nature biotechnology

Conditional targeted genome editing using somatically expressed TALENs in *C. elegans*

Ze Cheng[1–3], Peishan Yi[1,3], Xiangming Wang[1], Yongping Chai[1], Guoxin Feng[1], Yihong Yang[1], Xing Liang[1], Zhiwen Zhu[1], Wei Li[1] & Guangshuo Ou[1]

We have developed a method for the generation of conditional knockouts in *Caenorhabditis elegans* by expressing transcription activator–like effector nucleases (TALENs) in somatic cells. Using germline transformation with plasmids encoding TALENs under the control of an inducible or tissue-specific promoter, we observed effective gene modifications and resulting phenotypes in specific developmental stages and tissues. We further used this method to bypass the embryonic requirement of *cor-1*, which encodes the homolog of human severe combined immunodeficiency (SCID) protein coronin, and we determined its essential role in cell migration in larval Q-cell lineages. Our results show that TALENs expressed in the somatic cells of model organisms provide a versatile tool for functional genomics.

The nematode *C. elegans* has been a popular model system to study basic biology and human diseases for decades, and many genetic tools and resources are available for this organism[1–7]. Among the 20,377 predicted protein-coding genes, 6,764 currently have deletion or null mutations, which were generated primarily by random chemical mutagenesis or Mos1 transposon–based targeted gene deletion[1,3]. However, the inability to conditionally edit the wild-type *C. elegans* genome in a targeted fashion has limited the biological questions that can be addressed with current technology.

TALEN technology has recently been developed to generate locus-specific mutations in the genome[8,9]. TALENs comprise a nonspecific FokI nuclease domain fused to a customizable repeat domain that recognizes a predictable DNA sequence. The specific DNA recognition domain directs the nuclease to introduce DNA double-strand breaks at the target site, and the erroneous repair by nonhomologous end-joining often induces a mutagenic deletion or insertion at the breakpoint[8,9]. Although TALENs have been used to edit the genome in a wide range of organisms, their use has been limited to cultured cells or embryos, as well as the germ line in *C. elegans*[7–10]. So far it has been unknown whether TALENs can be directly applied to somatic cells of a multicellular organism. In this study, we report that a conditional knockout can be achieved in somatic lineages of *C. elegans* by expressing TALEN constructs with an inducible or tissue-specific promoter (hereafter referred to as somatic TALENs).

As a proof-of-principle experiment, we examined whether somatic TALENs could conditionally disrupt the *dpy-5* gene in the *C. elegans* genome. *dpy-5* encodes a cuticle collagen that affects body length[11], and its mutation causes a dumpy, short phenotype that can be easily scored. We first generated a TALEN pair targeting *dpy-5* by selecting a 52-base-pair (bp) region in the *dpy-5* coding sequence, which includes the left and right binding sites and a restriction site for SacI (**Fig. 1a**). We chose to use disruption by TALENs of a SacI site in the *dpy-5* sequence as a molecular assay. We used the 'unit assembly' method to construct the transcription activator-like effector repeats[10]. To achieve temporally controlled mutation of *dpy-5*, we used the promoter of the heat shock gene *hsp-16.2* (P*hsp*)[12] to express *dpy-5* TALEN constructs (**Fig. 1a**). We generated transgenic *C. elegans* by germline injection of TALEN plasmids expressing both TALEN left and right constructs and a selection marker[13]. After heat shock treatment of transgenic animals at the first larval L1 stage, we detected the DPY phenotype in 93 ± 4% of all adult animals (**Fig. 1b**,**c**; $n = 267$ from three generations). The DPY phenotype was not observed in the next generation of these animals because the heat shock promoter is not active in *C. elegans* germline cells[12].

We found that the DPY penetrance depended on the developmental stage during which the expression of the *dpy-5* TALENs was induced. Penetrance increased in the late embryos as the heat shock promoter (P*hsp*) was first activated[12] and reached its maximum in L1 larvae, which is consistent with previous work showing that *dpy-5* is expressed before secretion of new cuticle from the L1 larval stage[11]. We did not observe the DPY phenotype in transgenic animals without heat shock (**Fig. 1b**,**c**; $n > 100$) or in animals expressing only the left or the right portion of *dpy-5* TALENs ($n = 81$ or 90) and subjected to heat shock.

We examined the molecular lesions generated by *dpy-5* TALENs by detecting the disruption of the SacI site in the *dpy-5* sequence (**Fig. 1a**). We PCR-amplified a 402-bp genomic DNA fragment containing the target site from transgenic animals and digested the amplified DNA with SacI. In transgenic animals without heat shock, we showed that the PCR product was completely digested by SacI to a 239-bp and a 163-bp fragment (**Fig. 1d**; left). However, we detected intact DNA fragments after digestion in transgenic animals with heat shock, indicating mutations of the Sac I site were caused by *dpy-5* TALENs (**Fig. 1d**; right). We noticed that the molecular knockout

[1]National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Chaoyang District, Beijing, China. [2]School of Life Sciences, Peking University, Beijing, China. [3]These authors contributed equally to this work. Correspondence should be addressed to G.O. (guangshuo.ou@gmail.com) or W.L. (weili.oulab@gmail.com).

Received 26 March; accepted 22 July; published online 18 August 2013; doi:10.1038/nbt.2674

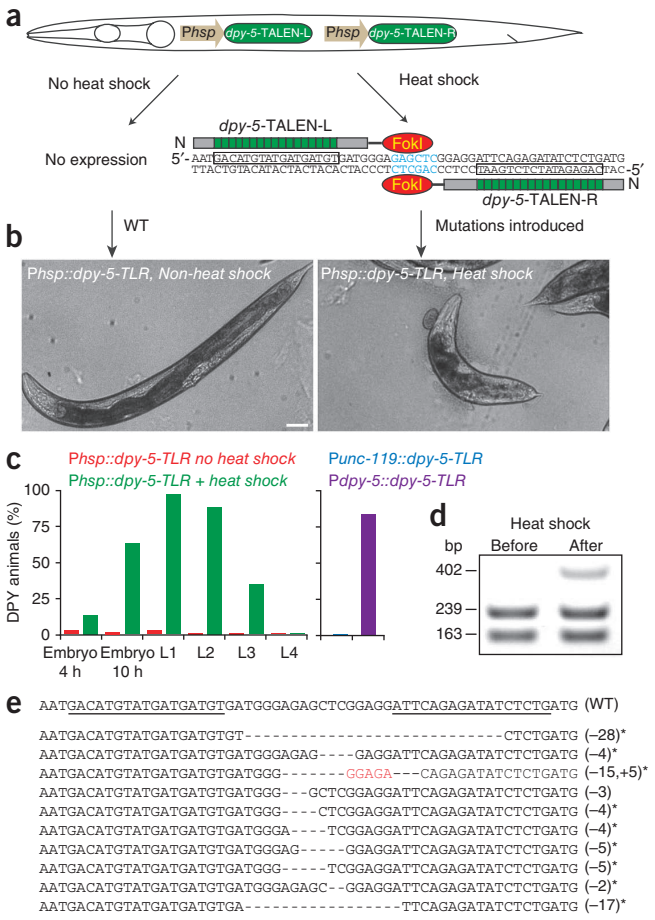

Figure 1 *dpy-5* was conditionally edited by spatially or temporally controlled somatic TALENs in *C. elegans*. (**a**) Schematic representation of TALEN-mediated temporal regulation of *dpy-5* function in *C. elegans*. Two plasmids that express the left and right recognition sites of *dpy-5* (green) and FokI (red) under the control of the heat shock promoter (gray, P*hsp*) transformed *C. elegans* by germline injection. Upon heat shock, FokI enzymes, fused with the *dpy-5* TALENs binding domain, were generated and mutated the *dyp-5* locus (right). The sequence cut by FokI contains a recognition site for the SacI restriction enzyme (blue). (**b**) Animals carrying P*hsp::dpy-5-TALENs* (TLR) constructs without (left) or with (right) heat shock treatment. Animals express both the left and right TALEN constructs. Scale bar, 50 µm. (**c**) Left; quantification of DPY animals carrying P*hsp::dpy-5-TALENs* (TLR) constructs without (red) or with (green) heat shock treatment at the indicated developmental stages. Right; the penetrance of the DPY phenotype in animals expressing P*unc-119::dpy-5-TLR* (blue) or P*dpy-5::dpy-5-TLR* (purple). $N = 92–267$ from three independent experiments. (**d**) A representative gel of the SacI restriction enzyme assay of *dpy-5* PCR products amplified from the genomic DNA in animals carrying P*hsp::dpy-5-TALENs* before (left) or after (right) heat shock. The full-length gel is presented in **Supplementary Figure 1c**. (**e**) DNA sequence of the *dpy-5* locus from animals carrying P*hsp::dpy-5-TALEN* constructs after heat shock. Dashed lines, deleted nucleotides. Red shows inserted nucleotides. Asterisks, the deletion or insertion causes a frameshift of the *dpy-5* open reading frame. Underlined sequences, binding sites of *dpy-5* TALENs.

frequency is low in the transgenic animals, which is probably because TALENs are only expressed in a subset of *C. elegans* tissues. Sequencing of intact fragments confirmed that different insertions and deletions had occurred at the target site (**Fig. 1e**).

To achieve spatially controlled mutation of *dpy-5*, we expressed *dpy-5* TALENs using tissue-specific promoters. Expression of *dpy-5* TALENs by either the *dpy-5* endogenous promoter (P*dpy-5*) or a neuronal promoter (P*unc-119*) caused mutations in the *dpy-5* locus (**Supplementary Fig. 1a**). However, only *dpy-5* TALENs expressed by the P*dpy-5* promoter, but not the P*unc-119* promoter, produced the DPY phenotype (83% DPY, $n = 76$ for P*dpy-5*; 0% DPY, $n = 82$ for P*unc-119*; **Fig. 1c**, right), consistent with the autonomous function of *dpy-5* in hypodermal cells[11]. We noticed that the expression of the *dpy-5* TALENs by the heat shock promoter generated a more pronounced DPY phenotype than expression by the *dpy-5* promoter, which is likely owing to the mutation of *dpy-5* in the hypodermal precursor cells by P*hsp* but not by P*dpy-5*.

We further examined whether somatic TALENs targeting *lon-2*, a member of the glypican family of heparan sulfate proteoglycans[14], could increase *C. elegans* body length. *lon-2* mutation causes the long phenotype[14]. We generated transgenic animals expressing *lon-2* TALENs under the control of P*hsp*. After the heat shock treatment, we found that 63% of animals expressing the *lon-2* TALENs ($n = 200$) developed the long phenotype (**Supplementary Fig. 2**), indicating that TALENs may be generally used to edit genes in *C. elegans* somatic tissues.

We next investigated whether somatic TALENs can disrupt multiple copies of a transgene in the genome. We used the P*gcy-32* promoter to express the *gfp* gene in oxygen-sensory URX, and AQR and PQR (A/PQR) neurons, and Q-cell asymmetric divisions generate A/PQR neurons.(**Supplementary Fig. 3a**,**b**). Transgenic animals were generated by standard germline injection and subsequent integration of extra-chromosomal arrays, which normally occurs with the insertion of multiple copies of transgenes into the genome[13]. We detected the GFP fluorescence from P*gcy-32::gfp* in the expected cells at low magnification (**Supplementary Fig. 3b**,**c**). We then expressed *gfp* TALENs in these animals using the *C. elegans* Q cell–specific P*egl-17* promoter. Because A/PQR neurons but not URX neurons are Q-cell descendants (**Supplementary Fig. 3a**), P*egl-17::gfp-TALENs* should mutate *gfp* in A/PQR neurons only (**Fig. 2a**). Indeed, we found that 57 ± 5% of the animals ($n = 85$ from three generations) lost GFP fluorescence in A/PQR but not in URX neurons. It is possible that TALENs cut the transgenes in multiple sites, which led to the instability or the loss of the transgene. None of the transgenic animals lost red fluorescence from P*gcy-32::mCherry* in any neurons ($n = 85$, **Fig. 2b** and **Supplementary Figs. 1b** and **3b**,**c**). These data not only reinforce our finding that the conditional editing by somatic TALENs can be achieved using cell lineage–specific promoters but also demonstrate that somatic TALENs can efficiently knock out multiple copies of a transgene.

We applied somatic TALENs to address the function of a *C. elegans* embryonic lethal gene during larval development. We chose to study *cor-1*, a worm homolog of the SCID gene coronin[15]. Coronins encode a conserved family of actin-binding proteins[16], and mouse genetics indicates that coronin 1 is essential for T-cell migration[17]. However, recent studies suggested that coronin 1 regulates T-cell function not through the actin cytoskeleton but by affecting cell viability[18]. *cor-1* has—to the best of our knowledge—not been studied in *C. elegans* previously. Q-cell development in L1 larvae can be a useful model to dissect the function of COR-1 because Q-cell migration, asymmetric division and apoptosis are actin-dependent processes and these events can be followed at single-cell resolution[19]. However, the embryonic lethality of *cor-1* mutants precluded an analysis of COR-1 in larvae.

To visualize the final position of Q-cell progenies after migration, we used P*mec-4::gfp* and P*gcy-32::mCherry* to mark the mechanosensory AVM and PVM (A/PVM) neurons, and A/PQR cells, respectively. In wild-type (WT) animals, the bilateral Q neuroblasts

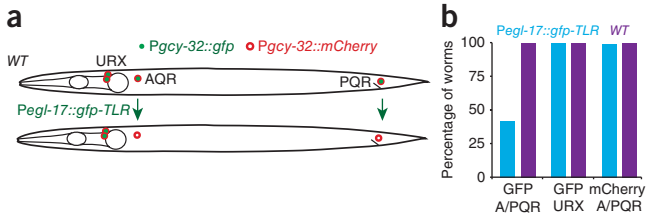

Figure 2 Somatic TALENs mutated *gfp* in *C. elegans* Q-cell lineages. (**a**) GFP (green) and mCherry (red) fluorescence in URX and AQR/PQR neurons due to the expression of P*gcy-32::gfp* and P*gcy-32::mCherry* in WT animals (upper) or worms expressing Q cell–specific *gfp* TALENs constructs driven by the *egl-17* gene promoter (P*egl-17*, lower). Still images are shown in **Supplementary Figure 3b**,**c**. (**b**) Quantification of GFP- and mCherry-positive AQR/PQR and URX neurons in WT (purple) or animals expressing P*egl-17::gfp-TLR* (blue). $N = 88–121$.

on the left (QL) and right side (QR) of the animal produce cells that migrate in opposite directions along the anterioposterior body axis[19]. QR descendants (QR.x) migrate anteriorly; AQR cells reach their final destination close to the posterior of URX cells, whereas AVM cells move past ALM cells (**Fig. 3a**,**b**). The QL-descendant PQR cells migrate posteriorly near PLMs, whereas PVM cells stay in the birthplace (**Fig. 3a**,**b**).

We first examined whether somatic TALENs of a gene, for which the null phenotype in Q-cell migration is already known and is not lethal, could reproduce the migration defects. *mab-5* encodes a homeodomain transcription factor, and with a *mab-5*(*e2088*) null allele, QL.x switches its posterior migration to the anterior whereas QR.x anterior migration is not affected (**Fig. 3c**)[20]. We generated transgenic animals expressing somatic *mab-5* TALEN constructs in Q cells using the P*egl-17* promoter. We found that QL.x (PQR and PVM cells) switched their posterior migration to the anterior in 61% of the *mab-5*-TALEN animals ($n = 71$) and that QR.x (AQR and AVM cells) migrate normally to the anterior in these animals (**Fig. 3d** and **Supplementary Fig. 4**). Our data demonstrate that somatic TALENs of *mab-5* can specifically recapitulate the QL.x migration phenotype in *mab-5* null allele.

We next used somatic TALENs to mutate *cor-1* within Q-cell lineages and then analyzed migration and other events in Q-cell descendants. In transgenic animals expressing P*egl-17::cor-1-TALENs* in Q-cell lineages, Q-cell descendants moved in the correct directions; however, their migration distances were substantially reduced (**Fig. 3e**,**f**). AQR cells went further in the posterior direction than URX cells, whereas AVM cells were in the posterior of ALM cells and PQR cells were close to the nonmigratory PVM cells (**Fig. 3e**,**f**). We showed that deletion/insertion mutations by *cor-1* TALENs caused the frameshift in the *cor-1* open reading frame (**Supplementary Fig. 5**). These data demonstrate that the disruption of *cor-1* in Q-cell lineages reduces Q-cell migration (**Fig. 3a**). Defects in Q-cell asymmetric division or apoptosis produce ectopic neurons, whereas the failure of Q-cell survival should cause a loss of neurons. Quantifying the number of A/PQR or A/PVM cells in these animals ($n = 83$), we did not find any gain or loss of these neurons. Our data demonstrate that COR-1 is only essential for cell migration and is not involved in cell proliferation or survival, at least, in *C. elegans* Q-cell lineages.

Compared to current *C. elegans* conditional gene-inactivating techniques or tools for reverse genetics, somatic TALENs have several advantages. First, the technique conditionally edits the WT *C. elegans* genome and can generate mutations in the specific somatic cell lineages or developmental stages, providing a versatile tool to address gene function. Alternative techniques such as the Cre/*LoxP* system[4] and FLP-FRT conditional system[2] were applied in *C. elegans*, but a null allele mutant of the target gene, which is not available for every *C. elegans* gene, is required to start with. Somatic TALENs can work on any genome, which will greatly facilitate functional genomics. Second, somatic TALENs can produce robust phenotypes with low variability, which can be explained by the same type of mutation (e.g., frameshift) that is generated during DNA repair in most cells (**Fig. 1e** and **Supplementary Figs. 1b** and **5**). A conditional RNA interference (RNAi) technique was developed using neuron-specific expression of the membrane protein SID-1 in the *sid-1* mutant background[21]. However, RNAi is sensitive to experimental conditions and its use can cause residual gene expression[6], causing high variability or low penetrance of the phenotype. Furthermore, *C. elegans* neurons are well-known for being refractory to RNAi[6]. Indeed, our attempts to knock down *gfp* in transgenic animals expressing P*gcy-32::gfp*, including the use of various RNAi-sensitive mutants or the direct expression

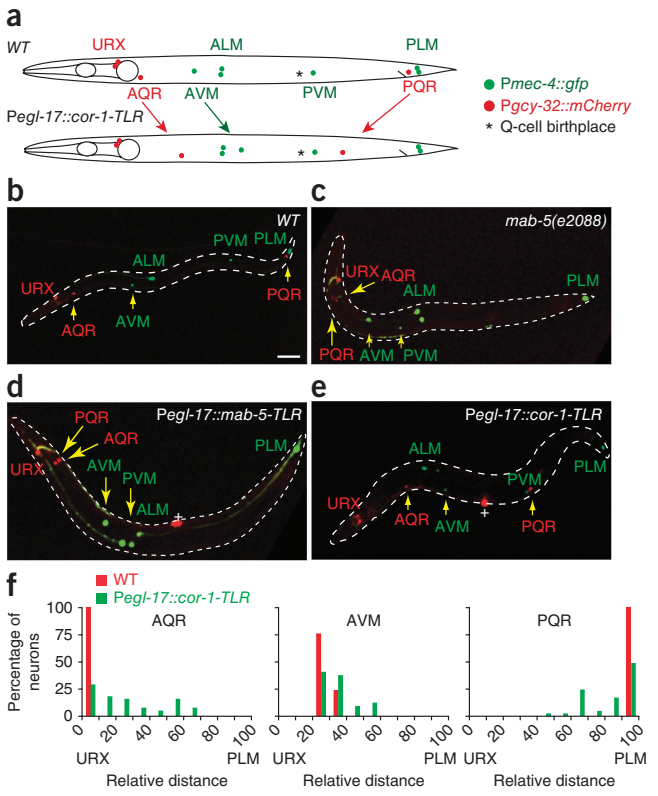

Figure 3 Somatic TALENs conditionally mutated *mab-5* and *cor-1* and revealed the function of COR-1 in *C. elegans* Q-cell migration. (**a**) The final position of mechanosensory neurons (green, marked by P*mec-4::gfp*) and oxygen sensory neurons (red, marked by P*gcy-32::mCherry*) in WT animals. Asterisks, Q-cell birthplace. AQR/PQR and AVM/PVM are derived from Q-cell lineages. (**b**–**e**) Still images show the final position of Q-cell progenies in WT animals (**b**) or *mab-5(e2088)* mutants (**c**) or in worms expressing Q cell–specific *mab-5* (**d**) or *cor-1* (**e**) TALENs constructs driven by P*egl-17*. Yellow arrows point to the position of AQR, AVM, PVM and PQR. URX and PLM were used as fiducial markers to quantify their positions. AQR and AVM are on the right side of the animal whereas PQR and PVM are on the left. In *mab-5* mutants or *mab-5* TALEN animals, PVM is always more posterior than AVM. Dashed lines show the animal periphery. +, vulva. Scale bar, 50 µm. (**f**) Quantification of AQR (left), AVM (middle) and PQR (right) positions in WT (red; $n = 25$) and animals with *cor-1* TALENs in Q cells (green; $n = 40$).

of double-stranded RNA in target neurons, have not been successful (data not shown). By contrast, somatic TALENs can efficiently eliminate GFP fluorescence from the P*gcy-32::gfp* transgene (**Fig. 2a**,**b**, and **Supplementary Figs. 1b** and **3b**,**c**). Temperature-sensitive alleles can be utilized to conditionally inactivate genes, but the penetrance is often limited because of only partial loss of function of the gene[5]. Moreover, producing a specific temperature-sensitive allele of every gene in *C. elegans* does not seem feasible. A third potential advantage is that the somatic TALENs technique is relatively fast and efficient. Starting from the experimental design, it might take only 3 weeks or less to obtain edited animals for phenotypic analysis.

This study focused on the local genome-editing changes induced by somatic TALENs, and more complete analyses will be required to assess potential off-target effects of TALENs. Prior deep sequencing studies on germ line–transmitting fish or human pluripotent cells did not uncover off-target effects caused by TALENs[10,22]. Our somatic TALENs used obligate heterodimer-based nuclease fusions, which can considerably increase the specificity of the gene-editing[23]. In our studies of *dpy-5*, *lon-2*, *mab-5*, *cor-1* and *gfp*, we did not notice additional morphological defects. Moreover, somatic TALENs of *cor-1* revealed its essential role in cell migration but not in other events, suggesting that somatic TALENs may not generally result in off-target effects.

In conclusion, somatic TALENs offer a useful tool for the *C. elegans* community. In principle, somatic TALENs should be useful in a variety of other organisms, including those in which conditional knockout techniques do not exist or are time consuming to implement. If the issue of off-target mutations can be addressed in the future, somatic TALENs may also provide an alternative strategy for gene therapy.

METHODS

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

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

ACKNOWLEDGMENTS

We thank B. Zhang and R. Jiao for TALEN reagents and W. Zhong, D. Huangfu, L. Cai and P. Gonczy for the helpful discussion. This work was supported by National Basic Research Program of China to W.L. and G.O. (973 Program, 2012CB966800, 2012CB945002 and 2013CB945600), the National Natural Science Foundation of China to G.O., X.W., Y.C., Y.Y. and W.L. (31201009, 31201048, 31222035, 31101002, 31100972, 31171295 and 31190063), the Natural Science Foundation of Beijing to X.W. (5123045), Chinese Academy of Sciences and the Junior Thousand Talents Program of China to G.O.

AUTHOR CONTRIBUTIONS

Z.C., P.Y., W.L. and G.O. designed experiments; Z.C., P.Y., X.W., Y.C., G.F., Y.Y., X.L., P.Y., Z.Z. and W.L. performed experiments; G.O. wrote the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. The *C. elegans* Deletion Mutant Consortium. Large-scale screening for targeted knockouts in the *Caenorhabditis elegans* genome. *G3* (*Bethesda*) **2**, 1415–1425 (2012).
2. Davis, M.W., Morton, J.J., Carroll, D. & Jorgensen, E.M. Gene activation using FLP recombinase in C. elegans. *PLoS Genet.* **4**, e1000028 (2008).
3. Frokjaer-Jensen, C. *et al.* Targeted gene deletions in *C. elegans* using transposon excision. *Nat. Methods* **7**, 451–453 (2010).
4. Garrison, J.L. *et al.* Oxytocin/vasopressin-related peptides have an ancient role in reproductive behavior. *Science* **338**, 540–543 (2012).
5. O'Rourke, S.M. *et al.* A survey of new temperature-sensitive, embryonic-lethal mutations in C. elegans: 24 alleles of thirteen genes. *PLoS ONE* **6**, e16644 (2011).
6. Pastrana, E. A toolset for the proficient geneticist. *Nat. Methods* **7**, 488–489 (2010).
7. Wood, A.J. *et al.* Targeted genome editing across species using ZFNs and TALENs. *Science* **333**, 307 (2011).
8. Joung, J.K. & Sander, J.D. TALENs: a widely applicable technology for targeted genome editing. *Nat. Rev. Mol. Cell Biol.* **14**, 49–55 (2013).
9. Miller, J.C. *et al.* A TALE nuclease architecture for efficient genome editing. *Nat. Biotechnol.* **29**, 143–148 (2011).
10. Huang, P. *et al.* Heritable gene targeting in zebrafish using customized TALENs. *Nat. Biotechnol.* **29**, 699–700 (2011).
11. Thacker, C., Sheps, J.A. & Rose, A.M. Caenorhabditis elegans dpy-5 is a cuticle procollagen processed by a proprotein convertase. *Cell Mol. Life Sci.* **63**, 1193–1204 (2006).
12. Stringham, E.G., Dixon, D.K., Jones, D. & Candido, E.P. Temporal and spatial expression patterns of the small heat shock (hsp16) genes in transgenic Caenorhabditis elegans. *Mol. Biol. Cell* **3**, 221–233 (1992).
13. Mello, C.C., Kramer, J.M., Stinchcomb, D. & Ambros, V. Efficient gene transfer in *C. elegans*: extrachromosomal maintenance and integration of transforming sequences. *EMBO J.* **10**, 3959–3970 (1991).
14. Gumienny, T.L. *et al.* Glypican LON-2 is a conserved negative regulator of BMP-like signaling in Caenorhabditis elegans. *Curr. Biol.* **17**, 159–164 (2007).
15. Shiow, L.R. *et al.* The actin regulator coronin 1A is mutant in a thymic egress-deficient mouse strain and in a patient with severe combined immunodeficiency. *Nat. Immunol.* **9**, 1307–1315 (2008).
16. Chan, K.T., Creed, S.J. & Bear, J.E. Unraveling the enigma: progress towards understanding the coronin family of actin regulators. *Trends Cell Biol.* **21**, 481–488 (2011).
17. Foger, N., Rangell, L., Danilenko, D.M. & Chan, A.C. Requirement for coronin 1 in T lymphocyte trafficking and cellular homeostasis. *Science* **313**, 839–842 (2006).
18. Mueller, P., Liu, X. & Pieters, J. Migration and homeostasis of naive T cells depends on coronin 1-mediated prosurvival signals and not on coronin 1-dependent filamentous actin modulation. *J. Immunol.* **186**, 4039–4050 (2011).
19. Chai, Y. *et al.* Live imaging of cellular dynamics during *Caenorhabditis elegans* postembryonic development. *Nat. Protoc.* **7**, 2090–2102 (2012).
20. Kenyon, C. A gene involved in the development of the posterior body region of C. elegans. *Cell* **46**, 477–487 (1986).
21. Calixto, A., Chelur, D., Topalidou, I., Chen, X. & Chalfie, M. Enhanced neuronal RNAi in C. elegans using SID-1. *Nat. Methods* **7**, 554–559 (2010).
22. Ding, Q. *et al.* A TALEN genome-editing system for generating human stem cell–based disease models. *Cell Stem Cell* **12**, 238–251 (2013).
23. Doyon, Y. *et al.* Enhancing zinc-finger-nuclease activity with improved obligate heterodimeric architectures. *Nat. Methods* **8**, 74–79 (2011).