File S1

Supplementary Online Methods

Nematode culture

Strains were cultured according to standard methods (BRENNER 1974). The *C. elegans* strains N2 (Bristol), CB120 *unc-4(e120)*, and PS252 *dpy-11(e224)* were used.

DNA cloning

The *hCas9* nuclease gene was amplified from plasmid hCas9 (MALI *et al.* 2013) (ID# 41815; Addgene, Cambridge, MA), modified to include 5' and 3' untranslated regions previously reported to function well in the *C. elegans* germline (WOOD *et al.* 2011), bounded on the 5' side with an SP6 phage RNA polymerase binding site and on the 3' side with a *Kpn*I restriction site. The resulting construct was cloned into a Bluescript vector (Agilent Technologies, Santa Clara, CA) to generate the plasmid SP6-hCas9-Ce-mRNA; this was cleaved with *Kpn*I and then used for the *in vitro* synthesis of *hCas9* mRNA using mMessage mMachine SP6 and Poly(A) tailing kits (Life Technologies, Grand Island, NY).

A scaffold for the cloning of targeting sequences to generate sgRNAs was constructed using as a template the plasmid "gRNA Empty Vector" (MALI *et al.* 2013) (ID# 41824; Addgene), The downstream gRNA fusion part of this vector was retained, along with the *Afl*II site for recombinational cloning, bounded on its 5' end with an SP6 phage RNA polymerase sequence on its 3' end with a *Kpn*I restriction site (replacing the PolIII terminator sequence found in the template plasmid). The resulting construct was cloned into Bluescript to generate the plasmid SP6-sgRNA-scaffold. Sequences intended for use in targeting CRISPR-Cas-mediated cleavage were amplified by PCR with flanking sequences suitable for recombination into the *Afl*II site of this plasmid using Gibson cloning (New England Biolabs, Beverly, MA). Clones were confirmed by sequencing, cleaved with *Kpn*I, and used for the *in vitro* synthesis of sgRNAs using MEGAscript SP6 (Life Technologies).

Sequences were selected for use in targeting for cleavage by CRISPR-Cas on the basis of their position within the gene (preferably in an early exon) and the absence of strong BLAST hits elsewhere in the target genome for the 3' half of the 20mer sequence. Sequences used were: *dpy-11*: GAGCTGGGCACCATGGAGCA; *unc-4*: GATATCGTCATCCGGTGACG.

CRISPR-Cas mutagenesis and mutant analysis

The germline syncytia of P_0 animals were injected with the synthetic hCas9 mRNA at a final concentration of 200 – 500 ng/µl together with the sgRNA of choice at a final concentration of 40 – 100 ng/µl, in water, using established microinjection techniques (MELLO and FIRE 1995). One to three injected animals were placed on a 6 cm or 10 cm Petri dish containing NGM agar and OP50 bacteria. Animals were transferred to new plates as food supplies became exhausted. F_2 progeny of injected P_0s were screened for the presence of phenotypic animals. Mutant strains were analyzed using PCR and DNA sequencing (Laragen, Culver City, CA) to determine the number

and nature of independent isolates, especially in the case of isolates arising from the same injected P_0 or the same pool of injected P_0s. In the case of strains where no amplicon was recovered using primers that amplified from the wild type a 400 – 600 bp product centered on the targeted cleavage site, additional primer pairs were used to test for the presence of sequences at intervals of 2 – 3 kb from the targeted cleavage site, in each direction. PCRs were also performed that in the wild type would amplify larger (5 or 10 kb) genomic sequences centered on the targeted cleavage site, using Expand Long polymerase mix (New England Biolabs). Selected candidates were tested for the presence of a reciprocal translocation by outcrossing and examining the self-progeny of animals heterozygous for the mutation for the presence of dead embryos that would result from aneuploid zygotes. In particular, *sy745* mutants contain all tested sequences near to the target site, but attempts to amplify across the target site using PCR were not successful; the self-progeny of *sy745*/+ heterozygotes did not include dead (aneuploid) embryos, indicating *sy745* is likely to be an inversion or a large insertion, rather than a reciprocal translocation. *sy750* mutants fail to complement *dpy-11(e224)*, but no mutation was found near the site targeted for cleavage by CRISPR-Cas. There may be an as-yet undiscovered mutation away from the cleavage site; homology-directed repair of double-strand breaks involves error-prone DNA synthesis, causing mutations at a significant remove from the site of the break (STRATHERN *et al.* 1995; DEEM *et al.* 2011); however, no coding change was found in *dpy-11* in *sy750* animals. Alternatively, the *dpy-11(sy750)* mutant may contain a deletion-duplication, including a wild-type copy of the locus surrounding the targeted cleavage site in the context of genomic rearrangements that disrupt the function of the *dpy-11* gene.

High-throughput sequencing and analysis

Genomic DNA was isolated from *dpy-11(sy740)* and *dpy-11(sy745)* mutant strains by thorough digestion using Proteinase K (Life Technologies) in the presence of β-mercaptoethanol, one round of phenol-chloroform extraction, chloroform extraction, and spooling from ethanol, followed by RNase treatment (Qiagen, Valencia, CA) and then a second round of extractions and spooling. Genomic DNA libraries were built using Illumina's standard paired-end protocol (BENTLEY *et al.* 2008), and 50mer unpaired reads were obtained from each library using an Illumina HiSeq (Illumina, San Diego, CA).

The sequence data were aligned to the *C. elegans* genome (WormBase release 235; ftp://ftp.wormbase.org/pub/wormbase/species/c_elegans) using BWA (LI and DURBIN 2009). SNPs and small insertions and deletions (indels) were identified using the GATK pipeline (DEPRISTO *et al.* 2011) following standard practices for variant detection. Larger deletions and insertions were identified using a custom pipeline implemented in Perl based on the split-read approach to define candidate indel locations followed by a refinement step based on the Smith-Waterman local alignment algorithm (SMITH and WATERMAN 1981) to determine precise indel structures.

Mismatches and small deletions and insertions predicted by GATK analysis to be unique to either the *dpy-11(sy740)* or the *dpy-11(sy745)* mutant strain and deletions and insertions predicted by split-read analysis to be present in either or both strains were manually curated by examining an alignment of the sequencing reads to the *C. elegans* genome generated using the Burrows-Wheeler

aligner (LI and DURBIN 2009) in Integrative Genomic Viewer (IGV) (ROBINSON *et al.* 2011). Examination of the reads aligned in this fashion could demonstrate the presence of wild-type sequence lacking the deletion; a predicted deletion could still be present in a subset of the sequenced DNA. The algorithm used to align reads for manual curation would not be able to align reads that indicated the presence of a large deletion, and so if such a deletion were present but not homozygous, inspection of the aligned reads would detect wild-type sequence at the locus and would not display reads substantiating the presence of the deletion. For this reason, we used PCR to test a subset of the predicted deletions for which inspection of the aligned reads demonstrated the presence of wild-type sequence. PCRs were performed using primers spanning selected predicted deletions shown by examination of the aligned reads not to be homozygous, to test for the possible presence of smaller bands from template carrying the deletion. The oligonucleotide sequences used in these PCRs are shown in Table S4. These PCR assays did not provide evidence for off-target effects of CRISPR-Cas mutagenesis (Table S2); the unconfirmed deletions may represent computational or sequencing artifacts.

Estimates of the false-negative rate for mutation detection for the GATK and split-read analyses were made by repeating the analyses using as reference genomes versions of the *C. elegans* genome with sequences inserted at known positions, such that a comparison should show deletions compared to the reference sequence. The results of these analyses are presented in Table S3.

Literature Cited

BENTLEY, D. R., S. BALASUBRAMANIAN, H. P. SWERDLOW, G. P. SMITH, J. MILTON *et al.*, 2008 Accurate whole human genome sequencing using reversible terminator chemistry. Nature **456:** 53-59.

BRENNER, S., 1974 The genetics of *Caenorhabditis elegans*. Genetics **77:** 71-94.

DEEM, A., A. KESZTHELYI, T. BLACKGROVE, A. VAYL, B. COFFEY *et al.*, 2011 Break-induced replication is highly inaccurate. PLoS Biol **9:** e1000594.

DEPRISTO, M. A., E. BANKS, R. POPLIN, K. V. GARIMELLA, J. R. MAGUIRE *et al.*, 2011 A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet **43:** 491-498.

LI, H., and R. DURBIN, 2009 Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics **25:** 1754-1760.

MALI, P., L. YANG, K. M. ESVELT, J. AACH, M. GUELL *et al.*, 2013 RNA-guided human genome engineering via Cas9. Science **339:** 823-826.

MELLO, C., and A. FIRE, 1995 DNA transformation. Methods Cell Biol **48:** 451-482.

ROBINSON, J. T., H. THORVALDSDOTTIR, W. WINCKLER, M. GUTTMAN, E. S. LANDER *et al.*, 2011 Integrative genomics viewer. Nat Biotechnol **29:** 24-26.

SMITH, T. F., and M. S. WATERMAN, 1981 Identification of common molecular subsequences. J Mol Biol **147:** 195-197.

STRATHERN, J. N., B. K. SHAFER and C. B. MCGILL, 1995 DNA synthesis errors associated with double-strand-break repair. Genetics **140:** 965-972.

WOOD, A. J., T. W. LO, B. ZEITLER, C. S. PICKLE, E. J. RALSTON *et al.*, 2011 Targeted genome editing across species using ZFNs and TALENs. Science **333:** 307.