1 **Advances and perspectives on the use of CRISPR/Cas9 systems in plant**
2 **genomics research**

3 Degao Liu, Rongbin Hu, Kaitlin J. Palla, Gerald A. Tuskan and Xiaohan Yang

4 **Addresses**

5 Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831-6422, USA

6

7 Corresponding author: Yang, Xiaohan (yangx@ornl.gov)

Genome editing with site-specific nucleases has become a powerful tool for functional characterization of plant genes and genetic improvement of agricultural crops. Among the various site-specific nuclease-based technologies available for genome editing, the clustered regularly interspaced short palindromic repeat (CRISPR)/CRISPR-associated protein 9 (Cas9) systems have shown the greatest potential for rapid and efficient editing of genomes in plant species. This article reviews the current status of application of CRISPR/Cas9 to plant genomics research, with a focus on loss-of-function and gain-of-function analysis of individual genes in the context of perennial plants and the potential application of CRISPR/Cas9 to perturbation of gene expression, and identification and analysis of gene modules as part of an accelerated domestication and synthetic biology effort.

Keywords: accelerated domestication, perennial plants, genome editing, genomics, genetic improvement

Introduction

Perennial plants uniquely provide ecosystems services that create keystone functions around which other organisms find habitat, physical structure and resource availability [1,2]. Moreover, in agronomic settings perennial plants typically require fewer chemical inputs, have reduced impacts on soil erosion and compaction [3], and contain increased wildlife diversity [4]. Perennial plants thus have been identified as promising sources of biomass for use in the fiber, fuels and feedstock arenas [5,6]. However, the recalcitrance of most perennial plants used in biomass production systems has not been domesticated [7]. In order to accelerate domestication, optimize feedstock quality, reduce environmental impacts and increase yield per unit time per unit area, modern genetics and genomics approaches need to be applied [8].

Two classical molecular genetic strategies have been established for plant functional genetics research: 1) loss-of-function analysis through physical and chemical mutagens, T-DNA random insertion [9] or RNA interference (RNAi) [10] and 2) gain-of-function analysis by the random activation of endogenous genes with transcriptional enhancers or the ectopic expression of individual transgenes [11]. Gene expression perturbation, especially spatiotemporal perturbation, is another important strategy for genomics research. Recently, a new genome engineering platform, using type II clustered regularly interspaced short palindromic repeat (CRISPR) and CRISPR-associated protein 9 (Cas9) from *Streptococcus pyogenes*, has become a more precise and powerful tool than traditional approaches for functional genomics research in plants [12]. CRISPR/Cas9 can target a specific genomic sequence, guided by an engineered 20-nt RNA sequence that binds to its DNA target by base-pairing rules [13] and holds great potential for analysis of loss-of-function, gain-of-function and gene expression. This article reviews recent advances in the utilization of CRISPR/Cas9 in plant genomics research with emphasis on applications to perennial plants, accelerated domestication and synthetic biology, at both single and multi-gene levels.

Characterization of individual genes with CRISPR/Cas9 systems

CRISPR/Cas9 has been applied to the functional characterization of individual genes through loss-of-function, gain-of-function and perturbation of gene expression.

Loss-of-function

Over the past decade, loss-of-function analysis was revolutionized by the development of tools that use the RNAi pathway for gene knock-down. However, the RNAi approach has several limitations such as incomplete loss-of-function and extensive off-target activities [14]. CRISPR/Cas9 has proven a powerful tool for loss-of-function analysis (Figure 1a), which is achieved by targeting a double-stranded break (DSB) to defined loci [15]. The DSBs introduced by CRISPR/Cas9 can be repaired via non-homologous

end-joining (NHEJ) or homology-directed repair (HDR) mechanisms. In most cases, NHEJ causes random insertion or deletion mutations (indels) of variable lengths resulting in knock-out mutants with frameshift mutations in protein-coding sequences or disruption of the *cis*-regulatory elements in promoters or enhancers. Compared with RNAi, CRISPR/Cas9 technology has several advantages, including 1) complete loss-of-function with relatively low off-target activities [16], 2) permanent (stable) and heritable knock-outs that will be important in domestication of perennial plants and for future breeding of such plants and 3) efficient and effective characterization of non-coding RNAs (ncRNAs) by disrupting their DNA coding sequences. Still, it should be noted that RNAi typically provides an allelic range with different degrees of knockdown, which can also be useful in functional genetics studies.

Successful applications of CRISPR/Cas9 to loss-of-function analysis across a range of plant species have been summarized in other review papers [17-19], and recently demonstrated in perennial plants, e.g., *Populus* [20^{••},21^{••}]. A big challenge in confirming loss-of-function of individual genes lies in the partial or total genetic redundancy of genes that belong to a well-conserved gene family [22]. In particular, a large proportion of retained paralogs resulted from recent genome-wide duplication in some perennial species, such as *Populus* [23] and *Eucalyptus* [24], imposes two types of challenges: difficulty in designing gene-specific gRNAs and genetic redundancy between paralogous genes with high similarity in DNA sequence. With simultaneous expression of multiple guide RNAs (sgRNAs), the CRISPR/Cas9 system allows for ‘multiplex genome editing’ as demonstrated in rice [25[•]], providing a powerful tool for addressing the genetic redundancy issue by quickly creating multiplex gene knockouts (e.g., double, triple, quadruple mutants). Figure 2 is provided as example of the applications of CRISPR/Cas9 to loss-of-function analysis in perennial plants.

Gain-of-function

In addition to the generation of knock-out mutants for loss-of-function analysis, the CRISPR/Cas9 systems can also be used for more sophisticated genome modifications; i.e., knock-in via either HDR or NHEJ to introduce specific point mutations or insert desired sequences through exogenously supplied DNA ‘donor templates’ (Figure 1b). HDR uses homologous DNA as a template for DSB repair, which is likely to have high fidelity and precision in plants [26-28,29^{••},30]. The problem with HDR-mediated knock-in in plants is the delivery of the donor template as single-stranded DNA (ssDNA) or double-stranded DNA (dsDNA). HDR-mediated knock-in has been achieved in annual plants such as *Oryza sativa* [26] and *Nicotiana benthamiana* [27] using protoplast transfection and in *Arabidopsis thaliana* using *Agrobacterium* transformation [28]. The application of CRISPR-based HDR to generation of knock-in mutants needs to be explored in perennials, several of which high efficient stable transformation protocols have been established, e.g., *Gossypium* [31], *Populus* [32] and *Ipomoea batatas* [33]. In

addition, the strategy of increasing the dsDNA donor copy number via geminivirus-based replicons (GVRs) has been successfully used to increase the frequency of targeted integration by HDR with zinc finger nucleases (ZFNs) in annual plants *Nicotiana tabacum* [29^{••}]. It is possible that the donor copy number is a limiting factor in the application of CRISPR-mediated knock-in in plants. Very recently, the same group reported that the GVRs harboring nuclease Cas9 and donor template can also enhance increase the frequency (i.e., tenfold increase vs. conventional *Agrobacterium tumefaciens* T-DNA) of targeted integration by HDR in *Solanum lycopersicum* [30]. The application of this strategy to perennial plants has not been reported yet, though this strategy offers the potential to increase the efficiency of CRISPR/Cas9-based knock-in in perennial plants. NHEJ, a highly active DNA repair mechanism within the somatic plant cells, can also be used for generating knock-in mutations. NHEJ mediated by transcription activator-like effector nucleases (TALENs) has been used to generate knock-in mutants in wheat [34[•]]. The efficient NHEJ pathway has been utilized to insert a 15-kb gene expression cassette at a defined locus with both ZFNs and TALENs in human cell lines [35]. Although the mechanistic of inducing DSBs to defined loci with CRISPR/Cas9 is different from that with ZFNs and TALENs, the DSBs are repaired in a similar manner. Thus, CRISPR/Cas9 may also be used for NHEJ-based knock-in. The potential of CRISPR/Cas9-mediated NHEJ strategy could be explored for introducing long DNA fragments into the plant genomes in the future.

Editing plant genomes without introducing foreign DNA into cells may alleviate regulatory concerns related to genetically modified plants. Unlike other plants, perennial plants have a long juvenile phase; therefore it is time-consuming to remove the CRISPR construct via backcrossing. Recently, two site-specific recombination systems were developed for excising T-DNA from transgene locus in rice [36] which may alleviate the need for backcrossing. More recently, a DNA-free genome editing method based on preassembled CRISPR-Cas9 ribonucleoproteins was established in *Arabidopsis*, tobacco, lettuce and rice by protoplast transfection [37]. This DNA-free genome editing method could be applied to perennial species, in which the protoplast transfection and regeneration system has been established, such as citrus [38] and apple [39].

Perturbation of gene expression

Target phenotypes for the domestication of perennial plants include modified height growth, increases in stem thickness, reduced branching, reduced apical dominance, optimized senescence/dormancy, reduced response to competition, limits to flowering and enhanced stress resistance [40]. One big challenge in the domestication of perennial plants is that the complete loss or strong constitutive expression of some gene (e.g., hormone metabolism and signaling-associated genes) causes numerous pleiotropic effects, even lethality. Perturbation of gene expression (i.e., modulating the level and pattern of gene expression at

defined developmental stages and tissue) could be helpful for addressing this challenge. It has been demonstrated that a catalytically inactive or “dead” Cas9 (dCas9), a variant bearing both the RuvC D10A and HNH H840A domain mutations that is unable to cleave DNA, could be recruited by sgRNAs to specific target DNA sites [41]. Thus, dCas9 can act as a general platform for blocking transcription (CRISPRi) [42] or recruitment of heterologous effector domains (e.g., transcriptional activation or repression domain, or demethylase domain) at defined genomic loci to regulate the expression of endogenous genes. For example, Piatek et al. [43] effectively regulated the expression of several genes in *Nicotiana benthamiana* leaves using dCas9 with its C-terminus fused either the activation domains of EDLL and TAL effectors or the SRDX repression domain. Furthermore, the CRISPR/Cas9 system has a great potential for controlling spatiotemporal patterns of gene expression. Recently, several targeted genome optogenetic tools based on CRISPR/Cas9 systems and the light-sensitive protein (e.g., cryptochrome 2, Vivid) were engineered in human cells [44^{••},45,46] (Figure 1c). These Cas9-based transcription systems allow rapid and reversible targeted gene activation/repression by light. Presumably such systems will function in a similar manner in light-penetrating tissue such as leaves and the outermost layer of the stem of the plants, but it may not work in inner tissue types such as xylem/woody tissues due to the limitation in light penetration. In addition, chemical-regulated gene expression systems based on CRISPR/Cas9 systems have been demonstrated in human cells [47] and mouse [48]. Using the CRISPR/Cas9 systems driven by tissue-specific promoters, gene expression has been controlled in a spatial manner as demonstrated in *Zebrafish* [49]. CRISPR/Cas9 systems can also be used to deliver specific cargos to targeted genomic loci. Shechner et al. [50] demonstrated that functional RNA cargos (e.g., natural long ncRNAs, protein-responsive riboswitches or temperature-responsive thermosensors) <4.8 kb could be inserted into sgRNA at multiple points and then delivered to defined genomic loci. These new Cas9-based systems could be applied to control spatiotemporal patterns of gene expression in perennial plants. Finally, replacing the endogenous promoter with a spatiotemporal promoter via knock-in could also be used in accelerated domestication research. For example, the effort of engineering crassulacean acid metabolism (CAM) into C₃ crops (e.g., *Populus*) will require a temporal reprogramming (e.g., diel-cycle shift) of the expression of genes shared between the C₃ and CAM pathways [5]. Replacing the promoter of C₃ endogenous gene with a diel-cycle shift promoter through CRISPR/Cas9 mediated knock-in technology could provide one new approach to this project, as illustrated in Figure 1c.

Visualization of spatiotemporal organization of the plant genome

Dynamic visualization of genomic loci is important for studying the spatiotemporal organization of the genome. DNA fluorescence in situ hybridization (FISH) has been widely used to visualize genomic loci for genomics research. Recently, Deng et al. [51[•]] demonstrated a CRISPR/Cas9-mediated FISH

(CASFISH) approach, which is less disruptive, more rapid, robust and cost-effective than DNA FISH, and can label repetitive and non-repetitive DNA elements in mouse. CASFISH can be advantageous in detecting “difficult” DNA FISH sequences as exemplified by G-rich telomeres, subtle DNA variations (i.e., SNPs) and large DNA fragments in plant genome research. In addition to using fluorescently labeled dCas9 assembled with various single-guide RNA (sgRNA) [51], an EGFP-dCas9 fusion has also been used to visualize repetitive genome sequences, such as telomeres, with a single sgRNA or non-repetitive sequences with 26 to 36 gRNAs tiled across a 5-kb genomic region in human cells [52]. These methods provide new approaches for visualizing the spatial organization of the plant genome, as illustrated in Figure 1d. Protoplasts or root hairs of plants may be used for such high-resolution detection.

Characterization of gene modules with CRISPR/Cas9 systems

Although functional characterization of individual genes has uncovered many of the underlying biological principles and the genomic basis of complex traits (e.g., signaling, regulatory and metabolic pathways), whole-genome domestication are too complex to be ascribed to a set of individual genes [8,53]. Analysis of gene modules, a group of genes or their products, which are related by one or more genetic or cellular interactions (e.g., co-regulation; co-expression or membership in a protein complex, a metabolic or signaling pathway, or a cellular aggregate) [54], could facilitate the dissection of the molecular basis of complex traits [55]. CRISPR/Cas9 systems have a great potential for identification and analysis of gene modules in perennial plants.

Genetic screens are powerful tools for assaying phenotypes and identifying causal genes of modules. Recently, genome-wide CRISPR screening has been performed using pooled library approaches coupled to positive or negative selection, or to alternatively arrayed libraries to dissect regulatory networks in humans [15,56], mouse [57,58**] and *Drosophila* [59]. This powerful tool has great potential for discovering plant gene modules, as illustrated in Figure 1e. For example, genome-wide CRISPR screening of plant and pathogen genes could be performed using *Agrobacterium*-mediated virus gene expression system, which was used for high-throughput functional screening in *Solanum lycopersicum* [60] and *Nicotiana tabacum* [61], harboring genome-wide sgRNAs targeting tens of thousands of annotated protein-coding and ncRNA genes. Also, the CRISPR screening based on protoplast transfection may be suitable for deciphering components of individual signaling or metabolic pathways and exploring mechanism of abiotic stresses. In addition to multiplex gene knockouts of a gene family, CRISPR/Cas9-mediated multiplex genome editing can also be used to create mutations for multiple genes that function as a module, as demonstrated in mouse [62]. Several toolboxes for multiplexed plant genome editing and transcriptional regulation have been developed in plants [63,64].

Identification of the components of gene modules is generally followed by reconstruction of the gene modules or networks. If the number of genes in a module exceeds the capacity of one-step reconstruction using direct DNA synthesis or *in vitro* DNA parts assembly technologies such as Gibson assembly and TNT cloning [65], an *in planta* gene stacking approach could be used to insert multiple transgenes into a single genomic locus [5]. Sequential transgene stacking mediated by ZFNs and meganuclease has been successfully established in *Zea mays* [66,67] and *Gossypium* [68], respectively. A novel CRISPR/Cas9-mediated gene stacking strategy *in planta* is illustrated in Figure 1f. We are in the process of testing proof-of-principle for using this gene stacking strategy to efficiently introduce gene module into plant genomes.

CRISPR/Cas9 can be used for dynamic analysis of gene modules in two ways: 1) conducting complex transcriptional programs of endogenous genes and 2) constructing synthetic gene circuits. Recently, Zalatan et al. [69**] successfully constructed a multigene transcriptional system composed of modular scaffold RNAs and dCas9 to facilitate control the expression of a five-gene pathway (i.e., bacterial violacein biosynthetic pathway) in yeast. This dCas9-mediated transcriptional reprogramming approach could be applied for characterizing multiple genes in a biosynthetic pathway in plants in general, as illustrated in Figure 1g. A synthetic gene circuit containing the dCas9-based regulators and IPTG sensors was described in *Bacteroides thetaiotaomicron* to dynamically manipulate bacterial processes [70**]. In addition, transcription activator-like effector (TALE) transcriptional repressors were successfully used to construct gene circuits in mammalian cells [71]. Thus, the dCas9-based regulators should have the potential to be integrated for higher-order computation (e.g., digital logic, memory, and analog circuits) for module dynamic analysis in plants, as illustrated in Figure 1h.

Conclusion

Leveraging whole-genome information in order to accelerate plant domestication will require loss-of-function and gain-of-function changes in genes and gene modules, along with modifications in spatiotemporal gene expression. The simplicity, robustness and versatility of the CRISPR/Cas9 systems make such systems attractive for plant genomics research in general and specifically to the process of accelerated domestication in perennial plants. CRISPR/Cas9-mediated gene knock-out and knock-in approaches hold promises for loss-of-function and gain-of-function analysis, respectively, of individual genes. Perturbation of gene expression based on CRISPR/Cas9 opens a new door to control spatiotemporal gene expression. Furthermore, CRISPR/Cas9 has a great potential for identifying and analyzing gene modules through genome-wide screening and multiplex genome editing. Gene stacking and genetic circuit construction strategy based on CRISPR/Cas9 could also be powerful tools for plant synthetic biology. Overall, the CRISPR/Cas9 systems could open a new door to plant functional genomics research leading to strategies for the accelerated domestication of many perennial plants.

Acknowledgments

This research is supported by the Department of Energy (DOE), Office of Science, Genomic Science Program under Award Number DE-SC0008834. The authors would like to thank Lee E. Gunter for critical review and comments on the manuscript. Oak Ridge National Laboratory is managed by UT-Battelle, LLC for the US DOE under Contract Number DE-AC05-00OR22725.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- of outstanding interest

1. Whitham TG, Young WP, Martinsen GD, Gehring CA, Schweitzer JA, Shuster SM, Wimp GM, Fischer DG, Bailey JK, Lindroth RL *et al.*: **Community and ecosystem genetics: a consequence of the extended phenotype.** *Ecology* 2003, **84**:559-573.
2. Weston DJ, Rogers A, Tschaplinski TJ, Gunter LE, Jawdy SA, Engle NL, Heady LE, Tuskan GA, Wullschlegel SD: **Scaling nitrogen and carbon interactions: what are the consequences of biological buffering?** *Ecol Evol* 2015, **5**:2839-2850.

- 233 3. Jansson C, Wullschleger SD, Kalluri UC, Tuskan GA: **Phytosequestration: carbon**
 234 **biosequestration by plants and the prospects of genetic engineering.** *Bioscience* 2010, **60**:685-
 235 696.
- 236 4. Lamit LJ, Lau MK, Næsborg RR, Wojtowicz T, Whitham TG, Gehring CA: **Genotype variation in**
 237 **bark texture drives lichen community assembly across multiple environments.** *Ecology* 2015,
 238 **96**:960-971.
- 239 5. Yang X, Cushman JC, Borland AM, Edwards EJ, Wullschleger SD, Tuskan GA, Owen NA, Griffiths
 240 H, Smith JAC, De Paoli HC *et al.*: **A roadmap for research on crassulacean acid metabolism**
 241 **(CAM) to enhance sustainable food and bioenergy production in a hotter, drier world.** *New*
 242 *Phytol* 2015, **207**: 491-504.
- 243 6. Ragauskas AJ, Beckham GT, Biddy MJ, Chandra R, Chen F, Davis MF, Davison BH, Dixon RA,
 244 Gilna P, Keller M *et al.*: **Lignin valorization: improving lignin processing in the biorefinery.**
 245 *Science* 2014, **344**:1246843.
- 246 7. Tuskan G: **Short-rotation woody crop supply systems in the United States: what do we know**
 247 **and what do we need to know?** *Biomass and Bioenergy* 1998, **14**:307-315.
- 248 8. Harfouche A, Meilan R, Kirst M, Morgante M, Boerjan W, Sabatti M, Mugnozza GS: **Accelerating**
 249 **the domestication of forest trees in a changing world.** *Trends Plant Sci* 2012, **17**:64-72.
- 250 9. Alonso JM, Stepanova AN, Leisse TJ, Kim CJ, Chen H, Shinn P, Stevenson DK, Zimmerman J,
 251 Barajas P, Cheuk R *et al.*: **Genome-wide insertional mutagenesis of *Arabidopsis thaliana*.** *Science*
 252 2003, **301**:653-657.
- 253 10. Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC: **Potent and specific genetic**
 254 **interference by double-stranded RNA in *Caenorhabditis elegans*.** *Nature* 1998, **391**:806-811.
- 255 11. Weigel D, Ahn JH, Blázquez MA, Borevitz JO, Christensen SK, Fankhauser C, Ferrándiz C,
 256 Kardailsky I, Malancharuvil EJ, Neff MM *et al.*: **Activation tagging in *Arabidopsis*.** *Plant Physiol*
 257 2000, **122**:1003-1014.
- 258 12. Belhaj K, Chaparro-Garcia A, Kamoun S, Patron NJ, Nekrasov V: **Editing plant genomes with**
 259 **CRISPR/Cas9.** *Curr Opin Biotechnol* 2015, **32**:76-84.
- 260 13. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E: **A programmable dual-RNA-**
 261 **guided DNA endonuclease in adaptive bacterial immunity.** *Science* 2012, **337**:816-821.

14. Xu P, Zhang Y, Kang L, Roossinck MJ, Mysore KS: **Computational estimation and experimental verification of off-target silencing during posttranscriptional gene silencing in plants.** *Plant Physiol* 2006, **142**:429-440.
 15. Shalem O, Sanjana NE, Zhang F: **High-throughput functional genomics using CRISPR-Cas9.** *Nat Rev Genet* 2015, **16**: 299-311.
 16. Barrangou R, Birmingham A, Wiemann S, Beijersbergen RL, Hornung V, van Brabant Smith A: **Advances in CRISPR-Cas9 genome engineering: lessons learned from RNA interference.** *Nucleic Acids Res* 2015:gkv226.
 17. Bortesi L, Fischer R: **The CRISPR/Cas9 system for plant genome editing and beyond.** *Biotechnol Adv* 2015, **33**:41-52.
 18. Belhaj K, Chaparro-Garcia A, Kamoun S, Nekrasov V: **Plant genome editing made easy: targeted mutagenesis in model and crop plants using the CRISPR/Cas system.** *Plant Methods* 2013, **9**:39.
 19. Schaeffer SM, Nakata PA: **CRISPR/Cas9-mediated genome editing and gene replacement in plants: Transitioning from lab to field.** *Plant Sci* 2015.
 20. Zhou X, Jacobs TB, Xue LJ, Harding SA, Tsai CJ: **Exploiting SNPs for biallelic CRISPR ●● mutations in the outcrossing woody perennial *Populus* reveals 4-coumarate: CoA ligase specificity and redundancy.** *New Phytol* 2015, **208**: 298-301.
- The authors demonstrated the application of CRISPR/Cas9 to generate knockout mutation of two genes involved in lignin and flavonoid biosynthesis in the woody perennial *Populus*. They achieved 100% knockout efficiency for the two targeted genes, with every transformant examined carrying biallelic modifications.
21. Fan D, Liu T, Li C, Jiao B, Li S, Hou Y, Luo K: **Efficient CRISPR/Cas9-mediated targeted ●● mutagenesis in *Populus* in the first generation.** *Sci Rep* 2015, **5**: 12217.
- With CRISPR/Cas9 system, homozygous knock-out mutation in predicted loci was generated in T₀ generation of transgenic *Populus*. This is very important for genetic modification of perennial plants with a long life span.
22. Jain M: **Functional genomics of abiotic stress tolerance in plants: a CRISPR approach.** *Front Plant Sci* 2015, **6**:375.
 23. Tuskan GA, Difazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S, Rombauts S, Salamov A *et al.*: **The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray).** *Science* 2006, **313**:1596-1604.

293 24. Myburg AA, Grattapaglia D, Tuskan GA, Hellsten U, Hayes RD, Grimwood J, Jenkins J, Lindquist E,
294 Tice H, Bauer D, *et al.*: **The genome of *Eucalyptus grandis***. *Nature* 2014, **510**:356-362.

295 25. Xie K, Minkenberg B, Yang Y: **Boosting CRISPR/Cas9 multiplex editing capability with the**
296 **• endogenous tRNA-processing system**. *Proc Natl Acad Sci U S A* 2015, **112**:3570-3575.

297 Using synthetic genes with tandemly arrayed tRNA–gRNA architecture, a high efficiency (up to 100%)
298 of multiplex genome editing and chromosomal-fragment deletion based on CRISPR/Cas9 was achieved
299 in rice.

300 26. Shan Q, Wang Y, Li J, Zhang Y, Chen K, Liang Z, Zhang K, Liu J, Xi JJ, Qiu J-L *et al.*: **Targeted**
301 **genome modification of crop plants using a CRISPR-Cas system**. *Nat Biotechnol* 2013, **31**:686-
302 688.

303 27. Li J-F, Norville JE, Aach J, McCormack M, Zhang D, Bush J, Church GM, Sheen J: **Multiplex and**
304 **homologous recombination-mediated genome editing in *Arabidopsis* and *Nicotiana***
305 ***benthamiana* using guide RNA and Cas9**. *Nat Biotechnol* 2013, **31**:688-691.

306 28. Schiml S, Fauser F, Puchta H: **The CRISPR/Cas system can be used as nuclease for in planta gene**
307 **targeting and as paired nickases for directed mutagenesis in *Arabidopsis* resulting in heritable**
308 **progeny**. *Plant J* 2014, **80**:1139-1150.

309 29. Baltes NJ, Gil-Humanes J, Cermak T, Atkins PA, Voytas DF: **DNA replicons for plant genome**
310 **•• engineering**. *Plant Cell* 2014, **26**:151-163.

311 Baltes et al. demonstrated that geminivirus-based replicons could enable highly efficient genome
312 engineering in tobacco. This strategy could be used to increase the efficiency of knock-in in plants.

313 30. Čermák T, Baltes NJ, Čegan R, Zhang Y, Voytas DF: **High-frequency, precise modification of the**
314 **tomato genome**. *Genome Biol* 2015, **16**:1-15.

315 31. Firoozabady E, DeBoer DL, Merlo DJ, Halk EL, Amerson LN, Rashka KE, Murray EE:
316 **Transformation of cotton (*Gossypium hirsutum* L.) by *Agrobacterium tumefaciens* and**
317 **regeneration of transgenic plants**. *Plant Mol Biol* 1987, **10**:105-116.

318 32. Cseke LJ, Cseke SB, Podila GK: **High efficiency poplar transformation**. *Plant Cell Rep* 2007,
319 **26**:1529-1538.

320 33. Yu B, Zhai H, Wang Y, Zang N, He S, Liu Q: **Efficient *Agrobacterium tumefaciens*-mediated**
321 **transformation using embryogenic suspension cultures in sweetpotato, *Ipomoea batatas* (L.)**
322 **Lam**. *Plant Cell Tiss Org* 2007, **90**:265-273.

323 34. Wang Y, Cheng X, Shan Q, Zhang Y, Liu J, Gao C, Qiu J-L: **Simultaneous editing of three**

- 324 • **homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew.**
 325 *Nat Biotechnol* 2014, **32**:947-951.
- 326 The authors demonstrated that TALEN and CRISPR/Cas9 tools could be efficiently applied to generate
 327 gene family (three homoeoalleles) knock-outs in a polyploid crop species. They also demonstrated the
 328 feasibility of engineering targeted DNA insertion in bread wheat through NHEJ of the DSBs caused by
 329 TALENs.
- 330 35. Maresca M, Lin VG, Guo N, Yang Y: **Obligate ligation-gated recombination (ObLiGaRe):**
 331 **custom-designed nuclease-mediated targeted integration through nonhomologous end joining.**
 332 *Genome Res* 2013, **23**:539-546.
- 333 36. Nandy S, Srivastava V: **Marker-free site-specific gene integration in rice based on the use of two**
 334 **recombination systems.** *Plant Biotechnol J* 2012, **10**:904-912.
- 335 37. Woo JW, Kim J, Kwon SI, Corvalán C, Cho SW, Kim H, Kim S-G, Kim S-T, Choe S, Kim J-S:
 336 **DNA-free genome editing in plants with preassembled CRISPR-Cas9 ribonucleoproteins.**
 337 *Nature Biotechnol* 2015.
- 338 38. Vardi A, Bleichman S, Aviv D: **Genetic transformation of *Citrus* protoplasts and regeneration of**
 339 **transgenic plants.** *Plant Sci* 1990, **69**:199-206.
- 340 39. Maddumage R, Fung RM, Weir I, Ding H, Simons JL, Allan AC: **Efficient transient transformation**
 341 **of suspension culture-derived apple protoplasts.** *Plant Cell Tiss Org* 2002, **70**:77-82.
- 342 40. Ragauskas AJ, Williams CK, Davison BH, Britovsek G, Cairney J, Eckert CA, Frederick WJ, Hallett
 343 JP, Leak DJ, Liotta CL *et al.*: **The path forward for biofuels and biomaterials.** *Science* 2006,
 344 **311**:484-489.
- 345 41. Sander JD, Joung JK: **CRISPR-Cas systems for editing, regulating and targeting genomes.** *Nat*
 346 *Biotechnol* 2014, **32**:347-355.
- 347 42. Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, Lim WA: **Repurposing**
 348 **CRISPR as an RNA-guided platform for sequence-specific control of gene expression.** *Cell*
 349 2013, **152**:1173-1183.
- 350 43 Piatek A, Ali Z, Baazim H, Li L, Abulfaraj A, Al-Shareef S, Aouida M, Mahfouz MM: **RNA-guided**
 351 **transcriptional regulation in planta via synthetic dCas9-based transcription factors.** *Plant*
 352 *Biotechnol J* 2015, **13**:578-589.
- 353 44. Nihongaki Y, Kawano F, Nakajima T, Sato M: **Photoactivatable CRISPR-Cas9 for optogenetic**
 354 **genome editing.** *Nat Biotechnol* 2015, **33**:755–760.

This article represents one of the examples of optogenetic tool based on CRISPR/Cas9 in human cells. The lightswitch could be used to control spatiotemporal patterns of gene expression in plants.

45. Polstein LR, Gersbach CA: **A light-inducible CRISPR-Cas9 system for control of endogenous gene activation.** *Nat Chem Biol* 2015, **11**:198-200.

46. Nihongaki Y, Yamamoto S, Kawano F, Suzuki H, Sato M: **CRISPR-Cas9-based photoactivatable transcription system.** *Chem Biol* 2015, **22**:169-174.

47. Zetsche B, Volz SE, Zhang F: **A split-Cas9 architecture for inducible genome editing and transcription modulation.** *Nat Biotechnol* 2015, **33**:139-142.

48. Dow LE, Fisher J, O'Rourke KP, Muley A, Kastenhuber ER, Livshits G, Tschaharganeh DF, Socci ND, Lowe SW: **Inducible *in vivo* genome editing with CRISPR-Cas9.** *Nat Biotechnol* 2015, **33**:390-394.

49. Ablain J, Durand EM, Yang S, Zhou Y, Zon LI: **A CRISPR/Cas9 vector system for tissue-specific gene disruption in zebrafish.** *Dev Cell* 2015, **32**:756-764.

50. Shechner DM, Hacisuleyman E, Younger ST, Rinn JL: **Multiplexable, locus-specific targeting of long RNAs with CRISPR-Display.** *Nature Methods* 2015, **12**:664-670.

51. Deng W, Shi X, Tjian R, Lionnet T, Singer RH: **CASFISH: CRISPR/Cas9-mediated *in situ* labeling of genomic loci in fixed cells.** *Proc Natl Acad Sci U S A* 2015, **112**:11870-11875.

This article presented a CASFISH technology based on fluorescently labeled dCas9 assembled with various sgRNA. This technology could be used to analyze the spatial relationships among the genetic elements in a plant genome and its association with gene expression.

52. Chen B, Gilbert LA, Cimini BA, Schnitzbauer J, Zhang W, Li G-W, Park J, Blackburn EH, Weissman JS, Qi LS *et al.*: **Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system.** *Cell* 2013, **155**:1479-1491.

53. Robson P, Jensen E, Hawkins S, White SR, Kenobi K, Clifton-Brown J, Donnison I, Farrar K: **Accelerating the domestication of a bioenergy crop: identifying and modelling morphological targets for sustainable yield increase in *Miscanthus*.** *J Exp Bot* 2013, **64**:4143-4155.

54. Tornow S, Mewes H: **Functional modules by relating protein interaction networks and gene expression.** *Nucleic Acids Res* 2003, **31**:6283-6289.

55. Wang X, Dalkic E, Wu M, Chan C: **Gene module level analysis: identification to networks and dynamics.** *Curr Opin Biotechnol* 2008, **19**:482-491.

56. Wang T, Wei JJ, Sabatini DM, Lander ES: **Genetic screens in human cells using the CRISPR-Cas9 system.** *Science* 2014, **343**:80-84.
 57. Koike-Yusa H, Li Y, Tan E-P, Velasco-Herrera MDC, Yusa K: **Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library.** *Nat Biotechnol* 2014, **32**:267-273.
 58. Parnas O, Jovanovic M, Eisenhaure TM, Herbst RH, Dixit A, Ye CJ, Przybylski D, Platt RJ, Tirosh I, Sanjana NE *et al.*: **A genome-wide CRISPR screen in primary immune cells to dissect regulatory networks.** *Cell* 2015, **162**:675-686.
- This study represents one of the first examples of dissecting functional modules with genome-wide CRISPR screening. This interesting strategy holds great potential for identification of plant gene modules.
59. Bassett AR, Kong L, Liu JL: **A genome-wide CRISPR library for high-throughput genetic screening in *Drosophila* cells.** *J Genet Genomics* 2015, **42**:301-309.
 60. Takken FL, Luderer R, Gabriëls SH, Westerink N, Lu R, De Wit PJ, Joosten MH: **A functional cloning strategy, based on a binary PVX-expression vector, to isolate HR-inducing cDNAs of plant pathogens.** *Plant J* 2000, **24**:275-283.
 61. Karrer EE, Beachy RN, Holt CA: **Cloning of tobacco genes that elicit the hypersensitive response.** *Plant Mol Biol* 1998, **36**:681-690.
 62. Heckl D, Kowalczyk MS, Yudovich D, Belizaire R, Puram RV, McConkey ME, Thielke A, Aster JC, Regev A, Ebert BL: **Generation of mouse models of myeloid malignancy with combinatorial genetic lesions using CRISPR-Cas9 genome editing.** *Nat Biotechnol* 2014, **32**:941-946.
 63. Lowder LG, Paul JW, Baltes NJ, Voytas DF, Zhang Y, Zhang D, Tang X, Zheng X, Hsieh T-F, Qi Y: **A CRISPR/Cas9 toolbox for multiplexed plant genome editing and transcriptional regulation.** *Plant physiol* 2015, **169**:971-985.
 64. Xing HL, Dong L, Wang ZP, Zhang HY, Han CY, Liu B, Wang XC, Chen QJ: **A CRISPR/Cas9 toolkit for multiplex genome editing in plants.** *BMC Plant Biol* 2014, **14**:327.
 65. De Paoli HC, Tuskan GA, Yang X: **An innovative platform for quick and flexible joining of assorted DNA fragments.** *Sci Rep-UK* 2016, **6**:19278.
 66. Ainley WM, Sastry-Dent L, Welter ME, Murray MG, Zeitler B, Amora R, Corbin DR, Miles RR, Arnold NL, Strange TL *et al.*: **Trait stacking via targeted genome editing.** *Plant Biotechnol J* 2013, **11**:1126-1134.

67. Kumar S, AlAbed D, Worden A, Novak S, Wu H, Ausmus C, Beck M, Robinson H, Minnick T, Hemingway D *et al.*: **A modular gene targeting system for sequential transgene stacking in plants.** *J Biotechnol* 2015, **207**:12-20.
 68. D'Halluin K, Vanderstraeten C, Hulle J, Rosolowska J, Den Brande I, Pennewaert A, D'Hont K, Bossut M, Jantz D, Ruiter R *et al.*: **Targeted molecular trait stacking in cotton through targeted double-strand break induction.** *Plant Biotechnol J* 2013, **11**:933-941.
 69. Zalatan JG, Lee ME, Almeida R, Gilbert LA, Whitehead EH, La Russa M, Tsai JC, Weissman JS, ●● Dueber JE, Qi LS *et al.*: **Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds.** *Cell* 2015, **160**:339-350.
- By extending gRNAs to include effector protein recruitment sites, the authors constructed modular scaffold RNAs (scRNAs) encoding both target locus and regulatory action. Sets of scRNAs, dCas9 and orthogonal dCas9 proteins can be used to generate synthetic multigene transcriptional programs, in which distinct target genes can be simultaneously activated or repressed. This article represents one of the second-generation CRISPR systems.
70. Mimee M, Tucker AC, Voigt CA, Lu TK: **Programming a human commensal bacterium, ●● *Bacteroides thetaiotaomicron*, to sense and respond to stimuli in the murine gut microbiota.** *Cell Systems* 2015, **1**:62-71.
- With the dCas9-based regulators and IPTG sensors, a synthetic gene circuit was constructed in *Bacteroides thetaiotaomicron* to dynamically manipulate bacterial processes. This research shows that CRISPR/Cas9 could be a facile toolbox for dynamic analysis of gene modules in plants.
71. Li Y, Jiang Y, Chen H, Liao W, Li Z, Weiss R, Xie Z: **Modular construction of mammalian gene circuits using TALE transcriptional repressors.** *Nat Chem Biol* 2015, **11**:207-213.
 72. Hu W-J, Kawaoka A, Tsai C-J, Lung J, Osakabe K, Ebinuma H, Chiang VL: **Compartmentalized expression of two structurally and functionally distinct 4-coumarate: CoA ligase genes in aspen (*Populus tremuloides*).** *Proc Natl Acad Sci U S A* 1998, **95**:5407-5412.
 73. Harding SA, Leshkevich J, Chiang VL, Tsai C-J: **Differential substrate inhibition couples kinetically distinct 4-coumarate: coenzyme A ligases with spatially distinct metabolic roles in quaking aspen.** *Plant Physiol* 2002, **128**:428-438.

Figure legends

Figure 1. Overview of potential CRISPR/Cas9-based applications in plant genomics research in either a synthetic biology and/or accelerated domesticaition effort. **(a, b)** Analysis of loss-of-function and gain-of-function carried out through CRISPR/Cas9-mediated knock-out and knock-in mutations, respectively. Two orange T's show that the insertion is flanked by transcriptional terminators. **(c)** Targeted genome optogenetic tools based on the catalytically inactive or “dead” Cas9 (dCas9) and the light-sensitive protein, and the strategy replacing the endogenous promoter with a spatiotemporal promoter by knock-in used to control spatiotemporal patterns of gene expression. **(d)** Fluorescently labeled dCas9 used to visualize plant genome organization and structure. **(e)** Genome-wide screening of sgRNA library using CRISPR/Cas9 for identification of plant gene modules, with the following workflow: construction of sgRNA library, delivery of the sgRNA library to plants through transformation, screening of phenotype and analysis of gene modules. **(f)** A CRISPR/Cas9-based gene stacking strategy used for *in planta* reconstruction of gene modules. First, a founder line is created by random integration of a T-DNA insert containing selection marker 1 (SM1), sgRNA1 recognition site (marked with green) and *Cas9*. Then, the founder line is transformed with donor 1 containing SM2, sgRNA1 (marked with green), sgRNA2 recognition site (marked with yellow), gene of interest 1 (GOI1) and part of homologous sequence of *Cas9*. HR-mediated knock-in will lead to replacement of SM1 with SM2 and stacking of GOI1. As such, multiple rounds of gene stacking can be achieved by alternating sgRNA recognition site and selection marker. **(g)** dCas9-mediated transcriptional reprograming approach used for characterizing gene modules in plants. This expression programs could be achieved in the same cell by harnessing orthogonal dCas9 proteins (the orange and light blue dCas9 in the figure) that recognize their scaffold RNAs (scRNAs) through distinct sequences. Each orthogonal dCas9 protein is placed under the control of different inducible promoters and could control a distinct set of scRNAs, allowing independent control over distinct gene expression programs. **(h)** Synthetic gene circuit containing the dCas9-based regulators applied for dynamic analysis of plant gene modules. dCas9 is driven by the inducible promoter, which could be repressed by the substance (black dot). Addition of inducer prompts the expression of dCas9, which complexes with constitutively expressed sgRNA targeting the coding sequence.

Figure 2. Examples of application of CRISPR/Cas9 to loss-of-function research in perennial plants. **(a)** CRISPR/Cas9-based knock-out of phytoene desaturase gene (*PDS*) in *Populus tomentosa* Carr. (clone 741) [20]. Obvious albino phenotype was observed in the knock-out mutants. **(b)** CRISPR/Cas9-mediated knock-out of 4-coumarate:CoA ligase 1 (*4CL1*) and *4CL2* involved in lignin and flavonoid biosynthesis in *Populus tremula* × *alba* (clone 717-1B4) [20]. Multiple 4CL isoforms with differential *in vitro* substrate specificities have been reported in several species. Two 4CL genes, *Pt4CL1* and *Pt4CL2*,

478 associated with lignin and flavonoid biosynthesis, respectively [72,73], were targeted for CRISPR/Cas9
479 editing. Here, four main results were found: 1) reddishbrown wood color was observed in each of the 36
480 mutant 4CL1 lines, 2) stem wood lignin content and syringyl-to-guaiacyl (S : G) monolignol ratio were
481 reduced in all 4CL1 mutants, but not changed in 4CL2 lines relative to the control, 3) root condensed
482 tannin levels were reduced in all 4CL2 mutants, and in some 4CL1 lines and 4) relative abundance of
483 chlorogenic acids in leaves was significantly lower in 4CL2 mutants, but unchanged in 4CL1 lines. C3H,
484 4-coumarate 3-hydroxylase; COMT, caffeate *O*-methyltransferase; CCoAOMT, caffeoyl-CoA *O*-
485 methyltransferase; CCR, cinnamoyl CoA reductase; CAld5H, coniferyl aldehyde 5-hydroxylase;
486 AldOMT, 5-hydroxyconiferyl aldehyde *O*-methyltransferase; F5H, ferulate 5-hydroxylase.

Figure 1

Figure 2