New and TALENted Genome Engineering Toolbox

Jarryd M. Campbell,* Katherine A. Hartjes,* Timothy J. Nelson, Xiaolei Xu, Stephen C. Ekker

Abstract: Recent advances in the burgeoning field of genome engineering are accelerating the realization of personalized therapeutics for cardiovascular disease. In the postgenomic era, sequence-specific gene-editing tools enable the functional analysis of genetic alterations implicated in disease. In partnership with high-throughput model systems, efficient gene manipulation provides an increasingly powerful toolkit to study phenotypes associated with patient-specific genetic defects. Herein, this review emphasizes the latest developments in genome engineering and how applications within the field are transforming our understanding of personalized medicine with an emphasis on cardiovascular diseases. (Circ Res. 2013;113:571-587.)

Key Words: cardiovascular disease modeling ■ genome engineering ■ induced pluripotent stem cells ■ TALENs ■ zebrafish

Cardiovascular disease affects ≈83.6 million American adults and is the leading cause of death in the United States.¹ Recent advances in the ease and affordability of genome sequencing have illuminated a spectrum of genetic underpinnings linked to various cardiovascular diseases.² ³ ⁴ From the initial discovery of genetic mutations, further study of the molecular pathways underlying cardiac dysfunction uses both in vitro and in vivo model systems. However, unprecedented numbers of candidate genes from whole-genome screens commonly overwhelm the capacity of downstream functional analyses. The prioritization of candidate mutations and identification of causative mutations require the continued development of high-throughput tools and platforms to assess candidate genes efficiently and effectively. Genomic engineering is a rapidly evolving field that can now facilitate gene manipulation in a sequence-specific manner in most genomic contexts. The disruptive nature of this technology, when coupled with high-throughput model systems, is accelerating the study of cardiovascular disease as well as the development of personalized therapeutics by revealing innate genotype-phenotype mechanisms.

With site-specific gene editing, researchers manipulate experimental model systems by modulating the genomic mutation of interest. Novel genetic toolkits can be used to generate targeted gene lesions or knockouts, genome-based reporter systems, and single nucleotide changes at gene loci across the genome. One prominent gene-editing strategy harnesses the sequence-specificity of transcription activator-like effector nucleases (TALENs) to introduce site-specific mutations in gene loci-of-interest.⁵ TALENs are attracting well-deserved attention for their simplicity and flexibility of design as well as their gene-targeting efficiency.⁶ TALEN technology can be maximized in the setting of personalized cardiovascular research when paired with high-sensitivity and high-throughput cell- and tissue-based model systems.

Zebrafish (Danio rerio) offer an established vertebrate model that is well positioned to capitalize on TALEN-based precision engineering. One pair of adult zebrafish can produce 100+ embryos weekly, and the relatively inexpensive maintenance of fish colonies facilitates a large number of specimens to be mutagenized, characterized, and raised. Once the appropriate mutations are introduced, the cardiac phenotype can be studied in fish larvae, as zebrafish undergo full cardiac development within 48 hours, or in the adult fish for cardiovascular diseases that develop over an extended time. Importantly, zebrafish bring the power of forward genetic screens into a vertebrate system.

TALEN technology also enables genetic manipulation of induced pluripotent stem (iPS) cells. Characterized by...
pluripotent capacity and unlimited self-renewal, iPS cells can provide a long-term cell culture platform with the ability to generate functional cardiomyocytes from primitive to mature developmental stages. Importantly, iPS cells can be derived from patients with a variety of cardiovascular diseases and, therefore, offer a novel way to model disease in vitro. Genetic manipulation of iPS cells via TALENs promises to further advance the use of this platform to interrogate cell-autonomous genotype-phenotype relationships in the absence of physiological compensation. In addition, gene correction of patient-derived iPS cells using TALENs could generate disease-free autologous cells with potential therapeutic promise.

In this review, we highlight the technical aspects of TALEN design and assembly as well as current applications of TALEN technology in the study of cardiovascular diseases. With a focus on zebrafish and pluripotent stem cell model systems, we explore the advantages of both disease-in-a-fish and disease-in-a-dish applications. Our objective is to highlight how TALEN-based genome engineering in partnership with biological model systems accelerates the advancement of personalized medicine for cardiovascular diseases.

**Origin of TALENs**

Transcription activator-like (TAL) proteins were originally discovered by studying secreted proteins of the bacterial plant pathogen *Xanthomonas* and named for their innate ability to activate transcription of endogenous plant genes essential for pathogenicity.7 The DNA-binding domain of TAL proteins was found to comprise several 33 to 35 amino acid repeats with the half-repeat will be included as a whole RVD in the final TAL protein targeting length (ie, 14.5 repeats=15 RVDs). Like zinc-finger nucleases (ZFNs) before,11 TAL proteins were fused to a type IIS endonuclease, FokI, assembling the TALENs, a new class of artificial targeted DNA nucleases (Figure 1).5,12

**Harnessing DNA Repair Mechanisms for Genome Engineering**

TALENs enable locus-specific mutagenesis by inducing targeted double-strand breaks after 2 TALEN arms each bring half of the FokI endonuclease within proximity to dimerize in the spacer, or the stretch of DNA between the TALEN arms (Figure 1).3 Eukaryotic cells repair the resulting double-strand breaks by either error-prone nonhomologous end joining (NHEJ), resulting in insertions or deletions (indels) at the cut site, or by homologous recombination (HR) or homology-directed repair (HDR) in the presence of an exogenous DNA donor (Figure 1).13 If targeted to a conserved exon, the indels from NHEJ may induce a frameshift mutation that can result in a protein knockout, a useful tool that has broad applications to any gene-of-interest.14

In contrast to NHEJ, both HR and HDR require exogenous DNA in the form of double-stranded DNA (dsDNA; plasmid, adenoviral vector, polymerase chain reaction amplicon, etc) or a synthesized single-stranded oligodeoxyribonucleotide (ssODN) to serve as a template for programming of site-specific knockins (Figure 1). The choice of this donor DNA depends on the experiment and model system. In general, to introduce larger stretches of DNA changes, such as coding exons or artificial expression cassettes, dsDNA with long homologous arms that will induce HR is effective.15 For example, researchers are able to make substantial changes to the genome that may include fluorescent labeling of a protein to better understand its function. In the context of cardiovascular disease modeling, increased understanding of protein function may provide insight into the mechanism of disease and inform the development of novel therapeutics. Efficacy can vary greatly between model systems or even between loci within the same system; furthermore, HR is a unanimously rarer event than NHEJ-based mutagenesis attributable to the much slower efficiency and speed of this mechanism of repair.16

Although smaller DNA changes may be accomplished using traditional HR, there is a growing appreciation for using ssODNs for this application (Figure 1).17 The efficiency of editing using ssODNs can be high enough to not necessarily require enrichment even when used in vivo.18 To date, ssODNs have been used to introduce short DNA fragments, such as single nucleotide polymorphisms (SNPs), small insertions, or epitope tags.18–21 Unlike dsDNA, which is used as a template in the canonical HR pathway, ssODNs seem to use a related, but distinct, HDR mechanism that is not well understood. One hallmark of this HDR pathway is the insertion imprecision that can arise by using ssODNs in zebrafish,19 mammalian cell culture,21 and mouse embryos.22 This asymmetrical repair process, however, has not been reported in all studies using ssODNs as HDR donors.20,21 Gaining a better understanding of this HDR mechanism, especially in vivo models, such as zebrafish, is an important next step in the field. In particular, given the ease of ssODN synthesis and delivery compared with

---

**Nonstandard Abbreviations and Acronyms**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cas</td>
<td>CRISPR-associated</td>
</tr>
<tr>
<td>CRISPR</td>
<td>clustered regularly interspaced short palindromic repeats</td>
</tr>
<tr>
<td>dsDNA</td>
<td>double-stranded DNA</td>
</tr>
<tr>
<td>GWAS</td>
<td>genome-wide association study</td>
</tr>
<tr>
<td>HDR</td>
<td>homology-directed repair</td>
</tr>
<tr>
<td>HR</td>
<td>homologous recombination</td>
</tr>
<tr>
<td>iPS cells</td>
<td>induced pluripotent stem cells</td>
</tr>
<tr>
<td>NHEJ</td>
<td>nonhomologous end joining</td>
</tr>
<tr>
<td>REAL</td>
<td>restriction enzyme and ligation</td>
</tr>
<tr>
<td>RVD</td>
<td>repeat variable di-residue</td>
</tr>
<tr>
<td>SNP</td>
<td>single nucleotide polymorphism</td>
</tr>
<tr>
<td>ssODN</td>
<td>single-stranded oligodeoxyribonucleotide</td>
</tr>
<tr>
<td>TALEN</td>
<td>transcription activator-like effector nuclease</td>
</tr>
<tr>
<td>ZFN</td>
<td>zinc-finger nuclease</td>
</tr>
</tbody>
</table>
dsDNA, ssODNs may be more amenable to the high-throughput genome editing required to screen the numerous candidate genes identified in whole-genome sequencing projects.

**TALEN Scaffold**

TALEN proteins can be conceptually broken down into the DNA-binding domain (composed of RVDs), the N- and C-termini of the protein, and the FokI endonuclease domain (Figure 1). The repeats that make up the RVD domain can be assembled in multiple ways, which will be discussed further below. For clarity, the N-, C-termini, and FokI domain together will be referred to as the TAL scaffold in this review.

Interestingly, there has been a divergence of TAL proteins chosen to make TALENs. The difference between the proteins is species specific; one is originally from *Xanthomonas oryzae*, a rice pathogen, and another is from *Xanthomonas axonopodis*, a soybean pathogen. Although both are effective mediators of their respective pathogens, there are minor amino acid substitutions between the proteins that may or may not impact artificial TALEN efficiencies. The natural TAL scaffold that originated from *X axonopodis* was originally nicknamed TALE13 and was the first to undergo an architecture change of the N- and C-termini, which was found to improve activity.26 The original scaffold using the *X oryzae* TAL protein, pTAL,27 has also undergone similar architectural changes. Table 1 describes commonly used TALEN scaffolds in each of these classes to provide insight into the critical characteristics of TALEN scaffolds.

The primary overt differences between scaffolds are based on a combination of 4 factors: the expression system, the epitope tag(s), the N- and C-termini sequences, and the FokI domain. It should also be noted that all TAL scaffolds have a nuclear localization sequence that traffics these artificial proteins to the nucleus. The choice of promoter will determine whether constitutive expression of a DNA plasmid post-gene transfer (cytomegalovirus promoter [CMV], or CMV enhancer with chicken β-actin promoter [CAG]) or in vitro transcription of RNA (T7, T3, or SP6) and subsequent delivery will be used in the model-of-interest. There is an emphasis placed on the promoter choice when deciding on an expression construct,
Table 1. TALEN Scaffolds

<table>
<thead>
<tr>
<th>Species</th>
<th>Architecture*</th>
<th>FokI‡</th>
<th>Epitope Tags</th>
<th>Promoters§</th>
<th>Addgene Name</th>
<th>Assembly Compatibility</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Xanthomonas axonopodis</td>
<td>NΔ152/C63</td>
<td>WT</td>
<td>3X Flag</td>
<td>CMV and T7</td>
<td>JDS 70, 71, 74, 78</td>
<td>FLASH, REAL, REAL-Fast</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>NΔ152/C63</td>
<td>EL/KK or ELD/KKR</td>
<td>3X Flag</td>
<td>CMV and T7</td>
<td>N/A</td>
<td>FLASH, REAL, REAL-Fast</td>
<td>106</td>
</tr>
<tr>
<td></td>
<td>NΔ152/C63</td>
<td>RR/DD  or AS/RR</td>
<td>3X Flag and HA</td>
<td>CMV and SP6</td>
<td>N/A</td>
<td>Unit assembly</td>
<td>44</td>
</tr>
<tr>
<td></td>
<td>NΔ152/C63</td>
<td>WT</td>
<td>3X Flag</td>
<td>CMV and T7</td>
<td>TALEN-ID12-A (C, T, and G)</td>
<td>Ligation-independent cloning</td>
<td>42</td>
</tr>
<tr>
<td>Xanthomonas oryzae</td>
<td>NΔ152/C63</td>
<td>WT</td>
<td>None</td>
<td>miniCAGGs or T3</td>
<td>pc-GoldyTALEN, RCliscript-GoldyTALEN</td>
<td>Golden Gate 2.0</td>
<td>18, 67</td>
</tr>
<tr>
<td>(pTAL)‡</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NΔ152/C63</td>
<td>RR/DD</td>
<td>HA or Flag</td>
<td>CMV and SP6</td>
<td>pCS2TAL3-RR, pCS2TAL3-DD</td>
<td>Golden Gate 2.0</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>NΔ134/C63</td>
<td>WT or ELD/KKR and AS/RR</td>
<td>None</td>
<td>CMV and T7</td>
<td>pCAG-T7-TALEN(Sangamo)-Destination, -FokI-KKR-Destination, -FokI-ELD-Destination</td>
<td>Golden Gate 2.0</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>NΔ134/C47</td>
<td>WT</td>
<td>Flag</td>
<td>CMV and T7 or CAG</td>
<td>pcDNA-TAL-NC2, pCAGGS-TAL-NC2</td>
<td>Golden Gate 2.0</td>
<td>43</td>
</tr>
<tr>
<td>C63†</td>
<td>WT</td>
<td></td>
<td>3X Flag</td>
<td>CMV</td>
<td>pTALEN_v2 (N), (NN), (HD), (Ni)</td>
<td>Hierarchical PCR/ligation</td>
<td>68</td>
</tr>
</tbody>
</table>

CAG indicates cytomegalovirus promoter enhancer with chicken β-actin promoter; CMV, cytomegalovirus promoter; FLASH, the fast ligation-based automatable solid-phase high-throughput; HA, hemagglutinin; PCR, polymerase chain reaction; REAL, restriction enzyme and ligation; and WT, wild-type.

*All TALEN scaffolds have a nuclear localization sequence at the N-terminus.

†N-terminal truncation not reported.

but it is important to remember there may also be different untranslated regions or polyadenylation signals that may impact TALEN expression. Traditionally, RNA injection is most common in zebrafish, whereas DNA transfection/transduction is most often used in mammalian cell lines. However, because of the high mutagenesis caused by nuclease activity, a transient burst of expression may be more desirable for TALEN applications and, therefore, RNA delivery to cells is an on-going area of investigation.

For functional studies to characterize protein activity, epitope tags are commonly added to the TAL scaffold. Although the primary reason for this addition is convention, it is unclear whether or not this feature impacts TALEN activity because the TAL scaffolds are often cloned into pre-existing plasmids that already include the epitope tag. The most commonly included epitopes are the hemagglutinin and flag (or 3x flag) tag (Table 1). The architecture of the TAL scaffold includes truncations of the N- and C-termini, most commonly a 152 amino acid truncation of the N-terminus with a 63 amino acid C-terminus (NA152/C63) that was initially established by Miller et al. There is a balance in how an architecture change impacts TALEN activity versus the specificity of binding. For example, shortening the C-terminus seems to restrict the TALEN spacer length required for activity (described below), which may increase the specificity of the TALEN. However, there have also been reports of a shorter C-terminus having more" and less" activity than the C63 design. Keeping this balance in mind, the preferred scaffold architecture for different models and applications is a continuing area of research.

An important consideration of any homing nuclease is off-target mutagenesis. TALENs seem to have minimal off-target effects," particularly when compared with ZFNs. Mutations that have been detected thus far are at loci with high homology to the target, and there is much work to be done to appreciate the rate of genomewide off-targeting of TALENs. One approach to further reduce off-targeting has been to alter the catalytic properties of the FokI endonuclease. Wild-type FokI requires homodimerization for activity and, in an attempt to improve specificity, obligate Fok1 heterodimers were developed by changing their amino acid composition. Today, there are several heterodimer FokI variants" and TALEN constructs with either homodimer or heterodimer FokI domains.

Although epitope tags facilitate protein analysis and the FokI heterodimer may reduce off-target effects, the impact of scaffold modifications on TALEN activity continues to be elucidated. The efficiency differences between architecture changes, notably the truncation of the C-terminus, have been well documented. The GoldyTALEN scaffold, for example, has been demonstrated to improve activity over the original, full-length pTAL scaffold. It should be noted, however, that it is difficult to compare activities of different scaffolds with the same architecture both within and between model systems because different groups often use different methods of quantification, and few direct activity comparisons have been made. Nonetheless, Table 2 compiles the TALEN scaffolds used and the reported efficacy in zebrafish and human stem cells to date. In zebrafish, it should be noted that both somatic and germline efficiencies using published TALEN scaffolds are sufficient to generate mutants. Furthermore, the efficiency of any HDR is markedly less efficient than NHEJ mutagenesis in this in vivo system. In stem cells, mutation rates tend to be
lower; however coselection using puromycin or green fluorescent protein (GFP) can drastically improve activity (Table 2). Perhaps the most important conclusion that can be drawn is that there is great variability in TALEN activity and, in general, each scaffold has the potential to produce high activity. However, it is becoming clear that the design of the TALEN in combination with the scaffold choice can be optimized to improve the average activity. While Table 2 provides a snapshot of current progress in this field, it also suggests there is room for further optimization.

### Considerations for TALEN Design

Although TALEN sequence targeting has few known absolute constraints, some best-practice rules governing the efficiency of mutagenesis have been identified. Early in silico data provided the first insights into the importance of some key components. For example, endogenous TAL protein sequence targets were preceded by a 5′ thymine (5′T) attributable to an interaction between the N-terminus of the TAL protein and the 5′T. There have been reports emphasizing and negating the importance of the 5′T when designing the targeted nuclease, however. Most current TALEN design tools recommend including the 5′T.

In addition, the length of both the TALEN arms and spacer region impacts efficiency. The length of the TALEN binding domain has a minimal requirement of 11 RVDs for activity. Depending on the scaffold, different groups have generated TALENs with binding arms ranging from 13 to 40 RVDs. To date, only 3 groups have systematically tested the effect of arm length on TALEN activity. Collectively, 14 to 20 RVDs are sufficient to maintain high activity with minimal toxicity. More specifically, one study suggests that 15-RVD TALEN pairs maintain high activity with a spacer length between 13 and 20 basepairs. There is currently no evidence that longer sequences reduce off-target effects; and structural evidence suggests that RVDs closer to the N-terminus may play a larger role in binding than more C-terminal RVDs. In contrast to the 5′T recommendation for the target sequence and the nucleotide binding length that may to be applicable to most TALEN systems, the optimal spacer length can be TALEN scaffold specific. Different backbones have different C-termini lengths that modulate the proximity between the Fok1 nuclease domains and impact dimerization. Conservatively, a spacer region from 14 to 20 basepairs will lead to efficient TALEN activity for scaffolds with a C-terminus length ≥ 63 amino acids. Other scaffold architectures with even more truncated C-termini, from 17 to 28 amino acids, function more efficiently with a shorter spacer of 12 to 14 basepairs.

RVD designs are based on several naturally occurring RVDs, each with differing specificities and binding strengths. Three of the 4 nucleotides (A, C, and T) have established RVDs that are commonly used (NI, HD, and NG, respectively). There is less consensus on how best to target guanine. NN, one of the first RVDs identified to target G, also recognizes and binds A. The other early identified RVD targeting G, NK, seems

### Table 2. Efficacy of TALEN Scaffolds

<table>
<thead>
<tr>
<th>Model System</th>
<th>TALEN Scaffold</th>
<th>Assembly Protocol</th>
<th>TALEN Efficiency</th>
<th>Notes</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zebrafish</td>
<td>JDS series</td>
<td>REAL assembly</td>
<td>0% to 55% somatic mutation rate, 10% to 100% germline efficiency (5 loci)</td>
<td></td>
<td>39</td>
</tr>
<tr>
<td>GoldyTALEN</td>
<td>Golden Gate</td>
<td>7% to 100% somatic mutation rate, 17% to 100% germline efficiency (10 loci)</td>
<td>18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GoldyTALEN</td>
<td>Golden Gate</td>
<td>24% to 86% somatic mutation rate, 18% to 100% germline efficiency (10 loci)</td>
<td>25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TALE13 N.A.152/C63 - Fok1 AS/RR</td>
<td>Unit assembly</td>
<td>9% to 33% germline efficiency (3 loci)</td>
<td>44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>JDS series</td>
<td>Iterative assembly</td>
<td>11% to 33% somatic mutation rate (4 loci)</td>
<td>158</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TALE13 N.A.152/C63 - Fok1 AS/RR</td>
<td>Unit assembly</td>
<td>70% to 98% somatic mutation rate (1 locus); HR</td>
<td>107</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pCS2TAL3-DD, pCS2TAL3-RR</td>
<td>Golden Gate</td>
<td>98% to 100% somatic positive embryos, 77–100 germline efficiency (4 loci)</td>
<td>29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>JDS series</td>
<td>FLASH</td>
<td>20% to 77% somatic mutation rate (6 loci)</td>
<td>159</td>
<td></td>
<td></td>
</tr>
<tr>
<td>JDS-Fok1 EL/KK or ELD/KKR</td>
<td>FLASH</td>
<td>0% to 76% somatic mutation rate, 8% to 63% (10 loci), germline efficiency (8 loci)</td>
<td>106</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Human stem cells</td>
<td>JDS series</td>
<td>REAL assembly</td>
<td>HR=100% with coselection* (1 loci)</td>
<td>152</td>
<td></td>
</tr>
<tr>
<td>pTAL-Δ152/C63-Fok1 EL/KKR and AS/RR</td>
<td>Modified Golden Gate</td>
<td>NHEJ=2% to 34%; HDR=1.6% without coselection (16 loci)</td>
<td>23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>JDS series</td>
<td>DNA Works160</td>
<td>HR=1% to 100% with coselection* (5 loci)</td>
<td>15</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

HDR indicates homology-directed repair; HR, homologous recombination; NHEJ, nonhomologous end joining; and REAL, restriction enzyme and ligation.

*HR-positive cells were enriched for by coselecting with puromycin or GFP.
to be more specific, but incorporation of NK instead of NN leads to a lower activity TALEN.\textsuperscript{26,44} Subsequently, NH was identified, which seems to be more specific to G than either NK or NN.\textsuperscript{45} The impact of using NH in place of NN or NK has not been thoroughly documented; however, one study suggests it has higher specificity but lower activity than NN, much like NK.\textsuperscript{46} An additional consideration is that TALEN activity can be more influenced by CpG methylation compared with ZFNs.\textsuperscript{39,47} To overcome this issue, NG, HG, and N* RVDs (deletion of residue 13) are able to bind 5-methyl cytosine efficiently and should be considered when designing TALENs to regions of high CpG methylation.\textsuperscript{48,49} To make identifying target TALEN sites simple, software that can probe DNA for potential target sequences using several of the rules governing TALEN design has been developed.\textsuperscript{50-52} After identifying the target region within the locus-of-interest, the RVD repeats can be systematically assembled using established protocols.

**TALEN DNA–Binding Domain Assembly**

Several methods have been developed to generate the RVD modules of the TAL protein DNA–binding domain before insertion into the TAL scaffold (for references, see Table 3). Broadly, the methodologies can be broken down into 4 unique classes: plasmid-based restriction digest and ligation, PCR-based restriction digest and ligation, ligation-independent assembly, and solid-state assembly (Figure 2). Each of the diverse assembly methods may encompass one or more of these classes, and the TAL scaffolds that complement each method are not always interchangeable; the assembly will dictate the selection of TAL scaffolds and vice versa (Table 1). Table 3 is a list of common assembly options, including the method used to build repeats, the minimum length of time to build, and the targeting length limitations. In addition, the availability of the method on the nonprofit plasmid repository, Addgene, and the corresponding name have been included for ease of identification.

The Cermak et al\textsuperscript{27} Golden Gate RVD assembly method is among the most popular approaches used to date and can be classified as a plasmid-based restriction digest and ligation reaction. Golden Gate reactions use a type II restriction endonuclease that cuts outside of its recognition sequence and leaves a predictable 4 basepair overhang, to which the next plasmid can be added sequentially by designing ends, which complement and eliminate the original restriction site.\textsuperscript{53} This allows a single enzyme to cut all the necessary plasmids in a one-pot reaction. The original Golden Gate method takes a minimum of 5 days to complete. Several groups have attempted to streamline the TALEN-building process for more rapid assembly using variations of an RVD multimer platform and a PCR-based restriction digest and ligation reaction to overcome longer individual steps like overnight bacterial cultures (Table 3; Figure 2). Importantly, not all Golden Gate-like methods are compatible with the same TAL scaffolds (Table 1).

Another iteration of the plasmid-based restriction digest and ligation assembly uses hierarchical digestion/ligation, but does not use the multifragment ligation reaction characteristic of Golden Gate. This assembly class includes the unit assembly method\textsuperscript{44} and the restriction enzyme and ligation (REAL) method.\textsuperscript{54} Unit assembly relies on simple isocaudamer restriction sites that leave complementary overhangs for each subsequent ligation. The REAL method uses a combination of type I and II restriction enzymes to make complementary overhangs without requiring a multistep reaction to assemble the final product.

As opposed to each of the steps thus far requiring ligation to assemble RVD repeats, a ligation-independent cloning reaction has been described.\textsuperscript{42} This method uses DNA T4 polymerase, a 3′→5′ exonuclease, which chews back free dsDNA ends and leave 5′ overhangs. To control the length of the overhangs, dNTPs can be added to compete with DNA T4 polymerase and stop its activity. For instance, DNA T4 polymerase in the dTTP chews back the free 3′ ends until the first thymine is reached. This method facilitates the synthesis of predictable complementary overhangs; however, it does not require specific site recognition characteristic of restriction digestion. In addition, the overhangs can be much longer for direct transformation of the fragments along with both an antibiotic resistance gene and an origin of replication into bacteria without in vitro ligation.

TALEN assembly on solid-state platforms has also been demonstrated. The fast ligation-based automatable solid-phase high-throughput (FLASH) assembly method uses a library of TAL effector repeat plasmids encoding 1 to 4 RVD repeats, which can be assembled in a directed fashion on solid-phase magnetic beads.\textsuperscript{50} The process was scaled up to a 96-well plate, from which 96 different TALENs could be made simultaneously. The REAL-Fast assembly\textsuperscript{44} complements both the REAL and the FLASH methods by enabling the FLASH plasmids to be assembled by basic restriction digest/ligation strategies. Like the FLASH method, iterative cap assembly was developed on a solid-state platform.\textsuperscript{58} Iterative cap assembly enables rapid construction of TALENs using RVD repeat monomers that are sequentially assembled on solid-state beads. Thus, this method is unique from the others in that it will be able to use advances in microarray printing technology to build TALENs.

Which TALEN assembly method to choose depends on the needs of each project. For example, the original Golden Gate and REAL assembly methods have translated well into practice for individual laboratories interested in making TALENs on an as-needed basis. These methods only require routine molecular cloning reagents and expertise, making it relatively straightforward to initiate without purchasing expensive equipment. However, unless modified, these methods are not able to scale-up like the FLASH or iterative cap assembly methods. Generally, there are limited kinds of TAL scaffolds (and often only 1) compatible with each assembly method, decreasing the flexibility of many methods for the user. One notable exception is the Golden Gate method, which has been popular with individual labs and, subsequently, additional compatible scaffolds and accessory systems have been established.

Although TALENs remain under active research and development, the promise of this technology to accelerate the field of cardiovascular biology hinges on the coupling of TALENs with in vivo and in vitro platforms. The complementary nature of whole animal models and cell culture systems allows researchers to interrogate questions of both cardiovascular disease physiology and cellular phenotypes. This review focuses...
primarily on TALEN applications within zebrafish and human iPS cells; however, these platforms are among many others that have been genetically engineered using TALENs. Recently, TALENS have been harnessed to engineer a variety of whole animal models via direct embryo injection or transfection of fibroblast cells followed by reproductive cloning. The use of this gene-editing tool has been demonstrated in diverse organisms including roundworms,55 silkworms,56 fruit flies,57 mosquitoes,58 crickets,59 frogs,60,61 mice,22,62–64 rats,65 pigs,66,67 and cows.67 In addition, TALENs have also shown efficacy in a variety of cellular model systems, including both primary and transformed human cells.26,31,37,40,68–70 When combined with these cell culture platforms, TALENs can facilitate a variety of disease modeling applications. Recent reports have demonstrated the efficacy of TALENs to induce mutations within human miRNAs70 and generate oncogenic translocations in human mesenchymal progenitors.71 With growing numbers of publications using TALEN technology, this genome editing tool is gaining well-deserved popularity within diverse fields of research. The following sections highlight the use of zebrafish and iPS cells to model cardiac diseases and facilitate the development of novel therapeutics by capitalizing on the power of TALENs.

### Zebrafish: Physiological Models of Cardiovascular Disease

Zebrafish occupy a unique niche for investigating cardiac diseases as a high-throughput in vivo vertebrate model system. With expression annotation via whole mount in situ data for one third of the genome,72–74 zebrafish are an increasingly well-characterized model system that can be further enriched through the power of TALEN technology. High fecundity, rapid cardiogenesis, and the ability to conduct forward genetic screens have already established zebrafish as the premier nonmammalian model for developmental genetic studies. Importantly, this system is quite economic with the maintenance cost of adult fish at one tenth or less than the maintenance cost of the same number of mice. As an inexpensive vertebrate model system, zebrafish are a quality platform to study late-onset or progressive cardiovascular diseases that require extended observation time. In addition, with impressive capacity to regenerate injured myocardium,

<table>
<thead>
<tr>
<th>Class/Name</th>
<th>Description</th>
<th>Minimum Length of Protocol*</th>
<th>Size RVD Assembly (bps)</th>
<th>Addgene Name</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Golden Gate (GG)</td>
<td>Original Golden Gate</td>
<td>5 days</td>
<td>12–31</td>
<td>Golden Gate 2.0</td>
<td>27</td>
</tr>
<tr>
<td>Modified Golden Gate</td>
<td>Original GG with FusB4 library</td>
<td>5 days</td>
<td>15</td>
<td>N/A</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>Original GG with 6-RVD FusA</td>
<td>5 days</td>
<td>12–31</td>
<td>TALEN construction and evaluation accessory pack</td>
<td>43</td>
</tr>
<tr>
<td></td>
<td>Original GG with tetramer/trimer library</td>
<td>3 days</td>
<td>15</td>
<td>N/A</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>GG/PCR assembly of multimers</td>
<td>3 days</td>
<td>14–19</td>
<td>N/A</td>
<td>156</td>
</tr>
<tr>
<td></td>
<td>GG tetramer/dimer/monomer library</td>
<td>3 days</td>
<td>15–19</td>
<td>N/A</td>
<td>14</td>
</tr>
<tr>
<td>Hierarchical PCR/ligation</td>
<td>GG followed by PCR/ligation (human codon optimized)</td>
<td>3 days</td>
<td>14–25</td>
<td>Zhang laboratory TALEN Toolbox</td>
<td>68</td>
</tr>
<tr>
<td>Modular Assembly</td>
<td>GG-like/8-mer assembly</td>
<td>5 days</td>
<td>16 or 24</td>
<td>N/A</td>
<td>157</td>
</tr>
<tr>
<td>Unit Assembly</td>
<td>Units assembled via icosaudamer restriction digest/ligation</td>
<td>7 days</td>
<td>10–17</td>
<td>N/A</td>
<td>44</td>
</tr>
<tr>
<td>FLASH</td>
<td>Libraries of 2-, 3-, and 4-mer repeats/solid-phase magnetic beads</td>
<td>3 days</td>
<td>Unlimited</td>
<td>N/A</td>
<td>40</td>
</tr>
<tr>
<td>REAL†</td>
<td>Restriction digest/ligation (non-GG)</td>
<td>9 days</td>
<td>13–17</td>
<td>Joung Laboratory REAL assembly TALEN kit</td>
<td>54</td>
</tr>
<tr>
<td>REAL-Fast‡</td>
<td>FLASH libraries/cloning hybrid</td>
<td>7 days</td>
<td>13–17</td>
<td>N/A</td>
<td>54</td>
</tr>
<tr>
<td>Ligation-independent cloning</td>
<td>T4 DNA polymerase endonuclease activity to assemble multimers</td>
<td>3 days</td>
<td>16 or 19</td>
<td>LIC TAL effector assembly kit</td>
<td>42</td>
</tr>
<tr>
<td>Iterative Cap Assembly</td>
<td>Sequential construction using monomers on solid-state beads</td>
<td>3 days</td>
<td>Unlimited</td>
<td>N/A</td>
<td>38</td>
</tr>
</tbody>
</table>

FLASH indicates the fast ligation-based automatable solid-phase high-throughput; PCR, polymerase chain reaction; REAL, restriction enzyme and ligation; and RVD, repeat variable di-residue.

* Differences in lengths a result of the number of overnight bacterial cultures.
† For TALENs longer than 17 RVDs, 2 additional days are needed.
Figure 2. Classes of transcription activator-like (TAL) repeat assembly methodologies. **Top**, Shows a plasmid-based restriction digest and ligation method that moves TAL repeat monomers (in blue, red, yellow, and green) in a predetermined order from one plasmid to a receiver plasmid by one round of restriction digest and ligation, and then into the destination TAL backbone by another round. The second panel from the top shows a PCR-based restriction digest and ligation method that is conceptually similar to a plasmid-based assembly, however a PCR amplicon replaces whole plasmids as the starting material, and PCR amplification can replace bacterial plasmid propagation. The restriction sites necessary for the digest and ligation reactions are built into the primer design. The next panel outlines a ligation-free assembly method, in which T4 DNA polymerase activity in the presence of dNTPs chews back free double-stranded DNA ends to make long single-stranded overhangs. These overhangs anneal efficiently, and bacteria ligate together the DNA backbone if a bacteria origin of replication (grey) and antibiotic resistance gene (black) are included in the reaction. This method eliminates the need for in vitro ligation. **Bottom**, Demonstrates how solid-state technology can facilitate high-throughput TAL repeat assembly. A biotinylated repeat monomer is first tethered to a plate coated with streptavidin beads, and then monomers are sequentially added until the desired repeat length is achieved. The final assembly is cloned into the final TAL scaffold. TALEN indicates transcription activator-like effector nuclease.
zebrafish offer an ideal model system to study cardiac regeneration.75–78 Although comprising a single atrium and a single ventricle, the 2-chamber zebrafish heart is an established in vivo platform for modeling several forms of human cardiovascular diseases.79–85 By means of positional cloning, pickwick and tnt2, truncation mutations of cardiac genes, titin and Tnt2,46 were considered the first 2 embryonic zebrafish models for cardiomyopathy. Later, a series of loss-of-function studies of known cardiomyopathy genes were performed, including actn2, mlc, rlc, cypher, and mlp, underscoring the value of embryonic zebrafish for annotating cardiac gene expression and function.83,87–89 Phenotypic analysis of ilk and nelixin mutants in zebrafish embryos supported the corresponding human genes as novel dilated cardiomyopathy-causative genes.84,90 Zebrafish embryos have also been used to model arrhythmia defects such as Long QT syndrome,91,92 as well as pathogenic events during atherosclerosis.93,94 These innovative, proof-of-principle studies have established zebrafish embryos as a unique platform capable of revealing novel insights into human cardiovascular diseases.95,96 A major incentive for generating disease models in zebrafish embryos is the feasibility to directly identify therapeutic compounds via high-throughput in vivo pharmaceutical screens because of their small size and ability to absorb small molecules from their environment. Several chemical screens have already been used using zebrafish as a platform to identify compounds of potential therapeutic value for treating Long QT syndrome and hypertrophic cardiomyopathy.97,98

Although recognized as a useful model for certain aspects of cardiac disease phenotypes, inherent limitations in embryonic zebrafish models continue to restrict their extensive use to predict genotype-phenotype relationships in humans. For example, many of the genes associated with human cardiomyopathies are completely depleted in embryonic zebrafish models. This makes it difficult to study cardiovascular diseases that develop later in life, especially those with autosomal dominant inheritance. In addition, the short developmental time restricts fish embryos from precisely recapitulating cardiac pathogenesis, which typically exhibits age-dependent penetrance and gradual progression to overt heart failure in adulthood. Given the intrinsic limitations of zebrafish embryos, several groups have explored adult fish as a model for cardiac diseases. For example, anemia imposes a high-output stress on the heart and can subsequently induce classical cardiomyopathy-like phenotypes in adult zebrafish, including myofibril disarray and reactivation of fetal gene transcription.80,81 Injection of doxorubicin, a widely used chemotherapeutic with cardiotoxic side effects,99,100 can also induce these cardiomyopathy-like phenotypes in adult fish.99 Cryoinjury of the adult zebrafish heart has recently been demonstrated as a useful model to study pathogenesis during myocardial infarction.76,101 As a complement to physically induced disease models, continued development in gene-editing technologies will further extend the capacity of this model system to recapitulate genetically linked cardiovascular diseases.

Zebrafish: Platform for Genomic Engineering

Classic genetic tools used in zebrafish include morpholino technology, targeted-induced local lesions in the genome,102,103 and zinc-finger nucleases.104 However, important caveats to each of these techniques limit their use as genomic tools. When morpholinos are used to knock down expression of a certain transcript, the downstream phenotype is largely restricted to the first few days of zebrafish development.105 Targeted-induced local lesions in the genome have been an important reverse genetics tool for identifying zebrafish mutants; however, it is a relatively expensive technique that can only feasibly be performed in large research centers, and the generated mutations carry a wide range of other, linked sequence changes as well.105 ZFNs have also notably contributed to genomic engineering in zebrafish, but the complexity of ZFN design and divergent range of activity have limited the use and adoption of this technology by most individual laboratories.106 In contrast, TALENs offer a more rapid and less expensive way to engineer stable mutant zebrafish lines genetically.29,107 TALEN technology coupled with zebrafish as a high-throughput, inexpensive in vivo discovery platform enables the ready generation of mutant lines for individual labs.

The experimental strategies that can be performed using TALENs differ based on which mechanism of genomic repair is induced downstream of the double-strand DNA break (Figure 3). When the genomic DNA is repaired by NHEJ, a frameshift mutation could result in a truncated gene product or abolish gene expression. The resulting mutant zebrafish lines can be easily screened for physiological genotype-phenotype relationships. In addition, TALENs can facilitate HR repair mechanisms using large dsDNA donors to insert selection cassettes or reporter systems downstream of candidate genes.107 A third application for TALEN technology in zebrafish is the use of small ssODNs to insert epitope tags, loxP sites, or alter single base pairs via HDR.18,19 Each of these 3 applications will be highlighted below in a discussion of current TALEN-based genome engineering in zebrafish.

Eliminate Gene Function: Truncational Mutations to Prioritize Candidate Genes From Human Genetic Studies

One application of TALENs in this model system is to generate numerous truncational, loss-of-function zebrafish lines that could prioritize candidate genes generated from whole-genome screens (Figure 3). As indicated in the Online Mendelian Inheritance in Man (OMIM) database,108 many human cardiac diseases are caused by nonsense mutations that result in truncated proteins. For such dominantly inherited alleles, the underpinning pathogenesis mechanism could be attributable to either haploinsufficiency or the generation of a dominant loss-of-function peptide, it is desirable to generate truncational mutations at the corresponding amino acid that is found in human patients. The flexibility of the TALEN-based system can effectively meet this challenge. A large number of disease-relevant truncational mutants will be generated via TALENs in the near future, which shall significantly expand the scope of cardiac diseases that adult zebrafish can model. This knowledge will lay the foundation for zebrafish as an in vivo functional assay to prioritize candidate genes identified from human cohort–based genetic studies, genome-wide association studies (GWAS), and quantitative trait locus analysis in rodents. As a complement to current mutagenesis techniques
in zebrafish, TALENs can be especially useful to probe the genes that were missed by previous nonspecific approaches, including mutagenesis screening and targeted-induced local lesions in the genome. In this way, TALEN-based gene editing contributes to the zebrafish phenome project as an efficient tool to generate novel mutant lines.

**Report Gene Function: Insertion of Reporter Systems to Selectively Study Candidate Genes**

The feasibility of TALEN-based gene editing in zebrafish will enable the insertion of reporter systems to any genomic loci, which can be used to tag proteins-of-interest (Figure 3). Both tissue-specific and subcellular localization of the tagged protein can be tracked during cardiac development or disease progression. The strength of this gene-editing strategy can be invaluable for interrogating the molecular underpinnings of cardiovascular diseases caused by the affected protein. The potential of inserting a targeted GFP reporter system in zebrafish via TALEN-mediated HR has been recently demonstrated. This method uses dsDNA with long homologous arms as a template to repair a double-strand break induced by TALENs and opens the door to the possibility of directing large fragments of DNA into the zebrafish genome.

**Alter Gene Function: Genetic Manipulation of Zebrafish Via HDR**

Smaller targeted modifications like loxP sites, epitope tags, or specific sequence changes can also be helpful when studying gene function (Figure 3). Recently, a loxP site was inserted into an intron using an ssODN as a donor for HDR in zebrafish. Although dsDNA may also facilitate this modification, ssODNs may be more simply synthesized in a high-throughput manner without requiring a time-consuming molecular cloning process. Approximately the same size as loxP sites, epitope tags may also be inserted using this same technique. Both of these genetic modifications could provide valuable tools for studying the function of cardiac genes. In addition, TALENs and ssODNs may also enable the insertion of SNPs into the zebrafish genome. By developing ways to modify single base-pairs in zebrafish, researchers can study numerous SNPs identified through whole-genome sequencing in an in vivo context.

There are several clinically relevant benefits to engineering SNPs in a high-throughput model like zebrafish. Unlike bigger vertebrates, zebrafish are amenable to large-scale drug screening efforts. Conceptually, zebrafish engineered with a patient-specific SNP can be subjected to FDA-approved drugs to search for those that ameliorate the physiological disease phenotype. This exciting prospect would allow for identification of drugs that are appropriate for that individual patient based on his/her personal genetic variability; true individualized medicine (Figure 3). Although this idea highlights the massive potential of using TALENs and zebrafish in cardiovascular medicine, there is much more work to be done to evolve precision genome engineering to meet such high demand.

**iPS Cells: Cell-Autonomous Models of Cardiovascular Disease**

Where high-throughput animal models are essential for interrogating physiological phenotypes of cardiac gene defects, questions about the cell-autonomous nature of cardiovascular diseases can now be probed within the native genomic context using patient-specific cell culture models. To address such clinical and patient-specific questions, iPS cells stand out as an ideal model system to harness TALEN technology and interrogate the cell-autonomous features of cardiac disease. Originally described by Takahashi and Yamanaka in 2006, iPS cells can be derived via nuclear reprogramming of mature cells back to the pluripotent stem cell state. Following the advent of nuclear reprogramming, it is now possible to derive patient-specific pluripotent stem cells that have the capacity to differentiate to any mature cell phenotype in the adult body, including functional cardiomyocytes. Thus, through in vitro cardiac differentiation, researchers can probe the molecular underpinnings of cardiac dysfunction as it relates to cell-autonomous genotype-phenotype perturbations. Perhaps the most exciting advantage of this platform is the capacity of iPS-derived cardiomyocytes to serve as a potential autologous cellular therapy for cardiovascular diseases.

Several unique qualities of iPS cells make them an unprecedented platform to harness TALEN technology for cardiovascular research (Figure 4). The pluripotent capacity of iPS cell lines enables unlimited in vitro generation of cardiac lineages that would otherwise be unavailable or inconvenient for cell culture. Primary cardiomyocytes are difficult to obtain from patients, have limited proliferative capacity, and can only be cultured for several days to weeks, whereas maintaining their contractile properties. In contrast, iPS cells can be derived from easily accessible cell sources, most often a skin biopsy, and can be expanded indefinitely as a renewable, high-throughput source of patient-specific cardiomyocytes. Also, by modeling normal developmental stages in the absence of physiological compensatory mechanisms, iPS cell platforms are well positioned to interrogate the cell-autonomous nature of genotype-phenotype relationships. Upon genetic manipulation of iPS cells, researchers can probe the impact of specific gene defects on cardiomyocytes, rather than on the heart as a whole organ. For cardiomyopathies characterized by variable penetrance, the generation of patient-specific iPS cells allows researchers to investigate genotype-phenotype relationships in the context of other unidentified modifier mutations, which may unmask the complexity of inheritance for these diseases. Importantly, cardiomyocytes derived from patient-specific iPS cells offer a novel way to model disease in vitro and test the efficacy and safety profile of pharmacological therapeutics on the patient’s own cells outside his/her body (Figure 4).

The value of patient-specific iPS cells as an autologous surrogate to model cardiac disease has been established by numerous studies to date. Although iPS-derived cardiac cells in vitro do not recapitulate all the structural organization of adult cardiomyocytes isolated directly from heart tissue, the functional characteristics of these bioengineered cells have demonstrated great potential as an in vitro disease-modeling platform. Several detailed reviews capture the current state of this field with an emphasis on arrhythmogenic cardiac diseases. Notably, iPS-based disease modeling has revolutionized the field of cardiovascular medicine by providing an
ex vivo drug screening platform. Applying high-throughput model systems, such as zebrafish and now patient-specific iPSC-derived cardiomyocytes, whole libraries of small molecules can be screened for functional improvement of cellular disease phenotypes. The patient-specific nature of these models promises to advance the field of personalized medicine and drug development. Previous studies using iPSC cells derived from patients with genetically linked arrhythmogenic cardiac diseases have demonstrated the use of this platform to recapitulate cellular physiology of gene-linked disease.

Figure 3. Genome engineering in zebrafish: modeling the physiology of gene-linked cardiovascular disease. 1, Whole-genome sequencing approaches reveal genetic mutations underlying cardiac disease in an individual patient. Identification of zebrafish orthologs enables transcription activator-like effector nuclease-based gene modification in zebrafish to study the patient-specific defect. 2, Error-prone or programmed gene repair mechanisms facilitate the generation of adult zebrafish with different experimental phenotypes. 3, Engineered zebrafish models provide patient-specific insights into the physiological mechanism of disease. 4, Drug screens offer the potential to identify novel therapeutic agents. 5, Pharmaceuticals are then tested and eventually delivered back to the patient to complete the individualized medicine loop.

Figure 4. Genome engineering in induced pluripotent (iPS) cells: modeling the cell-autonomous nature of gene-linked cardiovascular disease. 1, Fibroblasts are derived from skin biopsy of patient with an inherited cardiovascular disease. Nuclear reprogramming generates induced pluripotent stem cells that are genetically engineered via transcription activator-like effector nucleases. 2, Error-prone or programmed gene repair mechanisms followed by in vitro cardiac differentiation yield patient-specific cardiomyocytes with different experimental phenotypes. 3, Engineered cardiomyocytes provide patient-specific insights into the cell-autonomous mechanism of disease. 4a, Gene-corrected cardiomyocytes offer an autologous cellular therapy and (4b) safety profiling screens enable the identification of potential arrhythmogenic pharmaceutical agents. 5, Cellular therapies are delivered back to the patient to complete the individualized medicine loop.
disease phenotypes and pharmaceutical responses. iPS cell lines derived from patients with Long QT syndrome (types 1, 2, and 8) have been differentiated in vitro to cardiomyocytes that show prolonged action potentials attributable to mutations in KCNQ1, KCNH2, or CACNA1C, respectively.\textsuperscript{119–124} Mechanistically, these studies have shed light onto the cell-autonomous disease phenotype displayed by iPS-derived cardiomyocytes. In addition to Long QT syndrome, other arrhythmic cardiac diseases have been modeled in vitro using patient-specific iPS cells, including catecholaminergic polymorphic ventricular tachycardia, and an overlap syndrome caused by a defect in SCN5A.\textsuperscript{125–128} The use of iPS cells to model nonarrhythmic cardiac diseases has also been demonstrated in patient-derived iPS cells to model the pathognomonic features of LEOPARD syndrome and dilated cardiomyopathy.\textsuperscript{129,130}

Despite the documented capacity of iPS cells to model cardiac disease in vitro, an important limitation of current disease-modeling applications is the use of allogeneic healthy controls for comparison with patient-derived iPS cells. In addition to the clonal heterogeneity between iPS cell lines from a single individual, there is increased genetic and epigenetic variability when using allogeneic iPS cells.\textsuperscript{131–136} This genetic diversity may be unrelated to the disease phenotype and thereby confound comparative analysis of cardiac phenotypes derived from different iPS cells. To address this confounding variability, genome engineering strategies have been used to generate isogenic controls by inducing site-specific repair of disease-causing mutations in iPS cell lines.\textsuperscript{137} These gene-editing technologies ideally have minimal off-target effects to limit the mutational risk of unintentional genome modifications. The standard in vitro processing of iPS cells already poses a risk of altering the genotype of these cells, which must not be further increased by nonspecific mutations as a result of genomic engineering.

### iPS Cells: Platform for Genomic Engineering

Genome editing in iPS cells was pioneered by custom ZFNs. The use of this gene-editing tool in iPS cells was recently reviewed in Cheng et al.\textsuperscript{138} Briefly, ZFNs have been used to induce site-specific knockout,\textsuperscript{139–142} knockin,\textsuperscript{141,142} or correction of disease-related genes in human iPS cells.\textsuperscript{141,143–147} The power of genome engineering in iPS cells may be extended even further through the use of TALEN technology. With greater flexibility of design and ease of assembly, TALENs offer an increasingly popular alternative to ZFNs with high targeting efficiency in iPS cells.\textsuperscript{15,148} The capacity of TALENs to facilitate site-specific genome modifications in iPS cells through NHEJ, HR, and HDR repair mechanisms will be highlighted in the following sections.

### Eliminate Gene Function: Null Mutations to Probe Cell-Autonomous Genotype-Phenotype Relationships

A recent study by Ding et al.\textsuperscript{25} reveals the power of TALENs to induce genetic knockouts rapidly and efficiently as a strategy to study cell-autonomous genotype-phenotype relationships upon differentiation of human pluripotent stem cells. Although this group demonstrates similar efficiency of TALEN-targeted genetic mutations between human embryonic stem cells and iPS cells, the phenotypic characterization was restricted to genetically engineered human embryonic stem cells. The power and versatility of this platform is highlighted by the successful knockout of SORT1, a gene recently implicated in several disease states. Upon differentiation to hepatocyte-like cells, adipocytes, or motor neurons, the engineered cell lines displayed disease-related phenotypes when compared with isogenic controls.\textsuperscript{23} Through genetic manipulation at the pluripotent stage, this group modeled a variety of disease phenotypes by guiding the differentiation of human embryonic stem cells to different somatic lineages. Although this study focuses on noncardiac cell phenotypes, TALEN-based knockdown of critical cardiac genes would also provide mechanistic insights for cardiac disease modeling using iPS cells (Figure 4).

### Report Gene Function: Insertion of Reporter Systems to Identify and Study Cell Lineages

An additional application for TALEN-based engineering in iPS cells was recently demonstrated by Hockemeyer et al.\textsuperscript{15} This study demonstrates the use of TALENs to facilitate transgene insertion, including a GFP reporter downstream of pluripotency marker Oct4 in human embryonic stem and hiPS cells. The fidelity of this reporter system was demonstrated by the disappearance of GFP expression after in vitro differentiation. TALEN technology provides an effective genomic editing tool to label pluripotency genes selectively, which could be used for the isolation of residual iPS cells from a heterogeneous population of differentiating iPS progeny.\textsuperscript{149} By selectively purging the remaining pluripotent cells from a differentiating population, this strategy may complement other approaches to increase the safety profile of iPS-derived progeny for future therapeutic applications.\textsuperscript{150,151} In addition, TALENs could also be used to insert reporter constructs downstream of critical cardiac genes to follow their expression and localization during in vitro differentiation. Such a reporter system would also enable selection of transcriptionally defined cellular lineages from a heterogeneous population of differentiating iPS cells (Figure 4). A population of stage-matched cells would significantly improve the analytic power of this platform to study cell-autonomous mechanisms of cardiovascular disease.

### Alter Gene Function: Genetic Correction of Disease-Causing Mutations in iPS Cells

For monogenic diseases, TALEN-based gene correction can generate disease-free autologous iPS cells for potential delivery back to the patient (Figure 4). The efficiency of TALENs to facilitate gene correction in patient-derived iPS cells was recently demonstrated by Choi et al.\textsuperscript{152} In this study, traceless genomic repair of a disease-causing gene was reported in iPS cell lines derived from 2 patients with α1-antitrypsin-deficiency–related liver disease. TALENs targeting the mutated α1-antitrypsin efficiently induced biallelic HDR correction in 25% to 33% of puromycin-selected iPS cells. Upon differentiation to hepatocyte-like cells, the gene-corrected iPS cells displayed restoration of cellular function to the level of healthy controls. In addition, this group also performed a high-throughput pharmaceutical screen and identified 5
clinically relevant compounds that restored cellular function to the same level of rescue as the gene-corrected iPSCs. This study highlights the use of iPSCs cells for pharmaceutical screening and the efficiency of TALENs for genomic repair. In therapeutic applications, the gene-corrected cells may have increased viability that could translate to improved repair on reintroduction into the patient.

Overall, the unique advantages of TALEN-based genomic engineering paired with the unlimited potential of bioengineered patient-specific tissues with isogenic controls, reporter systems, and disease mapping mutagenesis have the potential to transform personalized cardiovascular research toward cell-based theranostics.

**Future Directions**

As TALEN technology continues to develop, there are questions being raised about the optimal methods of synthesis and distribution. One common option allows individual laboratories to assemble TALENs on an as-needed basis. Methods like Golden Gate assembly are ideal for this approach because they are straightforward for those comfortable with molecular cloning and readily accessible to the community. Although several labs have taken the approach, others aim to assemble more TALENs than feasible using standard cloning techniques and are primarily interested in downstream applications rather than the TALEN-building process. In response, scalable assembly platforms, such as FLASH or iterative cap assembly, may function well in core facilities at institutions in which TALENs are consistently being made in an efficient manner and distributed onsite. Several companies have also been established to assemble and distribute TALENs, and this is yet another avenue through which laboratories can obtain TALEN technology. With the simplicity of making TALENs, especially compared with ZFNs, the cost of TALENs is relatively low and continues to decrease as more companies enter the market.

Genetic engineering continues to be a fast-growing field, with new technologies regularly in development. The clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated (Cas9) system is one such technology. Unlike ZFNs and TALENs, which are protein-guided, the specificity of CRISPR/Cas9 targeting is RNA-guided. Although CRISPR/Cas9 applications are currently being actively investigated, its full range of activity, targeted and otherwise, is also being defined. Early studies indicate that on-site CRISPR/Cas9 targeting is roughly comparable with TALEN-based targeting; however, off-target activity is comparable or even greater than on-target activity in some cases. Just like ZFNs and TALENs before, there is likely room for optimization of the CRISPR/Cas9 system that could improve both specificity and efficacy.

Because disease modeling applications are rapidly emerging in both zebrafish and iPSC cell culture, genome engineering tools will also need to be optimized for both in vivo and in vitro model systems. For example, efficient donor templates will need to be developed to generate single basepair changes in zebrafish. Achieving this milestone will allow researchers to develop patient-specific zebrafish with subtle changes in the genome to model the physiological phenotype of specific SNPs. To further advance disease-modeling in iPSC cells, whose epigenetic profiles drastically change during in vitro differentiation, further studies will be needed to decipher the role that epigenetic environment of the target gene has on TALEN and CRISPR efficacy. Increased TALEN cutting efficiency at several gene loci has been demonstrated on treatment with a DNA methyltransferase inhibitor. This observation may indicate a decreased TALEN efficiency at somatic gene loci that are in a heterochromatic state in iPSC cells. It has been postulated that the helicase activity of Cas9 may facilitate more efficient gene editing at epigenetically silenced genes. Future studies that compare both gene editing tools will help address some of these remaining questions and further advance cardiac disease modeling in both zebrafish and iPSCs.

**Conclusions**

Here, we have highlighted the current state of TALEN technology as a novel genome editing strategy to advance cardiovascular research applications. The relative simplicity and flexibility of TALEN design and assembly enable efficient gene manipulation in a sequence-specific manner. When coupled with high-throughput model systems such as zebrafish and patient-specific iPSCs, TALENs can facilitate disease-modeling applications to study the cardiac phenotype and molecular underpinnings of individual genetic defects. The disruptive nature of this technology will accelerate the development of patient-specific cellular and pharmaceutical therapeutics for the personalized treatment of cardiovascular disease.

**Acknowledgments**

We thank Mayo Clinic Media Support Services for assistance in figure production and Alvin Ma, Han Lee, Hind Fadel, Patrick Blackburn, and Karl Clark for critical review of the article.

**Sources of Funding**

This project is supported by the State of Minnesota (University of Minnesota Mayo Clinic Partnership grant to S.C. Ekker), National Institutes of Health grant GM63904 to S.C. Ekker, National Institute of Diabetes and Digestive and Kidney Diseases grant P30DK084567 to S.C. Ekker, National Heart, Lung, and Blood Institute grants HL107304 and HL81753 to X. Xu, the Todd and Karen Wanek Family Program for Hypoplastic Left Heart Syndrome (T.J. Nelson), and the Mayo Foundation. J.M.C. was supported in part by Mayo CTSA Grant Number UL1TR000135 from the National Center for Advancing Translational Sciences.

**Disclosures**

T.J. Nelson and Mayo Clinic have a financial interest in the stem cell technology referenced in this review. The other authors have no conflict to report.

**References**

with NN and NK repeat variable di-residues.

Zhang Y, Carlson DF, Bradley P, Bogdanove AJ, Voytas DF. Targeting G.

2012;8:e1002861.

Curr Opin Microbiol.

ogous end joining and homologous recombination in human cells.

2011;29:731–734.

Lit T, Huang S, Wright D, Spalding MH, Weeks DF, Yang B. TAL


Li T, Huang S, Jiang WZ, Wright D, Spalding MH, Weeks DF, Yang B. TAL


Loehlein J, Stoddard BL. The crystal.

structure of TAL effector PhoX1 bound to its DNA target. Science.

2012;335:716–719.

Sun N, Liang J, Abil Z, Zhao H. Optimized TAL effector nucleases (TALENs) for use in treatment of sickle cell disease. Mol Bior.


2010;40;e117.

Chen S, Oikonomou G, Chiu CN, Niles BJ, Liu J, Lee DA, Antoshechkin I, Prober DA. 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;2799–2778.


2013;31;76–81.


2012;491;114–118.


2013;23;465–472.


2011;8;753–755.


2013;12;238–251.


2011;14;47–53.


2013;9;e80529.


2011;29;143–148.


2011;39;e82.


2012;7;e54383.


2012;8;1002861.


2013;2;363–367.


2011;39;9283–9293.


2007;25;778–785.


2008;26;695–701.


2010;400;96–107.


2011;8;74–79.

Mak AN, Bradley P, Cerdanova RA, Bogdanove AJ, Stoddard BL. The crystal.

structure of TAL effector PhoX1 bound to its DNA target. Science.

2012;335;716–719.


2011;29;699–700.

Nef KL, Argue DP, Ma AC, Lee HB, Clark KJ, Ekker SC. A TALEN design tool for genome editing applications. BMC Bioinformatics.

2013;14;1.


2012;40;5368–5377.


J Biol Chem.

2012;287;34827–34832.


2010;400:96–107.

Nakanishi K, Argue DP, Ma AC, Lee HB, Clark KJ, Ekker SC. MOJO HAND, a TALEN design tool for genome editing applications. BMC Bioinformatics.

2013;14;1.


2010;38;W462–W468.

Doyle EL, Booher NJ, Standage DS, Voytas DF, Brendel VP, Vandyk JK, Bogdanove AJ. TAL Effector-Nucleotide Targeter (TALE-N'T) 2.0.


New and TALENted Genome Engineering Toolbox
Jarryd M. Campbell, Katherine A. Hartjes, Timothy J. Nelson, Xiaolei Xu and Stephen C. Ekker

Circ Res. 2013;113:571-587
doi: 10.1161/CIRCRESAHA.113.301765

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circres.ahajournals.org/content/113/5/571

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Circulation Research can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Circulation Research is online at:
http://circres.ahajournals.org/subscriptions/