The CRISPR System—Keeping Zebrafish Gene Targeting Fresh

Patrick R. Blackburn, Jarryd M. Campbell, Karl J. Clark, and Stephen C. Ekker

Abstract

We are entering a new era in our ability to modify and edit the genomes of model organisms. Zinc finger nucleases (ZFNs) opened the door to the first custom nuclease-targeted genome engineering in the late 1990s. However, ZFNs remained out of reach for most research labs because of the difficulty of production, high costs, and modest efficacy in many applications. Transcription activator-like effector nucleases (TALENs) were built upon a DNA binding system discovered in a group of plant bacterial pathogens and broadened custom nuclease technology, showing significant improvements in both targeting flexibility and efficiency. Perhaps most importantly, TALENs are open source and easy to produce, providing zebrafish laboratories around the world with affordable tools that can be made in-house rapidly, at low cost, and with reliably high activity. Now a new system for targeted genome engineering derived from the CRISPR/Cas system in eubacteria and archaea promises to simplify this process further. Together, these tools will help overcome many of the bottlenecks that have constrained gene targeting in zebrafish, paving the way for advanced genome engineering applications in this model teleost.

Introduction

THE COMPLETED ZEBRAFISH (*Danio rerio*) genome provides researchers with a roadmap to study gene function in states of health and disease. There is a high level of diversity between laboratory strains and wild-type zebrafish compared to other vertebrate model organisms.¹ With the rapid emergence of genome editing tools and improvements in their targeting efficiencies within zebrafish, scientists are uniquely positioned to study how natural and engineered genetic heterogeneity impacts a range of phenotypes. In addition, novel uses of these tools will enable many engineering applications, including large deletions and complex rearrangements, targeted sequence insertions by exploiting homology directed repair (HDR) or homologous recombination (HR), as well as multiplex applications for concurrent modifications at multiple loci.

ZFNs and TALENs Make the Cut

Zinc finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs) both have a modular architecture to recognize a DNA sequence with varying specificities. ZFNs are composed of multiple three-base recognition motifs consisting of about 30 amino acids that are assembled together. ZFN binding is not completely intuitive, however, and relies on a number of different design and selection-based methods to generate good cutting reagents.²

TALENs are made up of repeat modules of approximately 34 amino acids that are organized into TAL effector arrays.³ A two-amino acid repeat variable di-residue (RVD) within each repeat module is responsible for one-to-one base pairing.³ The four most commonly used RVDs—NI, HD, NN, and NG—bind preferentially to one of the four DNA bases (A, C, G, and T, respectively).³ Where ZFN design has been restrictive due to nonintuitive binding rules, TALENs appear to have greater flexibility and more predictable design due to their simpler base recognition code.² Both ZFNs and TALENs have traditionally utilized a FokI nuclease homodimer fused to their DNA-binding domains to generate double-strand breaks. The FokI nuclease domain has been modified to reduce off-target effects (heterodimeric forms) and limit toxicity (modified nickases).^{4,5} The GoldyTALEN scaffold produces high efficacy TALENs, often resulting in high rates of biallelic conversion in somatic tissues within the zebrafish.^{6,7} In addition, the high efficiency cutting has been combined with single-stranded DNA oligonucleotides to modify sequences precisely or add a loxP site through an HDR-like mechanism using this system.⁶ A recent article by Zu et al. showed that dsDNA donors with approximately 900 base homology arms could induce ‘precise,’ targeted changes when co-injected with TALEN mRNAs.⁸ This strategy was used to integrate a long sequence donor (EGFP) at several loci within the zebrafish genome and could expand the current TALEN repertoire to include large sequence introduction or replacement.⁸

FIG. 1. The Cas9/guide RNA complex bound to a target sequence within the genome. The guide RNA is a chimera composed of both tracrRNA and crRNA connected by a four-base loop.¹¹ The target sequence within the protospacer is shown base-paired with the guide RNA. The Cas9 endonuclease active sites cut between 3 and 7 bases upstream of the PAM (protospacer adjacent motif) on both strands.¹¹ The first 8–12 bases directly upstream of the PAM seem to be essential for Cas9 cleavage.¹³ The importance of the remaining 8–12 bases is unclear.¹³

The CRISPR System—Keeping Things Fresh

Clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR-associated systems (Cas) are derived from eubacteria and archaea and constitute a form of adaptive immunity.⁹ The CRISPR system relies on the incorporation of exogenous DNA sequence from plasmids and phage into spacer regions bracketed by short palindromic direct repeats.⁹ These direct repeats and spacers are organized into arrays that are transcribed together, then cleaved and processed into CRISPR RNAs (crRNAs).⁹ The crRNAs contain sequences that recognize complementary target DNA sequences.⁹ The crRNAs complexed with Cas DNA nucleases bind to target DNA sequences and cleave both strands of DNA, thereby generating double-stranded breaks.⁹ There are three types of CRISPR-Cas systems.⁹ The type II CRISPR system is most likely to be useful for genome engineering applications and involves a transactivating CRISPR RNA (tracrRNA) and crRNA that assemble into a complex with Cas9.¹⁰ The crRNA contains a spacer sequence that determines the target DNA sequence and binds on either strand next to a protospacer adjacent motif (PAM) that is required for recognition of the

target sequence.¹¹ Cas9 then cuts both strands, resulting in a double-stranded break within the target sequence.¹¹

Recently, Jinek et al. showed that the type II CRISPR system could be programmed to recognize a specific genomic sequence using a single chimeric RNA molecule consisting of both crRNA- and tracrRNA-derived sequences connected by a four-base linker loop (Fig. 1).¹¹ A similar chimeric guide RNA was used to mutate zebrafish genes with a somatic targeting efficiency approaching levels seen in initial TALEN systems.¹² These promising early results strongly suggest that this CRISPR system may yield high efficacy germline targeting in zebrafish.

Perhaps the greatest advantage of the CRISPR system in zebrafish is the relatively simple requirement for a single unique guide RNA plus the Cas9 RNA. Guide RNA targeting any gene of interest can be synthesized using two 26 base overlapping oligonucleotides that are subcloned into a guide RNA expression vector.¹² However, there are some noteworthy limitations to current CRISPR gene editing systems. Only 8 to 12 bases of 20 in the spacer sequence and the NGG PAM sequence appear to be essential for efficient Cas9 cleavage.¹³ This has important implications for efficacy, as

TABLE 1. A COMPARISON BETWEEN CURRENT THREE CUSTOM RESTRICTION ENZYME SYSTEMS

System	Class	Protein coding size	Specificity (targeting sequence)	Zebrafish genome targeting coverage	Off-target effects	Technology adoption time and costs	Ongoing time and cost	Somatic DNA cutting efficiency	Germline efficacy
ZFN	Protein-DNA	1.1 to 1.2 kb per arm	18+bp (pair)	Target site every 140–400 bp ^{2,18,19}	Low	High	Moderate	Low (~2%) ^{19,20}	Low
TALEN	Protein-DNA	2.7 to 3 kb per arm	30+bp (pair)	Complete coverage (target site every 1–3 bp) ²	Very low	Moderate	Low	Moderate (~20%) to high (>50%) ^{2,4,7,20,21}	Moderate to high
CRISPR/Cas9	RNA-DNA	4.1 kb Cas9, 63–103 bp gRNA	12+bp?	Target site every 8–128 bp (NGG PAM with/without T7 promoter constraints) ^{12,13}	TBD	Low	Low	Moderate (~30%) ¹²	TBD

ZFN pioneered custom genome editing and engineering approaches; the smaller sizes of these proteins are still being used in applications where an upper limit on cargo capacity is critical. TALEN are a more accessible technology for most laboratories, compared to ZFN. The new CRISPR/Cas9 systems can target fewer sites within the zebrafish genome than TALEN. However, CRISPR/Cas9 enzymes promise further ease of design and implementation, potentially opening the door to use by even more zebrafish laboratories. TBD, to be determined.

well as the potential for off-target effects. In addition, the PAM sequence requirement constrains potential targeting sequences in the currently available CRISPR systems.¹² This exciting new RNA guided endonuclease system is markedly different from existing technologies and promises to open new doors for zebrafish researchers.

The Overflowing Toolbox

Several groups have published new or modified TALEN assembly systems, reducing time and cost for the production of a pair of TALENs. For example, the commonly used Golden Gate assembly system by Cermak et al. uses less than 50 total plasmids, an investment that is practical for even a small laboratory.¹⁴ The fast ligation-based automatable solid-phase high-throughput (FLASH) assembly system has been joined by several new protocols that modify the TALEN assembly process such that multimers are ligated together in a single step, reducing the generation time to two days.^{15–17} Upfront DNA preparation to generate such libraries (sometimes >800 individual plasmids) is significant, and is likely only practical for researchers who plan to generate large numbers of TALENs.

In contrast, the CRISPR/Cas9 system requires as few as two plasmids, one containing the Cas9 sequence and the other containing the single guide template sequence into which a 20 base spacer sequence is inserted for target recognition. The spacer sequence can be synthesized commercially as complementary oligonucleotides and rapidly integrated into the single guide template plasmid followed by RNA synthesis, with a total time investment of a few days. Large numbers of these single guide template plasmids can be assembled in parallel without substantial upfront investment in time or cost (Table 1).

A Sea Change for a Tiny Freshwater Fish

With such rapid, recent advances in genome engineering, the future is looking exceedingly bright for zebrafish genome engineering and genome science. Open technologies have allowed researchers across the globe to assemble and modify the systems discussed in this review at an exponential pace and in ways that were previously inconceivable. These innovations will bring swift change to the way we do science as we optimize our tools and learn to use them for new and exciting applications.

Acknowledgments

This article was supported by the State of Minnesota for the University of Minnesota Mayo Clinic Partnership Grant to SCE, NIH Grant GM63904 to SCE, NIDDK P30DK084567, DA032194 Grant to KJC, and the Mayo Foundation.

Disclosure Statement

No competing financial interests exist.

References

- Patowary A, Purkanti R, Singh M, et al. A sequence-based variation map of zebrafish. *Zebrafish* 2013;10:15–20.
- Moore FE, Reyon D, Sander JD, et al. Improved somatic mutagenesis in zebrafish using transcription activator-like effector nucleases (TALENs). *PLoS ONE* 2012;7:e37877.
- Bogdanove AJ, Voytas DF. TAL effectors: Customizable proteins for DNA targeting. *Science* 2011;333:1843–1846.
- Cade L, Reyon D, Hwang WY, et al. Highly efficient generation of heritable zebrafish gene mutations using homo- and heterodimeric TALENs. *Nucleic Acids Res* 2012;40:8001–8010.
- Ramirez CL, Certo MT, Mussolino C, et al. Engineered zinc finger nickases induce homology-directed repair with reduced mutagenic effects. *Nucleic Acids Res* 2012;40:5560–5568.
- Bedell VM, Wang Y, Campbell JM, et al. In vivo genome editing using a high-efficiency TALEN system. *Nature* 2012;491:114–118.
- Neff KL, Argue DP, Ma AC, Lee HB, Clark KJ. Mojo hand, a TALEN design tool for genome editing applications. *BMC Bioinformatics* 2013;14:1.
- Zu Y, Tong X, Wang Z, et al. TALEN-mediated precise genome modification by homologous recombination in zebrafish. *Nat Methods* Published Online First: 24 February 2013. doi:10.1038/nmeth.2374.
- Wiedenheft B, Sternberg SH, Doudna JA. RNA-guided genetic silencing systems in bacteria and archaea. *Nature* 2012;482:331–338.
- Deltcheva E, Chylinski K, Sharma CM, et al. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* 2011;471:602–607.
- Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 2012;337:816–821.
- Hwang WY, Fu Y, Reyon D, et al. Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat Biotechnol* 2013;1–3.
- Mali P, Yang L, Esvelt KM, et al. RNA-guided human genome engineering via Cas9. *Science* 2013;339:823–826.
- Cermak T, Doyle EL, Christian M, et al. Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res* 2011;39:82.
- Reyon D, Tsai SQ, Khayter C, Foden JA, Sander JD, Joung JK. FLASH assembly of TALENs for high-throughput genome editing. *Nat Biotechnol* 2012;30:460–465.
- Ding Q, Lee Y-K, Schaefer EAK, et al. A TALEN genome-editing system for generating human stem cell-based disease models. *Cell Stem Cell* 2013;12:238–251.
- Huang P, Xiao A, Zhou M, Zhu Z, Lin S, Zhang B. Heritable gene targeting in zebrafish using customized TALENs. *Nat Biotechnol* 2011;29:699–700.
- Gupta A, Christensen RG, Rayla AL, Lakshmanan A, Stormo GD, Wolfe SA. An optimized two-finger archive for ZFN-mediated gene targeting. *Nature Methods* 2012;9:588–590.
- Sander JD, Dahlborg EJ, Goodwin MJ, et al. Selection-free zinc-finger-nuclease engineering by context-dependent assembly (CoDA). *Nature Methods* 2010;8:67–69.
- Chen S, Oikonomou G, Chiu CN, et al. A large-scale in vivo analysis reveals that TALENs are significantly more mutagenic than ZFNs generated using context-dependent assembly. *Nucleic Acids Res* 2013;41:2769–2778.
- Sander JD, Cade L, Khayter C, et al. Targeted gene disruption in somatic zebrafish cells using engineered TALENs. *Nat Biotechnol* 2011;29:697–698.

Address correspondence to:

Stephen C. Ekker, PhD

Department of Biochemistry and Molecular Biology

Mayo Clinic

200 1st Street SW

Guggenheim 1342C

Rochester, MN 55905

E-mail: ekker.stephen@mayo.edu