HHS Public Access

Author manuscript

Nat Methods. Author manuscript; available in PMC 2015 December 02.

Published in final edited form as:

Nat Methods. 2015 June ; 12(6): 535–540. doi:10.1038/nmeth.3360.

Rapid reverse genetic screening using CRISPR in zebrafish

Arish N Shah[1], **Crystal F Davey**[1], **Alex C Whitebirch**[1], **Adam C Miller**[1,2], and **Cecilia B Moens**[1,2]

[1]Division of Basic Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA, USA

Abstract

Identifying genes involved in biological processes is critical for understanding the molecular building blocks of life. The effectiveness of engineered CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) to efficiently mutate specific loci coupled with the accessibility of zebrafish (*Danio rerio*) provides an opportunity to screen for genes involved in vertebrate biological processes. Injection of Cas9-encoding mRNA and an engineered, single guide RNA (sgRNA) can cause biallelic mutations in injected embryos that phenocopy known mutant phenotypes. We found that increasing CRISPR efficiency and multiplexing sgRNAs allowed for phenocopy of known mutants across many phenotypes. We performed a proof-of-concept screen examining 48 loci by intersecting, multiplexed pool injections, and identified two new genes involved in electrical synapse formation. By deep-sequencing target loci we found that 90% of the genes were effectively screened. We conclude that CRISPR can be used as a powerful reverse genetic screening strategy *in vivo* in a vertebrate system.

Introduction

While classical forward and reverse genetic approaches have identified key molecular pathways required for life they are generally limiting in the number of targets that can be assessed and are very time intensive. Invertebrate animal models and cell culture systems have employed reverse genetic screening techniques to quickly identify genes and pathways involved in many biological processes. However, identifying a robust and inexpensive method in an *in vivo* vertebrate model system has been challenging. The bacterial and archaea adaptive defense mechanism CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) has been engineered to work in zebrafish (*Danio rerio*) to efficiently target the genome and create a double-stranded DNA breaks[1,2]. The DNA break is repaired by one of two cellular mechanisms: homology directed repair that creates a precise copy of a

Corresponding authors. Adam Miller, amiller@fhcrc.org, Cecilia Moens, cmoens@fhcrc.org.
[2]These authors jointly directed this work.

Accession numbers
BioProject accession PRJNA273396
SRA accession SRS824980

Author Contributions
ANS, CFD, ACW, and ACM performed experiments, acquired and quantified data, and generated images for publication. ACM, ANS, and CBM wrote the manuscript. All authors edited the manuscript.

Statement of Competing Financial Interests
None.

complementary sequence, or non-homologous end joining, which often introduces insertion and deletions (InDels) that can disrupt gene function[3,4]. The latter type of repair has been used in many systems to introduce mutations into genes of interest that, once fixed and carried in the genome, can be assessed for their effect on a process of interest[3,4]. Type II CRISPR systems are particularly attractive as a reverse genetic tool because they require only two components: 1) the Cas9 enzyme that has two enzymatic sites that cleave DNA and 2) an engineered, single guide RNA (sgRNA) that is ~100 nucleotides long[5]. The sgRNA includes a 20 nucleotide target sequence, which provides sequence specificity via Watson-Crick base pairing, and an ~100 nucleotide tail that recruits the Cas9 enzyme. The target recognition site is constrained only by the Cas9 enzyme's requirement to bind an NGG sequence motif adjacent to the 3' end of the sgRNA target sequence[5]. In zebrafish the injection of Cas9-encoding mRNA with sgRNA into one-cell stage embryos induces mutations that are transmissible through the germline[1,5]. Additionally, by codon-optimizing Cas9 and adding an additional nuclear-localization sequence the frequency of biallelic disruption increased such that injected (F_0) embryos phenocopied known mutant phenotypes[6]. Because of the ease of engineering the targeting specificity of CRISPRs we wanted to examine whether the system could be used as an efficient reverse genetic screening tool to assess many genes for their function in processes of interest in zebrafish.

We reasoned that an effective reverse genetic screening strategy would require an increased CRISPR InDel generation efficiency, which would produce the greatest potential to observe phenotypes. By varying Cas9 and sgRNA concentrations across three orders of magnitude we identified conditions that provide high phenotypic penetrance and low toxicity. The robust efficiencies for detecting mutant phenotypes allowed us to perform a reverse genetic screen in CRISPR-injected embryos of candidate genes involved in electrical synapse formation. Electrical synapses are a type of connection used by all nervous systems yet the genes required for their formation are not well understood. To reduce the number of injections we screened by pooled sgRNA injection and identified several pools with synapse phenotypes. De-multiplexing of these pools identified two new candidate genes, both of which were confirmed in germline-transmitted mutant animals. Targeted deep sequencing of all screen targets revealed a large range of locus-dependent InDel rates but confirm that we effectively assayed ~90% of the targets in the screen. Our results show that CRISPR can be used as a powerful and efficient genetic screening strategy *in vivo* in a vertebrate system.

Results

Optimizing CRISPR for identifying phenotypic effects

Our goal was to use CRISPR to quickly and efficiently screen genes for their involvement in any developmental process in zebrafish. Injection of Cas9 and a single guide RNA (sgRNA) can reproduce known mutant phenotypes[6], suggesting that new genes involved in a process could be identified in injected (F_0) embryos or larvae. However, injecting sgRNA and Cas9-encoding mRNA leads to mosaic embryos with variable numbers of mutant cells in any given animal. This is due to a lag in Cas9 protein generation, with the first mutations being detectable only when there are more than 1000 cells in the embryo[7]. We therefore wanted to increase the mutational load in injected embryos to provide the best chance of identifying

whether a gene was involved in the process. To optimize conditions we used the *slc24a5* (*golden*) locus to examine the effects of varying Cas9 and sgRNA concentration for phenotypic identification and generalized toxicity. Mutation of the *slc24a5* gene leads to a loss of pigmentation, which is easily screenable in the retinal pigment epithelium of the eye (Supplementary Fig. 1a). We used Cas9, which was previously codon-optimized for zebrafish and has two nuclear localization sequences, and a sgRNA against *slc24a5* that efficiently causes InDels in the zebrafish genome (Supplementary Fig. 1a–c)[6]. We first varied the amount of Cas9-encoding mRNA from 75 to 2400 pg while using a constant 100 pg of *slc24a5* sgRNA in injections into one-cell stage embryos and tested for the amount of pigment loss in the eye. We found that in all experimental injection conditions there were embryos that had reduced pigmentation in the eye, but that there was variability in the phenotype with embryos ranging from no pigment loss to complete loss (Supplementary Fig. 1d). The reduction in pigmentation in all injections suggested a high mutational load; indeed, we found that 100% (24 of 24) of individually cloned and sequenced alleles from the 100/1200 pg *slc24a5/cas9* injection had InDels (Supplementary Fig. 1c). Increasing Cas9 concentration led to a greater number of embryos displaying a loss of pigment, and more embryos per injection with a complete loss of pigment, i.e. they phenocopied the homozygous mutant (Supplementary Fig. 1d). However, increasing the Cas9 concentration also led to increased toxicity with phenotypes ranging from death several hours after injection to general problems with heart (edema), nervous system (cell death), and axis formation (dorsalized and ventralized embryos) (Supplementary Fig. 1e); such phenotypes are a common side effect of injecting nucleic acids in zebrafish. Because a typical zebrafish clutch from a single female is 100 embryos, and in our experiments we typically combine multiple clutches and inject 120–150 embryos, we decided that 1200 pg of Cas9 gave a reasonable balance of phenotypic efficiency with acceptable toxicity and used this concentration to test the effect of varying sgRNA concentrations from 10 to 1000 pg. Increasing sgRNA levels increased both the loss of pigment (not shown) as well as the toxicity (Supplementary Fig. 1f). However, even at 1200 pg of Cas9-encoding mRNA and 1000 pg of *slc24a5* sgRNA we found that only ~30% of the clutch was unviable for phenotypic screening due to toxicity.

We wondered if the conditions we identified for CRISPR injection would be broadly applicable to interrogating gene function in more complex phenotypes. We examined genes involved in facial motor neuron migration (*pk1b, vangl2*) and mesodermal convergence and extension (*vangl2*) and were able to recapitulate published phenotypes (Supplementary Fig. 2)[8,9]. Our results validate the idea that despite the genetic mosaicism in injected embryos, CRISPR can be used to efficiently identify genes involved in a variety of biological processes of interest including both cell autonomous and non-cell autonomous phenotypes.

Electrical synapses as a model for CRISPR screening

Our increased phenotypic efficiencies using high Cas9 and sgRNA concentrations suggested we could now easily screen for new genes involved in any process of interest. We chose to optimize the CRISPR screening strategy for electrical synapse formation in the Mauthner (M) circuit of the spinal cord. This system is accessible, easily quantifiable, and little is known about the pathways involved in electrical synapse formation. Each of the ~60

identified electrical synapses in the spinal cord forms between the descending axon of M and a single segmentally-repeated CoLo neuron (Fig. 1a)[10]. This provides a simple readout of efficiency and mosaicism within CRISPR injected embryos, since biallelic loss of gene function in a single cell is expected to lead to loss of its electrical synapse. Electrical synapses are formed by hexamers of Connexins (Cxs) that form gap junction (GJ) channels between neurons[11]. The M circuit electrical synapses can be visualized by immunostaining with an antibody against the human Cx36 protein, which detects the zebrafish Cx36-related proteins (Fig. 1b)[12]. In a separate forward genetic screen we identified the gene *gap junction delta 1a* (*gjd1a*), which encodes zebrafish Cx34.1, as being required for M electrical synapse formation (Fig. 1c) (A.C.M., A.C.W., A.N.S., and C.B.M., unpublished data). *gjd1a* is a homologue of the mammalian *gjd2*, which encodes the major Cx protein expressed in the mammalian brain[11]. To confirm that we could use CRISPR/Cas9 to detect electrical synapse phenotypes, we first designed two sgRNAs targeting the *gjd1a locus* (crispr.mit.edu)[3], which overlapped each other in target sequence by 16 nucleotides. Each was injected independently at 100/1200 pg sgRNA/*cas9* mRNA and both caused the loss of electrical synapses in 100% of injected embryos (Fig. 1d–f). However, the frequency of synapse loss per embryo differed, with injection of the sgRNA against Target1 leading to 95% of embryos lacking more than 66% of synapses in the spinal cord, while Target2 led to 60% of embryos with similar synapse defects (Fig. 1f). We found that these phenotypic differences between sgRNA targets correlated with the efficiency of altering the genomic locus as assessed by qPCR directly from genomic DNA (Fig. 1g, see Methods). In this instance, higher efficiency correlated with higher purine content directly preceding the PAM sequence, as has been recently demonstrated on a genome-wide scale[13]. Finally, we tested the range of concentrations at which the *gjd1a* sgRNA target1 could produce electrical synapse phenotypes and found that it caused synapse loss after injection of as little as 0.19 pg of sgRNA, an amount that caused virtually no toxicity as compared to uninjected controls (Supplementary Fig. 3). We conclude that the electrical synapse phenotype is accessible to CRISPR screening.

Optimizing conditions for pooled sgRNA injection

Our goal was to efficiently screen a large number of candidate genes, so we tested the effectiveness of pooling several sgRNAs into a single injection to examine electrical synapse loss. Pooling of sgRNAs was previously shown to result in mutations at multiple genomic sites but also to decrease InDel efficiency at each locus[6]. To test the effect of pooling we combined sgRNAs for *gjd1a* and five other genes not involved in electrical synapse formation (*gfp, slc24a5, pk1b, sox10, hmcn1*; each was tested individually and none affected synapse formation; not shown). We tested a range of sgRNA concentrations from 0.39 to 400 pg for each individual sgRNA in the pool and coinjected with 1200 pg of Cas9-encoding mRNA. All injected embryos displayed a loss of electrical synapses with increasing concentrations resulting in increased phenotypic penetrance (Fig. 1h). Increasing the total amount of sgRNA injected increased toxicity with greater than 200 pg of each guide leading to the death of all injected embryos (Fig. 1i). However, injecting 12.5 pg of each sgRNA resulted in ~50% of embryos robustly lacking synapses yet led to only ~30% toxicity in embryos. We conclude that multiplexing sgRNAs into a single injection is a feasible way to reduce the number of injections necessary to screen multiple genes.

CRISPR screen using multiplexed pools of sgRNAs

We designed a set of 48 sgRNAs targeting genes potentially involved in synaptogenesis. Our list included all of the zebrafish *gjd2* homologues, genes encoding proteins that biochemically interact with Cxs, and the Neuroligin–Neurexin gene families and related members, which are involved in chemical synaptogenesis (Fig. 2a and Supplementary Table 1)[11,14]. Within the screen we included *gjd1a* as a positive control, but none of the other genes had a previously known requirement for electrical synaptogenesis. Pooled screening presents the problem that any phenotypically positive pool could be due to the loss of any of the target genes within the pool. Positive pools can be de-multiplexed by separate injections of each of the sgRNAs. However, we reasoned that we could reduce the number of injections overall if we arrayed our sgRNAs in 96-well plate format and injected sgRNAs pooled by row and column (Fig. 2a). In this way, each sgRNA is injected twice in a separate pool and the intersection of a positive hit in both the row and a column pools should reveal the most likely candidate gene. Whether row and column pooling reduces the number of total injections depends on the number of positive pools identified, but we felt that given the broad categories of genes included in the screen it was likely that we would only identify a small number of positives. When multiplexing sgRNAs care must be taken if the targets are on the same chromosome as large deletions between the sites can occur, with reported deletions up to 1 Mb in zebrafish[15]. The loss of multiple genes within a large deletion would complicate analysis and should be avoided in this context. Our final consideration in pooling was to plate the sgRNAs such that gene and family members were pooled in either a row or column. Given the high efficiency of CRISPR mutation in injected embryos, phenotypes requiring the removal of redundant genes could be revealed if present.

We used a constant 1200 pg of Cas9-encoding mRNA for all injections while column pools were at 12.5 pg of each sgRNA and row pools were at 16.5 pg each — this maintained the same amount of total sgRNA in each pool injection (100 pg total). Injection of the column and row pools resulted in groups of embryos that were grossly normal with levels of toxicity ranging from 9.3 to 33.6%. After Cx36 immunostaining we detected three row and three column pools that had synapse defects ranging from 30% to 100% of injected embryos (Fig. 2b). This set of row and column hits produced an intersection of nine genes that are the most likely to be causative for the synapse phenotypes (Fig. 2a). One of the overlaps was the positive control *gjd1a*, while the others represented two of the other *gjd2* homologues, the electrical synapse scaffold *tjp1b* (ZO-1), and two of the *neurexin* genes (Fig. 2a). We injected each of these sgRNAs individually and found a loss of electrical synapses when targeting *gjd2a* and *tjp1b* (targeting two different isoforms of *tjp1b, tjp1b_L and tjp1b_B*, gave similar synapse loss), while the others had no effect (Fig. 2c). We also examined whether the other genes within the positive rows and columns had phenotypes (i.e. the non-intersecting sgRNAs within the positive row/columns) by injecting them in pools and found that they did not affect synapses (not shown). We confirmed that both *gjd2a* and *tjp1b* are required for M electrical synapse formation by raising F_0 animals carrying deletions in each of the genes to adulthood and crossing carriers. We found that mutant embryos with biallelic frameshift mutations in either *gjd2a* or *tjp1b* lack electrical synapses in the spinal cord (Fig. 2d–h). Our 48-gene CRISPR screen, including sgRNA synthesis, pool injections, and de-

multiplexing, took a total of three weeks and identified new genes involved in electrical synaptogenesis.

Sequencing reveals most targets were successfully screened

To independently confirm that the genes we targeted with sgRNAs had been screened, we assessed the injected embryos' genotype to phenotype correlation by analyzing the frequency of InDels at all target sites from screened animals. We amplified each target from genomic DNA of 10 animals for each of the row and column pool injections (*e.g.* we amplified and barcoded *gjd1a* from column 1 and row A, etc.). We also amplified regions for the predicted off-target sites based on sequence relatedness to our phenotypically positive hits (Supplementary Table 1)[16]. We sequenced all amplicons on an Illumina MiSeq machine, obtaining 136,000 to 800,000 reads for each target. We assessed the frequency of InDels at each locus by counting the number of insertions, deletions, and wildtype alleles. We found that the majority of targets had InDel frequencies that ranged from 22% to 85% of the sequenced alleles, depending on target locus. The phenotypically positive hits were found at the top (85% non-wildtype reads; *gjd1a*) and near the bottom (25% non-wildtype reads; *tjp1b_L* and *gjd2a*) of this InDel frequency range (Fig. 2i). This suggests that most of the targets (42 of 48) were successfully screened.

We analyzed whether the broad range of mutational efficiencies observed at individual targets in our screen was influenced by recently identified sequence composition rules[13,17]. We found that increased GC content within the guide and purines directly upstream of the NGG correlated with increased mutational efficiency (Supplementary Fig. 4a–c). Further analysis of the NGS data revealed that across all injections, out-of-frame InDels ranged from 48 to 77% of non-wildtype reads (Supplementary Fig. 4d, avg. = 61.5 s.e.m. = 1.28). Deletions were centered near the NGG of the target sequence with a 5' bias in the location of most deletions (Supplementary Fig. 4e). The majority of deletions are less than 10 bp but they range from 1 to 198 bp (Supplementary Fig. 4f, 198 bp is the maximum detectable size of deletion using our amplicon sequencing and analysis). Importantly, we found that InDel frequency at off-target sites ranged from 0.16 to 3.17%, well below the frequency for our phenotypic hits (Fig. 2i and Supplementary Table 1). This suggests that off-target sites are unlikely to be causative for the phenotypes assessed.

The phenotypically-positive guides with low mutation rates (*tjp1b_L*, *gjd2a*) resulted in surprisingly high proportions of injected embryos displaying phenotypes (Fig. 2b). To address this discrepancy we examined the correlation between mutation load and the frequency of synapse loss at the *tjp1b* locus in individual embryos injected with 16.5/1200 pg *tjp1b_L/cas9* mRNA. We found that F_0 embryos had a wide range of mutation frequencies but that individual genotypes and electrical synapse phenotypes were highly correlated. Surprisingly, the phenotypic rate approached 100% in some embryos (Supplementary Fig. 5a). This suggests that phenotypes are manifest either with mono-allelic loss (heterozygous) or that in-frame deletions lead to a loss of gene function, or both. We found that heterozygosity at the *tjp1b* locus, with germline-transmitted out-of-frame or in-frame deletions, did not affect M electrical synapses (Supplementary Fig. 5b,c). By contrast, an in-frame deletion in trans to an out-of-frame deletion caused a loss of electrical

synapses (Supplementary Fig. 5d), confirming that in-frame changes at this *tjp1b* site are deleterious to function. For each of the genes identified in the screen that were required for electrical synapse formation the CRISPR target sites were in highly-conserved regions of the protein (>85% amino acid similarity compared to human), suggesting that such regions may be intolerant to in-frame InDels.

Six of our targets had InDel frequencies similar to the off target sites. We wondered whether this was due to inaccessibility of the Cas9-sgRNA complex to these genomic sites or instead to synthesis error during library preparation. We found that five of these failed sgRNAs had degraded during preparation (as determined subsequently on an Agilent TapeStation, not shown). We resynthesized each of them and found that all now effectively altered the genome based on qPCR (Supplementary Fig. 6a, none had synapse phenotypes, not shown). The one sgRNA that still failed to alter the genome had two single nucleotide polymorphisms in our designed sgRNA compared to the genome into which we injected (Supplementary Fig. 6b). We conclude that our optimized screening strategy robustly screened greater than ~90% of the targets even in the face of varying sgRNAs efficiencies.

Discussion

Here we show that CRISPR can be used as a reverse genetic screening tool to identify new genes for a wide range of biological processes in zebrafish. Reverse genetic screening methods such as RNAi have been powerful for defining genetic pathways in invertebrate model systems and cultured cells *in vitro*, but efficient methods have been missing from vertebrate model systems. The pilot CRISPR screen we performed took a total of three weeks and identified two previously unknown genes involved in electrical synaptogenesis, highlighting the usefulness of the approach. Our “row by column” multiplexed, intersectional-pooling strategy provided an efficient way to identify the most likely candidates. The pooling strategy provides addition benefit as, in principle, it could reveal phenotypes that require the removal of multiple genes. This property is particularly important given the partial genome duplication present in the zebrafish genome[18], but will be phenotype specific and dependent on the efficiency of mutagenesis at individual target loci. By sequencing all the targets in our screen we found a large degree of target-dependent variability in mutational efficiencies. However, because we found phenotypes at the peak and near the lowest efficiencies we are confident that we screened ~90% of our targets. While in this study we used Cas9 that had been engineered for zebrafish efficiency[6], future efforts that increase mutagenicity would be beneficial for screening. While both morpholinos and RNAi have been used in zebrafish to knockdown genes of interest, both technologies suffer from off-target effects that have limited their usefulness[19,20]. We and others have found that the rates of mutagenesis at potential off-target sites using CRISPR are low (1–3%)[21], suggesting they are unlikely to cause false positives within the screen. The breadth of accessible phenotypes and ease of engineering the sgRNAs to target unique genes suggest that the approach presented here will be broadly useful to identify candidates for many biological processes.

While CRISPR screening represents a powerful reverse genetic approach to identify candidate genes, it is important to note that the positive candidates should be viewed with

some caution. Due to the delay in Cas9 translation and folding, injected CRISPR animals are mosaic for multiple, independent genetic lesions with each F_0 animal having a unique spectrum of mutations. Genetic mosaicism can complicate phenotypic analysis; for example, competition or sorting between wildtype and normal cells can give rise to phenotypes not observed in fully mutant individuals[22]. Furthermore, even if an injected embryo had InDels introduced to each individual cell's target site, one third of InDels will be in-frame and therefore potentially hypomorphic or silent depending on the target site. Thus even with a highly effective sgRNA, less than half ($2/3^2 = 4/9$) of cells in injected embryos are expected to have bi-allelic frameshift mutations. While in-frame deletions might often be tolerated, we found that they can cause deleterious loss of function. Therefore care in selecting the targeted region of the protein, for example in conserved functional domains, may provide increased mutagenicity. However, given the mosaicism, potential for phenotypes in heterozygous cells, and the unknown genetic state of each cell within the animal examined, the phenotypes of F_0 injected animals should be interpreted with caution. Instead, affected F_0 embryos, or embryos injected with sub-phenotypic amounts of sgRNA, should be raised to generate offspring with stably inherited mutations, which can be assessed to confirm the phenotypes and extend the analysis of mutant defects. While there are caveats given the mosaic nature of the F_0 embryos, the CRISPR screen hits are likely to have some function for the process being investigated, and the system therefore presents a powerful first step in identifying new genes.

The success of our pilot screen paves the way towards applying CRISPR screening at a larger scale. Recent methods using lentiviral based Cas9-sgRNA libraries have been utilized to screen at a genome scale in cultured cells[13,23]. While genome scale screens would be challenging in zebrafish given the number of animals that would need to be injected, screens on the order of hundreds to thousands of genes are reasonable. Our method of sgRNA synthesis was simple, efficient, and inexpensive (~$20 per sgRNA injection). Within a large screen NGS analysis of injected embryos could be done through DNA collection and storage followed by subsequent sequencing to assess targets screened. Alternatively, sgRNA efficiency can be tested inexpensively by qPCR (see Methods). The CRISPR screening approach presented here will be particularly useful in identifying the highest priority candidates from high-throughput techniques such as RNA-seq and mass spectrometry, as well as for initial testing of candidate human disease genes identified by GWAS or exome sequencing. Finally, while we have focused on zebrafish, the recapitulation of known mutant phenotypes in CRISPR injected F_0 embryos has been observed in many genetic animal models[24–27]; thus approaches similar to those presented here are likely to work in a number of systems. Overall, CRISPR reverse genetic screening provides an efficient, cost effective approach that has broad application for the rapid identification of new genes involved in many biological processes.

Methods

sgRNA design and template synthesis

Single guide RNAs were designed using the design tool at crispr.mit.edu, which finds and ranks all 23 bp sgRNA sequences ending in the NGG motif[16]. It also outputs predictable off-

target sites based on sequence relatedness. When possible, sgRNAs chosen had a score of 90+ (Supplementary Table 1). Due to a restriction of the T7 RNA polymerase in our synthesis, only sgRNAs that began with a 5' G nucleotide were chosen. To synthesize the template DNA required for the *in vitro* transcription we employed a two oligo PCR method[24]. First, an oligo scaffold containing the RNA loop structure required for recognition by the Cas9 enzyme was synthesized — this was common to all of the sgRNAs generated. Next, a unique oligo containing a T7 binding site, the 20 nucleotides specific to the sgRNA, and 20 bases of homology to the scaffold oligo was synthesized. PCR was performed using these two oligos such that they template off one another and the full sgRNA sequence was created (Supplementary Fig. 7). The scaffold oligo sequence is: 5'[gatccgcaccgactcggtgccacttttcaagttgataacggactagccttattttaacttgctatttctagctctaaaac]3' and should be HPLC purified.

The gene specific oligo sequence is: 5'[aattaatacgactcactata(N20)gttttagagctagaaatagc]3' where (N20) refers to the 20 nucleotides of the sgRNA that will bind the genome (excluding the NGG motif). The full length PCR product then should be: 5'[aattaatacgactcactata(N20)gttttagagctagaaatagcaagttaaaataaggctagtccgttatcaacttgaaaaagtggcaccgagtcggtgcggatc]3'.

The PCR reaction included: 2.5 μL H2O, 12.5 μL 2× Phusion master mix (New England BioLabs, M0531L), 5 μL Scaffold Oligo (10 μM, synthesized at Eurofins Genomics), 5 μL sgRNA Oligo (10 μM, synthesized at Eurofins Genomics), and was run in a thermocycler under the following program: 95 °C for 30 sec; 40 cycles of 95 °C for 10 sec, 60 °C for 10 sec, 72 °C for 10 sec; followed by 72 °C for 5 min. This PCR product was purified and used as template for the *in vitro* transcription reaction. All DNA purifications were performed on columns (Zymo Research, D4014).

RNA *in vitro* transcription for sgRNA and Cas9

For Cas9-encoding mRNA synthesis, pT3Ts-nCas9n was obtained from the Chen lab via Addgene (46757)[6]. The plasmid was linearized using XbaI (New England BioLabs, R0145S) and DNA was purified. 1 μg of linear plasmid was used in an in vitro transcription reaction (T3 mMessage mMachine, Life Technologies, AM1348M). For sgRNA synthesis, 1 μg of guide template PCR product was used in a T7 in vitro transcription half reaction (MEGAscript T7, Life Technologies, AM1334M). Both RNA products were cleaned either by phenol-chloroform and isopropanol precipitation or by column (Zymo Research, R1016). A Nanodrop spectrophotometer was used to ensure purity and check the concentration. We recommend using an Agilent TapeStation or equivalent technology to assay the integrity of the synthesized RNAs as a further quality control measure.

Cas9-encoding mRNA and sgRNA injection

Unless otherwise stated, all injections contained 1200 ng/μL Cas9-encoding mRNA. All sgRNAs for the screen were diluted to 200 ng/μL prior to multiplexing. Therefore, column pools of 8 contained 12.5 ng/μL of each sgRNA and row pools of 6 contained 16.6 ng/μL of each sgRNA. 1 nL of Cas9 and sgRNA mix was used for all injections.

DNA isolation

Embryos were lysed in 30 µl alkaline lysis buffer (25 mM NaOH, 0.2 mM EDTA) and heated at 95 °C for 30 minutes. The solution was neutralized by adding an equal volume of neutralization buffer (40 mM Tris-HCl, pH 5.0). Samples were spun at 3,000 RPM for 5 minutes and the supernatant was transferred into new tubes.

qPCR quantification of CRISPR efficiency

We modified the method described by Yu *et. al.*[28] to use a genomic DNA template directly for qPCR, allowing us to quantify the efficiency of CRISPR-mediated mutagenesis at the target locus without requiring that the experimental and control template DNA be normalized. The method utilizes the fact that the binding of a primer overlapping the sgRNA site will be compromised in successfully mutagenized embryos, resulting in delayed amplification, while binding of a flanking primer pair will be unaffected. The "OUT" (flanking) primer pair encompasses at least 100 bp surrounding the sgRNA binding region. The "ON" (overlapping) primer pair uses one of the OUT primers, and another primer that binds the 20 bp of the sgRNA target sequence. The 3’ side of the “ON” primer binds just 5’ of the NGG as most InDels affect the −1 to −10 positions of the binding site (Supplementary Fig. 4e). At 2 days post fertilization (dpf), genomic DNA was isolated from five uninjected and five injected embryos (in triplicate). Using 1 µl of this template DNA, separate 20 µL OUT and ON qPCR reactions were run. qPCR was performed with SsoAdvanced SYBR Green Supermix (Bio-Rad, 172-5271) by using a Bio-Rad CFX96 Real-Time system. All primers are listed in Supplementary Table 1. The qPCR conditions were programmed as follows: 40 cycles of 95 °C for 10 sec and 60 °C for 20 sec after initial denaturing at 95 °C for 30 sec.

The ratio of the qPCR Cq values of the uninjected ON and uninjected OUT reflects the differences in amplification of the two primer pairs on uninjected template DNA. This may be due to inherent differences in amplification that exist even between perfectly complementary primer pairs. By contrast, the ON/OUT ratio in injected embryos reflects both this difference in amplification between the primer pairs and the loss of the ON binding site due to CRISPR-introduced InDels. A comparison of the ON/OUT Cq ratios of injected vs uninjected embryos thus reflects the efficiency of mutagenesis.

To examine how accurate the qPCR method estimates mutational efficiency we directly compared it to NGS analysis across eight independent screen targets. We used genomic DNA from the screen pools and determined the mutational efficiency using qPCR. We found that the qPCR method underestimates the mutation frequency as compared to the NGS analysis from 6% to 15% with an average underestimation of 10% (Supplementary Fig. 5, source data).

NGS prep and sequencing

We employed a 2-step PCR method to create the Illumina sequencing library. This requires special primer overhangs on the gene-specific primers in order to act as template for the Illumina barcode PCR. First, gene-specific primers were designed for each of the 48 targets as well as seven of the most likely off-target sites for the phenotypically positive hits

identified in the screen. The forward primers had the tag: 5' ACACTCTTTCCCTACACGACGCTCTTCCGATCT 3' appended to the 5' side of the gene-specific portion. The reverse primers had the tag: 5' GACTGGAGTTCAGACGTGTGCTCTTCCGATCT 3' appended to the 5' side of the gene-specific portion. Amplicons were in the range of 250–350 bp in length.

DNA was extracted from a pool of ten injected fish from each row and column in the CRISPR screen and used as a template for amplification of the region around the corresponding target site. The gene-specific PCRs were done for both the row and column pools. DNA concentration and amplification sizes were assessed on a quantifying gel. Equal amounts of each product were mixed in two separate row and column master pools containing all 55 PCR products. These two master pools were cleaned (Zymo Research, D4014) and used as template for the Illumina barcode PCR reaction (Nextera Index kit, Illumina, FC-121-1012). This second PCR product was cleaned and target libraries were prepared and then sequenced using Illumina's MiSeq Desktop Sequencer. Briefly, the target library was denatured, diluted to 15 pM, spiked with a premade PhiX control library at 5% (Illumina, FC-110-3001), loaded into a MiSeq v2 Reagent Kit (Illumina, MS-102-2003). Sequencing generated paired-end (2 × 250 bp) dual-indexed reads. Following sequencing, reads were de-multiplexed based on Illumina barcodes with the MiSeq Reporter software and stored as FASTQ files for downstream processing and analysis. All sequencing data has been submitted to NCBI Sequence Read Archive (http://www.ncbi.nlm.nih.gov/sra) and the project can be found with the BioProject accession PRJNA273396, and sequencing files can be found with the accession number SRS824980.

NGS analysis

SeqPrep was used to merge the paired-end reads (https://github.com/jstjohn/SeqPrep). Based on the amplicon length we chose for all targets, the forward and reverse reads from sequencing should be identical. We previously found that merging reads and keeping only those that matched 100% reduces the amount of sequencing error seen in the analysis [https://github.com/jstjohn/SeqPrep, -g -n 1.0 -s][29]. This removes all unpaired reads and collapses the forward and reverse pairs into a single read. Using NGSutils, each of the two resulting fastq files (column and row) was split into 55 individual FASTQ files corresponding to the 55 amplicons using the forward primer as a barcode sequence [-edit 1 -pos 1 -allow-revcomp -stats][30]. This resulted in 55 fastq files from the columns and 55 fastq files from the rows, each containing all the reads for a single amplicon. We used CRISPR-GA to align the reads and assess the percentage of NHEJ as well as to process information on the location and size of each deletion (http://54.80.152.219/)[31]. The data from CRISPR-GA was output in XML format and was parsed via the R programming language to assess and plot the InDel frequencies, the percent of out-of-frame deletions, and the positions and sizes of deletions.

Fish, lines, and maintenance

All animals were raised in an Institutional Animal Care and Use Committee (IACUC)-approved facility at the Fred Hutchinson Cancer Research Center. Zebrafish (*Danio rerio*) were bred and maintained as previously described[32]. Animal care is provided by Rachel

Garcia and veterinary care is provided by Dr. Rajesh *K*. Uthamanthil, DVM. Wildtype animals were from a mixed *AB/Tu background. The facial branchiomotor neurons were visualized using the transgenic reporter line *Tg(isl1:GFP)rw0*[33]. Separate from the CRISPR mutagenesis described herein, we created mutations in the *gjd2-family* of genes using classical forward genetic screening using ENU as a chemical mutagen *(gjd1a*^{fh360}*)*, TALENs (*gjd1a*^{fh436}, *gjd1b*^{fh435}, *gjd2a*^{fh437}) and TILLING (*gjd2b*^{fh329}). For *gjd1a*^{fh436}, heterozygous carriers were grown to adulthood and crossed to generate homozygous mutant embryos (8 bp deletion / 8 bp deletion); these mutants lacked electrical synapses, confirming *gjd1a's* requirement for electrical synapse formation in the M circuit (A.C.M., A.C.W., A.N.S., C.B.M. unpublished data). *gjd2a*^{fh437} was used to confirm the CRISPR-induced phenotype identified in the CRISPR screen described herein. The final two *gjd2*-family members (*gjd1b*^{fh435} and *gjd2b*^{fh329}) have no M-circuit phenotypes as homozygous mutants (not shown). For *tjp1b*, F_0 embryos carrying germline mutations were identified by crossing individual animals to wildtypes and analysis the sequence of progeny at the sgRNA's target site. Three male and three female animals were identified carrying mutations in their germline, each animal carrying its own unique spectrum of mutations. Animals carrying mutations (10, 5, and 18 bp deletions — *tjp1b*^{fh448}, *tjp1b*^{fh449}, *tjp1b*^{fh451}, respectively) were crossed and progeny were genotyped for changes to the *tjp1b* locus and phenotyped for electrical synapse defects.

Immunohistochemistry

Embryos were fixed and stained using standard procedures. Briefly, anesthetized embryos at 5 dpf were fixed in 2% trichloroacetic (TCA) acid for 3 hours. Fixed tissue was then washed in PBS + 0.5% TritonX100, followed by standard blocking and antibody incubations. Tissue was cleared step-wise in a 25%, 50%, 75% glycerol series and was dissected and mounted for imaging. Primary antibodies used were rabbit anti-human-Cx36 (Invitrogen, 36-4600, 1:200) and mouse anti-RMO44 (Life Technologies, 13-0500, 1:100). Secondary antibodies were goat-anti-rabbit and goat-anti-mouse conjugated to 555 and 633 fluorophores, respectively (Life Technologies, A-21428, A-21052, 1:250).

Phenotyping, imaging, and data analysis

Electrical synapse phenotypes were screened at 4 dpf using fixed and stained embryos on a Zeiss Cell Observer Spinning Disc confocal microscope. These were binned into phenotypic categories of wildtype (no loss), < 33% loss, 33–66% loss, and > 66% loss based on the presence/absence of electrical synapses along the length of the spinal cord. Retinal pigment was quantified using images of 2 dpf zebrafish taken on a Zeiss Lumar stereomicroscope with attached black and white digital camera. Using the Fiji imaging distribution of ImageJ[34], the mean pixel grey value of the eye was measured using a standard region of interest that encompassed the area of the eye. Values were normalized to the average wildtype and *slc24a5*$^{-/-}$ pixel intensity values. Motor neuron and convergent extension phenotypes were screened live at 2 dpf in the *Tg(isl1:GFP)rw0* transgenic background using a Leica stereomicroscope. Phenotypic analysis was not performed blind: in all cases phenotypes were visualized first followed by analysis of genotype. All animals were examined phenotypically for effects on toxicity. Animals with defects associated with toxicity were removed from further analysis. Synapse and motor neuron images presented

were collected on a Zeiss LSM 700 confocal microscope using the 488, 555, and 639 laser lines. The synapse images presented in Supplementary Fig. 5 were collected on a Zeiss Cell Observer Spinning Disc confocal microscope using the 555 and 633 laser lines. Figure images were created using Photoshop (Adobe) and Illustrator (Adobe).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank R. Garcia for superb animal care, the Moens lab for discussion and editing, J. Spiewak and the Parichy lab at the University of Washington for *slc24a5* mutant embryos, S. Peterson and the Postlethwait lab at the University of Oregon for advice on sgRNA cloning, and I. Phelps and D. Doherty for help and access to the MiSeq machine. Funding was provided by the National Institute of Health, R01HD076585 and R21NS076950 to C.B.M. and K99NS085035 to A.C.M.

References

1. Chang N, et al. Genome editing with RNA-guided Cas9 nuclease in zebrafish embryos. Cell Res. 2013; 23:465–472. [PubMed: 23528705]
2. Hwang WY, et al. Efficient genome editing in zebrafish using a CRISPR-Cas system. Nat Biotechnol. 2013; 31:227–229. [PubMed: 23360964]
3. Hsu PD, Lander ES, Zhang F. Development and applications of CRISPR-Cas9 for genome engineering. Cell. 2014; 157:1262–1278. [PubMed: 24906146]
4. Peng Y, et al. Making designer mutants in model organisms. Development. 2014; 141:4042–4054. [PubMed: 25336735]
5. Jinek M, et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science. 2012; 337:816–821. [PubMed: 22745249]
6. Jao L-E, Wente SR, Chen W. Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. Proc. Natl. Acad. Sci. U.S.A. 2013; 110:13904–13909. [PubMed: 23918387]
7. Sung YH, et al. Highly efficient gene knockout in mice and zebrafish with RNA-guided endonucleases. Genome Res. 2014; 24:125–131. [PubMed: 24253447]
8. Rohrschneider MR, Elsen GE, Prince VE. Zebrafish Hoxb1a regulates multiple downstream genes including prickle1b. Dev Biol. 2007; 309:358–372. [PubMed: 17651720]
9. Jessen JR, et al. Zebrafish trilobite identifies new roles for Strabismus in gastrulation and neuronal movements. Nat. Cell Biol. 2002; 4:610–615. [PubMed: 12105418]
10. Satou C, et al. Functional role of a specialized class of spinal commissural inhibitory neurons during fast escapes in zebrafish. J. Neurosci. 2009; 29:6780–6793. [PubMed: 19474306]
11. Pereda AE. Electrical synapses and theirfunctional interactions with chemicalsynapses. Nat. Rev. Neurosci. 2014; 15:250–263. [PubMed: 24619342]
12. Rash JE, et al. Molecular and Functional Asymmetry at a Vertebrate Electrical Synapse. Neuron. 2013; 79:957–969. [PubMed: 24012008]
13. Wang T, Wei JJ, Sabatini DM, Lander ES. Genetic screens in human cells using the CRISPR-Cas9 system. Science. 2014; 343:80–84. [PubMed: 24336569]
14. Craig AM, Kang Y. Neurexin-neuroligin signaling in synapse development. Curr. Opin. Neurobiol. 2007; 17:43–52. [PubMed: 17275284]
15. Xiao A, et al. Chromosomal deletions and inversions mediated by TALENs and CRISPR/Cas in zebrafish. Nucleic Acids Res. 2013; 41:e141. [PubMed: 23748566]
16. Hsu PD, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. Nat Biotechnol. 2013; 31:827–832. [PubMed: 23873081]
17. Gagnon JA, et al. Efficient mutagenesis by Cas9 protein-mediated oligonucleotide insertion and large-scale assessment of single-guide RNAs. PLoS ONE. 2014; 9:e98186. [PubMed: 24873830]

18. Howe K, et al. The zebrafish reference genome sequence and its relationship to the human genome. Nature. 2013; 496:498–503. [PubMed: 23594743]
19. Kelly A, Hurlstone AF. The use of RNAi technologies for gene knockdown in zebrafish. Brief Funct Genomics. 2011; 10:189–196. [PubMed: 21525144]
20. Schulte-Merker S, Stainier DYR. Out with the old, in with the new: reassessing morpholino knockdowns in light of genome editing technology. Development. 2014; 141:3103–3104. [PubMed: 25100652]
21. Hruscha A, et al. Efficient CRISPR/Cas9 genome editing with low off-target effects in zebrafish. Development. 2013; 140:4982–4987. [PubMed: 24257628]
22. Twigg SRF, et al. Cellular interference in craniofrontonasal syndrome: males mosaic for mutations in the X-linked EFNB1 gene are more severely affected than true hemizygotes. Hum. Mol. Genet. 2013; 22:1654–1662. [PubMed: 23335590]
23. Shalem O, et al. Genome-scale CRISPR-Cas9 knockout screening in human cells. Science. 2014; 343:84–87. [PubMed: 24336571]
24. Bassett AR, Tibbit C, Ponting CP, Liu J-L. Highly efficient targeted mutagenesis of Drosophila with the CRISPR/Cas9 system. Cell Rep. 2013; 4:220–228. [PubMed: 23827738]
25. Lo T-W, et al. Precise and heritable genome editing in evolutionarily diverse nematodes using TALENs and CRISPR/Cas9 to engineer insertions and deletions. Genetics. 2013; 195:331–348. [PubMed: 23934893]
26. Nakayama T, et al. Cas9-Based Genome Editing in Xenopus tropicalis. Meth. Enzymol. 2014; 546C:355–375. [PubMed: 25398349]
27. Yen S-T, et al. Somatic mosaicism and allele complexity induced by CRISPR/Cas9 RNA injections in mouse zygotes. Dev Biol. 2014; 393:3–9. [PubMed: 24984260]
28. Yu C, Zhang Y, Yao S, Wei Y. A PCR based protocol for detecting indel mutations induced by TALENs and CRISPR/Cas9 in zebrafish. PLoS ONE. 2014; 9:e98282. [PubMed: 24901507]
29. Pan L, et al. Rapid identification and recovery of ENU-induced mutations with next-generation sequencing and paired-end low-error analysis. BMC Genomics. 2015; 16
30. Breese MR, Liu Y. NGSUtils: a software suite for analyzing and manipulating next-generation sequencing datasets. Bioinformatics. 2013; 29:494–496. [PubMed: 23314324]
31. Güell M, Yang L, Church GM. Genome editing assessment using CRISPR Genome Analyzer (CRISPR-GA). Bioinformatics. 2014; 30:2968–2970. [PubMed: 24990609]
32. Kimmel CB, Ballard WW, Kimmel SR, Ullmann B, Schilling TF. Stages of embryonic development of the zebrafish. Dev. Dyn. 1995; 203:253–310. [PubMed: 8589427]
33. Higashijima S, Hotta Y, Okamoto H. Visualization of cranial motor neurons in live transgenic zebrafish expressing green fluorescent protein under the control of the islet-1 promoter/enhancer. J. Neurosci. 2000; 20:206–218. [PubMed: 10627598]
34. Schindelin J, et al. Fiji: an open-source platform for biological-image analysis. Nat. Methods. 2012; 9:676–682. [PubMed: 22743772]

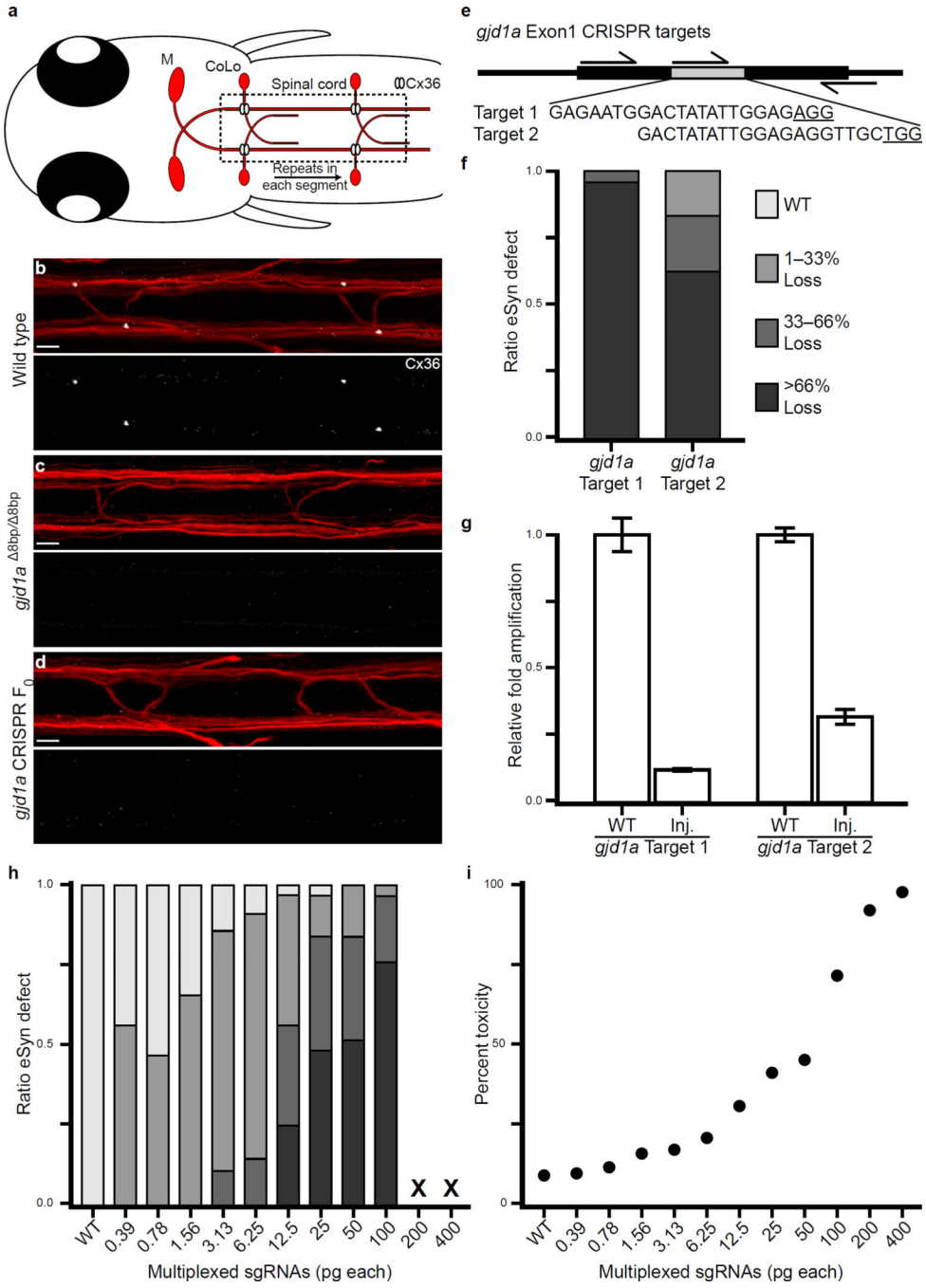

Figure 1.

Optimizing multiplexed CRISPR screening for electrical synapses. (**a**) Model of the Mauthner (M) circuit. Each M cell body resides in the hindbrain and sends an axon into the spinal cord making electrical synapses (Cx36) with repeating Commissural Local (CoLo) neurons. (**b–d**) Images are 15 μm dorsal view projections of two spinal cord segments at 5 dpf. Anterior is to the left. Scale bar = 10 μm. Larvae are stained for Connexin36 (Cx36, white) and neurofilaments (RMO44, red) to mark neuronal processes, including M and CoLo. Individual Cx36 channels are shown in neighboring panels. The Cx36 staining found

at M/CoLo synapses (**b**) is lost in *gjd1a*^{fh436} (8 bp deletion) mutant animals (**c**) and the phenotype can be recapitulated in injected CRISPR F_0 embryos (**d**, 100/1200 pg sgRNA/*cas9*). (**e**) Model of *gjd1a* exon1 and CRISPR targets. Arrows denote primers used for qPCR analysis of mutational efficiency. Underlined sequence denotes the NGG motif used by Cas9. (**f,g**) Quantitation of mosaic electrical synapse loss (**f**) and mutational efficiency (**g**) in injected embryos for each *gjd1a* target injected at 100/1200 pg sgRNA/*cas9*. The "ratio eSyn defect" is the proportion of electrical synapses missing from at least 30 sampled per animal. Mutational efficiency was assessed by qPCR and was done in triplicate. (**h,i**) Quantitation of mosaic electrical synapse loss (**h**) and toxicity (**i**) seen in injected embryos for *gjd1a* Target1 multiplexed with five other sgRNAs that have no effect on electrical synapses (*gfp, slc24a5, pk1b, sox10, hmcn1*) with 1200 pg of Cas9-encoding mRNA. "Toxicity" encompasses embryo death, edema, localized cell death, and general developmental defects. In **f** and **h**, N > 24 embryos for each bar. In **g**, each bar represents 5 embryos pooled in 3 replicates, error bars denote s.e.m. In **i**, N > 85 embryos for each point.

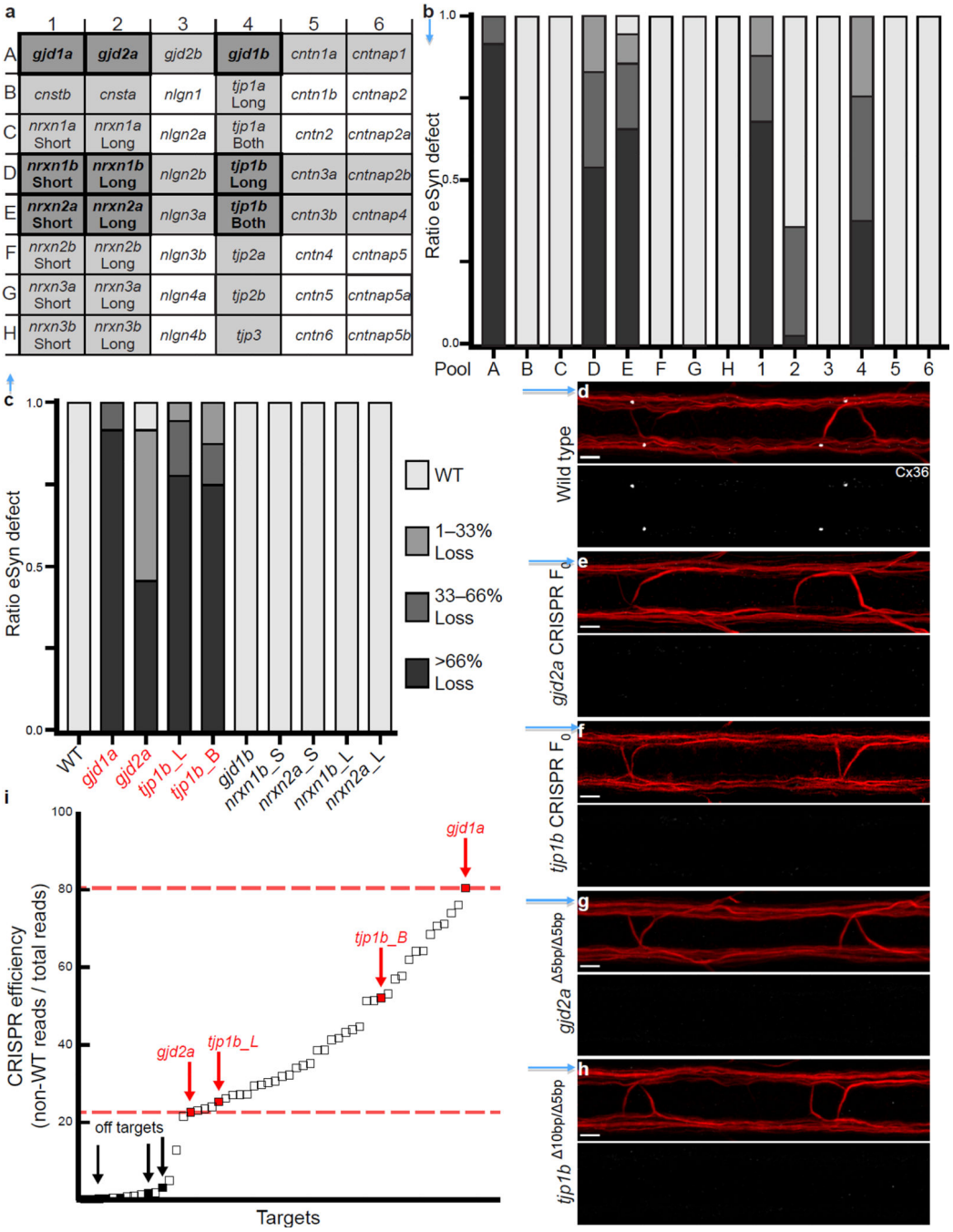

Figure 2.

CRISPR screening identifies new genes required for electrical synaptogenesis. (**a**) Layout of target genes for multiplexed row and column injections. Genes were arranged such that gene family members were targeted in either a row or column. Grey rows/columns denote multiplexed injections that caused a loss of electrical synapses. The dark-grey genes highlight the nine intersecting targets most likely to be causative for synapse loss. sgRNAs were designed against Short (S), Long (L), and if no unique site could be found, Both (B), isoforms of several genes. (**b**) Quantitation of mosaic electrical synapse loss in injected

embryos for multiplexed sgRNAs of the noted row or column. (**c**) Quantitation of mosaic electrical synapse loss in injected embryos for each of the intersecting sgRNAs identified individually injected at 100/1200 pg sgRNA/*cas9*. (**d–h**) Images are oriented, imaged, and stained as in Figure 1. The Cx36 staining found at M/CoLo synapses (**d**) is lost in injected CRISPR F_0 embryos targeting *gjd2a* or *tjp1b* (**e,f**). Homozygous mutant *gjd2a*^{fh437} (5 bp deletion) and trans-heterozygous *tjp1b* (10 bp deletion / 5 bp deletion) animals confirm each gene is required for electrical synapse formation (**g,h**). (**i**) Quantitation of InDel frequency at all target loci in the CRISPR screen based on deep sequencing. Each unfilled box represents an individual screen target. Filled black boxes represent the predicted off target loci for the phenotypically positive hits identified. In **b** and **c**, N > 24 embryos for each bar. In **i**, each point represents >136,000 reads.