TITLE

Multigene Editing via CRISPR/Cas9 Guided by a Single-sgRNA Seed in *Arabidopsis*

RUNNING TITLE

A Single sgRNA Seed for Multigene-editing

KEYWORDS

Single sgRNA Seed, CRISPR/Cas9, Multigene Editing, *AtRPL10*

Summary We report that a single-sgRNA seed is capable of guiding CRISPR/Cas9 to simultaneously edit multiple genes *AtRPL10A*, *AtRPL10B* and *AtRPL10C* in *Arabidopsis*. Our results also demonstrate that it is possible to use CRISPR/Cas9 technology to create *AtRPL10* triple mutants which otherwise cannot be generated by conventional genetic crossing. Compared to other conventional multiplex CRISPR/Cas systems, a single sgRNA seed has the advantage of reducing off-target gene-editing. Such a single sgRNA seed-induced gene editing system might be also applicable to modify other homologous genes or even less-homologous sequences for multiple gene-editing in plants and other organisms.

Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)/CRISPR -associated 9 (CRISRP/Cas9) is an adaptive immune mechanism that protects bacteria and archaea from extrachromosomal DNA and viral invasions (Jinek et al. 2012). CRISPR/Cas9 generates double-stranded breaks (DSBs) under specific guidance of a single guide RNA (sgRNA). These DSBs can then be repaired either by homologous recombination or predominantly by non-homologous end-joining, which leads to introduction of mutations such as nucleotide substitution, insertion or deletion into the targeted DNA molecules (Jinek et al. 2012; Cong et al. 2013). Such an ancient defense has been exploited for efficient genome/gene editing in organisms across kingdoms (Jinek et al. 2012; Cong et al. 2013; Fu et al. 2013; Mao et al. 2013; Gao and Zhao, 2014; Ma et al. 2015; Yan et al., 2015; Kim et al., 2016; Shen et al., 2016).

Moreover, multiplex CRISPR/Cas9-based gene editing can also be simultaneously achieved through the use of different sgRNAs in animals and plants (Cong et al. 2013; Feng et al. 2014; Wang et al. 2015; Yan et al. 2015). It remains to be elucidated, however, whether multigene-editing via CRISPR/Cas9 can be directed by a single sgRNA seed.

To address this, we searched the *Arabidopsis* genome database and identified the *AtRPL10* family that includes three homologous members *AtRPL10A* (AT1G14320), *AtRPL10B* (AT1G26910) and *AtRPL10C* (AT1G66580) coding for the Ribosomal Protein Large 10 subunits. The three *AtRPL10* genes reside at different loci on *Arabidopsis* chromosome 1 (Figure S1), sharing 81-88% nucleotide (nt) identities, and their protein products are 95-98% identical (Table S1) *AtRPL10A* and *AtRPL10B* are expressed in female and male reproductive organs whilst *AtRPL10C* is restricted to pollen grains. The three multifunctional genes are involved in protein translation and plant response to viral infection and abiotic stress (Falcone Ferreyra et al. 2013; Zorzatto et al. 2015). Homozygous *AtRPL10A* T-DNA insertion mutation is lethal and RNAi of *AtRPL10B* affects plant growth, although *AtRPL10C* knockout results in no phenotypic change (Falcone Ferreyra et al. 2010). Interestingly, genetic crosses can generate *AtRPL10A*, *AtRPL10B* or *AtRPL10C* heterologous double, but not triple, mutants in *Arabidopsis* (Falcone Ferreyra et al. 2013).

We generated an '*AtRPL10* sgRNA+CRISPR/Cas9' construct in pCAMBIA1300 (Figure 1A). The *AtRPL10* sgRNA consists of an identical 19 nucleotides (ATGTTGGTATGAAGAGGAA) targeting the three genes. However, the protospacer adjacent motif (PAM) is AGG in *AtRPL10B* and *AtRPL10C*, but GGG in *AtRPL10A* (Figure 1B). *A. thaliana* ecotype Col-0 was transformed with the binary vector via the floral dip method (Supplemental Materials and Methods). Four independent Line7, Line9, Line10 and Line11 were created. Transgenic T1 plants from Lines7, 9 and 10 showed severe growth retardation and delayed flowering whilst Line11 had slightly weaker growth compared to the wild-type Col-0 plants (Figure 1C, D). These lines

showed similar phenotypes to *AtRPL10B* RNAi plants, but differed from *AtRPL10A* T-DNA insertion mutants or *AtRPL10C* knockout plants. To detect potential multigene-editing events in these transgenic lines, we first analyzed the sgRNA-targeted sequences using a high-fidelity PCR-RFLP (restriction fragment length polymorphism) assay. An *Ear*I site is located 4-9 nucleotides upstream of the *AtRPL10* sgRNA PAM sequence (Figure 1B), the region in which CRISPR/Cas9-mediated DSBs frequently occur (Jinek et al. 2012). We extracted genomic DNA from transgenic and non-transformed Col-0 plant leaf tissues and amplified the *AtRPL10* target sequences using gene-specific primers (Table S2). Incomplete *Ear*I-digestion of the resultant PCR products suggests that *AtRPL10A* and *AtRPL10C* were successfully edited in Line9 (Figure 1E).

To further characterize multigene-editing in these transgenic lines, we cloned the PCR products into pMD19-T (Supplemental Materials and Methods). Sequencing analyses showed that nucleotide deletions and/or replacements were introduced into *AtRPL10A*, *AtRPL10B* and *AtRPL10C* in all transgenic lines (Figures S2-13; Table S3; Dataset S1). However, the efficiency of multigene-editing of all target sequences was relatively higher (Figure 1F; Figures S5-7) although varied among *AtRPL10A* (8.8%), *AtRPL10B* (3.8%) and *AtRPL10C* (23.6%) in Line9 (Table 1). Using an alternative assay, we identified 7 more (4 deletion and 3 substitution) mutations that were introduced into *AtRPL10B* in Line9 (Figure 1G-I). In Line7 (Figures S2-4) and Line10 (Figures S8-10), we detected nucleotide deletions in *AtRPL10A* or *AtRPL10C* but not in *AtRPL10B*, whilst only point mutations were found in the three *AtRPL10* genes in Line11 (Figure S11-13). In total we sequenced 1,222 clones and identified 75 different mutations, 37 of which were a deletion of 2 nucleotides. There were single cases of 1nt or 4nt-deletions, and 36 cases of 1nt-substitution (Table S3). Nevertheless, multiple deletion and/or point mutations introduced by a single-sgRNA seed-directed CRISPR/Cas9 were correlated with the abnormal phenotypes in the transgenic lines (Figure 1C).

Multiplex gene editing through CRISPR/Cas9 that is directed by a number of different sgRNAs has been previously reported in animals and plants (Cong et al. 2013; Feng et al. 2014; Wang et al. 2015; Yan et al. 2015). In this letter, we show that a single-sgRNA seed is capable of guiding CRISPR/Cas9 to edit multiple genes in *Arabidopsis*. Secondly, we demonstrate that it is possible to use CRISPR/Cas9 technology to create *AtRPL10A/B/C* triple mutants which otherwise cannot be generated by conventional genetic crossing. Thirdly, we observe that most of mutations resulted from the single-sgRNA seed-guided CRISPR-Cas9 are 2nt-deletion or 1nt-substitution within the sgRNA-target sequences. This differs from a previous report that mutations induced by CRISPR/Cas9 were predominantly 1nt-insertion and short deletions of nucleotides (Feng et al. 2014), but consistent with others (Wang et al. 2015; Yan et al., 2015). Fourthly, the different *AtRPL10A/B/C*-editing efficiencies (Table 1), particularly in Line9, suggest that chromosomal locations of genes along with the contexts of their surrounding-sequences, heterochromatin architectures and/or DNA/histone methylation may affect the CRISPR/Cas9 system for editing multiple homologous genes (Kleinstiver et al. 2015). Nevertheless, Line9 may prove to be a valuable model to investigate positional effects on the ability of single sgRNA-directed CRISPR/Cas9 to target and edit multiple genes in plants. Lines7, 10 and 11 may be also useful to explore why the single-sgRNA directed CRISPR/Cas9 system preferably causes nucleotide substitution, rather than deletion mutations in target genes. It is interesting to note that all deletion mutations created in our transgenic lines result from removal of 1, 2 or 4 nucleotides, causing frameshifts of the target genes. Compared to conventional multiplex CRISPR/Cas systems (Fu et al. 2013), a single-sgRNA seed has the advantage of reducing off-target gene-editing. This approach is also applicable for the modification of other homologous genes. Moreover, considering how CRISPR/Cas9 recognizes canonical or non-canonical PAMs such as NGG, NGA, NGCG, TTN and YTN (Kleinstiver et al. 2015; Zetsche et al. 2015; Fonfara et al. 2016) as well as how sgRNAs interact with their target sequences (Jinek et al. 2012), it should also be possible to design a single 'less-stringent' sgRNA seed that may target less-homologous sequences for

multigene-editing in plants and other organisms.

ACKNOWLEDGMENTS

We thank Jian-kang Zhu for his kind gifts of AtU6-26-sgRNA-SK and 35S-Cas9-SK plasmids. We are grateful to Chun Wang for her suggestions and technical support. This work was supported by grants from the National Science Foundation of China (NSFC, 31200913), the Ministry of Agriculture of the People's Republic of China National Transgenic Program (2016ZX08009001-004), the China Scholarship Council (201709645003), NSFC (31370180; 31601765) and the Zhejiang Provincial Natural Science Foundation (LY15C140006, LY14C010005).

Zhiming Yu[1,§], Qiyuan Chen[1,§], Weiwei Chen[1], Xian Zhang[1], Fengling Mei[1], Pengcheng Zhang[1], Mei Zhao[1], Xiaohong Wang[1], Nongnong Shi[1], Stephen Jackson[2], Yiguo Hong[1,2,3,*]

[1]Research Centre for Plant RNA Signaling, College of Life and Environmental Sciences, Hangzhou Normal University, Hangzhou 310036, China

[2]Warwick-Hangzhou Joint RNA Signaling Laboratory, School of Life Sciences, University of Warwick, Coventry CV4 7AL, UK

[3]Worcester-Hangzhou Joint Molecular Plant Health Laboratory, Institute of Science and the Environment, University of Worcester, WR2 6AJ, UK

[§]These authors contributed equally to this article.

[*]**Correspondence: Yiguo Hong** (yiguo.hong@hznu.edu.cn, yiguo.hong@warwick.ac.uk, and y.hong@worc.ac.uk)

AUTHOR CONTRIBUTIONS

Z.Y. and Y.H. designed experiments; Z.Y. and Q.C. performed all experiments; W.C. and X.Z. analyzed bioinformatics data; J.N., F.M., P.Z., M.Z., X.W. and N.S. performed research. S.J. analysed the data and helped write the paper; Z.Y. and Y.H. wrote the paper.

REFERENCES

Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F (2013) Multiplex genome engineering using CRISPR/Cas systems. **Science** 339:819-823.

Falcone Ferreyra ML, Casadevall R, Luciani MD, Pezza A, Casati P (2013) New evidence for differential roles of l10 ribosomal proteins from Arabidopsis. **Plant Physiol** 163:378-391.

Falcone Ferreyra ML, Pezza A, Biarc J, Burlingame AL, Casati P (2010) Plant L10 ribosomal proteins have different roles during development and translation under ultraviolet-B stress. **Plant Physiol** 153:1878-1894.

Feng Z, Mao Y, Xu N, Zhang B, Wei P, Yang DL, Wang Z, Zhang Z, Zheng R, Yang L, Zeng L, Liu X, Zhu JK (2014) Multigeneration analysis reveals the inheritance, specificity, and patterns of CRISPR/Cas-induced gene modifications in Arabidopsis. **Proc Natl Acad Sci USA** 111:4632-4637.

Fonfara I, Richter H, Bratovič M, Le Rhun A, Charpentier, E (2016) The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA. **Nature** 532:517-521.

Fu Y, Foden JA, Khayter C, Maeder ML, Reyon D, Joung JK, Sander JD (2013) High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. **Nat Biotechnol** 31:822-826.

Gao Y, Zhao Y (2014) Self-processing of ribozyme-flanked RNAs into guide RNAs in vitro and in vivo for CRISPR-mediated genome editing. **J Integr Plant Biol.** 56:343-349.

Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. **Science** 337:816-821.

Kim H, Kim ST, Ryu J, Choi MK, Kweon J, Kang BC, Ahn HM, Bae S, Kim J, Kim JS, Kim SG (2016) A simple, flexible and high-throughput cloning system for plant genome editing via CRISPR-Cas system. **J Integr Plant Biol.** 58:705-712.

Kleinstiver BP, Prew MS, Tsai SQ, Topkar VV, Nguyen NT, Zheng Z, Gonzales AP, Li Z, Peterson RT, Yeh JR, Aryee MJ, Joung JK (2015) Engineered CRISPR-Cas9 nucleases with altered PAM specificities. **Nature** 523:481-485.

Ma X, Zhang Q, Zhu Q, Liu W, Chen Y, Qiu R, Wang B, Yang Z, Li H, Lin Y, Xie Y, Shen R, Chen S, Wang Z, Chen Y, Guo J, Chen L, Zhao X, Dong Z, Liu YG (2015) A robust CRISPR/Cas9 system for convenient, high-efficiency multiplex genome editing in monocot and dicot plants. **Mol Plant** 8:1274-1284.

Mao Y, Zhang H, Xu N, Zhang B, Gou F, Zhu JK (2013) Application of the CRISPR-Cas system for efficient genome engineering in plants. **Mol Plant** 6:2008-2011.

Shen L, Wang C, Fu Y, Wang J, Liu Q, Zhang X, Yan C, Qian Q, Wang K (2016) QTL editing confers opposing yield performance in different rice varieties. **J Integr Plant Biol.** doi: 10.1111/jipb.12501.

Wang ZP, Xing HL, Dong L, Zhang HY, Han CY, Wang XC, Chen QJ (2015) Egg cell-specific promoter-controlled CRISPR/Cas9 efficiently generates homozygous mutants for multiple target genes in Arabidopsis in a single generation. **Genome Biol.** 16:144.

Yan L, Wei S, Wu Y, Hu R, Li H, Yang W, Xie Q (2015) High-Efficiency Genome Editing in Arabidopsis Using YAO Promoter-Driven CRISPR/Cas9 System. Mol Plant. 8:1820-1823.

Yan M, Zhou SR, Xue HW (2015) CRISPR Primer Designer: Design primers for knockout and chromosome imaging CRISPR-Cas system. **J Integr Plant Biol.** 57:613-617.

Zetsche B, Gootenberg JS, Abudayyeh OO, Slaymaker IM, Makarova KS, Essletzbichler P, Volz SE, Joung J, van der Oost J, Regev A, Koonin EV, Zhang F (2015) Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. **Cell** 163:759-771.

Zorzatto C, Machado JP, Lopes KV, Nascimento KJ, Pereira WA, Brustolini OJ, Reis PA, Calil IP, Deguchi M, Sachetto-Martins G, Gouveia BC, Loriato VA, Silva MA, Silva FF, Santos AA, Chory J, Fontes EP (2015) NIK1-mediated translation suppression functions as a plant antiviral immunity mechanism. **Nature** 520:679-682.

SUPPORTING INFORMATION

Additional Supporting Information is available.

Figure S1. Physical Positions of *AtRPL10A*, *AtRPL10B* and *AtRPL10C* in Arabidopsis Chromosome 1.

AtRPL10A (AT1G14320): 4,888,214 – 4,889,661;
AtRPL10B (AT1G26910): 9,321,650 – 9,322,965;
Chromosome Centromere: 14,899,838 – 14,906,596;
AtRPL10C (AT1G66580): 24,839,165 – 24,840,612.

Figure S2. Triple Gene Editing in Line7.

(A) *AtRPL10A* sequences of individual clones. PAM sequence is boxed and the edited nucleotides are highlighted red.
(B) Representative chromatograms of *AtRPL10A* sequences with edited nucleotides

Figure S3. Triple Gene Editing in Line7.

(A) *AtRPL10B* sequences of individual clones. PAM sequence is boxed and the edited nucleotides are highlighted red.
(B) Representative chromatograms of *AtRPL10B* sequences with edited nucleotides.

Figure S4. Triple Gene Editing in Line7.

(A) *AtRPL10C* sequences of individual clones. PAM sequence is boxed and the edited nucleotides are highlighted red.
(B) Representative chromatograms of *AtRPL10C* sequences with edited nucleotides.

Figure S5. Triple Gene Editing in Line9.

(A) *AtRPL10A* sequences of individual clones. PAM sequence is boxed and the edited nucleotides are highlighted red.
(B) Representative chromatograms of *AtRPL10A* sequences with edited nucleotides.

Figure S6. Triple Gene Editing in Line9.

(A) *AtRPL10B* sequences of individual clones. PAM sequence is boxed and the edited nucleotides are highlighted red.
(B) Representative chromatograms of *AtRPL10B* sequences with edited nucleotides.

Figure S7. Triple Gene Editing in Line9.

(A) *AtRPL10C* sequences of individual clones. PAM sequence is boxed and the edited nucleotides are highlighted red.
(B) Representative chromatograms of *AtRPL10C* sequences with edited nucleotides.

Figure S8. Triple Gene Editing in Line10.

(A) *AtRPL10A* sequences of individual clones. PAM sequence is boxed and the edited nucleotides are highlighted red.
(B) Representative chromatograms of *AtRPL10A* sequences with edited nucleotides.

Figure S9. Triple Gene Editing in Line10.

(A) *AtRPL10B* sequences of individual clones. PAM sequence is boxed and the edited nucleotides

are highlighted red.
(B) Representative chromatograms of *AtRPL10B* sequences with edited nucleotides.

Figure S10. Triple Gene Editing in Line10.
(A) *AtRPL10C* sequences of individual clones. PAM sequence is boxed and the edited nucleotides are highlighted red.
(B) Representative chromatograms of *AtRPL10C* sequences with edited nucleotides.

Figure S11. Triple Gene Editing in Line11.
(A) *AtRPL10A* sequences of individual clones. PAM sequence is boxed and the edited nucleotides are highlighted red.
(B) Representative chromatograms of *AtRPL10A* sequences with edited nucleotides.

Figure S12. Triple Gene Editing in Line11.
(A) *AtRPL10B* sequences of individual clones. PAM sequence is boxed and the edited nucleotides are highlighted red.
(B) Representative chromatograms of *AtRPL10B* sequences with edited nucleotides.

Figure S13. Triple Gene Editing in Line11.
(A) *AtRPL10C* sequences of individual clones. PAM sequence is boxed and the edited nucleotides are highlighted red.
(B) Representative chromatograms of *AtRPL10C* sequences with edited nucleotides.

Table S1. Comparisons of AtRPL10 Genes and Their Protein Products

Table S2. Primers Used in This Study

Table S3. Multigene Editing and Their Impacts on Protein Sequences in CRISPR/Cas9 Transgenic Lines

Dataset S1. Sequences of the PCR Products for the Three *AtRPL10* Genes.
(A-B) Restriction fragment length polymorphism (RFLP) analysis of *AtRPL10A* (**A**), *AtRPL10B* (**B**) *AtRPL10C* (**C**). Sequences corresponding to the 'seed' sgRNA are indicated in lowercase. The *Ear*I digestion site (|) is indicated and its recognition sequence is underlined.

FIGURE LEGEND

Figure 1. A Single sgRNA Seed Directs CRISPR/Cas9 to Simultaneously Edit Three *AtRPL10* Homologous Genes.

(A) Schematic of the single sgRNA seed and CRISPR/Cas9 consturct in the binary vector pCAMBIA1300. Nucleotides corresponding to the sgRNA seed sequence are underlined. The AtU6-26 promoter (arrow), sgRNA and the scaffold, enhanced 35S promoter (arrow), NLS (nuclear localization signal)-tagged Cas9, hygromycin (HYG) as well as the right and left borders (RB and LB) in the binary vector are indicated.

(**B**) Comparison of the sgRNA seed-targeted *AtRPL10* gene sequences. The *Ear*I site is

underlined. The PAM sequences are highlighted red. Nucleotide coordinates are indicated.

(**C)** Phenotypes of transgenic plants of four independent lines. Bar = 3cm in Line7, Line9, Line10 and Line11. Bar = 5cm in Col-0.

(D) Confirmation of plant transformation. The Cas9 gene was detected in four transgenic lines as indicated. A BM2000 DNA ladder (Marker) as well as the size and position of the Cas9 transgene PCR fragment are indicated.

(E) PCR-RFLP assay of multiple gene-editing in four transgenic lines. Gene-specific PCR products were digested with *Ear*I. Incomplete digestion shows three clear bands, indicating that successful editing of *AtRPL10A* and *AtRPL10C* in Line9. A BM2000 DNA ladder (Marker) was included in gel electrophoresis.

(F) Sequencing analysis of multiple gene-editing in Line9. Representative sequencings show indels in *AtRPL10A*, *AtRPL10B* and *AtRPL10C*. The sgRNA target sequences are shown in lowercase.

(G-I) PCR-RFLP and sequencing assays of *AtRPL10B* editing in Line9. After *Ear*I digestion, residual DNA in the postion of the red-box was extracted from the argarose gel **(G)** and subcloned for sequencing analysis **(H)**. A BM2000 DNA ladder (M) was included in gel electrophoresis. Sequences of 19 individual clones for *AtRPL10B* were aligned, and mutations with two nucleotide-deletion (red arrow) or single nucleotide-substitution (highlighted red) are indicated **(I)**. RD stands for restriction endonuclease digestion.

TABLE

Table 1. Summary of Multigene Editing Efficiency*.

Transgenic Lines	*AtRPL10A*		*AtRPL10B*		*AtRPL10C*	
	Deletion	Point Mutation	Deletion	Point Mutation	Deletion	Point Mutation
Line 7	1/95	2/95	0/99	5/99	0/112	6/112
Line 9	7/102	2/102	1/105	3/105	25/123	4/123
Line 10	0/106	3/106	0/101	2/101	1/96	1/96
Line 11	0/93	2/93	0/109	4/109	0/62	2/62

*The number of CRISPR/Cas9 edited sequences (clones) out of the total number of sequenced samples for the three *AtRPL10* genes in each of the transgenic lines.