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] & Cecilia B Moens[1,2]

Identifying genes involved in biological processes is critical for understanding the molecular building blocks of life. We used engineered CRISPR (clustered regularly interspaced short palindromic repeats) to efficiently mutate specific loci in zebrafish (*Danio rerio*) and screen for genes involved in vertebrate biological processes. We found that increasing CRISPR efficiency by injecting optimized amounts of Cas9-encoding mRNA and multiplexing single guide RNAs (sgRNAs) allowed for phenocopy of known mutants across many phenotypes in embryos. We performed a proof-of-concept screen in which we used intersecting, multiplexed pool injections to examine 48 loci 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.

© 2015 Nature America, Inc. All rights reserved.

Although classical forward and reverse genetic approaches have been used to identify key molecular pathways required for life, they are generally limiting in terms of the number of targets that can be assessed, and they are very time intensive. Reverse genetic screening techniques have been used with invertebrate animal models and cell culture systems to quickly identify genes and pathways involved in many biological processes; however, identifying a robust and inexpensive method for use in an *in vivo* vertebrate model system has been challenging. The bacterial and archaeal adaptive defense mechanism CRISPR has been engineered to work in zebrafish. CRISPR can be used to efficiently target the genome and introduce changes via homology-directed repair or nonhomologous end joining[1,2] that can disrupt gene function, thereby providing insight into a gene's effect on the process of interest[3,4]. Type II CRISPR systems are particularly attractive as reverse genetic tools because they require only the enzyme Cas9, which cleaves DNA, and an engineered sgRNA that is ~100 nucleotides long and includes a 20-nucleotide target sequence[5] that requires an 'NGG' protospacer-adjacent motif bound by Cas9. In zebrafish, the injection of Cas9-encoding mRNA with sgRNA into embryos at the one-cell stage induces mutations that are transmissible through the germline[1,5]. Additionally, with codon optimization of Cas9 and the addition of a nuclear-localization sequence, the frequency of biallelic disruption increased such that injected (F_0) embryos phenocopied known mutant phenotypes[6]. Here we optimized CRISPR insertion-deletion (indel) generation efficiency by varying Cas9 and sgRNA concentrations across three orders of magnitude and screened for candidate genes involved in electrical-synapse formation. Genes required for electrical synapses are not well understood, and we identified two new genes. Our results show that CRISPR can be used as a powerful and efficient reverse genetic screening strategy *in vivo* in a vertebrate system.

RESULTS

Optimizing CRISPR for identifying phenotypic effects

To increase the mutational load in embryos injected with Cas9-encoding mRNA and sgRNA, and to reduce genetic mosaicism due to a lag in Cas9 expression[6,7], we varied Cas9 and sgRNA concentrations to find optimal conditions. We targeted the *slc24a5* (*golden*) locus, which can be easily screened because mutation of *slc24a5* leads to a loss of pigmentation in the retinal epithelium (**Supplementary Fig. 1a**). We used Cas9 that had been previously codon-optimized for zebrafish and had two nuclear localization sequences, and we used an *slc24a5* sgRNA 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 2,400 pg while using a constant 100 pg of *slc24a5* sgRNA in injections into one-cell-stage embryos and tested for the amount of retinal pigment loss. We found that in all experimental injection conditions there were embryos that had reduced pigmentation in the eye, but there was variability in the phenotype, with embryos ranging from no pigment loss to complete loss (**Supplementary Fig. 1d**). The fact that there was a reduction in pigmentation associated with all injections suggested a high mutational load; indeed, we found that 100% (24 of 24) of individually cloned and sequenced alleles from the 100/1,200 pg *slc24a5*/Cas9 injection had indels (**Supplementary Fig. 1c**). Increasing the 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, meaning they had phenocopied the homozygous mutant (**Supplementary Fig. 1d**).

[1]Division of Basic Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA. [2]These authors jointly directed this work. Correspondence should be addressed to A.C.M. (amiller@fhcrc.org) or C.B.M. (cmoens@fhcrc.org).

RECEIVED 17 NOVEMBER 2014; ACCEPTED 17 MARCH 2015; PUBLISHED ONLINE 13 APRIL 2015; DOI:10.1038/NMETH.3360

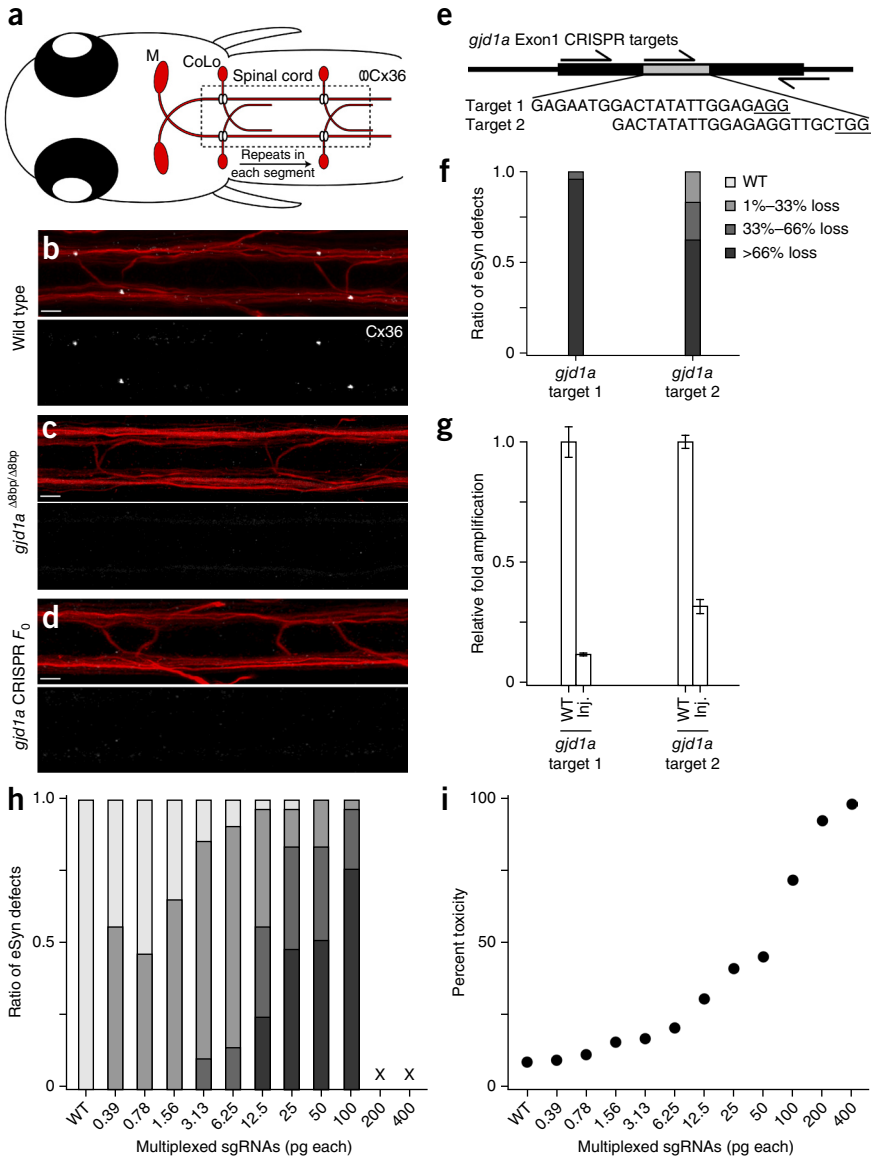


Figure 1 | Optimizing multiplexed CRISPR screening for electrical synapses. (**a**) Model of the Mauthner circuit (M). Each M cell body resides in the hindbrain and sends an axon into the spinal cord, making electrical synapses (Cx36) with repeating CoLo neurons. (**b**–**d**) Images of two spinal cord segments from specified genotypes at 5 d after fertilization stained for Cx36 (white) and neurofilaments (RMO44, red) to mark neuronal processes, including M and CoLo processes. Individual Cx36 channels are shown in the bottom panel for each. Anterior is to the left. Scale bars, 10 µm. Images are 15-µm dorsal-view projections. (**e**) Model of *gjd1a* Exon1 and CRISPR targets. Arrows denote primers used for qPCR analysis of mutational efficiency. Underlined sequences denote 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 100/1,200 pg sgRNA/Cas9-encoding mRNA. The "ratio of eSyn defects" is the proportion of electrical synapses missing from at least 30 sampled per animal. Mutational efficiency was assessed by qPCR in triplicate. WT, wild type; Inj., injected. (**h**,**i**) Quantitation of mosaic electrical-synapse loss (**h**) and toxicity (**i**) seen in embryos injected with *gjd1a* target 1 multiplexed with five other sgRNAs that have no effect on electrical synapses (*gfp*, *slc24a5*, *pk1b*, *sox10* and *hmcn1)* with 1,200 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 five embryos pooled in three replicates; error bars denote s.e.m. In **i**, $N > 85$ embryos for each point.

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 nucleic acid injection in zebrafish. Given that in our experiments we typically injected 120–150 embryos, we decided that 1,200 pg of Cas9 gave a reasonable balance between phenotypic efficiency and acceptable toxicity. We used this concentration to test the effect of varying sgRNA concentrations from 10 to 1,000 pg. Increasing sgRNA levels increased both the loss of pigment (not shown) and the toxicity (**Supplementary Fig. 1f**). However, we found that even at 1,200 pg of Cas9-encoding mRNA and 1,000 pg of *slc24a5* sgRNA, only ~30% of embryos could not be screened because of toxicity.

We next used these conditions to investigate 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

We optimized the CRISPR screening strategy for electrical-synapse formation in the Mauthner circuit (M) of the spinal cord. This system is accessible and 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 (commissural local) neuron (**Fig. 1a**)[10]. This provided a simple readout of efficiency and mosaicism within CRISPR injected embryos, as 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 channels between neurons[11]. The M electrical synapses can be visualized by immunostaining with an antibody to the human Cx36 protein, which detects zebrafish Cx36-related proteins (**Fig. 1b**)[12]. In a separate forward genetic screen we identified the gene *gap junction delta 1a*

© 2015 Nature America, Inc. All rights reserved.

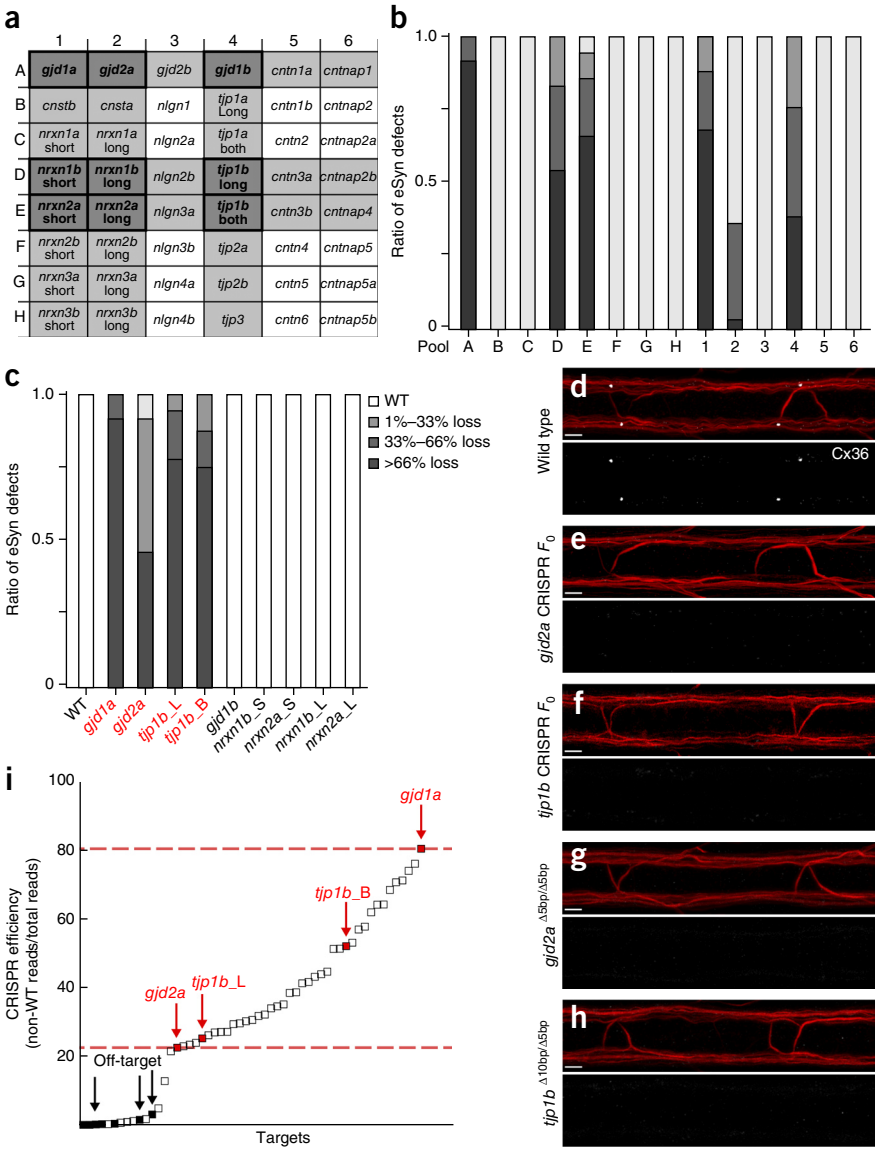


Figure 2 | CRISPR screening identified new genes required for electrical synaptogenesis. (**a**) Layout of target genes for multiplexed row and column injections. Genes were arranged such that gene-family members were targeted in either a row or a column. Gray rows and columns denote multiplexed injections that caused a loss of electrical synapses. The dark gray shading highlights the nine intersecting target genes most likely to be causative for synapse loss. sgRNAs were designed against short, long or, if no unique site could be found, both isoforms of several genes. (**b**) Quantitation of mosaic electrical-synapse loss in injected embryos for multiplexed sgRNAs of the indicated row or column. (**c**) Quantitation of mosaic electrical-synapse loss in injected embryos for each of the intersecting sgRNAs identified individually injected at 100/1,200 pg sgRNA/Cas9-encoding mRNA. L, long isoform; S, short isoform; B, both isoforms. (**d**–**h**) Images from specified genotypes, oriented, imaged and stained as described for **Figure 1**. (**i**) Quantitation of indel frequency at all target loci in the CRISPR screen based on deep sequencing. Each open box represents an individual screen target. Solid 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.

(*gjd1a*), which encodes zebrafish Cx34.1, as is 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 homolog of the mammalian *Gjd2*, which encodes the major Cx protein expressed in the mammalian brain[11]. To confirm that we could use CRISPR with Cas9 to detect electrical-synapse phenotypes, we first designed two sgRNAs targeting the *gjd1a* locus (http://crispr.mit.edu)[3] that overlapped in target sequence by 16 nucleotides. Each was injected independently at 100/1,200 pg sgRNA/Cas9-encoding 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 target 1 leading to 95% of embryos lacking more than 66% of synapses in the spinal cord, whereas injection for target 2 led to similar synapse defects in 60% of embryos (**Fig. 1f**). We found that these phenotypic differences between sgRNA targets correlated with the efficiency of genomic-locus alteration as assessed by quantitative PCR (qPCR) directly from genomic DNA (**Fig. 1g** and Online Methods). In this instance, higher efficiency was correlated with a higher purine content directly preceding the protospacer-adjacent motif sequence, as has recently been demonstrated on a genome-wide scale[13]. Finally, we examined the range of concentrations at which *gjd1a* sgRNA target 1 could produce electrical-synapse phenotypes and found that synapse loss resulted after the injection of as little as 0.19 pg of sgRNA, an amount that caused virtually no toxicity relative to uninjected controls (**Supplementary Fig. 3**). We concluded 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 it also decreased 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* and *hmcn1*); each was tested individually, and none affected synapse formation (data not shown). We tested a range of sgRNA concentrations from 0.39 to 400 pg for each individual sgRNA in the pool and coinjected them with 1,200 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 increased toxicity, and the injection of more than 200 pg of each guide led to the death of all injected embryos

© 2015 Nature America, Inc. All rights reserved.

(**Fig. 1i**). However, injecting 12.5 pg of each sgRNA resulted in a robust lack of synapses in ~50% of embryos but 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* homologs, genes encoding proteins that biochemically interact with Cxs, and the *neuroligin* and *neurexin* gene families and related members, which are involved in chemical synaptogenesis (**Fig. 2a** and **Supplementary Table 1**)[11,14]. We included *gjd1a* as a positive control in the screen, but none of the other genes had a previously known requirement for electrical synaptogenesis. To reduce the total number of injections, we arrayed our sgRNAs in a 96-well plate format and injected sgRNAs pooled by row and column (**Fig. 2a**). In this way, each sgRNA was injected twice in a separate pool, and it was assumed that the intersection of positive hits in both row and column pools would reveal the most likely candidate gene. Whether row-and-column pooling reduces the total number of 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 identify only a small number of positive pools. When multiplexing sgRNAs, care must be taken if the targets are on the same chromosome, as large deletions between sites can occur, with reported deletions of 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 a 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 1,200 pg of Cas9-encoding mRNA for all injections, with column pools at 12.5 pg of each sgRNA and row pools at 16.5 pg of each; this allowed us to maintain 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 in 30%–100% of injected embryos (**Fig. 2b**). This set of row and column hits produced an intersection of nine genes that were the most likely to be causative for the synapse phenotypes (**Fig. 2a**). One of these overlapping genes was the positive control *gjd1a*, and the others were two of the other *gjd2* homologs, two isoforms of the electrical-synapse scaffold gene *tjp1b* (ZO-1), and short and long isoforms of two different *neurexin*-family genes (**Fig. 2a**). We injected each of these sgRNAs individually and found a loss of electrical synapses when we targeted *gjd2a* and *tjp1b* (targeting two different isoforms of *tjp1b*, *tjp1b_L and tjp1b_B*, gave similar synapse loss), whereas the others had no effect (**Fig. 2c**). We also examined whether the other genes in the positive rows and columns had phenotypes (i.e., the non-intersecting sgRNAs in the positive rows or columns) by injecting them in pools and found that they did not affect synapses (not shown). We confirmed that both *gjd2a* and *tjp1b* were 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* lacked electrical synapses in the spinal cord (**Fig. 2d**–**h**). Our 48-gene CRISPR screen, including sgRNA synthesis, pool injections and demultiplexing, took a total of 3 weeks and identified new genes involved in electrical synaptogenesis.

Sequencing revealed 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 the genomic DNA of ten animals for each of the row and column pool injections (i.e., we amplified and barcoded *gjd1a* from column 1 and row A, etc.). We also amplified regions for the predicted off-target sites on the basis of sequence relatedness to our phenotypically positive hits (**Supplementary Table 1**)[16]. We sequenced all amplicons on an Illumina MiSeq machine and obtained 136,000–800,000 reads for each target. We assessed the frequency of indels at each locus by counting the number of insertions, deletions and wild-type alleles. We found that the majority of targets had indel frequencies that ranged from 22% to 85% of the sequenced alleles, depending on the target locus. The phenotypically positive hits were found at the top (85% non-wild-type reads; *gjd1a*) and near the bottom (25% non-wild-type 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 the broad range of mutational efficiencies observed at individual targets in our screen to see whether it was influenced by recently identified sequence-composition rules[13,17]. We found that an increased GC content within the guide and purines directly upstream of the NGG motif correlated with increased mutational efficiency (**Supplementary Fig. 4a**–**c**). Further analysis of the next-generation sequencing (NGS) data revealed that across all injections, out-of-frame indels accounted for 48%–77% of non-wild-type reads (**Supplementary Fig. 4d**; average = 61.5, s.e.m. = 1.28). Deletions were centered near the NGG motif of the target sequence with a 5′ bias in the location with the most deletions (**Supplementary Fig. 4e**). The majority of deletions were less than 10 bp, but they ranged from 1 to 198 bp (**Supplementary Fig. 4f**; 198 bp was the maximum deletion size detectable with 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 suggested that off-target sites were unlikely to be causative for the phenotypes assessed.

The phenotypically positive guides with low mutation rates (*tjp1b_L* and *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 16.5/1,200 pg *tjp1b_L*/Cas9-encoding 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 suggested that phenotypes manifest with monoallelic loss (heterozygous) or that in-frame deletions lead to a loss of gene

© 2015 Nature America, Inc. All rights reserved.

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 all 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 to the human protein), suggesting that such regions may be intolerant to in-frame indels.

Six of our targets had indel frequencies similar to those of the off-target sites (**Fig. 2i**). 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 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 relative to the genome into which it was injected (**Supplementary Fig. 6b**). We concluded that our optimized screening strategy robustly screened more than ~90% of the targets even in the face of varying sgRNA efficiencies.

DISCUSSION

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 lacking for *in vivo* vertebrate model systems. The pilot CRISPR screen we performed took a total of 3 weeks and identified two genes previously unknown to be involved in electrical synaptogenesis, highlighting the usefulness of the approach. Our 'row by column' multiplexed, intersectional-pooling strategy efficiently identified the most likely candidates. The pooling strategy can provide additional information, as it has the potential to 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 it will be phenotype specific and dependent on the efficiency of mutagenesis at individual target loci. In this study we used Cas9 that had been engineered for zebrafish efficiency[6], but future efforts that increase mutagenicity would be beneficial for screening. Additionally, using Cas9 protein, instead of mRNA, might provide improved mutagenicity and decreased genetic mosaicism[17]. Although both morpholinos and RNAi have been used in zebrafish to knock down 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 with CRISPR are low (1%–3%)[21], which suggests that 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 for identifying candidates for many biological processes.

Despite the success of the screen, it is important to view the positive candidates with some caution. Because of 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 wild-type 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 would 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 would be expected to have biallelic frameshift mutations. Although in-frame deletions might often be tolerated, we found that they can cause a deleterious loss of function. Therefore, care in selecting the targeted region of the protein—for example, in conserved functional domains—might provide increased mutagenicity. However, given the mosaicism, the 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 that can be assessed to confirm the phenotypes and extend the analysis of mutant defects.

The success of our pilot screen paves the way toward the application of CRISPR screening on a larger scale. Recent methods using lentiviral-based Cas9-sgRNA libraries have been used to screen at the genome scale in cultured cells[13,23]. Although 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. Within a large screen, NGS analysis of injected embryos could be done through DNA collection and storage followed by sequencing to assess the targets screened. Alternatively, sgRNA efficiency can be tested inexpensively by qPCR (Online 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 genome-wide association studies or exome sequencing. Finally, although we 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.

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes. BioProject: PRJNA273396; SRA: SRP052749.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

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 Institutes of Health (R01HD076585 and R21NS076950 to C.B.M. and K99NS085035 to A.C.M.).

© 2015 Nature America, Inc. All rights reserved.

AUTHOR CONTRIBUTIONS

A.N.S., C.F.D., A.C.W. and A.C.M. performed experiments, acquired and quantified data, and generated images for publication. A.C.M., A.N.S. and C.B.M. wrote the manuscript. All authors edited the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Chang, N. *et al.* Genome editing with RNA-guided Cas9 nuclease in zebrafish embryos. *Cell Res.* **23**, 465–472 (2013).
2. Hwang, W.Y. *et al.* Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat. Biotechnol.* **31**, 227–229 (2013).
3. Hsu, P.D., Lander, E.S. & Zhang, F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell* **157**, 1262–1278 (2014).
4. Peng, Y. *et al.* Making designer mutants in model organisms. *Development* **141**, 4042–4054 (2014).
5. Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012).
6. Jao, L.-E., Wente, S.R. & Chen, W. Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. *Proc. Natl. Acad. Sci. USA* **110**, 13904–13909 (2013).
7. Sung, Y.H. *et al.* Highly efficient gene knockout in mice and zebrafish with RNA-guided endonucleases. *Genome Res.* **24**, 125–131 (2014).
8. Rohrschneider, M.R., Elsen, G.E. & Prince, V.E. Zebrafish Hoxb1a regulates multiple downstream genes including prickle1b. *Dev. Biol.* **309**, 358–372 (2007).
9. Jessen, J.R. *et al.* Zebrafish trilobite identifies new roles for Strabismus in gastrulation and neuronal movements. *Nat. Cell Biol.* **4**, 610–615 (2002).
10. Satou, C. *et al.* Functional role of a specialized class of spinal commissural inhibitory neurons during fast escapes in zebrafish. *J. Neurosci.* **29**, 6780–6793 (2009).
11. Pereda, A.E. Electrical synapses and their functional interactions with chemical synapses. *Nat. Rev. Neurosci.* **15**, 250–263 (2014).
12. Rash, J.E. *et al.* Molecular and functional asymmetry at a vertebrate electrical synapse. *Neuron* **79**, 957–969 (2013).
13. Wang, T., Wei, J.J., Sabatini, D.M. & Lander, E.S. Genetic screens in human cells using the CRISPR-Cas9 system. *Science* **343**, 80–84 (2014).
14. Craig, A.M. & Kang, Y. Neurexin-neuroligin signaling in synapse development. *Curr. Opin. Neurobiol.* **17**, 43–52 (2007).
15. Xiao, A. *et al.* Chromosomal deletions and inversions mediated by TALENs and CRISPR/Cas in zebrafish. *Nucleic Acids Res.* **41**, e141 (2013).
16. Hsu, P.D. *et al.* DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* **31**, 827–832 (2013).
17. Gagnon, J.A. *et al.* Efficient mutagenesis by Cas9 protein-mediated oligonucleotide insertion and large-scale assessment of single-guide RNAs. *PLoS One* **9**, e98186 (2014).
18. Howe, K. *et al.* The zebrafish reference genome sequence and its relationship to the human genome. *Nature* **496**, 498–503 (2013).
19. Kelly, A. & Hurlstone, A.F. The use of RNAi technologies for gene knockdown in zebrafish. *Brief Funct. Genomics* **10**, 189–196 (2011).
20. Schulte-Merker, S. & Stainier, D.Y.R. Out with the old, in with the new: reassessing morpholino knockdowns in light of genome editing technology. *Development* **141**, 3103–3104 (2014).
21. Hruscha, A. *et al.* Efficient CRISPR/Cas9 genome editing with low off-target effects in zebrafish. *Development* **140**, 4982–4987 (2013).
22. Twigg, S.R.F. *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.* **22**, 1654–1662 (2013).
23. Shalem, O. *et al.* Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* **343**, 84–87 (2014).
24. Bassett, A.R., Tibbit, C., Ponting, C.P. & Liu, J.-L. Highly efficient targeted mutagenesis of *Drosophila* with the CRISPR/Cas9 system. *Cell Rep.* **4**, 220–228 (2013).
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* **195**, 331–348 (2013).
26. Nakayama, T. *et al.* Cas9-based genome editing in *Xenopus tropicalis*. *Methods Enzymol.* **546**, 355–375 (2014).
27. Yen, S.-T. *et al.* Somatic mosaicism and allele complexity induced by CRISPR/Cas9 RNA injections in mouse zygotes. *Dev. Biol.* **393**, 3–9 (2014).

© 2015 Nature America, Inc. All rights reserved.

ONLINE METHODS

sgRNA design and template synthesis. sgRNAs were designed using the design tool at http://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 on the basis of sequence relatedness. When possible, sgRNAs chosen had a score of 90+ (**Supplementary Table 1**). Because of 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 Cas9 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 templated off each other and the full sgRNA sequence was created (**Supplementary Fig. 7**). The scaffold oligo sequence was 5′-GATCCGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAAC-3′ and was HPLC purified. The gene-specific oligo sequence was 5′-AATTAATACGACTCACTATA(N20)GTTTTAGAGCTAGAAATAGC-3′, where (N20) refers to the 20 nucleotides of the sgRNA that bound the genome (excluding the NGG motif). The full-length PCR product then was 5′-AATTAATACGACTCACTATA(N20)GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGGATC-3′.

The PCR reaction included 2.5 µL H_2O, 12.5 µL 2× Phusion Master Mix (New England BioLabs, M0531L), 5 µL scaffold oligo (10 µM, synthesized at Eurofins Genomics), and 5 µL sgRNA oligo (10 µM, synthesized at Eurofins Genomics) and was run in a thermocycler under the following program: 95 °C for 30 s; 40 cycles of 95 °C for 10 s, 60 °C for 10 s, and 72 °C for 10 s; and 72 °C for 5 min. This PCR product was purified and used as a 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 synthesis of Cas9-encoding mRNA, we used pT3Ts-nCas9n from the Chen lab, obtained 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 concentrations. We recommend using an Agilent TapeStation or equivalent technology to assay the integrity of the synthesized RNAs as a further quality-control measure.

Injection of Cas9-encoding mRNA and sgRNA. Unless otherwise stated, all injections contained 1,200 ng/µL Cas9-encoding mRNA. All sgRNAs for the screen were diluted to 200 ng/µL before multiplexing. Therefore, column pools of eight contained 12.5 ng/µL of each sgRNA, and row pools of six contained 16.6 ng/µL of each sgRNA. 1 nL of Cas9-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 min. The solution was neutralized by an equal volume of neutralization buffer (40 mM Tris-HCl, pH 5.0). Samples were spun at 3,000 RPM for 5 min, 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, which allowed 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 made use of the fact that the binding of a primer overlapping the sgRNA site was compromised in successfully mutagenized embryos, resulting in delayed amplification, whereas binding of a flanking primer pair was unaffected. The 'OUT' (flanking) primer pair encompassed at least 100 bp surrounding the sgRNA binding region. The 'ON' (overlapping) primer pair used one of the OUT primers and another primer that bound the 20 bp of the sgRNA target sequence. The 3′ side of the ON primer bound just 5′ of the NGG motif, as most indels affected the -1 to -10 positions of the binding site (**Supplementary Fig. 4e**). At 2 d post-fertilization (dpf), genomic DNA was isolated from five uninjected and five injected embryos (in triplicate). Using 1 µl of this template DNA, we ran separate 20-µL OUT and ON qPCR reactions. qPCR was performed with SsoAdvanced SYBR Green Supermix (Bio-Rad, 172-5271) 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 s and 60 °C for 20 s after initial denaturing at 95 °C for 30 s.

The ratio of the qPCR quantification-cycle values for the uninjected ON and uninjected OUT primers reflected the differences in amplification of the two primer pairs on uninjected template DNA. This might have been due to inherent differences in amplification that exist even between perfectly complementary primer pairs. In contrast, the ON/OUT ratio in injected embryos reflected 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 quantification-cycle ratios of injected versus uninjected embryos thus reflected the efficiency of mutagenesis.

To examine how accurately this 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 underestimated the mutation frequency relative to the NGS analysis by between 6% and 15%, with an average underestimation of 10% (**Supplementary Fig. 5** and Source Data).

NGS preparation and sequencing. We employed a two-step PCR method to create the Illumina sequencing library. This required special primer overhangs on the gene-specific primers to act as a template for the Illumina barcode PCR. First, gene-specific primers were designed for each of the 48 targets, as well as for seven of

© 2015 Nature America, Inc. All rights reserved.

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 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 templates 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), and loaded into a MiSeq v2 Reagent Kit (Illumina, MS-102-2003). Sequencing generated paired-end (2 × 250 bp) dual-indexed reads. After sequencing, reads were demultiplexed on the basis of Illumina barcodes with the MiSeq Reporter software and stored as FASTQ files for downstream processing and analysis. All sequencing data have been submitted to the NCBI Sequence Read Archive (http://www.ncbi.nlm.nih.gov/sra); the project can be found with the BioProject accession PRJNA273396, and sequencing files can be found in the SRA database with identifier SRP052749.

NGS analysis. SeqPrep was used to merge the paired-end reads (https://github.com/jstjohn/SeqPrep). Given the amplicon lengths we chose for all targets, we expected the forward and reverse reads from sequencing to be identical. We had previously found that merging reads and keeping only those that matched 100% reduces the amount of sequencing error seen in analysis (https://github.com/jstjohn/SeqPrep, -g -n 1.0 -s)[29]. This method removes all unpaired reads and collapses the forward and reverse pairs into a single read. Using NGSutils, we split each of the two resulting FASTQ files (column and row) 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 non-wild-type reads resulting from nonhomologous end-joining, 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 were output in XML format and parsed via the R programming language to assess and plot the indel frequencies, the percentage 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–approved facility at the Fred Hutchinson Cancer Research Center. Zebrafish (*Danio rerio*) were bred and maintained as previously described[32]. Animal care was provided by R. Garcia, and veterinary care was provided by R.K. Uthamanthil. Wild-type animals were from a mixed *AB/Tu background. The facial branchiomotor neurons were visualized using the transgenic reporter line *Tg(isl1:GFP)rw0* (ref. 33). Separate from the CRISPR mutagenesis described herein, we created mutations in the *gjd2* family of genes using classical forward genetic screening using ENU (*N*-ethyl-*N*-nitrosourea) as a chemical mutagen (*gjd1a*^{fh360}), transcription activator–like effector nucleases (TALENs) (*gjd1a*^{fh436}, *gjd1b*^{fh435}, *gjd2a*^{fh437}) and targeting of induced local lesions in genomes (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 the requirement of *gjd1a* for electrical-synapse formation in the Mauthner circuit (M) (A.C.M., A.C.W., A.N.S. and 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 phenotypes as homozygous mutants (data not shown). For *tjp1b*, we identified F_0 embryos carrying germline mutations by crossing individual animals to wild types and analyzing the sequence of progeny at the sgRNA's target site. Three male and three female animals were identified carrying mutations in their germline, and each animal carried its own unique spectrum of mutations. Animals carrying mutations (10-, 5- and 18-bp deletions—*tjp1b*^{fh448}, *tjp1b*^{fh449} and *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 4 dpf were fixed in 2% trichloroacetic acid for 3 h. Fixed tissue was then washed in PBS + 0.5% Triton X-100, after which it was subjected to standard blocking and antibody incubations. Tissue was cleared stepwise 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 wild type (no loss), <33% loss, 33%–66% loss, and >66% loss on the basis of the presence or absence of at least 30 electrical synapses per animal along the length of the spinal cord. Retinal pigment was quantified using images of 2-dpf zebrafish taken on a Zeiss Lumar stereomicroscope with an attached black-and-white digital camera. Using the Fiji imaging distribution of ImageJ[34], the mean pixel gray 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 wild-type 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, we first visualized phenotypes and then analyzed

© 2015 Nature America, Inc. All rights reserved.

doi:10.1038/nmeth.3360

genotypes. All animals were examined phenotypically for effects of 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 Figure 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).

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* **9**, e98282 (2014).
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* **16**, 1263 (2015).
30. Breese, M.R. & Liu, Y. NGSUtils: a software suite for analyzing and manipulating next-generation sequencing datasets. *Bioinformatics* **29**, 494–496 (2013).
31. Güell, M., Yang, L. & Church, G.M. Genome editing assessment using CRISPR Genome Analyzer (CRISPR-GA). *Bioinformatics* **30**, 2968–2970 (2014).
32. Kimmel, C.B., Ballard, W.W., Kimmel, S.R., Ullmann, B. & Schilling, T.F. Stages of embryonic development of the zebrafish. *Dev. Dyn.* **203**, 253–310 (1995).
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.* **20**, 206–218 (2000).
34. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).

© 2015 Nature America, Inc. All rights reserved.