GENE EDITING

Re-writing genomes with targetable nucleases

We are in the midst of a revolution in genome engineering, based on reagents that can be designed to cut chromosomal DNA at arbitrary sites. These targetable nucleases allow creation of new mutations at specific sites, introduction of designed sequence changes and production of larger alterations, as desired by the experimenter. They are being used to explore gene function and create disease models and they hold great promise for human gene therapy.

By Dr Dana Carroll

One of the major advances in DNA technology in the last several years has been the invention of methods that allow rapid, massively-parallel DNA sequencing. With these technologies, the complete genome sequences of thousands of organisms have been determined, including those of many individual humans. For many scientists, the next step is to begin to manipulate these genome sequences for a variety of specific goals, including analysis of gene function, creating models of human disease, generating better models for drug discovery and, ultimately, human gene therapy.

The new tools that have opened the door to facile manipulation of genomic DNA sequences are enzymes that can be designed to cut DNA specifically at essentially any arbitrarily chosen sequence. When a break is made in genomic DNA, cells hasten to repair the damage using two major pathways (**Figure 1**). In one, sequences related to those at the break are used as a template for repair by a process called homologous recombination (HR). An experimenter can provide such a template that has been manipulated in the laboratory to carry desired sequence changes that will then be incorporated in and around the break. A second pathway, called nonhomologous end joining (NHEJ) often fuses the broken ends inaccurately, causing mutations at the break site. Thus, a targeted break in DNA can lead to localised mutations or engineered sequence changes. The challenge we faced 15 years ago was how to make targeted breaks in DNA both efficiently and specifically. This has now been achieved at a remarkable level.

Three flavours of targetable nucleases

The first of the targetable cleavage reagents to be developed were the zinc-finger nucleases (ZFNs)[1], which consist of a non-specific cleavage domain from a restriction endonuclease linked to a DNA-recognition domain made up of zinc fingers. The latter are DNA-binding modules found in natural transcription factor proteins. Each finger contacts primarily three base pairs of DNA, and there are natural and derived fingers that recognise many of the 64 possible triplets. It takes two ZFN monomers to make a break, each of which carries three to six fingers chosen to bind the desired genomic target (**Table 1**). The requirement for a total of 6-12 fingers, each of which binds three

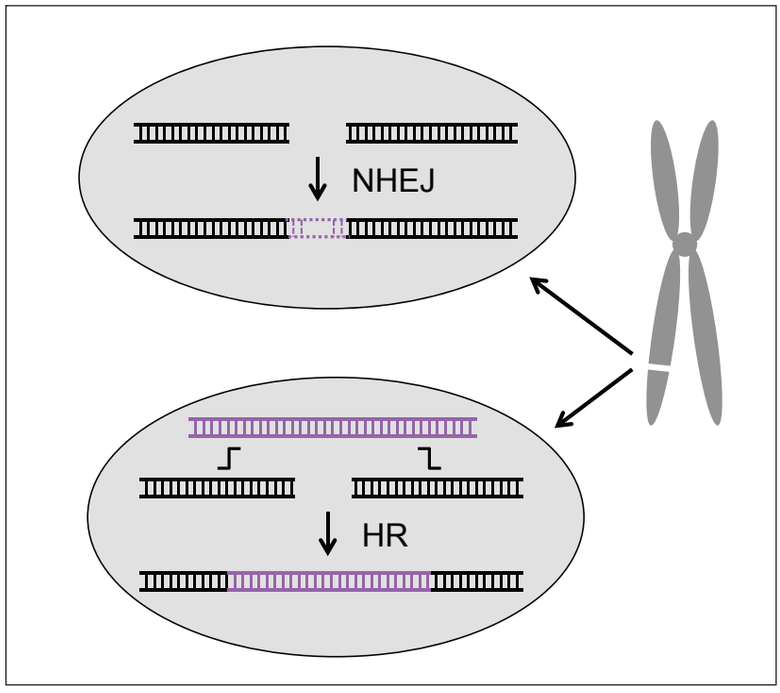

Figure 1
A break in chromosomal DNA can be repaired by homologous recombination (HR) with a related DNA sequence, or by error-prone nonhomologous end joining (NHEJ). Both types of repair allow introduction of sequence changes precisely at the break

base pairs, provides enough specificity to address a unique sequence in a genome as complex as the three billion base pairs of human DNA.

ZFNs have been used successfully to introduce sequence alterations in a wide range of organisms and cell types. They have proved somewhat difficult to engineer for new targets, however, because an individual finger may not bind its triplet well in all possible sequence contexts. A simpler alternative to zinc fingers emerged with the characterisation of transcription activator-like effectors (TALEs) in plant pathogenic bacteria of the genus Xanthomonas. DNA recognition in these proteins is provided by modules, each of which binds a single base pair, and there is a robust code to address each of the four base pairs[2,3]. Linking an array of such modules to the same cleavage domain used in ZFNs led to the development of TALE nucleases, or TALENs. Like the ZFNs, two TALEN monomers are required to achieve cleavage at a single site. The number of recognition modules in each monomer can be quite variable, but arrays of 14-20 are most common. The simplicity of design and the publication of several modular assembly schemes have led to the rapid adoption of TALENs for genome engineering.

Quite recently a third style of targetable nuclease has emerged, based on an innate immunity system in micro-organisms, called CRISPR[4]. In this scheme, recognition is mediated by a small guide RNA that forms base pairs with the target DNA. Formation of a specific RNA-DNA hybrid directs cleavage by a single protein, Cas9. To address a new target, one need only know the Watson-Crick base pairing rules, and the same protein is used for each case. Typical guide RNAs have 20 base pairs of homology to the DNA target, but some mismatches are tolerated, particularly on one end. This leaves room for recognition and cleavage of DNA sequences related, but not identical, to the desired target and, therefore, to concerns about specificity. The simplicity of the CRISPR-Cas approach has made it wildly popular in a very short time. In addition, the reliance on a single, constant protein makes this system ripe for multiplexing in ways that are not apparent for ZFNs or TALENs. Several groups have shown that introducing multiple guide RNAs along with the Cas9 protein can result in simultaneous modification of multiple targets.

ZFNs, TALENs and the CRISPR-Cas reagents can be prepared and tested in individual laboratories – the CRISPR reagents are particularly easy to engineer – but they are also commercially available. Sigma-Aldrich uses a proprietary database held by Sangamo Biosciences, Inc, to produce ZFNs (called CompoZr)[5] for client-specified targets, and these are tested extensively before release. Custom and off-the-shelf TALENs are offered by Cellectis[6], ToolGen[7], Transposagen[8] and GeneCopoeia[9]. The CRISPR-Cas materials are provided by ToolGen[7] and System Biosciences[10]. Some of these companies, and Horizon Discovery[11], will provide cells that have been edited and Sage Research Labs[12] offers engineered mice and rats. There are probably additional suppliers of which I am not aware. Quite a number of academic websites offer help with design and construction of the nucleases, and many of the necessary reagents are available for nominal prices from AddGene[13].

Applications of targetable nucleases

The targetable nucleases are relatively new on the scene, but we can already see how they can be used in a wide variety of applications. For example, the rat has long been a favoured model organism for drug testing and for physiological studies; but, compared to mice, many fewer genetic tools were available. With the advent of ZFNs, and then TALENs and CRISPR-Cas, this has changed significantly. It is now possible, using embryo injection, to produce rats with targeted knockouts or human disease mutations in just two generations[14].

An organism very popular with developmental

biologists – the zebrafish – is now being used for drug screening. It has the advantages of small size and relatively short generation time, while it shares many developmental, physiological and even some behavioural characteristics with higher vertebrates. These features are being exploited in high-throughput screening procedures[15]. All three of the targetable nuclease platforms have been used successfully in zebrafish, so it is quite feasible to produce animals that can be used to screen compounds for efficacy in a range of genetic backgrounds.

The utility of gene editing tools has been recognised by the agricultural industry as well. DowAgrosciences adopted the ZFNs as tools for crop improvement in an initiative they call EXZACT Precision Technology[16]. While no commercial products have yet emerged, the efficacy of ZFNs, TALENs and CRISPR-Cas has been demonstrated by many different investigators in a range of species, including maize, rice, soybean, wheat, barley and sorghum. One significant advantage of nuclease-mediated gene editing is that beneficial traits can be introduced without activating some of the fears associated with GMOs (genetically modified organisms). Targeted gene modifications allow the creation of simple gene knockouts or other endogenous sequence alterations without inserting any genetic material from another organism.

This same approach is being applied to farm animals – cows, pigs, sheep, goats – as a means of improving food production or livestock characteristics[17,18]. Examples include genetic dehorning of dairy cattle and increasing skeletal muscle mass (ie, meat) in beef cattle and pigs. Gene editing technology allows introduction of well-characterised mutations, again without adding foreign material. Large animals can also serve as excellent models for human disease and potentially as organ donors, and these applications are being pursued with nuclease-mediated gene editing.

Human therapy

Perhaps most exciting are applications of the gene editing technology directly to humans, and several

References

1 Carroll, D. Genome engineering with zinc-finger nucleases. Genetics 188, 773-782 (2011).
2 Bogdanove, AJ and Voytas, DF. TAL effectors: customizable proteins for DNA targeting. Science 333, 1843-1846 (2011).
3 Joung, JK and Sander, JD. TALENs: a widely applicable technology for targeted genome editing. Nature Reviews Mol. Cell Biol. 14, 49-55 (2013).
4 Pennisi, E. The CRISPR craze. Science 341, 833-836 (2013).
5 http://www.sigmaaldrich.com.
6 http://www.cellectis-bioresearch.com/talen-solutions.
7 http://www.toolgen.com.
8 http://transposagenbio.com.
9 http://www.genecopoeia.com.
10 http://www.systembio.com.
11 http://www.horizon discovery.com.
12 http://www.sageresearch labs.com.
13 http://www.addgene.org.
14 Jacob, HJ et al. Gene targeting in the rat: advances and opportunities. Trends Genet. 26, 510-518 (2010).
15 Peterson, RT and Fishman, MC. Designing zebrafish chemical screens. Methods Cell Biol. 105, 525-541 (2011).
16 http://www.exzactprecision technology.com.
17 Tan, WS et al. Precision editing of large animal genomes. Adv. Genet. 80, 37-97 (2012).
18 http://www.recombinetics.com.
19 Urnov, FD et al. Genome editing with engineered zinc finber nucleases. Nature Reviews Genet. 11, 636-646 (2010).
20 Li, L et al. Genomic editing of the HIV-1 coreceptor CCR5 in adult hematopoietic stem cells and progenitor cells using zinc finger nucleases. Molecular Therapy 21, 1259-1269 (2013).
21 Reyon, D et al. FLASH assembly of TALENs for high-throughput genome editing. Nature Biotechnol. 30, 460-465 (2012).

Continued on page 52

Table 1: Characteristics of the targetable nucleases

NUCLEASE PLATFORM	CLEAVAGE ELEMENT	DNA RECOGNITION	STOICHIOMETRY
ZFN	FokI	Zinc fingers	1 finger/3bp
TALEN	FokI	TALE modules	1 module/1bp
CRISPR-Cas	Cas9	RNA-DNA base pairing	1 base/1bp

The cleavage element of ZFNs and TALENs is the nuclease domain of the restriction endonuclease, FokI, which must form dimers to cut DNA. The Cas9 protein in the CRISPR system is monomeric and has two separate nuclease active sites. bp, DNA base pair(s)

recent examples illustrate the possibilities. A current clinical trial uses ZFNs targeted to the CCR5 gene as a treatment – and perhaps prevention – for AIDS[19,20]. The CCR5 gene product is the most common co-receptor for HIV-1. Natural human mutants who lack CCR5 are resistant to the development of AIDS, even after exposure to the virus. One lucky patient, who received a bone marrow transplant from a CCR5 mutant donor, was cured of both leukaemia and AIDS. In the current clinical trial, T cells derived from patients are treated *ex vivo* with ZFNs targeted to the CCR5 gene. The treated cells, many of which now have CCR5 inactivated by NHEJ, are given back to the same patient. While results of this trial have not yet been published, early reports are encouraging.

Other applications involve the use of embryonic stem (ES) cells or induced pluripotent stem (iPS) cells. These cells can be manipulated in culture, then induced to differentiate along a multitude of pathways. This allows exploration of disease-related defects at the cellular level. In addition, skin cells from a patient can be turned into iPS cells and gene-corrected in culture, then potentially provide material for tissue repair and regeneration. Several groups have shown that targetable nucleases are effective in stimulating gene editing in stem cells.

Challenges and concerns

When applying ZFNs, TALENs and CRISPR-Cas nucleases to different organisms and situations, one always has to be aware of potential problems with delivery of the reagents. Delivery to cultured cells is usually straightforward, and even primary cells have been edited effectively. Getting the reagents into whole organisms can be quite challenging, however. An effective approach with many experimental animals has been direct embryo injection of messenger RNAs for the nuclease proteins, along with long or short donor DNAs. This has been used with many types of mammals, including livestock. Such an approach in humans would, of course, result in germ line modifications that raise serious ethical issues.

All of these reagents have very high specificity for their intended targets, but even with careful design, they are not perfect. There is some concern that secondary sites within the genome may be mutated while an intentional target is being attacked. Modifications have been made in the design and use of the nucleases that have greatly improved discrimination against related sites. In the case of applications to humans, however, little or no off-target mutagenesis will be tolerated, and only experience will tell whether additional safety measures will be required.

The efficiency of *in vivo* DNA cleavage by the targetable nucleases can be truly remarkable, approaching 100%, but it can also be quite variable. Some of this has to do with the design of each of the reagents. As noted above, zinc fingers can be somewhat temperamental in their recognition characteristics, and there are not good fingers for all possible DNA triplets. TALENs are generally better behaved, and the success rate in attacking new targets is impressive[21]. Nonetheless, some TALEN designs fail for unknown reasons. The CRISPR-Cas nucleases are also showing excellent success rates, with variability that is not obviously due to differences in design. A question that has not been addressed adequately is whether some genomic targets are more accessible than others due to inherent differences in chromatin structure or other parameters.

Continued from page 50

22 Kim, YG, Cha, J and Chandrasegarn, S. Hybrid restriction enzymes: zinc finger fusions to FokI cleavage domain. Proc. Natl. Acad. Sci. USA 93, 1156-1160 (1996).
23 Wiedenheft, B, Sternberg, S and Doudna, JA. RNA-guided genetic silencing systems in bacteria and archaea. Nature 482, 331-338 (2012).

Looking ahead

It is worth noting that all three nuclease platforms were developed from unexpected sources. With ZFNs it was the characterisation of natural transcription factors and of a bacterial restriction enzyme that led to construction of the useful hybrid proteins[22]. The DNA-recognition domain of TALENs was discovered in bacteria that cause disease in crop plants[2]. The CRISPR system is native to many bacterial and archaeal species, and the mechanism it uses to provide protection against viral infection was only recently elucidated[23]. The application of these reagents to gene editing built on earlier studies that characterised pathways of DNA break repair. This sort of convergence of apparently unrelated research threads is common to the majority of scientific advances. The lesson is that only very broadly based research into the natural world provides the raw materials for new applications.

Research with the targetable nucleases is still accelerating. New applications are being identified and pursued at an astounding rate. As noted above, these include manipulations of crop plants for improved nutritional and agricultural characteristics; introduction of favourable mutations into food animals; establishment of model organisms and cell cultures with novel genetic characteristics; and, of course, applications to human therapy. Exciting things lie ahead in the future of this technology. **DDW**

Dr Dana Carroll *is Professor and former Chair of the Department of Biochemistry at the University of Utah School of Medicine, Salt Lake City, UT, USA. He has a PhD in Chemistry from the University of California, Berkeley, and did postdoctoral research at the Beatson Institute for Cancer Research in Glasgow, Scotland, and at the Carnegie Institution of Washington Department of Embryology in Baltimore, MD. He pioneered the use of zinc-finger nucleases as tools for genome engineering, and has more recently worked with TALENs and the CRISPR-Cas reagents as well. Using the fruit fly,* Drosophila melanogaster, *as a model organism, the Carroll lab established many of the parameters that influence ZFN efficacy. He has collaborated with others in applying the targetable nucleases to several different organisms, including nematodes, plants, and silkworm. Technology developed in his lab for targeted mutagenesis with ZFNs has been licensed to Sangamo Biosciences, Inc.*