 2005). *Syne1*; *Syne2* double-knockout mice die of respiratory failure shortly after birth, suggesting that these genes' role in anchoring myonuclei is crucial for respiratory muscle innervation and function (Banerjee et al., 2014; Zhang et al., 2007b).

Other mutations in the *SYNE1* 3' region (exons 81–85) have been implicated in an upper motor neuron disease, ARCA1 (Gros-Louis et al., 2007). The disease is characterized by progressive movement, coordination, and balance problems caused by disrupted Purkinje cell function and their impaired signaling to other cerebellar neurons (Gros-Louis et al., 2007). To date it is still unclear how the *SYNE1* gene mutations lead to the cerebellar cellular phenotypes contributing to ARCA1. Interestingly, a large *SYNE1* transcript has recently been identified in mouse cerebellum, which encodes a nesprin-1 isoform that lacks the KASH domain for nuclear anchorage and instead localizes to glomeruli of cerebellar mossy fibers (Razafsky and Hodzic, 2015). We speculate that the larger transcript size band observed in the cerebellum in our Northern blotting analysis might belong to human homologs of the cerebellum specific transcript identified in mice (Razafsky and Hodzic, 2015) that potentially harbor the cerebellar ataxia associated SNPs (Gros-Louis et al., 2007).

A second group of mutations in *SYNE1* have been implicated in synapse-associated psychiatric disorders, ASD, BD, and major depression, which display considerable comorbidity, and diagnostic boundaries that are often difficult to define (Cross-Disorder Group of the Psychiatric Genomics, 2013). While family studies have repeatedly demonstrated that genetic risk factors are important in the causality of all these diseases, the role of specific genes remain unclear (Berrettini, 2000). The mutations in *SYNE1* associated with psychiatric disorders map to the part of the gene that spans the rat *Cpg2* homologous locus and the nearby 3' and 5' regions. In recent meta-analyses of GWAS data (Green et al., 2013a; Psychiatric, 2011; Xu et al., 2014), the markers in *SYNE1* that showed genome-wide statistically significant association with BD tag (via linkage disequilibrium) a missense SNP that alters the protein encoded by *CPG2*. Several other missense SNPs are in close proximity, and exome sequencing suggests additional polymorphisms in the *CPG2* coding region, which could potentially alter *CPG2* protein function.

The ASD associated mutations fall outside the *CPG2* region that we mapped in this study. While they are close enough (exon 9 and exon 60) that we initially considered the possibility that they would be contained within a human *CPG2* transcript longer than the rat homolog, our data suggest that other as yet unidentified transcripts are more likely to exist in these regions such as transcripts encoding the N-terminal nesprin isoforms identified in mammalian cell lines with suggested functions in the Golgi complex (Gough et al., 2003).

To further investigate the *CPG2* locus of *SYNE1* as the potential risk region for neuropsychiatric disorders, we used an integrated approach combining several complementary gene-mapping strategies, overcoming their individual limitations. Our data verified several human *CPG2* transcripts, confidently determining their 3' and 5' ends. We cloned a full-length human *CPG2* transcript as well as two shorter forms. The complex transcriptional pattern from the *SYNE1/CPG2* locus with several potential splice variants suggests multiple layers of transcriptional regulation. The shorter transcripts could potentially have regulatory functions at the mRNA level through competitive binding to microRNAs

or to the translational machinery. Notably, a large number of microRNAs have been identified that interfere with expression of synaptic proteins such as NMDA receptor and calcium/calmodulin-dependent kinase II (CaMKII), and might influence neuronal plasticity (Bredy et al., 2011). Such functions may be impaired in the case of specific *SYNE1* mutations/polymorphisms, either by regulation of transcript abundance or by alternative RNA splicing. Alternatively, truncated *CPG2*-like proteins could regulate *CPG2* function through competitive association with cellular binding partners as naturally occurring dominant negatives. While our inability to detect the two *CPG2* short forms in human brain extracts suggests the former possibility as more likely, we cannot exclude that this is due to limitations of our antibodies or to limited probe sensitivity.

The full-length human *CPG2* protein is expressed in CNS tissue with the highest expression in the neocortex, hippocampus and striatum, consistent with mRNA localization patterns of the rat *Cpg2* homolog (Cottrell et al., 2004). In the rat, the *CPG2* protein localizes primarily to the postsynaptic endocytic zone of excitatory synapses, where it is required for the internalization of glutamate receptors (Cottrell et al., 2004). Here we show the existence of human *CPG2* transcripts and their protein products in the human brain, and conserved human *CPG2* function in glutamate receptor internalization. Interestingly, several independent lines of evidence, including GWAS have implicated abnormal glutamatergic activity in the neuronal impairments affecting patients with BD (Ferreira et al., 2008; Nurnberger et al., 2014; Psychiatric, 2011; Sanacora et al., 2008). Notably, SNPs in *GRIA2*, which encode the GluA2 subunit of AMPARs, have been associated with time to recurrence of mood episodes in BD patients on lithium (Perlis et al., 2009). Studies have also shown differences in glutamate levels as well as glutamate receptor expression or function between individuals with mood disorders and control subjects (Beneyto and Meador-Woodruff, 2006; McCullumsmith et al., 2007; Meador-Woodruff et al., 2001; Nudmamud-Thanoi and Reynolds, 2004; Scarr et al., 2003). The other major BD risk gene identified by GWAS (Cross-Disorder Group of the Psychiatric Genomics, 2013; Ferreira et al., 2008), *ANK3*, encodes the actin cytoskeleton-associated adaptor protein ankyrin-G, which is also implicated in glutamate receptor-mediated synaptic transmission as well as the maintenance of mature spine morphology (Smith et al., 2014). Furthermore, *Ank3* has been shown to regulate lithium-attenuated psychiatric-related behaviors in mice (Leussis et al., 2013).

Transcriptional mapping of the *SYNE1* gene, focusing on the *CPG2* region, is a first step towards clarifying which of its transcripts and protein products are likely to play a role in BD. Identification of a full-length *CPG2* transcript expressed in human brain encoding a protein that is functionally conserved between rat and human, provides a platform for future testing of missense SNPs identified by BD patient exome sequencing for a role in glutamatergic transmission.

## Funding

The work was supported by the Gail Steele Fund and the Jeffrey M. and Barbara Picower Foundation (EN); NIH grants RC2-MH089952 and P50-MH106933 (MG); and by grants from the Carlsberg Foundation 2013\_01\_0416, Lundbeck Foundation R151-2013-14457, and the Danish Council for Independent Research 4004-00046B (MR).

## Conflicts of interest

The authors declare no conflicts of interest.

## Acknowledgements

Human brain tissue samples were generously provided by the Massachusetts Alzheimer's Disease Research Center and from the Harvard Brain Tissue Resource Center. Furthermore, we acknowledge our colleagues at the Nedivi Laboratory for valuable editing of the manuscript.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.mcn.2015.12.007>.

## References

- Attali, R., Warwar, N., Israel, A., Gurt, I., McNally, E., Puckelwartz, M., Glick, B., Nevo, Y., Ben-Neriah, Z., Melki, J., 2009. Mutation of SYNE-1, encoding an essential component of the nuclear lamina, is responsible for autosomal recessive arthrogryposis. *Hum. Mol. Genet.* 18, 3462–3469.
- Banerjee, I., Zhang, J., Moore-Morris, T., Pfeiffer, E., Buchholz, K.S., Liu, A., Ouyang, K., Stroud, M.J., Gerace, L., Evans, S.M., et al., 2014. Targeted ablation of nesprin 1 and nesprin 2 from murine myocardium results in cardiomyopathy, altered nuclear morphology and inhibition of the biomechanical gene response. *PLoS Genet.* 10, e1004114.
- Beneyto, M., Meador-Woodruff, J.H., 2006. Lamina-specific abnormalities of AMPA receptor trafficking and signaling molecule transcripts in the prefrontal cortex in schizophrenia. *Synapse* 60, 585–598.
- Berrettini, W.H., 2000. Genetics of psychiatric disease. *Annu. Rev. Med.* 51, 465–479.
- Bredy, T.W., Lin, Q., Wei, W., Baker-Andersen, D., Mattick, J.S., 2011. MicroRNA regulation of neural plasticity and memory. *Neurobiol. Learn. Mem.* 96, 89–94.
- Cottrell, J.R., Borok, E., Horvath, T.L., Nedivi, E., 2004. CPG2: a brain- and synapse-specific protein that regulates the endocytosis of glutamate receptors. *Neuron* 44, 677–690.
- Craddock, N., Sklar, P., 2009. Genetics of bipolar disorder: successful start to a long journey. *Trends Genet.* 25, 99–105.
- Craddock, N., Sklar, P., 2013. Genetics of bipolar disorder. *Lancet* 381, 1654–1662.
- Cross-Disorder Group of the Psychiatric Genomics, C., 2013. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *Lancet* 381, 1371–1379.
- Ferreira, M.A., O'Donovan, M.C., Meng, Y.A., Jones, I.R., Ruderfer, D.M., Jones, L., Fan, J., Kirov, G., Perlis, R.H., Green, E.K., et al., 2008. Collaborative genome-wide association analysis supports a role for ANK3 and CACNA1C in bipolar disorder. *Nat. Genet.* 40, 1056–1058.
- Gough, L.L., Fan, J., Chu, S., Winnick, S., Beck, K.A., 2003. Golgi localization of Syne-1. *Mol. Biol. Cell* 14, 2410–2424.
- Green, E.K., Grozeva, D., Forty, L., Gordon-Smith, K., Russell, E., Farmer, A., Hamshere, M., Jones, I.R., Jones, L., McGuffin, P., et al., 2013a. Association at SYNE1 in both bipolar disorder and recurrent major depression. *Mol. Psychiatry* 18, 614–617.
- Green, E.K., Hamshere, M., Forty, L., Gordon-Smith, K., Fraser, C., Russell, E., Grozeva, D., Kirov, G., Holmans, P., Moran, J.L., et al., 2013b. Replication of bipolar disorder susceptibility alleles and identification of two novel genome-wide significant associations in a new bipolar disorder case-control sample. *Mol. Psychiatry* 18, 1302–1307.
- Gros-Louis, F., Dupre, N., Dion, P., Fox, M.A., Laurent, S., Verreault, S., Sanes, J.R., Bouchard, J.P., Rouleau, G.A., 2007. Mutations in SYNE1 lead to a newly discovered form of autosomal recessive cerebellar ataxia. *Nat. Genet.* 39, 80–85.
- Lee, J.C., Espeli, M., Anderson, C.A., Linterman, M.A., Pocock, J.M., Williams, N.J., Roberts, R., Viatte, S., Fu, B., Peshu, N., et al., 2013. Human SNP links differential outcomes in inflammatory and infectious disease to a FOXO3-regulated pathway. *Cell* 155, 57–69.
- Leussis, M.P., Berry-Scott, E.M., Saito, M., Jhuang, H., de Haan, G., Alkan, O., Luce, C.J., Madison, J.M., Sklar, P., Serre, T., et al., 2013. The ANK3 bipolar disorder gene regulates psychiatric-related behaviors that are modulated by lithium and stress. *Biol. Psychiatry* 73, 683–690.
- Loeblich, S., Nedivi, E., 2009. The function of activity-regulated genes in the nervous system. *Physiol. Rev.* 89, 1079–1103.
- Loeblich, S., Djukic, B., Tong, Z.J., Cottrell, J.R., Turrigiano, G.G., Nedivi, E., 2013. Regulation of glutamate receptor internalization by the spine cytoskeleton is mediated by its PKA-dependent association with CPG2. *Proc. Natl. Acad. Sci. U. S. A.* 110, E4548–E4556.
- Manolio, T.A., Collins, F.S., Cox, N.J., Goldstein, D.B., Hindorf, L.A., Hunter, D.J., McCarthy, M.I., Ramos, E.M., Cardon, L.R., Chakravarti, A., et al., 2009. Finding the missing heritability of complex diseases. *Nature* 461, 747–753.
- McCullumsmith, R.E., Kristiansen, L.V., Beneyto, M., Scarr, E., Dean, B., Meador-Woodruff, J.H., 2007. Decreased NR1, NR2A, and SAP102 transcript expression in the hippocampus in bipolar disorder. *Brain Res.* 1127, 108–118.
- Meador-Woodruff, J.H., Hogg Jr., A.J., Smith, R.E., 2001. Striatal ionotropic glutamate receptor expression in schizophrenia, bipolar disorder, and major depressive disorder. *Brain Res. Bull.* 55, 631–640.
- Nedivi, E., Hevroni, D., Naot, D., Israeli, D., Citri, Y., 1993. Numerous candidate plasticity-related genes revealed by differential cDNA cloning. *Nature* 363, 718–722.
- Nedivi, E., Fieldust, S., Theill, L.E., Hevron, D., 1996. A set of genes expressed in response to light in the adult cerebral cortex and regulated during development. *Proc. Natl. Acad. Sci. U. S. A.* 93, 2048–2053.
- Nejentsev, S., Walker, N., Riches, D., Egholm, M., Todd, J.A., 2009. Rare variants of IFIH1, a gene implicated in antiviral responses, protect against type 1 diabetes. *Science* 324, 387–389.
- Nudmamud-Thanoi, S., Reynolds, G.P., 2004. The NR1 subunit of the glutamate/NMDA receptor in the superior temporal cortex in schizophrenia and affective disorders. *Neurosci. Lett.* 372, 173–177.
- Nurnberger Jr., J.J., Koller, D.L., Jung, J., Edenberg, H.J., Foroud, T., Guella, I., Vawter, M.P., Kelsoe, J.R., Psychiatric Genomics Consortium Bipolar, G., 2014. Identification of pathways for bipolar disorder: a meta-analysis. *JAMA Psychiatry* 71, 657–664.
- O'Roak, B.J., Deriziotis, P., Lee, C., Vives, L., Schwartz, J.J., Girirajan, S., Karakoc, E., Mackenzie, A.P., Ng, S.B., Baker, C., et al., 2011. Exome sequencing in sporadic autism spectrum disorders identifies severe de novo mutations. *Nat. Genet.* 43, 585–589.
- Perlis, R.H., Smoller, J.W., Ferreira, M.A., McQuillin, A., Bass, N., Lawrence, J., Sachs, G.S., Nimgaonkar, V., Scolnick, E.M., Gurling, H., et al., 2009. A genome-wide association study of response to lithium for prevention of recurrence in bipolar disorder. *Am. J. Psychiatry* 166, 718–725.
- Psychiatric, G.C.B.D.W.G., 2011. Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. *Nat. Genet.* 43, 977–983.
- Puckelwartz, M.J., Kessler, E., Zhang, Y., Hodzic, D., Randles, K.N., Morris, G., Earley, J.U., Hadhazy, M., Holaska, J.M., Mewborn, S.K., et al., 2009. Disruption of nesprin-1 produces an Emery Dreifuss muscular dystrophy-like phenotype in mice. *Hum. Mol. Genet.* 18, 607–620.
- Rajgor, D., Shanahan, C.M., 2013. Nesprins: from the nuclear envelope and beyond. *Expert Rev. Mol. Med.* 15, e5.
- Rajgor, D., Mellad, J.A., Autore, F., Zhang, Q., Shanahan, C.M., 2012. Multiple novel nesprin-1 and nesprin-2 variants act as versatile tissue-specific intracellular scaffolds. *PLoS One* 7, e40098.
- Rathje, M., Fang, H., Bachman, J.L., Anggono, V., Gether, U., Haganir, R.L., Madsen, K.L., 2013. AMPA receptor pHluorin-GluA2 reports NMDA receptor-induced intracellular acidification in hippocampal neurons. *Proc. Natl. Acad. Sci. U. S. A.* 110, 14426–14431.
- Razafsky, D., Hodzic, D., 2015. A variant of Nesprin1 giant devoid of KASH domain underlies the molecular etiology of autosomal recessive cerebellar ataxia type I. *Neurobiol. Dis.* 78, 57–67.
- Sanacora, G., Zarate, C.A., Krystal, J.H., Manji, H.K., 2008. Targeting the glutamatergic system to develop novel, improved therapeutics for mood disorders. *Nat. Rev. Drug Discov.* 7, 426–437.
- Scarr, E., Pavey, G., Sundram, S., MacKinnon, A., Dean, B., 2003. Decreased hippocampal NMDA, but not kainate or AMPA receptors in bipolar disorder. *Bipolar Disord.* 5, 257–264.
- Smith, K.R., Kopeikina, K.J., Fawcett-Patel, J.M., Leaderbrand, K., Gao, R., Schurmann, B., Myczek, K., Radulovic, J., Swanson, G.T., Penzes, P., 2014. Psychiatric risk factor ANK3/ankyrin-G nanodomains regulate the structure and function of glutamatergic synapses. *Neuron* 84, 399–415.
- Xu, W., Cohen-Woods, S., Chen, Q., Noor, A., Knight, J., Hosang, G., Parikh, S.V., De Luca, V., Tozzi, F., Muglia, P., et al., 2014. Genome-wide association study of bipolar disorder in Canadian and UK populations corroborates disease loci including SYNE1 and CSMD1. *BMC Med. Genet.* 15, 2.
- Yu, T.W., Chahrour, M.H., Coulter, M.E., Jiralerspong, S., Okamura-Ikeda, K., Ataman, B., Schmitz-Abe, K., Harmin, D.A., Adli, M., Malik, A.N., et al., 2013. Using whole-exome sequencing to identify inherited causes of autism. *Neuron* 77, 259–273.
- Zhang, Q., Skepper, J.N., Yang, F., Davies, J.D., Hegyi, L., Roberts, R.G., Weissberg, P.L., Ellis, J.A., Shanahan, C.M., 2001. Nesprins: a novel family of spectrin-repeat-containing proteins that localize to the nuclear membrane in multiple tissues. *J. Cell Sci.* 114, 4485–4498.
- Zhang, Q., Ragnauth, C.D., Skepper, J.N., Worth, N.F., Warren, D.T., Roberts, R.G., Weissberg, P.L., Ellis, J.A., Shanahan, C.M., 2005. Nesprin-2 is a multi-isoform protein that binds lamin and emerin at the nuclear envelope and forms a subcellular network in skeletal muscle. *J. Cell Sci.* 118, 673–687.
- Zhang, Q., Bethmann, C., Worth, N.F., Davies, J.D., Wasner, C., Feuer, A., Ragnauth, C.D., Yi, Q., Mellad, J.A., Warren, D.T., et al., 2007a. Nesprin-1 and -2 are involved in the pathogenesis of Emery Dreifuss muscular dystrophy and are critical for nuclear envelope integrity. *Hum. Mol. Genet.* 16, 2816–2833.
- Zhang, X., Xu, R., Zhu, B., Yang, X., Ding, X., Duan, S., Xu, T., Zhuang, Y., Han, M., 2007b. Syne-1 and Syne-2 play crucial roles in myonuclear anchorage and motor neuron innervation. *Development* 134, 901–908.
- Zhang, J., Felder, A., Liu, Y., Guo, L.T., Lange, S., Dalton, N.D., Gu, Y., Peterson, K.L., Mizisin, A.P., Shelton, G.D., et al., 2010. Nesprin 1 is critical for nuclear positioning and anchorage. *Hum. Mol. Genet.* 19, 329–341.
- Zylka, M.J., Simon, J.M., Philpot, B.D., 2015. Gene length matters in neurons. *Neuron* 86, 353–355.