Selection-free zinc-finger-nuclease engineering by context-dependent assembly (CoDA)

Jeffry D Sander[1–3], Elizabeth J Dahlborg[1,2], Mathew J Goodwin[1,2], Lindsay Cade[4], Feng Zhang[5], Daniel Cifuentes[6], Shaun J Curtin[7], Jessica S Blackburn[1,3], Stacey Thibodeau-Beganny[1,2], Yiping Qi[5], Christopher J Pierick[5], Ellen Hoffman[6], Morgan L Maeder[1,2,8], Cyd Khayter[1,2], Deepak Reyon[9], Drena Dobbs[9], David M Langenau[1,3,8], Robert M Stupar[7], Antonio J Giraldez[6], Daniel F Voytas[5], Randall T Peterson[4,10,11], Jing-Ruey J Yeh[4,10] & J Keith Joung[1–3,8]

Engineered zinc-finger nucleases (ZFNs) enable targeted genome modification. Here we describe context-dependent assembly (CoDA), a platform for engineering ZFNs using only standard cloning techniques or custom DNA synthesis. Using CoDA-generated ZFNs, we rapidly altered 20 genes in *Danio rerio*, *Arabidopsis thaliana* and *Glycine max*. The simplicity and efficacy of CoDA will enable broad adoption of ZFN technology and make possible large-scale projects focused on multigene pathways or genome-wide alterations.

Engineered zinc-finger nucleases (ZFNs) can be used to introduce targeted alterations into genomes of model organisms, plants and human cells[1,2]. Repair of ZFN-induced double-strand breaks by error-prone nonhomologous end joining leads to efficient introduction of insertion or deletion mutations at the site of the double-strand break. Alternatively, repair of a double-strand break by homology-directed repair with an exogenously introduced donor template can promote efficient introduction of alterations or insertions at or near the break site.

Widespread adoption and large-scale use of ZFN technology have been hindered by continued lack of a robust, easy-to-use and publicly available method for engineering zinc-finger arrays. In one approach, known as modular assembly, preselected zinc-finger modules are joined into arrays[3], a procedure simple enough for any researcher to implement. Some recent reports have indicated a high failure rate for this method[4,5], although the consequent need to construct and test large numbers of ZFNs for any given target gene can be mitigated by using a more limited subset of modules[6]. We recently described a robust selection-based method known as oligomerized pool engineering (OPEN)[7], but the labor and expertise required to screen combinatorial libraries have limited its broad adoption[3]. Researchers at Sangamo BioSciences have also developed a platform for engineering ZFNs; although some details of this method have been published[8], implementation requires access to a proprietary archive of engineered zinc-finger units[9]. Researchers may purchase customized ZFNs made by the Sangamo approach through the Sigma-Aldrich CompoZr service, but the cost of these proteins[9] limits the scale and scope of projects that can be performed.

Here we describe context-dependent assembly (CoDA), a publicly available platform of reagents and software that is simple to practice and has a success rate for generating active zinc-finger arrays comparable to that of selection-based methods such as OPEN. With the CoDA approach, three-finger arrays are assembled using N- and C-terminal fingers that have been previously identified in other arrays containing a common middle finger (F2 units) (**Fig. 1**). CoDA can be implemented by using a large archive consisting of 319 N-terminal-end fingers (F1 units) and 344 C-terminal-end fingers (F3 units) (**Supplementary Tables 1** and **2**) engineered to function well when positioned adjacent to one of 18 fixed F2 units (Online Methods). Thus, in contrast to modular assembly, CoDA does not treat fingers as independent modules but instead explicitly accounts for context-dependent effects between adjacent fingers[10,11], thereby increasing the probability that a multifinger array will function well. CoDA is rapid and requires neither specialized expertise nor labor-intensive selections; dozens of multifinger arrays can be constructed in 1–2 weeks or less using standard cloning techniques or commercial DNA synthesis.

To test the CoDA approach, we assembled 181 three-finger arrays and evaluated each for its ability to bind its cognate DNA target site using an established bacterial two-hybrid (B2H) reporter assay[4,7]. Previous work has shown that three-finger arrays that do not activate transcription by more than 1.57-fold in the

[1]Molecular Pathology Unit and Center for Cancer Research, Massachusetts General Hospital, Charlestown, Massachusetts, USA. [2]Center for Computational and Integrative Biology, Massachusetts General Hospital, Charlestown, Massachusetts, USA. [3]Department of Pathology, Harvard Medical School, Boston, Massachusetts, USA. [4]Cardiovascular Research Center, Massachusetts General Hospital, Charlestown, Massachusetts, USA. [5]Department of Genetics, Cell Biology and Development and Center for Genome Engineering, University of Minnesota, Minneapolis, Minnesota, USA. [6]Genetics Department, Yale University School of Medicine, New Haven, Connecticut, USA. [7]Department of Agronomy and Plant Genetics, University of Minnesota, St. Paul, Minnesota, USA. [8]Biological and Biomedical Sciences Program, Harvard Medical School, Boston, Massachusetts, USA. [9]Department of Genetics, Development and Cell Biology and Bioinformatics and Computational Biology Program, Iowa State University, Ames, Iowa, USA. [10]Department of Medicine, Harvard Medical School, Boston, Massachusetts, USA. [11]Broad Institute, Cambridge, Massachusetts, USA. Correspondence should be addressed to J.K.J. (jjoung@partners.org).

RECEIVED 17 JUNE; ACCEPTED 16 NOVEMBER; PUBLISHED ONLINE 12 DECEMBER 2010; DOI:10.1038/NMETH.1542

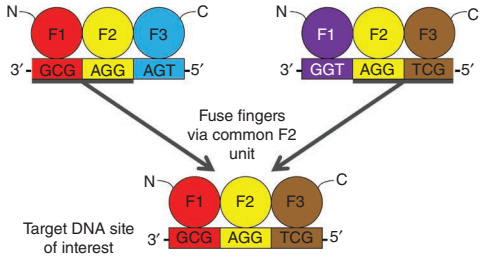

Figure 1 | Schematic overview of CoDA. Zinc fingers (units F1–F3) and their respective 3-bp DNA 'subsites' are shown. Two different three-zinc-finger arrays, each engineered to bind different 9-bp target sites and that have in common a middle F2 unit, can be used to create a three-finger array with a new specificity by joining together the F1 unit from the first array, the F2 unit, and the F3 unit from the second array.

B2H reporter assay are likely to be inactive as ZFNs in mammalian cells[4] and those that activate by threefold or more have a high probability of functioning efficiently as ZFNs in zebrafish, plant and human cells[7,12–15]. Of the 181 CoDA-generated arrays we tested using the B2H reporter assay, <8% (14 arrays) activated transcription by <1.57-fold and >76% (139 arrays) activated transcription by greater than threefold (**Supplementary Fig. 1** and **Supplementary Table 3**). These 'failure' and 'success' rates for DNA-binding activity (as predicted by the B2H reporter assay) are comparable to what we have previously observed with three-finger arrays made by OPEN[7]. Because so few (<25%) of the CoDA-generated arrays we tested gave less than threefold activation in the B2H reporter assay, our results suggest that one could skip the B2H reporter assay step and instead directly test the arrays in the cell type of interest.

We compared the efficacy of CoDA with that of modular assembly by using both approaches to construct three-finger arrays for 26 different 9–base-pair (bp) sites and by testing these proteins for DNA-binding activity in the B2H reporter assay (**Supplementary Table 4**). For these sites that we could target by both methods, CoDA-generated zinc-finger arrays performed better than those generated by modular assembly as judged by multiple comparisons of their DNA-binding activities (**Supplementary Fig. 2** and **Supplementary Discussion**). The most likely explanation for the relatively higher rates of generating active arrays by CoDA is its explicit consideration of context-dependent activities between fingers[10,11]. This difference will be more pronounced when one targets a ZFN site because two functional arrays must be engineered to create a ZFN dimer.

We used CoDA to engineer ZFNs for endogenous gene targets in zebrafish and plants. Using CoDA-generated zinc-finger arrays that activated transcription at least threefold in the B2H reporter assay, we constructed ZFN pairs for 24 gene targets in zebrafish, 13 gene targets in *Arabidopsis thaliana* and one target present in two duplicated genes in soybean (**Table 1**). CoDA-generated ZFNs induced targeted insertion or deletion mutations with high efficiencies in 12 of 24 zebrafish target sites (mutation frequencies, ≤1% to 16.7%), in 6 of 13 *Arabidopsis* gene targets (mutation frequencies, ≤1% to 8.4%) and in a target site present in two duplicated soybean genes in transformed root tissue (mutation frequencies, 18.8% and 10.7%) (**Table 1** and **Supplementary Figs. 3** and **4**).

Our overall per-target success rate for obtaining mutations with CoDA-generated ZFNs is 50% (19 of 38 target sites) in zebrafish and plants, a frequency comparable to our success rates of ~67% (16 of 24 target sites) with OPEN-generated ZFNs in zebrafish, plants and human cells (refs. 7,12–15 and unpublished data). For CoDA, success rates for obtaining mutations as calculated per ZFN pair and per ZFN target site are the same because only

Table 1 | Endogenous zebrafish and plant genes targeted by CoDA-generated ZFNs

Gene symbol	Organism	ZFN target site[a]	Mutant alleles/ total alleles	Mutation frequency (%)
Dcl4a	Soybean	TGCTTCATCacaatGGAGATGAT	6/32	18.8
Dcl4b	Soybean	TGCTTCATCacaatGGAGATGAT	3/28	10.7
MPK8	Thale cress	CTCCACAACatcagGATGACGAA	7/83	8.4
MPK11	Thale cress	CTCTTCGTCctatcgGCAGAGGCG	3/90	3.3
MKK9	Thale cress	GCCAGCGACggtggtGGTGGTGGC	3/95	3.2
MPK15	Thale cress	TTCTTCATCcagatGTTGTTGAG	2/73	2.7
MAPKKK18	Thale cress	CCCTTCCACaacaacGGAGAAGCT	2/75	2.7
GA3OX2	Thale cress	AGCTACGCCgtagccGGAGACGCC	1/94	≤1
MAPKKK1	Thale cress	GGCACCTCCgatttcGTGGAGGAA	0/190	0
MAPKKK12	Thale cress	TCCTCCACCgaatcGACGGCGCT	0/187	0
MAPKKK12	Thale cress	TTCCTCCACcgaatcGACGGCGCT	0/186	0
MAPKKK4	Thale cress	GTCTCCGCCtaggaGATGCAGAC	0/190	0
MPK15	Thale cress	TGCTTCTTCatccaGATGTTGTT	0/94	0
MPK4	Thale cress	CTCTTCGTCctatcgGTAGAGGCG	0/190	0
TZP	Thale cress	TTCGTCTTCgagtcGTCGTTGTT	0/141	0
actn1	Zebrafish	GCCTTCTCCggggcGCAGAAGGT	10/60	16.7
rag2	Zebrafish	ATCTTCTGCtccaggGGTGAAGGT	4/52	7.7
gad2	Zebrafish	AGCCGCAGCtctcgGCTGTAGAC	3/43	7
lmna	Zebrafish	CTCTTCTCCcccagaGCTGTGGAG	2/41	4.9
apoeb	Zebrafish	CCCCTCAGCccagaTGGGAGGAG	3/64	4.7
trpm7	Zebrafish	CACACCTGCacacaGATGCTGCT	2/55	3.6
grip1	Zebrafish	GGCCACCTCcaccaGCAGCGGGC	3/90	3.3
pclo	Zebrafish	CCCCTCTCCtcaaaGCAGATGCA	3/96	3.1
jak3	Zebrafish	GGCCCCACCaagcctGCTGGAGGA	1/71	≤1
ago1	Zebrafish	CTCTGCCGCcacctaGAGGATGGT	1/96	≤1
slitrk1	Zebrafish	GCCCACAGCaatggcGGAGCCGCC	1/96	≤1
bmpr2a	Zebrafish	GACTTCCTCtctgtGCAGTCGGC	1/117	≤1
bmpr2a	Zebrafish	ACCTCCTGCagtgtGAGGTTGTC	0/156	0
cnot1	Zebrafish	GGCGTCCACgtacgaGCGGAGGAG	0/93	0
ctcf	Zebrafish	TTCCTCCTCctgatGCGGAGGCT	0/96	0
dicer1	Zebrafish	TTCTGCAGCtcaatGGAGATGGT	0/96	0
dicer1	Zebrafish	AGCTTCCTCcgccgGAAGTTGAG	0/96	0
drosha	Zebrafish	GTCCTCCTCatggcgGTCGATGGT	0/96	0
g6pcb	Zebrafish	TCCCACTGCtgattGTAGGTGGA	0/134	0
nedd4l	Zebrafish	AACCGCACCacacaGTGGAAGAG	0/86	0
nod2	Zebrafish	AACTACAACattaggGCTGGAGGA	0/103	0
rag1	Zebrafish	GTCCTCCCCttcaaGTCGAATAG	0/91	0
th2	Zebrafish	CTCCTCCTCaaacacGAAGCTGTC	0/142	0
tp53	Zebrafish	AGCAGCTGCatgggGGGGATGAA	0/107	0

[a]Target sites in each gene are written 5′ to 3′ with the two half-sites targeted by the zinc-finger arrays (uppercase letters) and the intervening spacer sequence (lowercase letters).

a single CoDA-generated ZFN pair is tested per ZFN target site. Although we do not know why some CoDA- and OPEN-generated ZFNs do not induce mutations, we hypothesize that chromatin state or DNA methylation of the site, or protein stability or folding might be responsible. Regardless of the precise mechanism, we recommend that users of CoDA plan to make ZFNs for at least two target sites per gene of interest to increase the likelihood that at least one pair will introduce mutations.

CoDA still has some limitations compared to existing methods. Although modular assembly was less efficient than CoDA in our direct comparisons, modular assembly can potentially be used to target sites that CoDA currently cannot target[5,6], and one recent report demonstrated a comparable success rate of 23% for modular assembly using a more limited subset of modules[6]. In addition, although CoDA accounts for context-dependence between adjacent fingers, it also has some limitations relative to selection-based methods such as OPEN. For example, CoDA constrains the identity of the F2 unit and does not 'balance' the effects of all three fingers on affinity and specificity of the final array. In addition, CoDA in its current form guides assembly of arrays to 9-bp target sites, ignoring the identities of the adjacent upstream and downstream bases. Thus, for highly demanding therapeutic applications (for example, introduction of alterations into human pluripotent stem cells[13]), ZFNs made by OPEN may still be preferable to those made by CoDA, and it may be necessary to engineer zinc-finger arrays with greater specificities. Nonetheless, our overall results demonstrate that CoDA is a method for assembling zinc-finger arrays that accounts for context-dependent effects, is easier to perform than OPEN selections and yields ZFNs that efficiently modify genes.

With the current archive of CoDA units, a potential ZFN target site can be found approximately once in every 500 bp of random sequence (**Supplementary Discussion**). However, actual targeting range can be higher, depending on genomic sequences. For example, ~81% of 27,305 unique protein-coding transcripts in the zebrafish genome (Ensembl Zv8.57 database) contain one or more potentially targetable ZFN sites (mean, 4.37 sites), a frequency equivalent to one potential site every ~400 bp of transcript-coding sequence. By contrast, ~63% of 33,200 unique protein coding transcripts in the *Arabidopsis* genome (The Arabidopsis Information Resource 9 release) contain one or more potential ZFN target sites (mean, 2.45 sites), a frequency equal to one potential site every ~790 bp of transcript-coding sequence. We updated our publicly available web-based zinc finger targeter (ZiFiT) program (http://bindr.gdcb.iastate.edu/ZiFiT/ or http://www.zincfingers.org/software-tools.htm) to enable users to identify potential CoDA ZFN target sites in any given gene sequence (**Supplementary Fig. 5** and **Supplementary Discussion**).

In summary, CoDA is an effective alternative method for using publicly available reagents to engineer ZFNs. The rapidity and high success rate of CoDA enabled us to mutate 20 endogenous genes in three different organisms. CoDA will foster broader adoption of ZFN technology and also enable large-scale ZFN projects focused on multigene pathways or genome-wide alterations that are difficult to implement using existing methodologies.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturemethods/.

Note: Supplementary information is available on the Nature Methods website.

ACKNOWLEDGMENTS

This work was supported by US National Institutes of Health grants R01 GM088040 (J.K.J. and R.T.P.), R01 GM069906 (J.K.J.), R24 GM078369 (J.K.J.), R01 CA140188 (J.-R.J.Y.), R01 GM081602 (A.J.G.), RC2 MH089956 (A.J.G.), K01 AG031300 (J.-R.J.Y.), K01 AR055619 (D.M.L.) and T32 CA009216 (J.D.S. and J.S.B.); by National Science Foundation grant DBI 0923827 (D.F.V., D.D. and J.K.J.); by the Claflin Distinguished Scholar Award (J.-R.J.Y.); by the Minnesota Soybean Research and Promotion Council (S.J.C. and R.M.S.); and by Alex's Lemonade Stand and the Leukemia Research Foundation (D.M.L.).

AUTHOR CONTRIBUTIONS

J.D.S. and J.K.J. conceived of the CoDA engineering method; J.D.S., S.J.C., D.M.L., R.M.S., A.J.G., D.F.V., R.T.P., J.-R.J.Y. and J.K.J. designed research; J.D.S., E.J.D., M.J.G., L.C., F.Z., D.C., S.J.C., J.S.B., S.T.-B., Y.Q., C.J.P., E.H., M.L.M. and C.K. performed experiments; J.D.S., D.R. and D.D. identified potential genomic CoDA target sites; and J.D.S., R.M.S., D.F.V., R.T.P., J.-R.J.Y. and J.K.J. wrote the paper.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details accompany the full-text HTML version of the paper at http://www.nature.com/naturemethods/.

Published online at http://www.nature.com/naturemethods/.
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/.

1. Carroll, D. *Gene Ther.* **15**, 1463–1468 (2008).
2. Cathomen, T. & Joung, J.K. *Mol. Ther.* **16**, 1200–1207 (2008).
3. Kim, J.S., Lee, H.J. & Carroll, D. *Nat. Methods* **7**, 91 (author reply, 91–92) (2010).
4. Ramirez, C.L. *et al. Nat. Methods* **5**, 374–375 (2008).
5. Kim, H.J., Lee, H.J., Kim, H., Cho, S.W. & Kim, J.S. *Genome Res.* **19**, 1279–1288 (2009).
6. Lee, H.J., Kim, E. & Kim, J.S. *Genome Res.* **20**, 81–89 (2010).
7. Maeder, M.L. *et al. Mol. Cell* **31**, 294–301 (2008).
8. Doyon, Y. *et al. Nat. Biotechnol.* **26**, 702–708 (2008).
9. Pearson, H. *Nature* **455**, 160–164 (2008).
10. Isalan, M., Choo, Y. & Klug, A. *Proc. Natl. Acad. Sci. USA* **94**, 5617–5621 (1997).
11. Isalan, M., Klug, A. & Choo, Y. *Biochemistry* **37**, 12026–12033 (1998).
12. Foley, J.E. *et al. PLoS ONE* **4**, e4348 (2009).
13. Zou, J. *et al. Cell Stem Cell* **5**, 97–110 (2009).
14. Zhang, F. *et al. Proc. Natl. Acad. Sci. USA* **107**, 12028–12033 (2010).
15. Townsend, J.A. *et al. Nature* **459**, 442–445 (2009).

ONLINE METHODS

Identification of finger units for practicing CoDA. To identify 'fixed' F2 fingers for various 3-bp target subsites, we analyzed the amino acid sequences of F2 units from a collection of three-finger arrays previously identified from OPEN selections performed for over 130 different 9-bp sites (refs. 7,12–15 and M. Maeder *et al.*, unpublished data). From this analysis, we identified F2 units for 18 different 3-bp subsites that occurred in at least two or more different contexts. The F1 and F3 units found adjacent to these F2 units were also chosen as CoDA units because they had been selected to work well together. To obtain additional F1 and F3 CoDA units for other 3-bp subsites, we performed OPEN selections in which we interrogated combinatorial three-finger array libraries composed of a fixed F2 unit and randomized F1 and F3 units for binding to specific 9-bp target sequences. From these selections, we analyzed the amino acid sequences of three-finger arrays that activated transcription threefold or more in the B2H reporter assay to identify additional F1 and F3 finger units that worked well when positioned next to a specific fixed F2 CoDA unit. For selections that yielded multiple three-finger array clones, we chose F1 and F3 units that occurred the most frequently in multiple distinct arrays and/or that were found in three-finger arrays that gave the highest-fold activation in the B2H reporter assay. OPEN selections were performed essentially as described previously[7,16] but with the modification that a beta-lactamase antibiotic resistance gene was used for selection instead of the *HIS3* gene. This modified version of OPEN enabled selections to be performed with higher throughput (M.J.G. *et al.*, unpublished data). Each of the three-finger arrays from which the F1, F2 and F3 units were derived was determined to be active in a B2H reporter assay.

Construction of zinc-finger arrays by modular assembly. Construction of plasmids encoding the modularly assembled zinc-finger arrays used in this study has been described previously[4].

Construction of zinc-finger arrays by CoDA. To assemble CoDA zinc-finger arrays, DNA fragments encoding an F1–F2 cassette or an F3 cassette were amplified by PCR from plasmids using primer pairs OK1424 and OK1427 or OK1428 and OK1429, respectively. (Sequences of all primers are listed in **Supplementary Table 5**). The resulting PCR products were digested with DpnI (New England Biolabs) to degrade template plasmid DNA and purified using a Qiagen PCR purification kit. The cassettes were then fused together and amplified in a single PCR step using primers OK1430 and OK1432. PCR products encoding a three-finger array were then purified using a Qiagen PCR purification kit, treated with Pfu polymerase (Stratagene) in the presence of dTTP nucleotide to create overhangs, phosphorylated with T4 polynucleotide kinase (New England Biolabs) and ligated to a B2H expression plasmid (pMG414) in which the zinc-finger array is expressed as a fusion to a fragment of the yeast Gal11P protein[16]. All plasmids were sequenced using primer OK61.

B2H reporter assay. Zinc-finger arrays made by modular assembly or CoDA were each tested for binding to its cognate target site by measuring its ability to activate transcription in the B2H reporter assay as described previously[16,17]. All assays were performed in triplicate.

Zebrafish gene mutation analysis. Injection of zebrafish embryos, isolation of genomic DNA, limited-cycle PCR amplification of the locus of interest, cloning of PCR fragments using the TOPO TA Cloning kit (Invitrogen) and transformation of *Escherichia coli* were performed as described previously[12,18]. Resulting colonies were assessed for gene mutations by one of two methods: (i) direct sequencing of individual clones or (ii) screening of three pooled clones for alterations in PCR fragment size using fluorescence-based analysis as described previously[18], followed by identification of specific mutations by direct sequencing.

***Arabidopsis* gene mutation analysis.** ZFN transgene expression constructs, *Arabidopsis* transformation methods, induction of ZFN expression in *Arabidopsis* seedlings by β-estradiol and isolation of *Arabidopsis* genomic DNA were done as described previously[14]. ZFN recognition sites in the *Arabidopsis* genomic DNA were amplified by PCR, the resulting fragments were cloned using the TOPO TA Cloning kit, and DNA from individual colonies was sequenced to identify mutations at the ZFN recognition site.

Soybean gene mutation analysis. Cotyledons of the soybean variety Bert were transformed using a previously described hairy root transformation protocol[19]. The ZFN transgene was induced by application of 10 μM of β-estradiol (Sigma) on tissue culture medium. Hairy root DNA was isolated using the Qiagen DNeasy kit. Transformed roots were screened for the ZFN transgene using primers (forward primer, 5′-TGGATATGTATATGGTGGTAATGC-3′ and reverse primer, 5′-TTGAGCTTGTGGCGCAGCTCG-3′). Roots containing the transgene were then screened for mutations by a cleaved amplified polymorphic sequence (CAPS) analysis (forward primer, 5′-GTAAAAGATGTTGAAAGAAAGTTGG-3′ and reverse primer, 5′-GCTTTTGACTTGAGCATGATGG-3′) using restriction enzyme MslI, which digests the nucleotide sequence targeted for mutagenesis. A single root was identified as carrying putative mutations in the *Dcl4a* and *Dcl4b* genes. The targeted regions of *Dcl4a* and *Dcl4b* were amplified by PCR from this root using the CAPS primers. PCR fragments were cloned in pGem T-easy (Promega) and colony PCR products for 60 clones were subsequently sequenced. Mutations were identified via sequence alignments using MEGA 4.1 (ref. 20).

Identification of potential CoDA ZFN sites in *D. rerio* and *Arabidopsis*. Potential ZFN target sites in *D. rerio* and *Arabidopsis* were identified from the Ensembl (Zv8.57) and The *Arabidopsis* information resource (TAIR9) chromosomal assemblies and gene table files. Potential ZFN target sites were defined as those that could be targeted using the CoDA reagents described here and that had a spacer sequence of 5, 6 or 7 nucleotides that was entirely within an exon.

16. Maeder, M.L., Thibodeau-Beganny, S., Sander, J.D., Voytas, D.F. & Joung, J.K. *Nat. Protoc.* **4**, 1471–1501 (2009).
17. Wright, D.A. *et al. Nat. Protoc.* **1**, 1637–1652 (2006).
18. Foley, J.E. *et al. Nat. Protoc.* **4**, 1855–1867 (2009).
19. Govindarajulu, M., Elmore, J.M., Fester, T. & Taylor, C.G. *Mol. Plant Microbe Interact.* **21**, 1027–1035 (2008).
20. Tamura, K., Dudley, J., Nei, M. & Kumar, S. *Mol. Biol. Evol.* **24**, 1596–1599 (2007).

doi:10.1038/nmeth.1542