Large donor CRISPR for whole-CDS replacement of cell adhesion molecule LRRTM2

Authors and Affiliations:
Stephanie L Pollitt[1,2], Aaron D Levy[1,2], Michael C Anderson[1,2,3], Thomas A Blanpied[1,2,3*]

[1] Department of Physiology, University of Maryland School of Medicine, Baltimore, MD, USA

[2] University of Maryland-Medicine Institute for Neuroscience Discovery (UM-MIND), University of Maryland School of Medicine, Baltimore, MD, USA

[3] Program in Neuroscience, University of Maryland School of Medicine, Baltimore, MD, USA
[*]corresponding author

ORCID SLP: 0000-0002-9297-6275
ORCID ADL: 0000-0002-5856-8294
ORCID MCA: 0000-0001-8746-4681
ORCID TAB: 0000-0003-4957-557X

Contributions:
The project was conceptualized by SLP and experiments designed by SLP and TAB. TKIT CRISPR constructs were designed and created by SLP and ADL. Experimental conditions and reagents were optimized by SLP and MCA. Data were collected and analyzed by SLP. TAB supervised the study. Figures were prepared and the manuscript written by SLP with constructive review and editing by TAB, MCA, and ADL.

Funding:
SLP – F32MH130106
MCA – F31MH12428330
TAB – R37MH080046, R01MH119826

Abstract

The cell adhesion molecule LRRTM2 is crucial for synapse development and function. However, our understanding of its endogenous trafficking has been limited due to difficulties in manipulating its coding sequence (CDS) using standard genome editing techniques. We instead replaced the whole LRRTM2 CDS by adapting the recent CRISPR method Targeted Knock-in using Two Guides (TKIT), enabling complete control of LRRTM2. In primary rat hippocampal cultures, N-terminally tagged, endogenous LRRTM2 was found in 80% of synapses, and synaptic LRRTM2 content correlated with PSD-95 and AMPAR levels. LRRTM2 was also enriched with AMPARs outside synapses, demonstrating the sensitivity of this method to detect relevant new biology. Finally, we leveraged total genomic control to increase the synaptic levels of LRRTM2 via simultaneous mutation of its C-terminal domain, which did not correspondingly increase AMPAR enrichment. The coding region of thousands of genes span lengths suitable for whole-CDS replacement, suggesting this simple approach will enable straightforward structure-function analysis in diverse cellular systems.

Introduction

Synaptic organizing molecules, or cell adhesion molecules (CAMs), are unique transmembrane proteins that reach across the synaptic cleft, where they play critical roles in both cis- and transsynaptic biological processes from synaptogenesis to receptor trafficking and plasticity. The trafficking and expression of CAMs are tightly regulated, and mutations in CAM genes can severely disrupt neuronal function and are linked to neurological disease[1–3]. Accordingly, alterations of CAM expression levels, such as due to the use of exogenous expression for visualization and mutational analysis, can alter neuronal function and cloud interpretations of CAM biology. It is therefore necessary to measure and manipulate the endogenous protein itself to advance our understanding of CAMs. However, endogenous structure-function studies have been challenging, particularly because one critical approach is to mutate CAM sequences at multiple locations simultaneously. For example, N-terminal, extracellular tagging is ideal for most CAMs to visualize surface-expressed protein, but many critical protein interaction sites are quite distant from the N-terminus or even intracellular, as CAM C-tails play important roles in their anchoring and synaptic signaling. Standard genome editing techniques have largely been insufficient to make these changes simultaneously.

One example of a CAM for which new tools are needed is the synaptic CAM Leucine-Rich Repeat Transmembrane Protein 2 (LRRTM2). LRRTM2 resides in the post-synaptic membrane, where it binds to PSD-95 at its C-terminus and interacts trans-synaptically with presynaptic neurexins[4,5]. LRRTM2 exerts considerable power over synaptic function at mature synapses, largely as a result of its influence over AMPA receptor (AMPAR) anchoring and abundance at synapses. Knockdown of LRRTM2 affects basal AMPAR synaptic enrichment and synaptic strength[4], and dual knockdown or knockout of LRRTM2 and its sister protein LRRTM1 abolishes long term potentiation (LTP) in hippocampal CA1 neurons [6,7]. Furthermore, enzymatic cleavage of the LRRTM2 extracellular domain in a knockdown-rescue context disorganizes AMPAR subsynaptic distribution within minutes, and reduces total synaptic AMPAR content over the following hours[8]. LRRTM2 also plays a dose-dependent role in synaptogenesis, where over- or under- expression leads to corresponding changes in synapse density[4,9]. These results highlight the central role LRRTM2 plays in regulating synaptic strength at mature synapses and motivate detailed analysis of its molecular mechanisms.

As is true for many proteins, our understanding of LRRTM2 cellular functions has relied on genetic knockout and overexpression or antibody-based immunocytochemistry. Each of these approaches has considerable shortcomings. In particular, methods that manipulate LRRTM2 expression levels dramatically alter synaptic development and can engage compensatory mechanisms from related CAMs[4,9]. The use of knockdown-rescue

can reduce overexpression phenotypes[4,8], but expression via exogenous promoters can nevertheless cause diverse effects that are difficult to adequately monitor. While some studies have utilized antibody detection of endogenous LRRTM2, the low signal-to-noise ratio and lack of possibility to manipulate the protein sequence limits their scope[10,11]. Knock-in mouse models could bridge this gap; however LRRTM2 has numerous binding partners, and making a separate mouse model for each mutation is expensive and inefficient. Disappointingly, traditional CRISPR knock-in approaches to tag the LRRTM2 N-terminus have proven infeasible due to a lack of suitable PAM sites in the protein encoding region, and N-terminal knock-ins could threaten the integrity of the reading frame due to possible indels. C-terminal tagging is less than ideal due to inability to accurately identify protein present on the cell surface, and likely disrupts the PDZ-binding motif. Recently, the CRISPR approach TKIT (Targeted Knock-In with Two guides) was used to replace a protein-coding exon using two intron-localized guide RNA sequences[12]. This approach allowed N-terminal tagging of proteins without shifting the reading frame, as any Cas9-induced indels occur in intronic regions. We considered whether this strategy could be expanded and adapted to permit tagging and mutagenesis anywhere within a protein where the entire protein coding sequence (CDS) can be replaced. To test the feasibility of this approach while investigating LRRTM2 trafficking, we have applied whole-CDS replacement TKIT to the rat *Lrrtm2* gene.

In this work, we demonstrate whole-CDS replacement in neurons as well as the simultaneous tagging and mutagenesis of the *Lrrtm2* coding region. We utilized endogenously tagged LRRTM2 to measure its synaptic enrichment as well as its relationship with AMPAR enrichment at synapses. To enable rapid identification and measurement of knock-in cells, we also developed a strategy to express a marker conditional on successful knock-in, enabling new and flexible experimental designs. Finally, we demonstrated the power of this technique by combining N-terminal tagging with distal point mutations to increase synaptic surface enrichment of LRRTM2. Together, our findings provide context for our understanding of LRRTM2 function as well as a novel method for future structure-function studies in neurons and other post-mitotic cells.

Results

Two-guide CRISPR approach TKIT can be adapted for whole-CDS replacement of LRRTM2

To decipher the trafficking patterns of the key synaptic cell adhesion molecule LRRTM2, we designed an approach to simultaneously measure and manipulate the protein while maintaining its endogenous transcriptional regulation and expression level. HITI is a common CRISPR genome editing method used in post-mitotic cells such as

130 neurons that allows for DNA insertions that could be useful for tagging native LRRTM2[13].
131 However, HITI is limited to insertion at single sites, which for LRRTM2 would only allow
132 either an N-terminal tag or protein mutagenesis, not both. In addition, HITI targets PAMs
133 to make double stranded breaks within the coding region, and therefore may generate
134 indels that alter the reading frame unpredictably (Figure 1A, left). Therefore, we
135 considered whether we could edit the entire mature LRRTM2 CDS by adapting the exon-
136 replacement technique Targeted Knock-In with Two guides (TKIT)[12]. Instead of targeting
137 a single PAM site within the coding region, TKIT utilizes two guide RNAs targeting PAMs
138 in the non-coding regions flanking an exon, which directs Cas9 to excise the exon entirely
139 and enables its replacement with a synthetic exon (Figure 1A, middle). This has the
140 advantage of improving knock-in efficiency since multiple guide RNAs and PAM sites can
141 be tested throughout the intronic sequences, and it reduces the effects of indels as they
142 will occur outside the coding region. TKIT has previously been used to replace the 5'-
143 most coding exon of diverse neurotransmitter receptors with a synthetic version that
144 includes an N-terminal fluorescent tag[12]. We wondered whether this approach could be
145 modified such that the entire coding region of a gene, in our case LRRTM2, could be
146 replaced instead (Figure 1A, right). Notably, nearly the entire LRRTM2 protein is encoded
147 in a single exon, with the 5' UTR and only 4 bases of signal peptide in exon 1, and the
148 rest of the coding region, plus the 3' UTR, in exon 2 (Figure 1B). This genomic
149 organization offered the attractive possibility of replacing the entire protein-coding
150 sequence of LRRTM2 with an entirely customizable one using TKIT. Therefore, we
151 modified the TKIT approach to replace much of the second exon, including all the coding
152 sequence contained therein. Given that the signal peptide is cleaved co-translationally,
153 this approach would allow us to replace the entire mature LRRTM2 protein with a donor
154 sequence of our choosing.
155 To replace the LRRTM2 coding sequence using TKIT, we designed a 5' guide RNA
156 (guide 1) targeting a PAM site within intron 1 of *Rattus norvegicus Lrrtm2* and a 3' guide
157 RNA (guide 2) targeting a PAM site after the *Lrrtm2* stop codon within the 3'UTR, resulting
158 in a large span of replacement that exceeds those of previous efforts[12,14,15]. The
159 replacement donor sequence was engineered using rat genomic DNA and subcloned to
160 include an epitope tag between the signal peptide and LRRTM2 N-terminal regions. The
161 donor was also flanked by flipped and reversed guide sequences to allow reversed
162 integration to be fixed, as used in both HITI and TKIT[12,13] (Figure 1B). We tested several
163 N-terminal tags for both specific labeling intensity and non-specific background from the
164 affinity reagent using exogenous LRRTM2 expression in HEK cells (Supplemental Figure
165 1). We found that the ALFA nanobody[16] displayed lower non-specific signal than an HA
166 antibody plus secondary antibody. Further, we were able to increase specific staining of
167 the ALFA epitope by adding a triple ALFA tag without interstitial linkers, which had
168 approximately 50% higher fluorescence intensity than either the single ALFA tag or a
169 triple ALFA tag with linkers. Therefore, we utilized a triple ALFA tag without linkers in the

170 donor sequence for fluorescent labeling (note that for genomic PCR and sequencing
171 experiments, a triple HA tag was used instead for convenience). As transfection efficiency
172 of cultured neurons is low, we delivered the necessary DNA with two lentiviruses, one to
173 express the guides and donor sequence and a second to express Cas9.
174 We first validated the success of LRRTM2 whole-CDS replacement by infecting
175 dissociated embryonic rat hippocampal neurons with either Cas9 and LRRTM2
176 guides/donor lentiviruses or with guides/donor virus only as a control, then sequenced
177 genomic DNA after three weeks in culture. Briefly, we utilized PCR to isolate the *Lrrtm2*
178 gene and amplified knock-in-positive alleles using sequential PCRs (Figure 1B, bottom)
179 and sequenced the resulting amplicons via Sanger sequencing. Positive sequences for
180 the 3xHA tag were identified in the knock-in condition (Figure 1C, middle) but not in the
181 control condition lacking Cas9 (Figure 1C, bottom), indicating that the tag was
182 successfully incorporated into the correct gene locus and that knock-in was successful.
183 Note that we did not introduce indels at the signal peptide-tag junction or at the tag-protein
184 junction, as this region was contained entirely within the engineered donor sequence as
185 is true for TKIT. We next evaluated whether the knocked-in 3xALFA tag would allow us
186 to visualize the endogenous expression of LRRTM2. We performed surface fluorescent
187 labeling of 3xALFA-LRRTM2 on live, dissociated hippocampal cultured neurons infected
188 three weeks prior with the knock-in constructs, then fixed the cells and additionally
189 immunolabeled for PSD-95 and the presynaptic scaffold protein Bassoon to visualize
190 synapses. Fluorescent confocal imaging showed isolated cells with anti-ALFA labeling in
191 discrete puncta along the cell body and dendrites (Figure 1D). Close inspection revealed
192 these puncta typically occurred at synapses, as indicated by colocalization with both
193 Bassoon and PSD-95, consistent with previous studies of LRRTM2[4,10] (Figure 1E,
194 especially arrowheads). In line with previous TKIT-based knock-ins, we commonly
195 observed 20 to 30 ALFA-stained cells per coverslip with this method; note that this may
196 be an underestimate of the knock-in efficiency as some cells may be knocked-in but not
197 express LRRTM2, and some cells may not be co-infected by both viruses. These results
198 together demonstrate that our knock-in was successful and could be used to visualize
199 endogenous LRRTM2 in cultured neurons.
200 The previous best practice for visualizing LRRTM2 has been knockdown-rescue
201 (KDR), which is frequently assumed to minimize overexpression due to the knockdown
202 of endogenous expression. However, this method still typically relies on exogenous or
203 unregulated promoters, which often express proteins at higher than endogenous levels.
204 To compare knock-in LRRTM2 expression levels with those following KDR, we performed
205 live surface ALFA labeling on cells infected with knock-in constructs and on cells
206 transfected with a knockdown-rescue 3xALFA-LRRTM2 construct. As expected, ALFA
207 staining was clearly far higher in the KDR condition than the knock-in (Figure 1F). Further,
208 while the expression pattern of both was largely synaptic, LRRTM2 signal in the KDR
209 condition showed higher and more consistent levels of LRRTM2 at each spine (Figure

1F, line scan, dark arrows), and also showed higher ALFA labeling along the dendrite outside of dendritic spines (Figure 1F, line scan, white arrows). This indicates that overexpression of LRRTM2, even following knockdown, can alter its surface and subcellular trafficking, highlighting the importance of labeling and measuring endogenous protein.

Rapid identification of whole-CDS knock-in cells

As is true for many endogenously labeled proteins, the low endogenous expression levels of LRRTM2 made visualizing knock-in cells or tracing their morphology for analysis difficult. We therefore leveraged our whole-CDS replacement system to design a knock-in conditional marker. The conditional effect was achieved by adding an Internal Ribosome Entry Sequence (IRES) and Cre gene after the stop codon in the LRRTM2 donor containing the 3xALFA tag. This donor could allow for expression of a FLExible marker protein[17] to be conditional upon successful knock-in of 3xALFA-LRRTM2 (Figure 2A). To test the utility of this approach, we infected neurons with three lentiviruses expressing the 3xALFA-LRRTM2-IRES-Cre donor and guide RNAs, Cas9, and FLEx-mTagBFP2, expecting that when the 3xALFA-LRRTM2-IRES-Cre donor is knocked into the *Lrrtm2* locus and the mRNA translated, Cre will also be expressed and able to induce FLEx-mTagBFP2 expression (Figure 2A, bottom). This strategy successfully produced cells expressing both knock-in 3xALFA-LRRTM2 and knock-in-dependent mTagBFP2 cell fill, facilitating rapid visual identification on the microscope (Figure 2B). Notably, the 3xALFA tag and IRES-Cre sequences are 1449 bp away from one another on the donor sequence, a simultaneous knock-in impossible to make by other methods. The donor used was also 3925 bp long, over two times larger than the largest TKIT donor previously reported[12] and among the largest we are aware of successfully being knocked in using NHEJ-based CRISPR methods. We observed a fraction of neurons that were positive for mTagBFP2 without detectable surface ALFA labeling (16%), presumably reflecting cell types with successful knock-in that do not express or only transiently expressed LRRTM2. Cells in this category were excluded from further analysis. A small number of neurons (8%) were positive for surface ALFA staining without detectible mTagBFP2 expression, potentially due to a lack of coinfection by all three viruses. Nevertheless, expression of the marker protein was routinely high enough to rapidly identify presumed knock-in cells for image acquisition. An important advantage of this approach is the potential to use alternative conditional reporters. To illustrate this, we utilized a FLEx-IRES-EGFP lentivirus in place of the FLEx-mTagBFP2 and observed EGFP-positive knock-in cells (Figure 2C). Conceivably, any FLEx fluorescent protein, sensor, or optogenetic tool could be used to accommodate numerous experimental approaches, demonstrating the utility of our whole-CDS control.

249 ***Endogenous LRRTM2 labeling enables investigation of LRRTM2 trafficking and***
250 ***synaptic enrichment***
251 We next set out to deploy our LRRTM2 knock-in and conditional reporter system to
252 characterize LRRTM2 expression and synaptic trafficking at endogenous levels. We first
253 quantified the average synaptic LRRTM2 expression level per cell and found a wide range
254 of intensities across cells (Figure 3A). This suggests that LRRTM2 expression varies
255 within the transcriptomic profile of neurons in these CA1-enriched hippocampal cultures,
256 indicating possible cell-intrinsic control of LRRTM2 expression. We considered that some
257 of the cells with higher LRRTM2 intensity might be homozygous for the 3xALFA-LRRTM2
258 knock-in, and those with lower expression might be heterozygous. However, we observed
259 a smooth, rather than bimodal, distribution of average synaptic intensity (Figure 3B),
260 consistent with regulation of LRRTM2 expression levels by activity history or other cell-
261 specific mechanism.
262 LRRTM2 is synaptically localized and plays a key role in synaptic function[4,6,7].
263 However, our observation that LRRTM2 puncta were more discrete and variable with our
264 knock-in than with KDR due to overexpression artifacts (Figure 1F) indicated its
265 endogenous distribution may be more nuanced. We therefore investigated in detail the
266 variability of endogenous LRRTM2 content at individual synapses in 3xALFA-LRRTM2
267 knock-in neurons, immunolabelled for PSD-95 with a fluorescently labeled nanobody
268 (Figure 3C). Despite prominent colocalization at synapses (Figure 3C, white arrows), we
269 observed that the quantity of LRRTM2 varied substantially between synapses (coefficient
270 of variation 1.514) and that, in fact, not every synapse along dendrites from knock-in cells
271 contained LRRTM2 (Figure 3C, blue arrow). To better understand the synaptic variation
272 of LRRTM2 content while accounting for cellular variability of LRRTM2 expression levels
273 (Figure 3A), we measured LRRTM2 intensity within synapses and normalized for variable
274 cellular expression by calculating cell-based intensity Z-scores for each synapse. We
275 identified synaptic puncta as PSD-95 regions of interest (ROIs) using a semi-automated
276 synapse detection analysis[18,19] of images from 54 knock-in neurons across 3 culture
277 replicates and calculated Z-scores for each synapse. We measured the area as well as
278 the fluorescence intensity of each labeled protein within these synapse ROIs, and
279 normalized for their variable cellular expression by calculating cell-based intensity Z-
280 scores for each synapse using the cellular average and standard deviation for LRRTM2
281 or PSD-95, respectively (Figure 3D). This analysis clearly showed that synaptic LRRTM2
282 content varied considerably within individual neurons, even when normalized for cellular
283 variability.
284 To define a synapse as lacking LRRTM2, we sought to establish a baseline by
285 examining the off-target ALFA staining. This was done by measuring ALFA intensity at
286 neighboring synapses from non-knock-in neurons, reflecting non-specific staining and
287 noise, which as expected had Z-scores near 0 (Figure 3E, Neighboring Synapses). To
288 compare these two distributions, we recalculated the Z-scores of knock-in LRRTM2

intensity using the mean and standard deviation of ALFA signal in neighboring PSDs, and found the synapses along the knock-in dendrite were, as expected, far brighter in ALFA signal than the neighboring background staining. (Figure 3E). We then used the 95th percentile of non-knock-in synapse ALFA staining z-scores to delineate a cutoff between LRRTM2-containing and LRRTM2-lacking synapses within the knock-in dataset. Using this cutoff, approximately 80% of the synapses along knock-in dendrites contained LRRTM2 (Figure 3F). This is slightly higher than a previous estimate utilizing LRRTM2 antibody staining in DIV13-15 mouse hippocampal cultures[11], which could represent a developmental or species difference, but also could be a reflection of our ability to more precisely estimate background noise from cells that unequivocally lacked the knock-in and therefore accurately identify true signal.

Our observation of a synapse subtype lacking LRRTM2 raises additional questions about what other differences these synapses exhibit relative to their LRRTM2-containing neighbors. Given that LRRTM2 plays a role in synaptic plasticity and binds to PSD-95, one explanation of its synaptic variability is that synapses with higher LRRTM2 levels might be larger and contain more PSD-95. Indeed, we found that synapses with LRRTM2 exhibited larger PSDs than those without (Wilcox test: $p < 2.2e-16$, r-statistic: -0.142; Figure 3G). Further, within the population of LRRTM2-containing synapses, larger synapses contained more LRRTM2 protein (slope: 7.061, R^2: 0.226; Figure 3H). Our interpretation is further supported by examining PSD-95 content – synapses with LRRTM2 contained more PSD-95 than those without LRRTM2 (Wilcox: $p < 2.2e-16$, r-statistic: -0.149; Figure 3I), and PSD-95 and LRRTM2 intensities were positively correlated (Slope: 0.442, R-squared: 0.183; Figure 3J). These data indicate that the subcellular trafficking of LRRTM2 to synapses is correlated with both synapse size and the amount of PSD-95 present. While previous studies have indicated that overexpressed LRRTM2 does not require PSD-95 binding to be enriched at synapses[10], our data indicates that nonetheless LRRTM2 trafficking scales with these key markers of synaptic strength.

Synaptic AMPA receptor content scales with LRRTM2 content

Previous studies utilizing LRRTM1/2 KDR or LRRTM1/2 KO replacement with LRRTM2 have established that LRRTM2 stabilizes AMPARs in synapses at baseline and anchors new AMPARs after long-term potentiation[6,7]. Despite these findings, it has been difficult to disentangle how AMPAR trafficking compares with endogenous expression levels of LRRTM2. Furthermore, our ability to visualize endogenous LRRTM2 and identify synapses lacking it gives us the opportunity to compare AMPAR content between these two groups of synapses. We hypothesized that synapses lacking endogenous LRRTM2 would also have reduced AMPAR content relative to synapses with LRRTM2. To examine the relationship between the synaptic enrichment of LRRTM2 and AMPARs, we immunolabeled surface AMPA receptors simultaneously with ALFA-LRRTM2, fixed and

329 stained the neurons for PSD-95 to identify synaptic puncta, and measured synaptic
330 LRRTM2 and AMPAR intensity from the same synapses (Figure 4A). As in Figure 3, we
331 normalized both the LRRTM2 and AMPAR raw integrated density signals using Z-scores
332 based on respective means in each cell, then assessed synapses along knock-in
333 dendrites for the presence or absence of LRRTM2. Consistent with the role of LRRTM2
334 in AMPAR anchoring, LRRTM2-lacking synapses contained less surface AMPAR signal
335 than synapses with LRRTM2 (Wilcox: $p < 2.2e-16$, r-statistic: -0.136; Figure 4B). While
336 LRRTM2-lacking synapses still exhibited AMPAR staining, this is likely due to stabilization
337 by other mechanisms. When LRRTM2 was present, synapses with more LRRTM2 also
338 had more AMPAR content on the surface (Slope: 0.3513, R-squared 0.1326; Figure 4C).
339 While these results do not support a model where endogenous LRRTM2 is solely
340 responsible for synaptic AMPAR anchoring, the correlation between AMPAR content and
341 LRRTM2 content supports the hypothesis that LRRTM2 plays a strong role in control of
342 AMPAR trafficking[6,7].
343 Surprisingly, close examination of these triply-labeled knock-in cells revealed
344 occasional LRRTM2 puncta that fell outside of PSD-95-demarcated synapses (Figure 4D,
345 red vs white arrows). It is well documented that extrasynaptic AMPAR pools play a role
346 in regulating synapse strength, particularly during long-term potentiation[20,21]. Given the
347 role of LRRTM2 in LTP and AMPAR stabilization, we therefore wondered whether these
348 LRRTM2 puncta outside of PSD-95 might also contain AMPARs. To address this, we
349 selected isolated dendrites from 23 knock-in cells across three culture replicates and
350 identified puncta in each channel (surface AMPARs, PSD-95, and 3xALFA-LRRTM2) as
351 ROIs for measurement. As the semi-automated detection was trained to detect puncta
352 that are sized and shaped like synapses and ignore nonsynaptic labeling, these ROIs
353 were hand-selected. We then measured the raw integrated density within these ROIs in
354 all three channels, and qualitatively labeled each as positive or negative for the other
355 labeled proteins based on its intensity, where intensities above the first local minimum of
356 the distribution were described as “positive” for that corresponding protein (Figure 4E,
357 left). Consistent with our previous automated analysis, when the hand-selected PSD-95
358 puncta were pooled together, we found similar numbers of LRRTM2- positive and -
359 negative synapses (by hand: 79.7% positive and 20.2% negative; automated: 80.5%
360 positive and 19.5% negative), validating our approaches. We then broke down the non-
361 PSD-95-containing puncta into groups based on their protein expression (Figure 4E,
362 right). Most such puncta contained both LRRTM2 and AMPARs together (52%), more
363 than the puncta with either protein on its own (29% AMPAR only and 19% LRRTM2 only).
364 The existence of LRRTM2/AMPAR puncta outside PSD-95 synapses could be partially
365 explained to be synapses that do not contain PSD-95 and instead have another scaffold
366 such as SAP-102 or PSD-93[36,37]. However, given that such synapses are relatively rare
367 in mature hippocampal neurons, it’s likely that at least some of them are genuinely
368 extrasynaptic. This raises the exciting possibility that along with its known localization

369 within synapses, LRRTM2 could co-diffuse with extrasynaptic AMPARs, and possibly play
370 a role in their trafficking there.
371
372 ***Simultaneous mutagenesis and tagging enabled by whole-CDS replacement***
373 ***permits analysis of mutation-induced changes in LRRTM2 surface expression***
374 A major advantage of whole-CDS replacement is the potential it offers to easily edit a
375 protein of interest at multiple sequence points, as we have already demonstrated by
376 adding IRES-Cre at the C-terminus simultaneously with the N-terminal 3xALFA tag. To
377 demonstrate the power of whole-CDS replacement of LRRTM2, we modified our synthetic
378 donor sequence to include, along with the IRES-Cre, a functional mutation near the C-
379 terminus expected to alter LRRTM2 trafficking. Previous work has shown that surface
380 enrichment of exogenously expressed LRRTM2 can be manipulated through two point
381 mutations in the C-terminal domain, Y501A and C504A[22,23], which are hypothesized to
382 alter the membrane-targeting trafficking mechanisms of LRRTM2[23]. Termed "YACA",
383 these mutations induce a 400% increase in surface LRRTM2 levels in a knockdown-
384 rescue context[22]; however as we established, the KDR protein trafficking pattern may not
385 be the same as endogenous. Our whole-CDS replacement model offers the ideal platform
386 on which to test this. We introduced the YACA mutations into our donor construct to create
387 a 3xALFA-LRRTM2-YACA-IRES-Cre knock-in (Figure 5A), infected neurons with
388 lentivirus to generate wild type or YACA knock-ins, then measured AMPAR and LRRTM2
389 content in PSD-95-labeled synapses as above (54 wild type and 64 YACA neurons across
390 3 culture replicates; Figure 5B). On a per cell basis, average synaptic LRRTM2-YACA
391 content was 24.6% higher than wildtype (Wilcox: $p < 0.005$, r-statistic = 0.280; Figure 5C).
392 When we examined the per synapse distribution of LRRTM2 intensities between the two
393 conditions, we found a similar trend where the distribution of YACA-containing synapses
394 was right-shifted relative to wildtype (Wilcox: $p < 2.2e-16$; r-statistic = 0.102; Figure 5D).
395 Therefore, while the YACA mutations did increase endogenous LRRTM2 surface
396 trafficking as expected, the percent change was dramatically smaller than that observed
397 in a KDR model. These findings reinforce the value of our whole-CDS replacement
398 approach and demonstrate the power of simultaneous genomic editing at multiple sites
399 within a protein.
400 Previous studies have established that removing LRRTM2 reduces synaptic AMPAR
401 content[6–8], and our data indicate that AMPAR content scales with LRRTM2 content at
402 synapses. While we have shown that the YACA mutations moderately increase LRRTM2
403 content, it is unclear whether this small change would follow the pattern in our
404 correlational data and drive an increase in AMPAR content. We therefore measured
405 synaptic AMPAR staining in each condition and found that the YACA mutations did not
406 have a positive effect on synaptic AMPAR content (Wilcox: $p = 1$, r-statistic = -0.035;
407 Figure 5E). This suggests that despite the correlation between LRRTM2 synaptic content

408 and AMPAR synaptic content, increasing the LRRTM2 synaptic content by 24.6% via
409 YACA is insufficient or unable to induce a similar increase in AMPAR content.
410
411 ***Applying whole-CDS replacement to other genes***
412 We conclude that whole-CDS replacement TKIT will be a valuable approach for
413 protein structure-function analysis. The structure of the rat *Lrrtm2* gene, being essentially
414 contained on a single exon, is undoubtedly highly advantageous for the technique, but
415 other genes with different exon/intron structures may still be amenable for the same
416 approach. We presume that the length of the CDS would have to be somewhat modest,
417 but note that the insertion sizes utilized here (3925 bp) are much greater than typically
418 amenable to knock-in via HITI[13]. In fact, the length rather than the content or functional
419 complexity of what is being replaced appears to be most critical for the success of the
420 knock-in. We explored the Santa Cruz Genome Browser for rat genes smaller than a
421 practical limit of 10,000 bases and found 39,459 coding sequences (including splice
422 variants) that had a distance from start to stop codon (CDS spanning region) that is the
423 same or shorter than that of *Lrrtm2* (1882 bp, though the span between our selected
424 guides is 2258 bp) (Figure 6A). Out of the rat genes whose CDS spanning regions are
425 under 10kb, approximately 42.6% are smaller than *Lrrtm2* (Figure 6B). While of course
426 many important genes have long genomic spans, this simple analysis shows that
427 numerous genes may be targetable for whole-CDS replacement. Our conservative
428 estimate reflects the fact that we have not systematically explored the upper range of
429 donor sizes amenable to our approach, suggesting that the actual number of suitable
430 genes may in fact be much higher. Furthermore, while guide selection is undoubtedly
431 critical to knock-in success, the flexibility of guide locations as adapted from TKIT permits
432 more successful knock-in of large donors. Overall, the power of whole-CDS replacement
433 will enable new research on endogenous protein localization, trafficking, and function in
434 neurons avoiding the constraints of overexpression and single-locus CRISPR editing
435 techniques.
436
437
438 ## Discussion
439
440 In this work, we have demonstrated whole-CDS replacement in neurons, and shown
441 its power to simultaneously tag and mutate a protein at widely separated points of its
442 sequence while maintaining native genetic regulation. The protein-coding portion of
443 thousands of genes span lengths suitable for whole-CDS replacement, suggesting that
444 this simple approach for total control over protein sequence will be a straightforward
445 method for structure-function analysis in diverse systems. We used the approach here to
446 identify new characteristics of expression and trafficking of the critical synaptic adhesion
447 molecule LRRTM2. LRRTM2 levels in cultured hippocampal neurons were unexpectedly

variable between neurons, and while its expression correlated with both PSD-95 and AMPAR content, the protein appeared absent from 20% of synapses. We also observed LRRTM2 outside of synapses at puncta that contained AMPARs but lacked PSD-95, suggesting a previously unappreciated role for LRRTM2 outside of PSD-95-containing synapses. Finally, utilizing our ability to manipulate the endogenous LRRTM2 sequence, we were able to increase synaptic LRRTM2 content without affecting AMPAR content, suggesting that this relationship may not be intrinsic or bi-directional.

Whole-CDS replacement is a general and flexible approach for simultaneous multi-site manipulations

While CDS replacement strategies have been demonstrated in dividing cells[24], this work is to our knowledge the first demonstration of whole-CDS replacement in neurons. A major advantage of this method is the ability to make multiple, simultaneous modifications to a gene at disparate locations along it in a single editing step. In theory, other CRISPR technologies that allow for DNA replacement could be used for whole-CDS replacement, though likely TKIT, which we relied upon here, is the most generalizable approach. For example, CDS replacement has been achieved in dividing cells for a coding region of similar size to LRRTM2[24]. However, this relied on homology-directed DNA repair, which makes it intractable in post-mitotic cell types such as neurons. This is due to the necessary reliance in post-mitotic cells on Non-Homologous End Joining (NHEJ), a method of donor DNA incorporation that religates blunt ends of DNA together. PRIME editing, which utilizes neither nonhomologous end joining nor homology-directed repair[25], allows for short substitutions (~100 bp[26]), but this is too small to replace the vast majority of genes. Designer exon approaches such as CRISPIE[27] could also be used to essentially “knock-in” an engineered CDS in place of a gene’s first exon and bypassing the later exons, though this would eliminate key intronic regulation and alternative splicing. Finally, while making two separate edits using CRISPR is possible[26,28], this remains extremely challenging and less efficient than whole-CDS replacement. Further, there is as yet no path for expanding to more than two editing sites, whereas whole-CDS replacement offers near-infinite flexibility within the large donor sequence. Due to the permanent nature of genetic modifications, this whole-CDS replacement approach could simplify the process of generating multiple stem cell lines or animals with multiple genetic mutations in the same gene, as there would be no need to select and validate new guides for each modification that is introduced. Given these factors and the proven efficiency and accuracy of TKIT[12], we are confident it presents a flexible tool to manipulate entire coding sequences in diverse systems.

The limit on donor size for TKIT is not known, though previous research would suggest that large donor sizes have generally been a struggle for CRISPR in neurons and other models reliant on NHEJ. Previous studies suggest NHEJ may have greater difficulty with larger DNA sizes[29], though it is also more efficient[13]. It is also conceivable that larger

488 replacements may be facilitated by fusing repair machinery onto Cas9, which has been
489 recently demonstrated to impact its efficiency[30]. Practically, donor size and efficiency
490 likely depend in part on the genetic and epigenetic context as well as practical limits such
491 as plasmid size or viral packaging and must therefore be determined on an individual
492 basis. The former can be addressed with TKIT due to the flexibility of guide locations
493 within the intronic/noncoding regions, which permits higher efficiency guides with lower
494 off-target effects to be selected. We were able to efficiently knock-in a large donor
495 sequence (3925 bp, including IRES-Cre), and while we have demonstrated that a
496 replacement of this size is feasible, replacement of larger regions may certainly also be
497 possible. Furthermore, we have identified a sizeable group of genes whose CDS-
498 spanning region size is smaller than this and thus appear amenable to whole-CDS
499 replacement. Note that for many genes of interest, the difficulty of applying the approach
500 will not be the size of the translated protein but the presence of introns so large as to be
501 impractical to supply in a donor. In such cases, it may still be possible to adapt this
502 technique to replace regions on one side or the other of a large intron, including multiple
503 small, contiguous exons and introns, without modifying the entire CDS. This approach
504 would largely preserve the regulation intrinsic to intronic sequences such as transcription
505 and alternative splicing.
506
507 ***Whole-CDS replacement permits versatile tagging and reporter strategies***
508 Ready identification of knock-in cells is key to experimental performance in many
509 techniques, including microscopy and flow cytometry. However, knocked-in tags can be
510 difficult to visualize because endogenous protein expression is often far lower than typical
511 overexpression, and reliance on abundant target protein translation for tag detection is a
512 potential source of bias. While empirical determination of the brightest protein tags can
513 help improve utility, a knock-in conditional marker independent of the target protein's
514 expression level is highly useful, and essential in many cases. By integrating existing Cre-
515 FLEx conditional marker technologies into non-coding regions simultaneously with our
516 modifications to the LRRTM2 coding sequence, we were able to identify, tag, and
517 manipulate knock-in cells in a single step. Importantly, this strategy has previously only
518 been available for C-terminal knock-ins, or N-terminal using a Cre-P2A sequence[31,32],
519 whereas whole-CDS replacement allows combining this manipulation with other
520 mutations throughout the CDS. The FLEx system is highly adaptable, and the cell fills we
521 utilized here can be exchanged for any genetically-encoded protein marker, sensor, or
522 optogenetic tool desired. This combination of powerful tools in an efficient CRISPR
523 environment should permit elegant, high-throughput studies of protein function in their
524 cellular context in the future.
525

526 ***Endogenous LRRTM2 expression varies across and within neurons***
527 One surprising result from this study was the cellular variability of LRRTM2 expression
528 even within the relatively restricted range of neuron types found in hippocampal culture.
529 LRRTM2 mRNA levels have been found to vary between cell types in neocortex and
530 hippocampus[33,34]. The correlation in single cells between mRNA and protein levels is
531 frequently poor, yet analysis of population protein expression levels at single-cell
532 resolution is challenging. Here, sparse knock-in allowed systematic evaluation of single-
533 neuron expression as well as the measurement of expression levels of mutant protein.
534 Note that such analysis is aided substantially by the presence of a knock-in dependent
535 marker, which we expect will be particularly useful for analysis *in vivo* where the range of
536 expression levels may be even greater. The mechanisms of cell-specific LRRTM2
537 expression level remain unknown, but one attractive possibility is that the activity history
538 of the neuron drives regulation of LRRTM2 as the number of AMPARs and synapses is
539 modulated.
540 Our data have shown conclusively that LRRTM2 is present at higher levels in larger
541 synapses. LRRTM2 binds directly with PSD-95 via its PDZ-binding motif[4,10], and
542 endogenous LRRTM2 was indeed most abundant at synapses with more PSD-95
543 content. Furthermore, we were able to increase the synaptic content of LRRTM2 via the
544 YACA mutations. Notably, this increase in synaptic content was smaller than expected
545 from mutant overexpression[22], highlighting the importance of manipulating LRRTM2
546 trafficking, and protein trafficking more generally, with endogenous regulations intact.
547 While the mechanisms by which the YACA mutations influence endogenous LRRTM2
548 trafficking are unclear, the intracellular location of the YACA mutations suggests
549 intracellular interactions may be involved in trafficking LRRTM2, including possibly PSD-
550 95 whose interaction domain is close by. However, we also found that LRRTM2 was not
551 present at nearly 20% of PSD-95-containing synapses and was also present at dendritic
552 locations lacking PSD-95, suggesting there are additional mechanisms beyond binding
553 to PSD-95 responsible for LRRTM2 localization. Several overexpression studies have
554 found that removing the PDZ-binding motif does not affect the ability of LRRTM2 to traffic
555 to synapses or facilitate LTP[7,10], suggesting it is not exclusively responsible for LRRTM2
556 trafficking. Furthermore, it appears that these potentially extrasynaptic LRRTM2 puncta
557 frequently contained AMPARs. What would prompt enrichment of LRRTM2 at these
558 points is unclear. The neurexin-binding domain of LRRTM2 is required for rescuing the
559 effects of removing LRRTM2 on LTP via overexpression[6,7], but it is unclear whether
560 neurexins, or indeed any presynaptic scaffold, would be present outside of PSD-95
561 synapses. One hypothesis is that these LRRTM2 puncta are indeed synapses but lack
562 PSD-95 and contain other synaptic scaffolds whose interaction with LRRTM2 has not
563 been investigated. Given the homology in PDZ motifs across synaptic scaffolds, it is
564 certainly likely that the LRRTM2 PDZ-binding motif also interacts with other scaffolds such
565 as SAP-102 or PSD-93[35]. However, previous research has found that SAP-102-only

566 synapses are relatively rare past early postnatal development[36,37], and as such are
567 unlikely to account for very many of these LRRTM2 puncta. Alternatively, it is possible
568 that these puncta represent a diffusible pool of LRRTM2 that can be recruited to the
569 synapse as needed, much the way that AMPARs are[20,21]. This model would include two
570 pools of LRRTM2 protein, one synaptic and one extrasynaptic, where the synaptic pool
571 would be either non-functional or diminished in size when unable to bind presynaptic
572 neurexins but not PSD-95 or other scaffolds.
573 LRRTM2 and its sister protein LRRTM1 share high sequence homology and are
574 expressed in the same broad regions of hippocampus[38], and many studies have relied on
575 their simultaneous deletion[6,7]. This is done because CAMs frequently compensate for one
576 another, and dual knockouts can better illustrate their roles than either knockout alone,
577 as has been recently shown for CAMs LRRTM1 and SynCAM1 in hippocampal
578 synapses[39]. Interestingly, the effects of dual knockout can be rescued with
579 overexpression of LRRTM2 alone[6,7]. We found endogenous LRRTM2 in 80% of
580 synapses, which raises the possibility that the remaining 20% may be LRRTM1-positive.
581 This is potentially an interesting case of differential subcellular trafficking of highly
582 homologous proteins. A more extreme segregation between members of a single family
583 of CAMs has been observed previously *in vivo*; for example the neuregulin proteins 1 and
584 3 are differentially sorted to the somatic and axonal domains of pyramidal neurons,
585 respectively[40]. The functional implications of such segregation are unknown at this point,
586 but given that many studies have relied on their simultaneous deletion, future work is
587 needed to tease out their individual roles. While it is currently extremely challenging to
588 simultaneously target multiple genes via CRISPR, future developments could make it
589 possible to tag and manipulate both LRRTM1 and LRRTM2 in the same cells to further
590 investigate their relative trafficking.
591
592 ***Relationship of endogenous LRRTM2 with AMPARs***
593 Numerous previous studies have established that LRRTM2 controls synaptic AMPAR
594 content. One possibility is that changing LRRTM2 levels would be determinative in
595 establishing AMPAR levels. Our data show that synapses lacking LRRTM2 are smaller
596 and contain fewer AMPARs than those with LRRTM2, though the latter correlation was
597 weaker than predicted by this hypothesis. However, when we increased the content of
598 LRRTM2 with the YACA mutations, there did not appear to be a similar increase in
599 synaptic AMPAR content. Several possible explanations exist for this finding. One
600 possibility is that while YACA mutations may increase LRRTM2 surface content by
601 altering membrane-targeting trafficking mechanisms[23], they also block the as yet
602 unknown mechanisms by which LRRTM2 anchors AMPARs at the surface. However,
603 given that previous studies have removed the entire C-terminal domain of LRRTM2 in a
604 replacement context with no effect on LTP[6,7], this appears unlikely. Another possibility is
605 that the YACA mutations, which lead to a cell-wide increase in LRRTM2 content, may not

606 significantly alter AMPAR trafficking at the cellular level due to potential compensatory
607 mechanisms. Instead, it's plausible that AMPAR trafficking could be influenced by
608 localized increases in LRRTM2 content on a synapse-by-synapse basis. Further, it is
609 possible that the mechanism connecting AMPAR enrichment to LRRTM2 levels is not
610 sensitive enough to respond to the small increase that the YACA mutations induce in
611 LRRTM2 content. Unfortunately, these latter two possibilities are difficult to test without a
612 greater understanding of what induces LRRTM2 trafficking to synapses so that it can be
613 manipulated more directly and in a physiologically relevant manner. Finally, it is important
614 to note that we measured the LRRTM2 YACA mutation at a static baseline state, whereas
615 it is possible that the role of the domain containing the YACA mutants is engaged
616 primarily during synaptic potentiation.
617 Previous studies have established that surface diffusion of AMPARs to synapses is
618 critical during plasticity[20,21]. Our data have shown that LRRTM2 and AMPARs localize to
619 puncta outside of PSD-95 and appear together more frequently than either protein alone.
620 This suggests the exciting possibility that LRRTM2 may traffic together with, or potentially
621 shepherd, extrasynaptic AMPARs as they are trafficked to synapses during plasticity. This
622 would give LRRTM2 the ability to retain new AMPARs at synapses via its interaction with
623 presynaptic neurexins in a timely fashion. If LRRTM2 recruits AMPARs to the synapse
624 via this surface trafficking mechanism, it would explain why a whole-cell increase in
625 LRRTM2 content, as induced via the YACA mutation, is insufficient to induce a similar
626 increase in synaptic AMPAR content. LRRTM2 facilitates synaptogenesis[4], and
627 extrasynaptic LRRTM2-AMPAR puncta suggest a means by which LRRTM2 could enable
628 AMPAR trafficking to new synapses. The presence of LRRTM2 clusters outside of PSD-
629 95 synapses, and their potential role in local AMPAR trafficking, represents an intriguing
630 new development in our understanding of LRRTM2 function and requires further
631 investigation to fully understand.
632 Together, our approach for whole-CDS replacement facilitates labeling, imaging, and
633 manipulating endogenous proteins, including LRRTM2, and represents a potent
634 methodological advance in the field of cellular neurobiology. While we have illustrated
635 many new findings regarding the endogenous trafficking of LRRTM2, the approach of
636 whole-CDS replacement will assist in determining the mechanisms by which this cell
637 adhesion molecule controls such critical processes as synaptic AMPAR retention and
638 plasticity.
639
640
641 **Methods**
642
643 ***CRISPR Design and Plasmids***
644 The NCBI Rnor 6.0 database LRRTM2 sequence (NC_005117.4) was used as a
645 reference for both guide and donor design. Guides positioned at least 50bp from the

intron-exon splicing boundary (5' guide) or over 50bp into the 3'UTR (3' guide) were identified using the Benchling CRISPR guide ranking tool, targeting those with the best on-target and off-target scores based on previous publications[41,42], then were cloned into the pX330 backbone as previously described[12]. Guide sequences: 5' GTTTTAATCTCTCTTATACA 3' (guide 1, anneals to Intron 1) and 5' CTTTTTAAGTAGGAAGCCAG 3' (guide 2, anneals to the 3'UTR). Once cloned into the pX330 vector under identical U6 promoters, the guide construct was grown in NEB STABLE cells at a reduced temperature of 30°C to prevent bacterial recombination of the promoters. Guide 2 was cloned with a guanine at the 5' end after the U6 promoter, as described in Fang et al. 2021. 3xALFA epitope tag and IRES2-Cre insertions as well as YACA (Y501A/C504A) mutations were added to the donor by either Gibson NEB HIFI Assembly or NEB Q5 site-directed mutagenesis as appropriate. The 3xALFA tag was inserted after the signal peptide along with a small linker (TS) after the tag, and IRES2-Cre was added immediately after the stop codon. For lentiviral expression, the guides and donor were combined into the pFW lentiviral backbone[18] by NEB HIFI Assembly. FLEx-mTagBFP2 was generated by NEB HIFI Assembly, replacing the mCherry-KASH in Addgene #139652 (a gift from Harold MacGillavry; http://n2t.net/addgene:139652; RRID:Addgene_139652) with the mTagBFP2 gene and replacing the hSyn promoter with the CAG promoter for improved expression levels. FLEx-IRES-EGFP was made similarly, with a CMV promoter in place of CAG. HA-spCas9 was a gift from Harold MacGillavry (Addgene plasmid # 131506; http://n2t.net/addgene:131506; RRID:Addgene_131506), constructs to make lentivirus (pMD2.G and psPAX2) were gifts of Didier Trono (Addgene plasmid # 12259; http://n2t.net/addgene:12259; RRID:Addgene_12259; Addgene plasmid # 12260; http://n2t.net/addgene:12260; RRID:Addgene_12260). ALFA and HA tagged LRRTM2 KDR variants for transient expression in HEK cells were generated by replacing the GFP tag in KDR GFP-LRRTM2 previously described[8] using NEB HIFI Gibson cloning. All sequences were confirmed by whole plasmid sequencing (Plasmidsaurus) using Oxford Nanopore Technology with custom analysis and annotation. These plasmid sequences will be deposited with public databases upon publication.

Note that we originally tried to combine the FLEx fluorescent protein with the guides/donor lentiviral plasmid, but full plasmid sequencing revealed that growing a construct with both IRES2-Cre and FLEx genes in bacteria caused a reversal of the flipped gene. Based on previous literature, we suspect that this could be due to bacteria recognizing the IRES2 sequence and expressing the Cre protein themselves[43]. Thus, we elected to separate these constructs into two viruses, with the Cas9 virus as a third virus due to its large size and the limitations of lentiviral packaging. The guides and Cas9 were cloned into separate viral constructs to avoid Cas9 activity at the *Lrrtm2* locus in the absence of donor sequence to replace it.

686 ***Genomic sequencing***
687 Cultured hippocampal neurons were infected with either knock-in viruses (Cas9 +
688 guides/donor) or guides/donor alone at DIV1. DNA was extracted and purified at DIV21
689 using the Wizard Genomic DNA Purification Kit (Promega) according to manufacturer's
690 instructions. The LRRTM2 gene was amplified using primers that anneal on either side of
691 the guide1 Cas9 cut sites (5' AGCCAGTGAATTCCCGTTTT 3', 5'
692 AGGCGAACTGGGATAGTCCGCA 3'). This PCR product was gel-purified, then PCR
693 amplified again using a reverse primer that anneals specifically to the knock-in HA
694 sequence (5' CATTTAGGTGGACAACTAGTAGCGTAGTCTGGTACATCAT 3') and the
695 same forward primer as in the first PCR run. Gel purified PCR products from the round 1
696 PCR control condition and the round 2 PCR knock-in condition (no product was made
697 from the control condition due to the lack of knock-in HA sequence in the *Lrrtm2* locus)
698 were Sanger sequenced (Genewiz/Azenta) to confirm knock-in in the correct genetic
699 locus.
700
701 ***Lentivirus Production***
702 HEK293T cells (ATCC CRL-3216) maintained in DMEM + 10% FBS supplemented with
703 penicillin/streptomycin were plated at high density and transfected using PEI as we have
704 described[18]. Cells were incubated for 6 hours followed by a media change to standard
705 neuronal culture media. Lentivirus was allowed to accumulate in the media for 2 days
706 before harvesting, and debris were removed by centrifugation at 1000 x RPM for 5
707 minutes. Virus aliquots were used fresh or else frozen at -80°C. Based on prior titrations
708 and to achieve high levels of co-infection, neurons were infected with 100μL of each virus.
709
710 ***Immunostaining***
711 Dissociated, mixed-sex hippocampal cultures were prepared as previously described[18].
712 Knock-in preparations were infected at DIV1 with 100 ul each unconcentrated Cas9,
713 guide/donor, and marker viruses. To immunostain LRRTM2 and synaptic proteins,
714 coverslips were removed from their culture media at DIV21 and blocked for 5 minutes in
715 1% BSA in Tyrode's buffer, then live-labeled with ALFA-AlexaFluor647 nanobody
716 (NanoTag) and mouse anti-pan GluA (Synaptic Systems) at 1:500 and 1:200,
717 respectively, in the same solution composition for 12 minutes at room temperature. To
718 avoid background from aggregation, aliquots of nanobody are vortexed thoroughly for 2
719 minutes followed by centrifugation at max speed for 1 minute. Coverslips were then rinsed
720 twice briefly in phosphate buffered saline (PBS) and fixed in 4% paraformaldehyde (PFA)
721 + 4% sucrose in PBS for 9 minutes at room temperature. The cells were washed with
722 PBS-100mM Glycine (PBS-Glycine) three times for five minutes each, permeabilized with
723 0.3% TritonX-100 in PBS-Glycine for 15 minutes at room temperature, then incubated
724 with PSD-95 nanobody conjugated to AZ568 dye (NanoTag) at 1:100 and goat anti-rabbit

725 AlexaFluor488 at 1:750 in 4% TBS-milk for 45 minutes at room temperature. Coverslips
726 were then rinsed in TBS three times for five minutes prior to imaging.
727 To test different knock-in tags, we transfected HEK cells using Lipofectamine 2000
728 (ThermoFisher) with an EGFP cell fill to normalize for transfection as well as the
729 respective KDR LRRTM2 constructs with 1xALFA, 3xALFA with interstitial linkers, and a
730 3xALFA without linkers. A 1xHA tag visualized with a traditional primary-secondary
731 antibody approach was included for comparison. We then immunolabeled, fixed, and
732 stained as above. Cells were fixed and surface stained with either the ALFA nanobody or
733 an HA primary antibody followed by fluorescent secondary. Both approaches utilized the
734 AlexaFluor 647 dye to detect LRRTM2 as well as identical imaging parameters. Images
735 were analyzed in ImageJ for 647 fluorescence intensity normalized to GFP as a
736 transfection control.
737
738 ***Microscopy***
739 Images were acquired on an Andor Dragonfly spinning disc confocal (Andor) attached to
740 a Nikon Ti2 Eclipse inverted microscope base with a 60x Plan Apo 1.49 NA objective.
741 Excitation laser light (405, 488, 561, or 638 nm) from an Andor ILE, flattened by an Andor
742 Beam Conditioning Unit, was passed to the sample by a 405/488/561/640 quadband
743 polychroic (Chroma). Emission light was passed through an appropriate bandpass filter
744 (FF02-447/60-25 (Semrock), ET525/50, ET600/50 (Chroma), or Em01-R442/647
745 (Semrock), for 405nm, 488nm, 561nm, and 638nm emission, respectively) and collected
746 on a Zyla4.2 sCMOS camera. Cells of interest were imaged with confocal z-stacks with
747 0.5 µm z-steps at 25-90% laser power with 200 ms exposure (400 ms for 638nm channel),
748 with each channel imaged sequentially.
749
750 ***Analysis***
751 Image processing was performed using macros and plugins in Fiji/ImageJ[44], and image
752 file names were blinded during analysis. All z-stacks were converted to maximum
753 intensity projections, then *xy* chromatic aberrations were corrected using the *Register*
754 *Channels* tool of the NanoJ – Core plugin[45] (and 4-color TetraSpek bead images acquired
755 prior to imaging as a standard.
756 We used a semi-automated ImageJ macro similar to that described in Dharmasri
757 et al[18] to detect synapses. In brief, the macro provides user-guided image cropping,
758 followed by automated thresholding to isolate ALFA-LRRTM2 signal, then uses the
759 puncta detection plugin SynQuant[19] to detect synapses on knock-in neurons from the
760 PSD-95 staining. Detected synapses were converted to ROIs and used to measure
761 synapse area and PSD-95 intensity, as well as intensity of LRRTM2 and/or AMPARs
762 within those puncta. Due to high cellular variability, we took the cellular average for each
763 channel and assigned each synapse ROI a z-score based on their relative intensity within
764 that channel to the cellular average.

765 $$Z - score = \frac{(x-\mu)}{\sigma},$$

766 where x is the intensity at the individual synapse, μ represents the mean synaptic intensity
767 across the neuron, and σ represents the standard deviation. This was performed for two
768 measures of intensity readout: Mean Intensity (Figure 3C) and Raw Integrated Density
769 (everything else), the latter of which is superior for establishing protein content
770 correlations as it does not normalize for synapse size, which is biologically relevant. We
771 observed infrequent outliers on the upper end of the intensity distributions across all three
772 staining conditions (LRRTM2, PSD-95, and AMPARs), presumably due to clumping of
773 the antibody or nanobody. To eliminate these synapses from our dataset in a way that
774 was agnostic to the normality of the distribution, we utilized a cutoff of 150% of the
775 interquartile range (IQR) above the third quartile (Q3 + 1.5*IQR) in each of the three
776 channels' distributions. These outlier calculations were performed using the z-scored data
777 to avoid artificially removing synapses from cells with high expression levels of any of the
778 measured proteins, particularly LRRTM2. Synapses above the calculated cutoff were
779 removed completely from the datasets.

780 Synapses along knock-in dendrites were categorized as containing or lacking
781 LRRTM2 by comparing the distributions of ALFA labeling along knock-in dendrites or
782 neighboring dendrites (acquired simultaneously and analyzed with the same pipeline as
783 above in parallel).

784 $$Relative\ ALFA\ Z - score = \frac{(i-\mu_n)}{\sigma_n},$$

785 where i is the synaptic ALFA intensity (KI dendrite or neighboring region), μ_n represents
786 the mean synaptic ALFA intensity of the adjacent neighboring region, and σ_n represents
787 the standard deviation of the adjacent neighboring region. The 95th percentile of the ALFA
788 intensity distribution on neighboring dendrites was used as the cutoff between "LRRTM2-
789 lacking" and "LRRTM2-containing" synapses along knock-in dendrites.

790 For detection of LRRTM2 and/or AMPAR-containing puncta within and outside
791 PSD-95 puncta, we hand-picked elliptical ROIs around puncta in each channel using FIJI
792 ROI selection tools, then quantified fluorescence intensities as above. Raw Integrated
793 Density was measured within each hand-drawn ROI, followed by background subtraction
794 using similarly-sized ROIs drawn nearby in the background region of each image. Each
795 puncta was categorized as "positive" or "negative" for each of LRRTM2, PSD-95, and
796 AMPARs utilizing a first local-minimum cutoff in the intensity distribution of each protein
797 across the dataset. These categories were plotted as pie charts using a combination of
798 RStudio and GraphPad Prism.
799

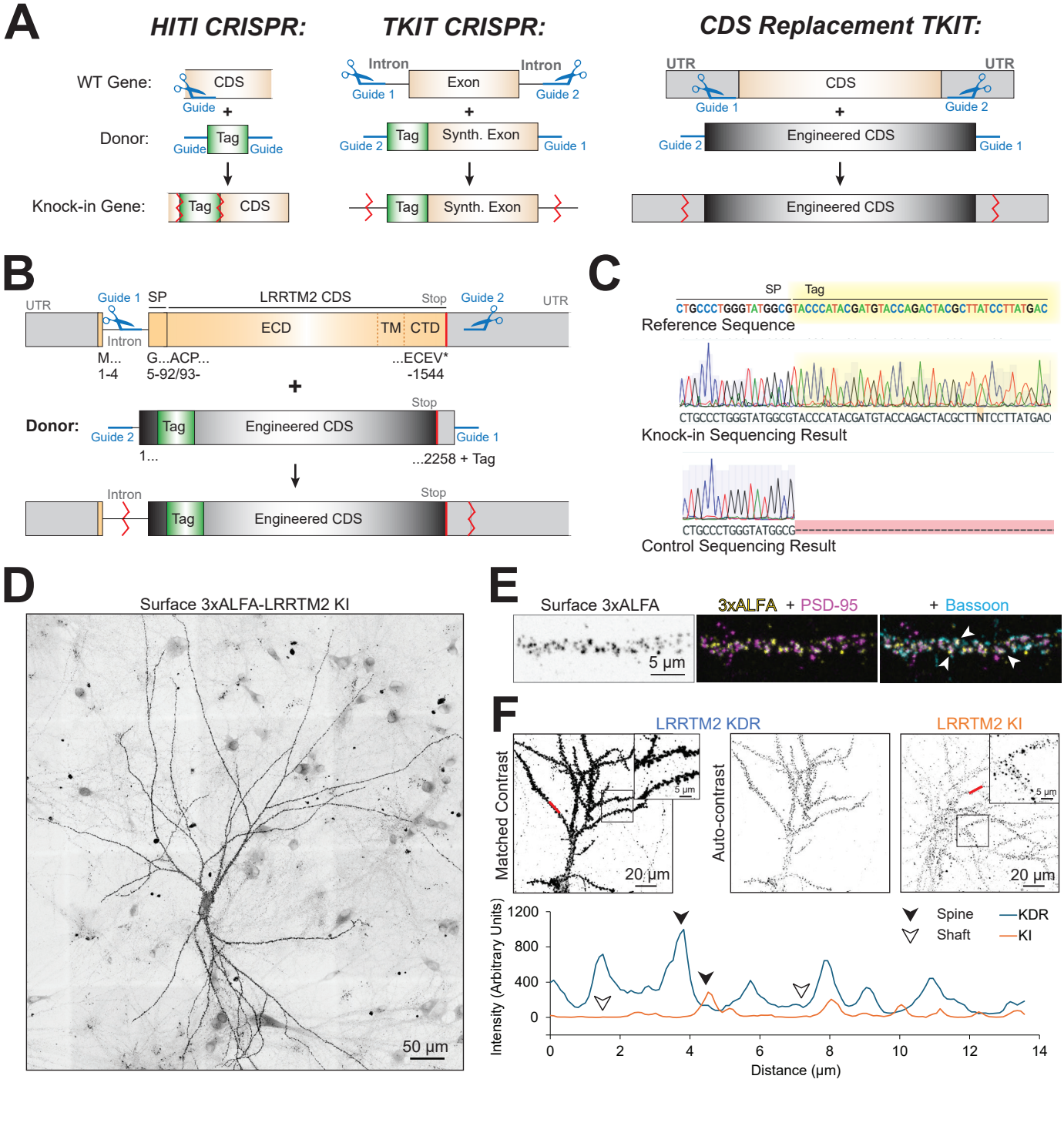
800 ***Bioinformatics***

801 To assess how applicable a whole-CDS replacement of a similar size to our approach
802 would be to other genes, we turned to a bioinformatics approach. Genetic data on *Rattus*

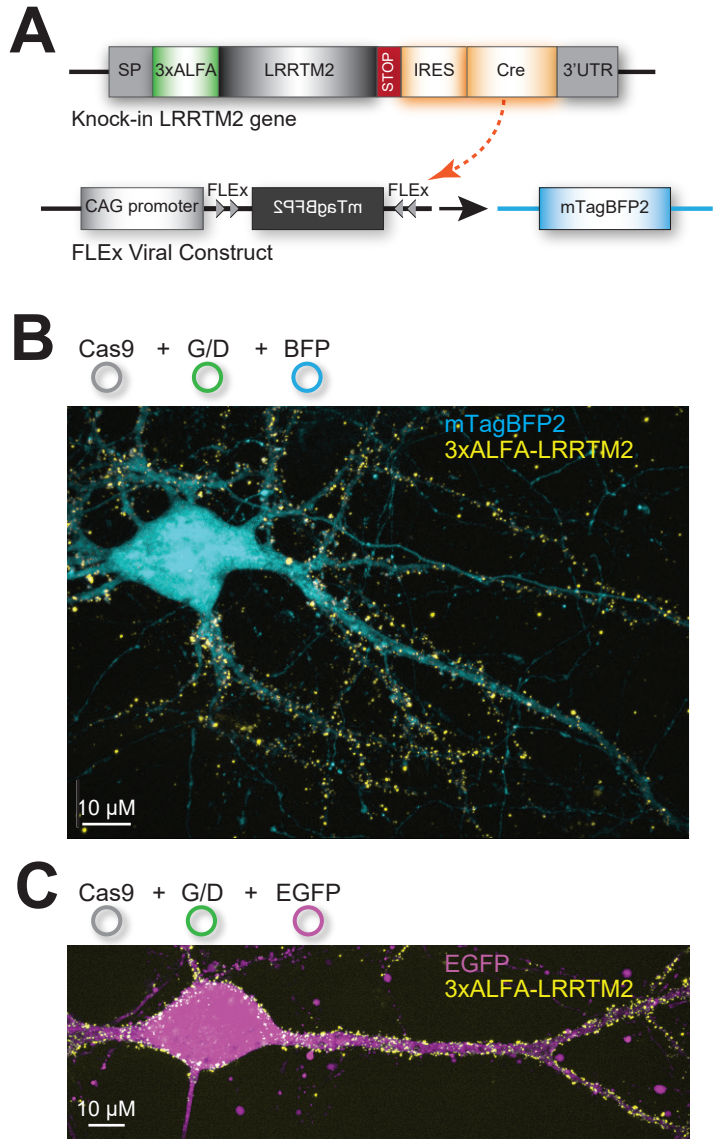
norvegicus was pulled from the search function of the University of California Santa Cruz Genome Database (https://genome.ucsc.edu/cgi-bin/hgGateway). To measure the CDS Spanning Region along the genome for each gene and splice isoform, we subtracted the stop codon position from the start codon position. Practically, replacement size will depend on the proximity of suitable PAM sites and guide sequences on either side of the CDS. For practicality of graphing, we set the maximum CDS Spanning Region size as 10,000bp. Search parameters entered into the UCSCGD were as follows: absolute value(endCDS – startCDS) <= 10000bp. Alternative splice isoforms of the same gene were included as separate genes because they likely differ in genomic span and carry differing biological significance. Histogram was produced in RStudio, cumulative plot in GraphPad Prism.

Statistics

Data from ImageJ measurements was processed and statistics calculated using RStudio (Posit). Frequency distributions were produced for each of the synaptic intensity measurements to evaluate normality. These distributions were not normally distributed as evaluated with a Shapiro-Wilks test, and we therefore elected to utilize Wilcoxon tests to compare LRRTM2-containing and -lacking synapses, as well as YACA and wildtype LRRTM2 mutant conditions. Graphs were produced in RStudio or GraphPad Prism (Dotmatics). Due to the high throughput nature of our semi-automated synapse detection method, we pruned the number of points displayed in scatter plots by averaging every 5 synapses together based on ranked data (PruneRows, Prism). While this had very little effect on the slope of the regression lines, it does influence the R-squared values so we have reported the values of regressions performed on the raw data in the text.


A
HITI CRISPR:
TKIT CRISPR:
CDS Replacement TKIT:
WT Gene:
CDS
Guide
Donor:
Tag
Knock-in Gene:
Intron
Exon
Guide 1
Guide 2
Synth. Exon
UTR
Engineered CDS
B
SP
LRRTM2 CDS
Stop
ECD
TM
CTD
M...
1-4
G...ACP...
5-92/93-
...ECEV*
-1544
1...
...2258 + Tag
C
Reference Sequence
Knock-in Sequencing Result
Control Sequencing Result
D
Surface 3xALFA-LRRTM2 KI
50 µm
E
Surface 3xALFA
3xALFA + PSD-95
+ Bassoon
5 µm
F
LRRTM2 KDR
LRRTM2 KI
Matched Contrast
Auto-contrast
20 µm
Spine
Shaft
KDR
KI
Intensity (Arbitrary Units)
Distance (µm)

827 **Figure 1**: Successful whole-CDS replacement of LRRTM2
828 **A)** Diagram of traditional HITI, TKIT, and whole-CDS replacement CRISPR methods.
829 Locations of possible Cas9-induced indels are indicated with red zigzag lines. **B)** Diagram
830 of the *Rattus norvegicus Lrrtm2* gene with CRISPR guide locations. Note that the entire
831 coding sequence (orange), minus 4 bases of the signal peptide that is cleaved from the
832 mature protein, is confined to exon 2. Guide 1 anneals to the intronic region and guide 2
833 to the 3'UTR. The replacement donor (Engineered CDS, black) contains an N-terminal
834 epitope tag between the signal peptide and mature protein. **C)** Genomic sequencing
835 shows positive knock-in result. Samples of genomic DNA from either the guides/donor
836 virus alone (Control) or the guides/donor and Cas9 viruses (Knock-in, KI) underwent two
837 rounds of PCR amplification, first to amplify the *Lrrtm2* gene, then again to amplify the
838 knock-in tag (no product in the control condition). Sanger sequencing shows the 3xHA
839 tag sequence in the KI condition, without indels. **D)** Image of a cultured hippocampal
840 knock-in neuron, stained for 3xALFA-LRRTM2. Sparse knock-in permits identification of
841 synaptic puncta along knock-in vs neighboring dendrites. **E)** Images of a dendrite from a
842 knock-in neuron, stained for 3xALFA-LRRTM2 (yellow), PSD-95 (magenta), and pre-
843 synaptic Bassoon (cyan). LRRTM2 puncta are largely synaptic (white arrows). **F)**
844 Endogenous LRRTM2 (KI, right) is expressed at far lower levels than typical exogenous
845 expression via knockdown-rescue (KDR, left). Matched staining, imaging, and look-up
846 tables show that while both KDR and KI show synaptic localization, knockdown-rescue
847 ALFA labeling is considerably higher than knock-in both at synapses and between (blue
848 and orange, respectively, line scan, lower right). Dark arrows indicate spine enrichment,
849 white arrows dendritic signal. Line scan locations shown with red lines in the respective
850 images.


A
SP
3xALFA
LRRTM2
STOP
IRES
Cre
3'UTR
Knock-in LRRTM2 gene
CAG promoter
FLEx
FLEx
mTagBFP2
FLEx Viral Construct
B
Cas9 + G/D + BFP
mTagBFP2
3xALFA-LRRTM2
10 μM
C
Cas9 + G/D + EGFP
EGFP
3xALFA-LRRTM2
10 μM

851 **Figure 2**: Rapid identification of whole-CDS knock-in cells
852 **A)** Schematic of knock-in conditional marker implementation. An Internal Ribosome Entry
853 Site (IRES) sequence and a Cre gene were inserted into the donor construct between the
854 LRRTM2 STOP codon and the 3'UTR (IRES-Cre, orange). Thus, Cre is expressed once
855 the donor is successfully incorporated into the LRRTM2 gene. A separate lentiviral
856 construct was used to deliver FLEx-mTagBFP2 cell fill, which can be targeted by the
857 knocked-in Cre. **B)** Knock-in neuron expresses the mTagBFP2 cell fill. Surface
858 immunostained 3xALFA-LRRTM2 (yellow), cytosolic mTagBFP2 (cyan). **C)** The IRES-
859 Cre approach permits flexibility in the conditional marker. Due to the flexible multi-virus
860 approach, it is simple to swap mTagBFP2 for any other FLEx protein marker, such as
861 EGFP (magenta). Three viral constructs were used: Cas9, Guides/Donor, and FLEx
862 marker.
863

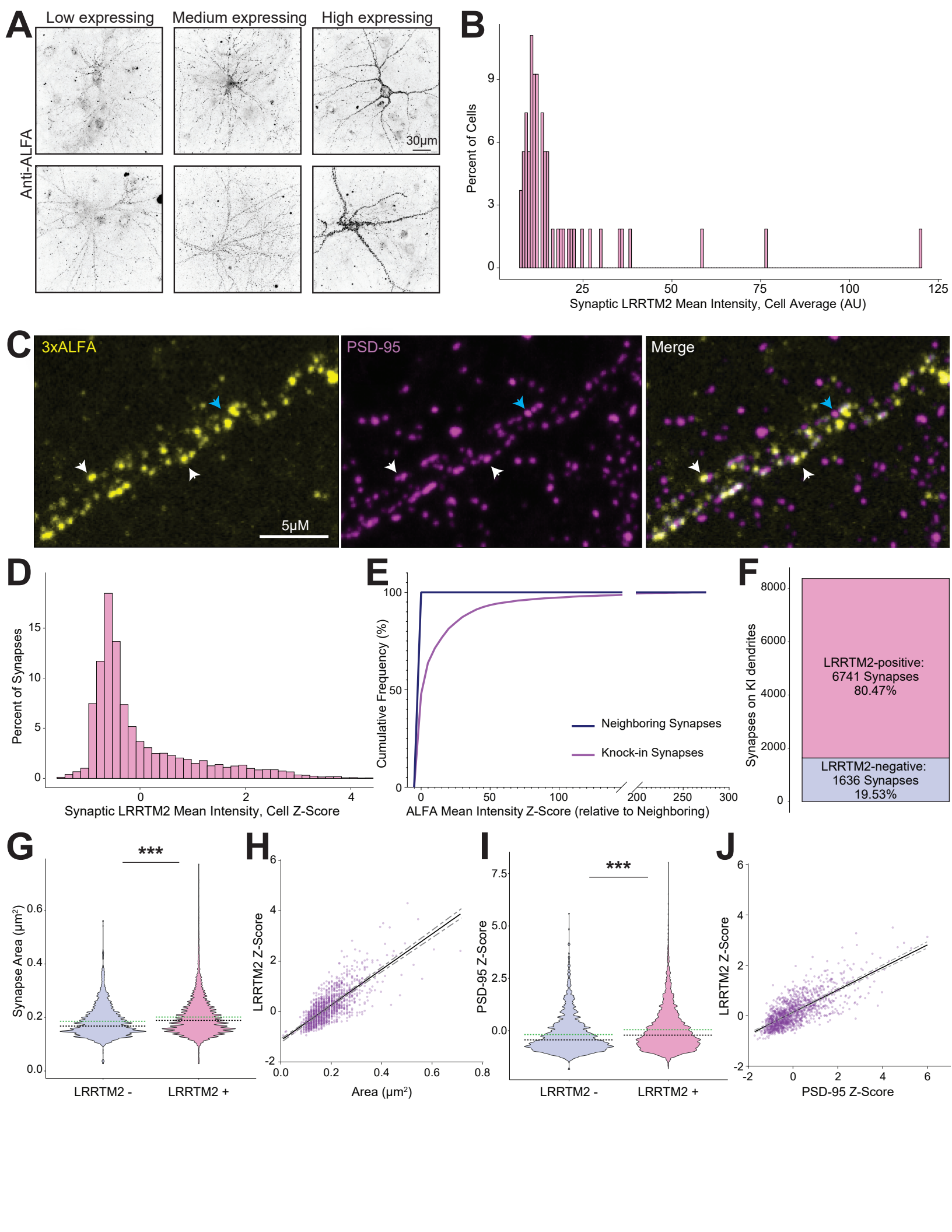

A
Low expressing
Medium expressing
High expressing
Anti-ALFA
30µm
B
Percent of Cells
Synaptic LRRTM2 Mean Intensity, Cell Average (AU)
C
3xALFA
PSD-95
Merge
5µM
D
Percent of Synapses
Synaptic LRRTM2 Mean Intensity, Cell Z-Score
E
Cumulative Frequency (%)
Neighboring Synapses
Knock-in Synapses
ALFA Mean Intensity Z-Score (relative to Neighboring)
F
Synapses on KI dendrites
LRRTM2-positive:
6741 Synapses
80.47%
LRRTM2-negative:
1636 Synapses
19.53%
G
***
Synapse Area (µm²)
LRRTM2 -
LRRTM2 +
H
LRRTM2 Z-Score
Area (µm²)
I
***
PSD-95 Z-Score
LRRTM2 -
LRRTM2 +
J
LRRTM2 Z-Score
PSD-95 Z-Score

Figure 3: Variation in endogenous synaptic LRRTM2 content reflects key markers of synaptic strength.

A) Exemplar hippocampal neurons expressing 3xALFA-LRRTM2 at low, medium, and high levels. **B)** Histogram of synaptic LRRTM2 expression across the dendrite, averaged by cell. **C)** Image of an example stretch of dendrite expressing knock-in 3xALFA-LRRTM2 (yellow) immunolabeled for PSD-95 (magenta). LRRTM2 is highly synaptic in localization (exemplars, white arrows), but does not appear in every synapse on the dendrite (blue arrow). **D)** Distribution of synaptic LRRTM2 mean intensity along knock-in dendrites, normalized for cellular variability using a per-cell Z-scoring method. **E)** Cumulative frequency plot of ALFA mean intensity at synapses in knock-in neurons (pink) or neighboring non-knock-in dendrites (blue). Z-scores for LRRTM2 mean intensity measures were calculated from the mean and standard deviation of synaptic ALFA signal within neighboring non-knock-in synapses. **F)** Bar plot of synapses along knock-in dendrites categorized as containing (pink) or lacking (blue) LRRTM2. The 95th percentile of non-knock-in mean intensity Z-scores was used to delineate the cutoff between LRRTM2-containing and -lacking synapses. **G)** Violin plot comparing synapse area along knock-in dendrites that do or do not contain LRRTM2 (pink or blue, respectively). ($p < 2.2^{-16}$) **H)** Scatter plot correlating synapse area with synaptic LRRTM2 content (raw integrated density Z-score) in LRRTM2-containing synapses. **I)** Violin plot comparing PSD-95 content within synapses along the knock-in dendrites that do or do not contain LRRTM2. ($p < 2.2^{-16}$) **J)** Scatter plot of PSD-95 content with synaptic LRRTM2 content at LRRTM2-containing synapses. In violin plots, green lines represent means and black lines represent medians for each distribution. Scatter plots have been averaged across every 5 ranked data points for visibility; statistics are calculated from raw data. Lines show linear regression with gray dotted lines representing 95% confidence intervals.

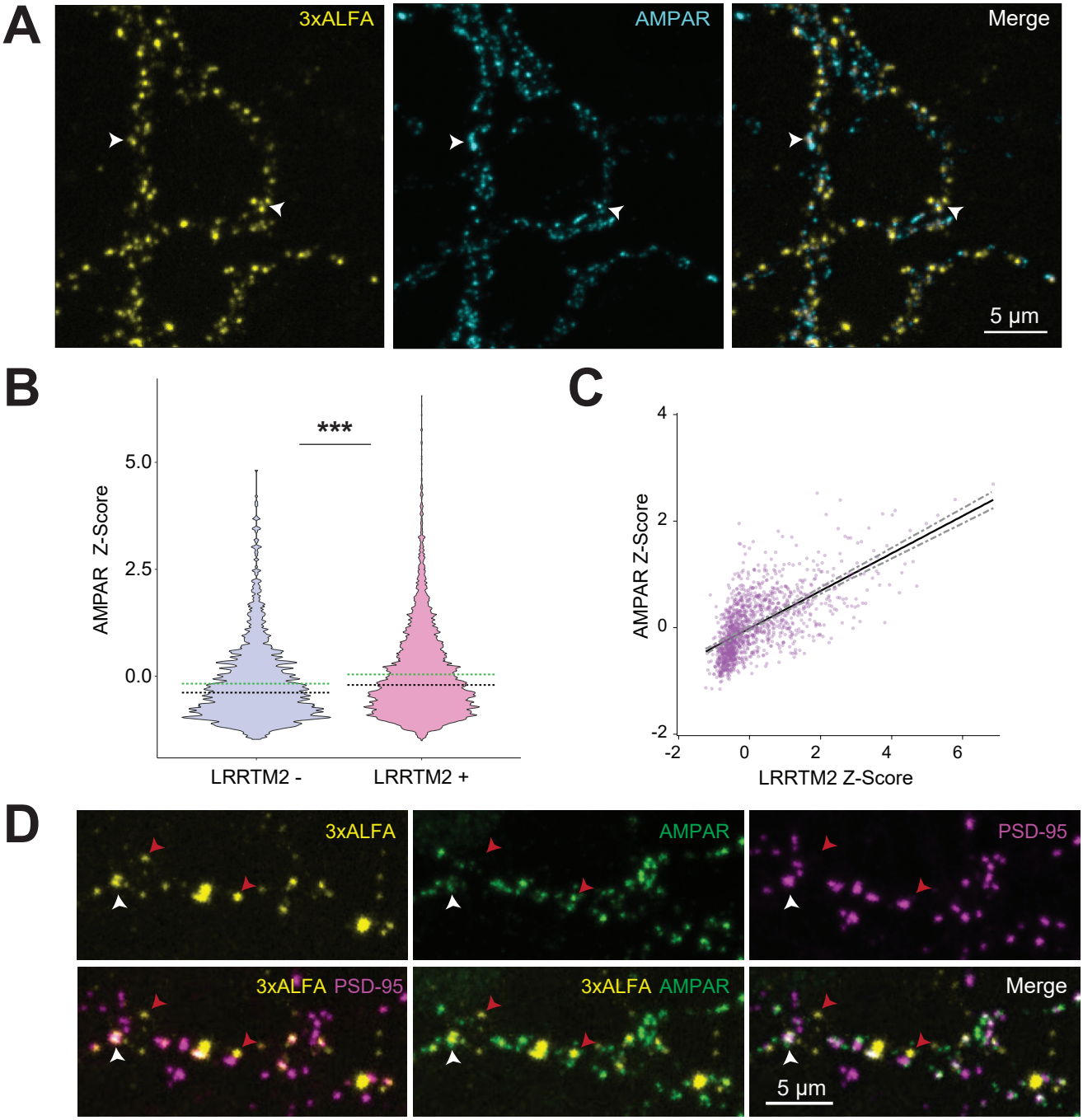
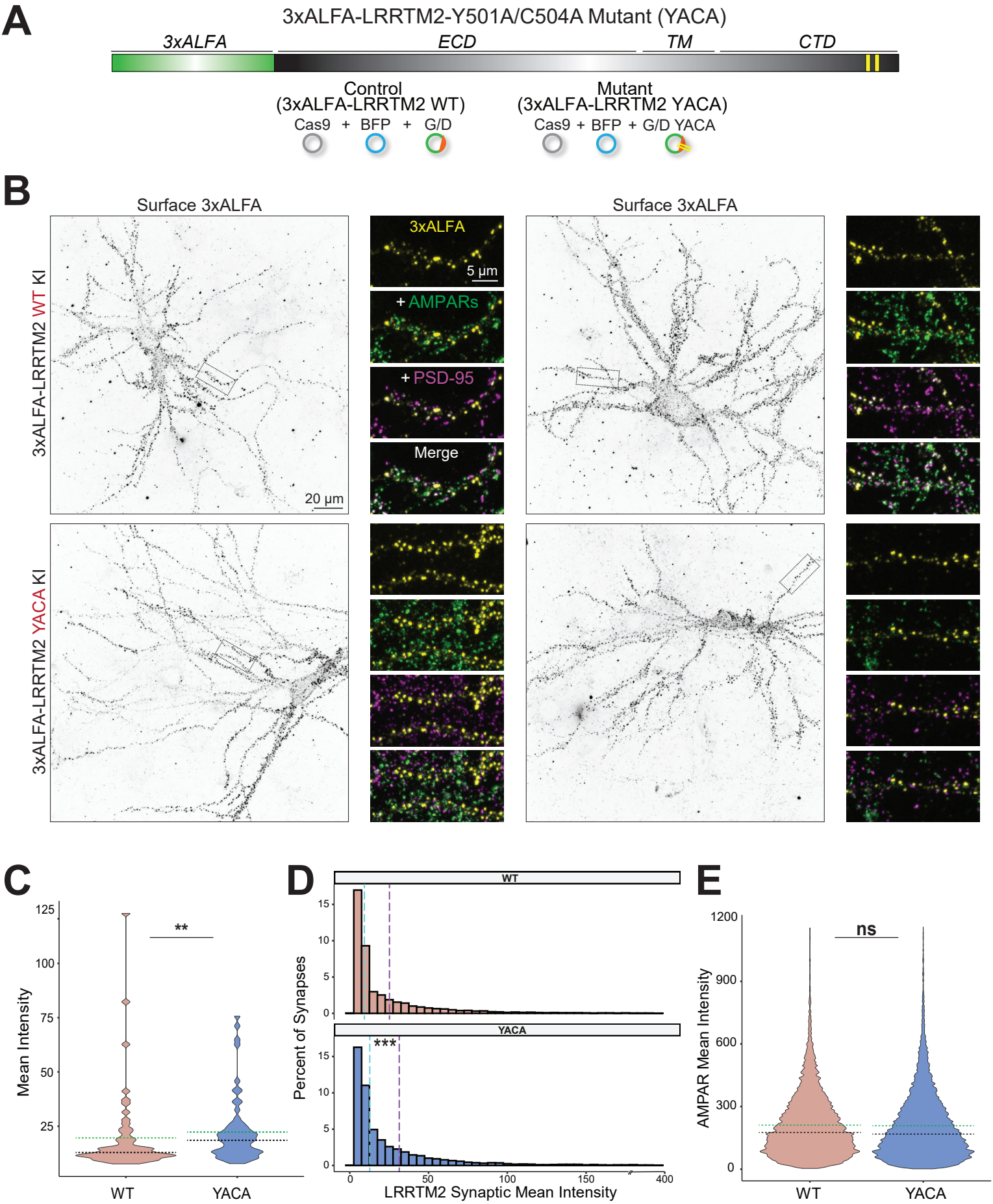
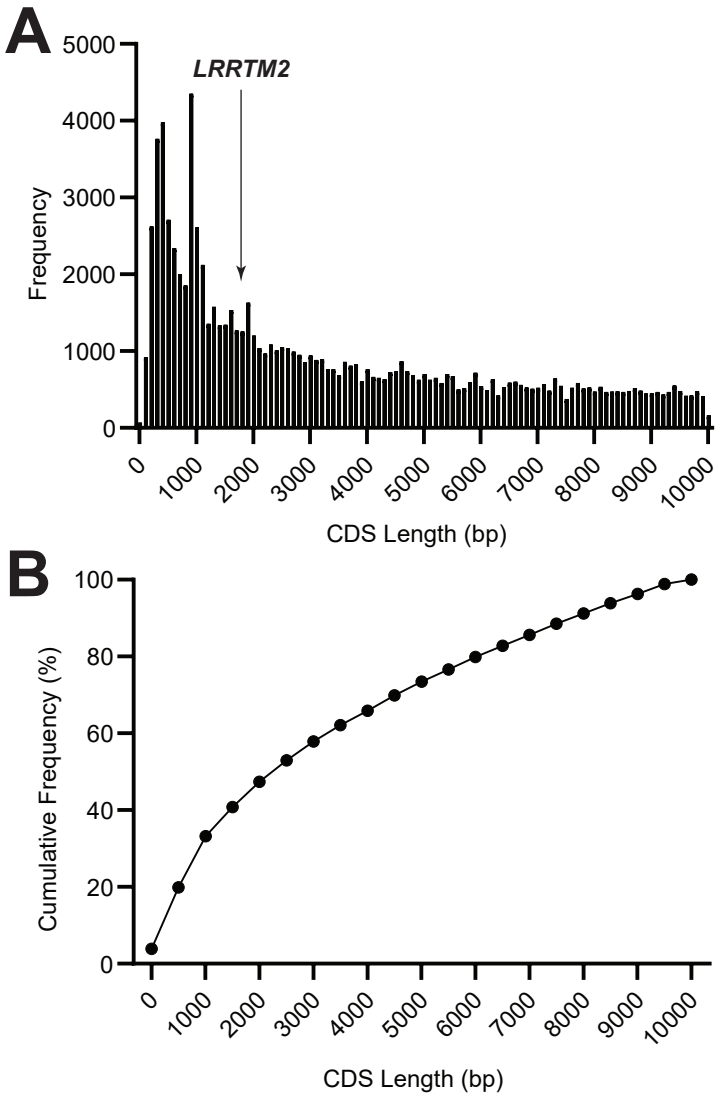

A
3xALFA
AMPAR
Merge
5 µm
B
***
AMPAR Z-Score
5.0
2.5
0.0
LRRTM2 -
LRRTM2 +
C
AMPAR Z-Score
4
2
0
-2
-2
0
2
4
6
LRRTM2 Z-Score
D
3xALFA
AMPAR
PSD-95
3xALFA PSD-95
3xALFA AMPAR
Merge
5 µm
E
All Puncta
4.7%
0.3%
3.0%
0.2%
16.9%
8.4%
66.4%
AMPAR only
PSD-95 only
LRRTM2 only
LRRTM2 + PSD-95 only
AMPAR + PSD-95 only
AMPAR + LRRTM2 only
LRRTM2 + PSD-95 + AMPAR
Puncta without PSD-95
18.9%
52.0%
29.2%

Figure 4: Synaptic AMPAR content scales with LRRTM2 content

A) Images of a knock-in neuron labeled for 3xALFA-LRRTM2 (yellow) and surface AMPARs (cyan). Arrows show high co-enrichment of AMPARs with LRRTM2. **B)** Violin plot comparing synaptic surface AMPAR content along the knock-in dendrites that do or do not contain LRRTM2 (pink and blue, respectively). $p < 2.2^{-16}$ Green lines represent means and black lines represent medians for each distribution. **C)** Scatter plot correlating AMPAR content (Raw Integrated Density z-score) with synaptic LRRTM2 content (Raw Integrated Density z-score) at LRRTM2-containing synapses. Scatter plot has been averaged across every 5 ranked data points for visibility; statistics are calculated from raw data. Lines show linear regression with gray dotted lines representing 95% confidence intervals. **D)** Images of triply labeled neurons with 3xALFA-LRRTM2 (yellow), PSD-95 (magenta), and surface AMPARs (green). Arrows indicate LRRTM2 clusters that do (white) or do not (red) co-enrich with PSD-95. **D)** Left: Pie chart depicting the proportions of LRRTM2, AMPAR, and/or PSD-95-containing puncta along knock-in dendrites. Right: Subset of left, AMPAR and/or LRRTM2 clusters without PSD-95 enrichment. Notably, LRRTM2 puncta outside of PSD-95 often also contain AMPAR clusters.


A
3xALFA-LRRTM2-Y501A/C504A Mutant (YACA)
3xALFA
ECD
TM
CTD
Control
(3xALFA-LRRTM2 WT)
Cas9 + BFP + G/D
Mutant
(3xALFA-LRRTM2 YACA)
Cas9 + BFP + G/D YACA
B
Surface 3xALFA
Surface 3xALFA
3xALFA-LRRTM2 WT KI
3xALFA-LRRTM2 YACA KI
3xALFA
5 µm
+ AMPARs
+ PSD-95
Merge
20 µm
C
Mean Intensity
125
100
75
50
25
**
WT
YACA
D
WT
YACA
Percent of Synapses
15
10
5
0
***
0
50
100
150
400
LRRTM2 Synaptic Mean Intensity
E
AMPAR Mean Intensity
1200
900
600
300
0
ns
WT
YACA

906 **Figure 5**: Simultaneous mutagenesis and tagging enabled by whole-CDS TKIT permits
907 analysis of mutation-induced changes in LRRTM2 surface expression
908 **A)** N-terminal 3xALFA tags were inserted in LRRTM2 as in Figure 2, as well as two point
909 mutations in the C-terminus (Y501A + C504A, or YACA). Control condition knock-ins
910 contain tagged wildtype (WT) sequence, YACA condition knock-ins contain both the tag
911 and the two point mutations. **B)** Exemplar images of WT or YACA-mutated LRRTM2
912 (yellow), co-stained for PSD-95 (magenta) and surface AMPARs (green). **C)** The YACA
913 mutation (navy) of LRRTM2 induces higher synaptic surface intensity than wildtype (tan),
914 averaged by cell ($p < 0.005$). Green lines represent means and black lines represent
915 medians for each distribution. **D)** Distribution of synaptic enrichment of WT or YACA
916 LRRTM2, showing a rightward shift of the YACA distribution (navy) relative to the WT
917 (tan) ($p < 2.2^{-16}$). Purple lines represent the mean, cyan the median of the respective
918 distributions. **E)** Synaptic surface AMPAR enrichment is not significantly higher in the
919 YACA-mutated condition relative to WT. Green lines represent means and black lines
920 represent medians for each distribution.


A
LRRTM2
Frequency
5000
4000
3000
2000
1000
0
0
1000
2000
3000
4000
5000
6000
7000
8000
9000
10000
CDS Length (bp)
B
Cumulative Frequency (%)
100
80
60
40
20
0
0
1000
2000
3000
4000
5000
6000
7000
8000
9000
10000
CDS Length (bp)

921 **Figure 6**: Applying whole-CDS replacement to other genes in *Rattus norvegicus*
922 **A)** Bioinformatics analysis of coding sequence span within the *Rattus norvegicus*
923 genome. Data procured from the Rnor6.0 sequence via UCSC Genome Browser. Results
924 were restricted to complete CDS sequences whose total span across the genome does
925 not exceed 10,000 bases. Arrow shows size of the LRRTM2 CDS, which spans a total of
926 1882bp from start codon in exon 1 to stop codon in exon 2, with our CRISPR replacement
927 of 2171 bp from guide 1 (intron 1) to guide 2 (3'UTR). **B)** Cumulative frequency distribution
928 of data in A, showing a considerable proportion of the sequences with a CDS spanning
929 region below 10kb are smaller than that of LRRTM2 and therefore potential targets for
930 this method.

931 ## References

932

933 1. Schroeder, A., and De Wit, J. (2018). Leucine-rich repeat-containing synaptic
934 adhesion molecules as organizers of synaptic specificity and diversity. Exp. Mol.
935 Med. *50*. 10.1038/s12276-017-0023-8.
936 2. Leshchyns'Ka, I., and Sytnyk, V. (2016). Synaptic Cell Adhesion Molecules in
937 Alzheimer's Disease. Neural Plast. *2016*. 10.1155/2016/6427537.
938 3. Südhof, T.C. (2008). Neuroligins and neurexins link synaptic function to cognitive
939 disease. Nature *455*, 903–911. 10.1038/nature07456.
940 4. de Wit, J., Sylwestrak, E., O'Sullivan, M.L., Otto, S., Tiglio, K., Savas, J.N., Yates,
941 J.R., Comoletti, D., Taylor, P., and Ghosh, A. (2009). LRRTM2 Interacts with
942 Neurexin1 and Regulates Excitatory Synapse Formation. Neuron *64*, 799–806.
943 10.1016/j.neuron.2009.12.019.
944 5. Ko, J., Fuccillo, M. V., Malenka, R.C., and Südhof, T.C. (2009). LRRTM2 Functions
945 as a Neurexin Ligand in Promoting Excitatory Synapse Formation. Neuron *64*, 791–
946 798. 10.1016/j.neuron.2009.12.012.
947 6. Soler-Llavina, G.J., Arstikaitis, P., Morishita, W., Ahmad, M., Südhof, T.C., and
948 Malenka, R.C. (2013). Leucine-rich repeat transmembrane proteins are essential
949 for maintenance of long-term potentiation. Neuron *79*, 439–446.
950 10.1016/j.neuron.2013.06.007.
951 7. Bhouri, M., Morishita, W., Temkin, P., Goswami, D., Kawabe, H., Brose, N., Südhof,
952 T.C., Craig, A.M., Siddiqui, T.J., and Malenka, R. (2018). Deletion of LRRTM1 and
953 LRRTM2 in adult mice impairs basal AMPA receptor transmission and LTP in
954 hippocampal CA1 pyramidal neurons. Proc. Natl. Acad. Sci. U. S. A. *115*, E5382–
955 E5389. 10.1073/pnas.1803280115.
956 8. Ramsey, A.M., Tang, A.H., LeGates, T.A., Gou, X.Z., Carbone, B.E., Thompson,
957 S.M., Biederer, T., and Blanpied, T.A. (2021). Subsynaptic positioning of AMPARs
958 by LRRTM2 controls synaptic strength. Sci. Adv. *7*, 1–19. 10.1126/sciadv.abf3126.
959 9. Ko, J., Soler-Llavina, G.J., Fuccillo, M. V., Malenka, R.C., and Südhof, T.C. (2011).
960 Neuroligins/LRRTMs prevent activity- and Ca 2+/calmodulin-dependent synapse
961 elimination in cultured neurons. J. Cell Biol. *194*, 323–334. 10.1083/jcb.201101072.
962 10. Linhoff, M.W., Laurén, J., Cassidy, R.M., Dobie, F.A., Takahashi, H., Nygaard, H.B.,
963 Airaksinen, M.S., Strittmatter, S.M., and Craig, A.M. (2009). An Unbiased
964 Expression Screen for Synaptogenic Proteins Identifies the LRRTM Protein Family
965 as Synaptic Organizers. Neuron *61*, 734–749. 10.1016/j.neuron.2009.01.017.
966 11. Lloyd, B.A., Han, Y., Roth, R., Zhang, B., and Aoto, J. (2023). Neurexin-3
967 subsynaptic densities are spatially distinct from Neurexin-1 and essential for
968 excitatory synapse nanoscale organization in the hippocampus. Nat. Commun. *14*.
969 10.1038/s41467-023-40419-2.
970 12. Fang, H., Bygrave, A.M., Roth, R.H., Johnson, R.C., and Huganir, R.L. (2021). An

971 optimized crispr/cas9 approach for precise genome editing in neurons. Elife *10*, 1–
972 25. 10.7554/eLife.65202.
973 13. Suzuki, K., Tsunekawa, Y., Hernandez-Benitez, R., Wu, J., Zhu, J., Kim, E.J.,
974 Hatanaka, F., Yamamoto, M., Araoka, T., Li, Z., et al. (2016). In vivo genome editing
975 via CRISPR/Cas9 mediated homology-independent targeted integration. Nature
976 *540*, 144–149. 10.1038/nature20565.
977 14. Arsić, A., Hagemann, C., Stajković, N., Schubert, T., and Nikić-Spiegel, I. (2022).
978 Minimal genetically encoded tags for fluorescent protein labeling in living neurons.
979 Nat. Commun. *13*. 10.1038/s41467-022-27956-y.
980 15. Ziak, J., Dorskind, J.M., Trigg, B., Sudarsanam, S., Jin, X.O., Hand, R.A., and
981 Kolodkin, A.L. (2024). Microtubule-binding protein MAP1B regulates interstitial
982 axon branching of cortical neurons via the tubulin tyrosination cycle. EMBO J. *43*.
983 10.1038/s44318-024-00050-3.
984 16. Götzke, H., Kilisch, M., Martínez-Carranza, M., Sograte-Idrissi, S., Rajavel, A.,
985 Schlichthaerle, T., Engels, N., Jungmann, R., Stenmark, P., Opazo, F., et al. (2019).
986 The ALFA-tag is a highly versatile tool for nanobody-based bioscience applications.
987 Nat. Commun. *10*, 1–12. 10.1038/s41467-019-12301-7.
988 17. Schnütgen, F., Doerflinger, N., Calléja, C., Wendling, O., Chambon, P., and
989 Ghyselinck, N.B. (2003). A directional strategy for monitoring Cre-mediated
990 recombination at the cellular level in the mouse. Nat. Biotechnol. *21*, 562–565.
991 10.1038/nbt811.
992 18. Dharmasri, P.A., Levy, A.D., and Blanpied, T.A. (2023). Differential nanoscale
993 organization of excitatory synapses onto excitatory vs inhibitory neurons. bioRxiv.
994 doi.org/10.1101/2023.09.06.556279.
995 19. Wang, Y., Wang, C., Ranefall, P., Broussard, G.J., Wang, Y., Shi, G., Lyu, B., Wu,
996 C.T., Wang, Y., Tian, L., et al. (2020). SynQuant: An automatic tool to quantify
997 synapses from microscopy images. Bioinformatics *36*, 1599–1606.
998 10.1093/bioinformatics/btz760.
999 20. Penn, A.C., Zhang, C.L., Georges, F., Royer, L., Breillat, C., Hosy, E., Petersen,
1000 J.D., Humeau, Y., and Choquet, D. (2017). Hippocampal LTP and contextual
1001 learning require surface diffusion of AMPA receptors. Nature *549*, 384–388.
1002 10.1038/nature23658.
1003 21. Choquet, D. (2018). Linking nanoscale dynamics of AMPA receptor organization to
1004 plasticity of excitatory synapses and learning. J. Neurosci. *38*, 9318–9329.
1005 10.1523/JNEUROSCI.2119-18.2018.
1006 22. Liouta, K., Chabbert, J., Benquet, S., Tessier, B., Studer, V., Sainlos, M., De Wit,
1007 J., Thoumine, O., and Chamma, I. (2021). Role of regulatory C-terminal motifs in
1008 synaptic confinement of LRRTM2. Biol. Cell, 1–24. 10.1111/boc.202100026.
1009 23. Minatohara, K., Murata, Y., Fujiyoshi, Y., and Doi, T. (2015). An intracellular domain
1010 with a novel sequence regulates cell surface expression and synaptic clustering of

1011 leucine-rich repeat transmembrane proteins in hippocampal neurons. J.
1012 Neurochem. *134*, 618–628. 10.1111/jnc.13159.
1013 24. Allen, D., Knop, O., Itkowitz, B., Kalter, N., Rosenberg, M., Iancu, O., Beider, K.,
1014 Lee, Y.N., Nagler, A., Somech, R., et al. (2023). CRISPR-Cas9 engineering of the
1015 RAG2 locus via complete coding sequence replacement for therapeutic
1016 applications. Nat. Commun. *14*. 10.1038/s41467-023-42036-5.
1017 25. Anzalone, A. V., Randolph, P.B., Davis, J.R., Sousa, A.A., Koblan, L.W., Levy, J.M.,
1018 Chen, P.J., Wilson, C., Newby, G.A., Raguram, A., et al. (2019). Search-and-
1019 replace genome editing without double-strand breaks or donor DNA. Nature *576*,
1020 149–157. 10.1038/s41586-019-1711-4.
1021 26. Anzalone, A. V., Gao, X.D., Podracky, C.J., Nelson, A.T., Koblan, L.W., Raguram,
1022 A., Levy, J.M., Mercer, J.A.M., and Liu, D.R. (2022). Programmable deletion,
1023 replacement, integration and inversion of large DNA sequences with twin prime
1024 editing. Nat. Biotechnol. *40*, 731–740. 10.1038/s41587-021-01133-w.
1025 27. Zhong, H., Ceballos, C.C., Massengill, C.I., Muniak, M.A., Ma, L., Qin, M., Petrie,
1026 S.K., and Mao, T. (2021). High-fidelity, efficient, and reversible labeling of
1027 endogenous proteins using crispr-based designer exon insertion. Elife *10*, 1–21.
1028 10.7554/eLife.64911.
1029 28. Droogers, W.J., Willems, J., Macgillavry, H.D., and de Jong, A.P.H. (2022). Duplex
1030 Labeling and Manipulation of Neuronal Proteins Using Sequential CRISPR/Cas9
1031 Gene Editing. eNeuro *9*. 10.1523/ENEURO.0056-22.2022.
1032 29. Marks, D., Bachmann, L., Villarejo, L.G., Geidies, A., Heinen, N., Anstatt, J., and
1033 Müller, T. (2021). The method of choice to knock-in large inserts via CRISPR.
1034 bioRxiv.
1035 30. Richardson, R.R., Steyert, M., Khim, S.N., Crutcher, G.W., Brandenburg, C.,
1036 Robertson, C.D., Romanowski, A.J., Inen, J., Altas, B., and Poulopoulos, A. (2023).
1037 Enhancing Precision and Efficiency of Cas9-Mediated Knockin Through
1038 Combinatorial Fusions of DNA Repair Proteins. Cris. J. *6*, 447–461.
1039 10.1089/crispr.2023.0036.
1040 31. Willems, J., de Jong, A.P.H., Scheefhals, N., Mertens, E., Catsburg, L.A.E.,
1041 Poorthuis, R.B., de Winter, F., Verhaagen, J., Meye, F.J., and MacGillavry, H.D.
1042 (2020). Orange: A CRISPR/Cas9-based genome editing toolbox for epitope tagging
1043 of endogenous proteins in neurons 10.1371/journal.pbio.3000665.
1044 32. Gao, Y., Hisey, E., Bradshaw, T.W.A., Erata, E., Brown, W.E., Courtland, J.L.,
1045 Uezu, A., Xiang, Y., Diao, Y., and Soderling, S.H. (2019). Plug-and-Play Protein
1046 Modification Using Homology-Independent Universal Genome Engineering.
1047 Neuron *103*, 583-597.e8. 10.1016/j.neuron.2019.05.047.
1048 33. Paul, A., Crow, M., Raudales, R., He, M., Gillis, J., and Huang, Z.J. (2017).
1049 Transcriptional Architecture of Synaptic Communication Delineates GABAergic
1050 Neuron Identity. Cell *171*, 522-539.e20. 10.1016/j.cell.2017.08.032.

1051 34. Földy, C., Darmanis, S., Aoto, J., Malenka, R.C., Quake, S.R., and Südhof, T.C.
1052 (2016). Single-cell RNAseq reveals cell adhesion molecule profiles in
1053 electrophysiologically defined neurons. Proc. Natl. Acad. Sci. U. S. A. *113*, E5222–
1054 E5231. 10.1073/pnas.1610155113.
1055 35. Lim, I.A., Hall, D.D., and Hell, J.W. (2002). Selectivity and promiscuity of the first
1056 and second PDZ domains of PSD-95 and synapse-associated protein 102. J. Biol.
1057 Chem. *277*, 21697–21711. 10.1074/jbc.M112339200.
1058 36. Metzbower, S.R., Dharmasri, P.A., Levy, A.D., Anderson, M.C., and Blanpied, T.A.
1059 (2023). Distinct SAP102 and PSD-95 nano-organization defines multiple types of
1060 synaptic scaffold protein domains at single synapses. bioRxiv.
1061 doi.org/10.1101/2023.09.12.557372.
1062 37. Elias, G.M., Funke, L., Stein, V., Grant, S.G., Bredt, D.S., and Nicoll, R.A. (2006).
1063 Synapse-Specific and Developmentally Regulated Targeting of AMPA Receptors
1064 by a Family of MAGUK Scaffolding Proteins. Neuron *52*, 307–320.
1065 10.1016/j.neuron.2006.09.012.
1066 38. Lein, E.S., Hawrylycz, M.J., Ao, N., Ayres, M., Bensinger, A., Bernard, A., Boe,
1067 A.F., Boguski, M.S., Brockway, K.S., Byrnes, E.J., et al. (2007). Genome-wide atlas
1068 of gene expression in the adult mouse brain. Nature *445*, 168–176.
1069 10.1038/nature05453.
1070 39. de Arce, K.P., Ribic, A., Chowdhury, D., Watters, K., Thompson, G.J.,
1071 Sanganahalli, B.G., Lippard, E.T.C., Rohlmann, A., Strittmatter, S.M., Missler, M.,
1072 et al. (2023). Concerted roles of LRRTM1 and SynCAM 1 in organizing prefrontal
1073 cortex synapses and cognitive functions. Nat. Commun. *14*. 10.1038/s41467-023-
1074 36042-w.
1075 40. Exposito-Alonso, D., Osório, C., Bernard, C., Pascual-García, S., Del Pino, I.,
1076 Marín, O., and Rico, B. (2020). Subcellular sorting of neuregulins controls the
1077 assembly of excitatory-inhibitory cortical circuits. Elife *9*, 1–28.
1078 10.7554/eLife.57000.
1079 41. Doench, J.G., Fusi, N., Sullender, M., Hegde, M., Vaimberg, E.W., Donovan, K.F.,
1080 Smith, I., Tothova, Z., Wilen, C., Orchard, R., et al. (2016). Optimized sgRNA design
1081 to maximize activity and minimize off-target effects of CRISPR-Cas9. Nat.
1082 Biotechnol. *34*, 184–191. 10.1038/nbt.3437.
1083 42. Hsu, P.D., Scott, D.A., Weinstein, J.A., Ran, F.A., Konermann, S., Agarwala, V., Li,
1084 Y., Fine, E.J., Wu, X., Shalem, O., et al. (2013). DNA targeting specificity of RNA-
1085 guided Cas9 nucleases. Nat. Biotechnol. *31*, 827–832. 10.1038/nbt.2647.
1086 43. Colussi, T.M., Costantino, D.A., Zhu, J., Donohue, J.P., Korostelev, A.A., Jaafar,
1087 Z.A., Plank, T.D.M., Noller, H.F., and Kieft, J.S. (2015). Initiation of translation in
1088 bacteria by a structured eukaryotic IRES RNA. Nature *519*, 110–113.
1089 10.1038/nature14219.
1090 44. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T.,

Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: An open-source platform for biological-image analysis. Nat. Methods *9*, 676–682. 10.1038/nmeth.2019.

45. Laine, R.F., Tosheva, K.L., Gustafsson, N., Gray, R.D.M., Almada, P., Albrecht, D., Risa, G.T., Hurtig, F., Lindås, A.C., Baum, B., et al. (2019). NanoJ: A high-performance open-source super-resolution microscopy toolbox. J. Phys. D. Appl. Phys. *52*. 10.1088/1361-6463/ab0261.