Genome editing with engineered zinc finger nucleases

Fyodor D. Urnov, Edward J. Rebar, Michael C. Holmes, H. Steve Zhang and Philip D. Gregory

Abstract | Reverse genetics in model organisms such as *Drosophila melanogaster*, *Arabidopsis thaliana*, zebrafish and rats, efficient genome engineering in human embryonic stem and induced pluripotent stem cells, targeted integration in crop plants, and HIV resistance in immune cells — this broad range of outcomes has resulted from the application of the same core technology: targeted genome cleavage by engineered, sequence-specific zinc finger nucleases followed by gene modification during subsequent repair. Such 'genome editing' is now established in human cells and a number of model organisms, thus opening the door to a range of new experimental and therapeutic possibilities.

Species with experimentally manipulable genomes dominate our ability to investigate the role of genes in biology and disease. In organisms such as yeast and mice, the ability to specifically add or delete genetic information enables an unmatched level of precision in studies of gene function, and consequently more is known about biological mechanisms in these two species than in any other within their respective taxonomic group. A technique known as 'genome editing'[1] (BOX 1), which was initially applied to *Drosophila melanogaster*[2,3], promises to extend this capability to cells and entire organisms from potentially any species. This approach enables efficient and precise genetic modification via the induction of a double-strand break (DSB) in a specific genomic target sequence, followed by the generation of desired modifications during subsequent DNA break repair. The DSB is induced by a 'zinc finger nuclease' (ZFN)[4,5], which is a designed, sequence-specific endonuclease that can be customized to cleave a user-chosen DNA target. Since the most recent comprehensive review of the subject[1], this approach has been used to disrupt native loci in model organisms such as rats and *Arabidopsis thaliana*, to drive trait stacking in a crop species, to engineer HIV-resistant human T cells and haematopoietic stem cells (HSCs), and to drive targeted integration in human embryonic stem (ES) cells, induced pluripotent stem (iPS) cells and mesenchymal stem cells.

Here we review the development of site-specific ZFNs and the application of these reagents to genome editing. We describe methods for engineering ZFNs and discuss strategies for using these proteins to introduce targeted modifications into endogenous loci. These modifications include gene disruption (the targeted induction of minor insertions and deletions), 'gene correction' (the introduction of discrete base substitutions specified by a homologous donor DNA construct) and targeted gene addition (the transfer of entire transgenes into a native genomic locus). Finally, we explore therapeutic applications for this technology.

Zinc finger nucleases

To be useful for genome engineering, an endonuclease must exhibit an extraordinary combination of qualities: specific recognition of long target sequences (ideally, long enough for unique occurrence in a eukaryotic genome) coupled with sufficient adaptability for retargeting to user-defined sequences. The ZFN architecture (FIG. 1a) meets these specifications by linking the DNA-binding domain of a versatile class of eukaryotic transcription factors — zinc finger proteins (ZFPs) — with the nuclease domain of the *Fok*I restriction enzyme. By virtue of their structure, ZFNs combine the favourable qualities of both components — the DNA binding specificity and flexibility of ZFPs and a cleavage activity that is robust but restrained in the absence of a specific binding event — while retaining functional modularity. As a consequence, both the DNA-binding and catalytic domains can be optimized in isolation, which simplifies retargeting and platform improvement efforts.

Sangamo BioSciences Inc., 501 Canal Boulevard, Suite A100, Richmond, California 94804, USA. Correspondence to P.D.G. e-mail: pgregory@sangamo.com
doi:10.1038/nrg2842

Box 1 | **Gene targeting and genome editing**

The work of Mario Capecchi, Oliver Smithies and colleagues on 'gene targeting' has enabled the disruption of a gene via positive–negative selection using two selectable markers and the transfer of additional genes adjacent to one of the markers to an investigator-specified locus. The title of the original paper[43] indicates what the term means: "Targeting of genes to specific sites in the mammalian genome". The targeting in this method is driven by the inclusion of extended homologous sequences flanking the gene of interest in the donor plasmid and selecting for rare homologous recombination events. Double-strand-break-assisted genome editing is often not 'gene targeting' in this sense: gene disruption mediated by non-homologous end joining (NHEJ) requires no donor or drug selection, and the resulting allele carries no promoter-transcription unit encoding a drug resistance marker, so this is not the targeting of a gene to a site. In the case of genome editing using donor constructs, drug selection is often also not required, and the mutations created can be as small as a single-base-pair change (a process referred to as 'gene correction'). In situations in which zinc finger nucleases are used to drive targeted integration, selection is also often not required, prompting the use of the term 'gene addition'. It seems timely to consider the adoption of terminology that more accurately reflects what is being done. In addition, the impact on biomedical research of classical 'gene targeting' as developed by Capecchi and colleagues deserves recognition through the preservation of the original meaning of the term and its use to describe the process that was originally developed and that is still widely used.

The zinc finger — a platform for the design of novel DNA binding domains. The ZFP region provides a ZFN with the ability to bind a discrete base sequence. This region contains a tandem array of Cys2-His2 fingers[6,7], each recognizing approximately 3 bp of DNA (FIG. 1a). In early studies individual ZFNs used three fingers to bind a 9-bp target, which enabled ZFN dimers (the active species) to specify 18 bp of DNA per cleavage site. More recent studies have added more fingers (up to six per ZFN) to specify longer and rarer cleavage targets.

A variety of strategies have been described for making ZFPs with new, user-chosen binding specificities. The first emerged from observations of the initial ZFP–DNA co-crystal structure, which suggested a substantial degree of functional autonomy in the interaction of individual fingers with DNA[8]. The approach, which has been termed 'modular assembly'[9], generates candidate ZFPs for a given target sequence by identifying fingers for each component triplet and linking them into a multifinger peptide targeted to the corresponding composite sequence (FIG. 1b). Fingers used for modular assembly have been developed for most triplet sequences[10–16]. The method has been used to develop the zinc finger component of active ZFNs for a number of endogenous targets in higher eukaryotic cells[17] (see the entries labelled 'modular assembly' in TABLE 1).

Besides modular assembly, several alternative strategies for making ZFPs have been developed. These newer methods were designed to accommodate the deviations from strict functional modularity observed for many zinc fingers. Many designed and natural fingers, for example, can contact neighbouring fingers as well as bases outside their proximal DNA triplet[18–23]. Although such interactions can enable more selective binding, they can also complicate efforts to produce new ZFPs through modular design[24]. BOX 2 summarizes these alternative strategies for making ZFPs with new sequence specificities.

Whatever the design method, the production of a DNA binding module evaluated *in vitro* for affinity and specificity towards its intended target provides only the first step towards use *in vivo*. Indeed ZFNs assembled from *in vitro* 'validated' ZFPs often fail to drive genome editing at the endogenous locus when tested in living cells. One factor is specificity: complex genomes often contain naturally occurring multiple copies of a sequence that is identical or highly related to the intended target (for example, paralogues or pseudogenes), and these copies can act as additional targets for ZFNs. Researchers have addressed this issue by building up a detailed understanding of the rules governing protein–DNA interactions and by exploiting minor sequence divergences between related genomic regions (see below). An additional problem is the chromatin structure at target sites, which may not be amenable to cleavage. This hurdle can be overcome by tiling the target region with a large ZFN panel, followed by screening for activity directly at the endogenous locus and — when necessary — by iteratively optimizing the protein–DNA interface (see below).

The* FokI *catalytic domain. The *Fok*I domain has been crucial to the success of ZFNs, as it possesses several characteristics that support the goal of targeted cleavage within complex genomes. For example, *Fok*I must dimerize to cleave DNA[25]. As this interaction is weak[25], cleavage by *Fok*I as part of a ZFN requires two adjacent and independent binding events, which must occur in both the correct orientation and with appropriate spacing to permit dimer formation (FIG. 1a). The requirement for two DNA binding events enables specific targeting of long and potentially unique recognition sites (from 18–36 bp). Moreover, the dependence on productive dimerization has spurred the development of variants that cleave only as a heterodimer pair, thus improving specificity via the elimination of unwanted homodimers[26,27] (FIG. 2). In other studies and in ongoing efforts, the ZFP–*Fok* linker has been modified as means of developing ZFN dimers with novel spacing requirements for the two monomer binding events[28,29]. *Fok*I catalytic domain variants with enhanced cleavage activities have also been reported[30].

Resolving a DNA double-strand break

The applications of genome editing using ZFNs are based on the introduction of a site-specific DNA DSB into the locus of interest. All eukaryotic cells efficiently repair DSBs via the homology-directed repair (HDR) or non-homologous end joining (NHEJ) pathways[31–33]. These highly conserved pathways can be harnessed to generate defined genetic outcomes across a wide range of cell types and species (FIG. 3). NHEJ repair, for example, rapidly and efficiently ligates the two broken ends, with the occasional gain or loss of genetic information; it can therefore be used to introduce small insertions and/or deletions at the site of the break, an outcome that can be exploited to disrupt a target gene. Alternatively, if an investigator-designed homologous donor DNA is provided in combination with the ZFNs, information encoded on this template can be used to repair the DSB, thus resulting in gene correction (a few nucleotides

Paralogues
Two or more genes within a given species that originated from a single parent gene via duplication events, usually with a subsequent, sometimes subtle, divergence of function.

Pseudogenes
Non-functional paralogues that often lack promoter or intron sequences.

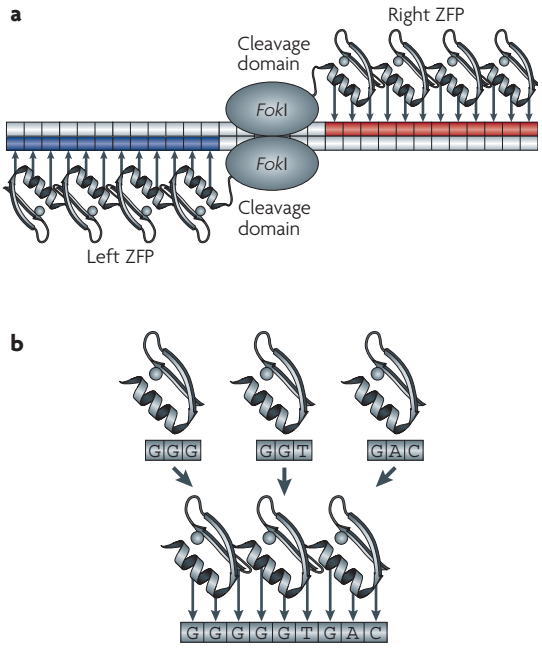

Figure 1 | **Structure and design of zinc finger nucleases.** **a** | Sketch of a zinc finger nuclease (ZFN) dimer bound to its target. Each ZFN contains the cleavage domain of *FokI* linked to an array of three to six zinc fingers (four are shown here) that have been designed to specifically recognize sequences (blue and red boxes) that flank the cleavage site. A small number of bases (typically five or six) separate the ZFN targets. **b** | Modular assembly of a three-finger protein from individual fingers. To generate a zinc finger protein (ZFP) with specificity for the sequence GGGGGTGAC, three fingers are identified that each bind a component triplet. These fingers are then linked. Part **a** is adapted, with permission, from REF. 26 © (2007) Macmillan Publishers Ltd. All rights reserved. Part **b** is adapted, with permission, from REF. 91 © (2004) PanGlobal Media.

changed at the endogenous site) or the addition of a new gene at the site of the break (FIG. 3; most available evidence argues for a synthesis-dependent strand annealing process in homology-based repair of breaks, which for simplicity we refer to as HDR). Below we discuss each type of genome editing using ZFNs and how they have been applied for experimental, biotechnological and therapeutic purposes (TABLE 1).

Gene disruption

The simplest means of genome editing, gene disruption, takes advantage of errors introduced during DNA repair to disrupt or abolish the function of a gene or genomic region (FIG. 2). This approach has been applied in various species and cellular contexts to knock out user-specified genes in a single step and without selection for the desired event. Examples of genome editing using engineered ZFNs are provided below.

Gene disruption in model organisms. To disrupt a gene in *D. melanogaster*, ZFNs targeting exonic sequences can be delivered via mRNA injection into the early fly embryo; up to 10% of the progeny produced by the resulting adult flies is mutated for the gene of interest[34]. For one gene (*coilin*), six different alleles were obtained, all carrying small frameshift-inducing deletions. Animals homozygous for each allele completely lacked expression of the protein encoded by the target gene. Injection of mRNA encoding engineered ZFNs into embryos has also been used to generate zebrafish carrying desired genetic lesions. In four separate studies up to 50% germline mosaicism at the targeted genes was obtained[35–37,93].

For gene disruption in rats, in which early development proceeds much slower than in insects or in fish, engineered ZFNs with extended recognition sites were used. This produced knockout animals for two separate endogenous genes, and ZFN-treated founders transmitted disrupted alleles at a frequency of 10–100%[38]. A rat model of severe combined immune deficiency (SCID) was also generated[39]. In systems in which mRNA microinjection is at present not an option (for example, the model plant *A. thaliana*), stable transgenesis of an inducible ZFN expression cassette allows gene disruption[40–42].

As the path to a knockout organism is now one generation long, ZFN-driven gene disruption, even when allowing for a period of ZFN development, compares favourably in terms of duration and screening effort with other strategies for generating targeted knockouts (such as classical gene targeting in mouse ES cells).

Gene disruption in mammalian somatic cells. ZFN-driven gene disruption has also been used for mammalian somatic cell genetics, in which the ZFN is transiently expressed followed by analysis of single-cell-derived clones.

Classical gene targeting combined with positive and negative selection strategies is a powerful tool for gene knockout in mouse ES cells[43], and the use of engineered adeno-associated viruses (AAVs)[44] has allowed its application in transformed and primary human cells[45,46]. ZFNs obviate the need for drug selection, extend the application of gene knockout to potentially any cell type and species for which transient DNA or mRNA delivery is available, and result in knockouts in 1–50% of all cells. Recently, transient hypothermia has been shown to further increase ZFN-driven gene disruption frequency in transformed and primary cells by two- to fivefold[47].

The first published example of the use of engineered ZFNs to disrupt an endogenous locus[48] in a mammalian cell involved a knockout of the dihydrofolate reductase (*Dhfr*) gene in Chinese hamster ovary (CHO) cells. A plasmid encoding the ZFNs was introduced by transient transfection, which resulted in disruption frequencies of up to 15% of alleles in the cell population. Limiting dilution and genotyping yielded two clones (out of ~60) in which the *Dhfr* gene was biallelically disrupted and which lacked measurable DHFR protein expression. Subsequently, the sequential or simultaneous application of locus-specific ZFNs has been used to efficiently make double[49] and triple[50] locus gene knockouts in CHO and K562 cells. As we describe below, ZFN-driven gene knockout technology has

Adeno-associated viruses (AAVs). As defined by Flint and colleagues in *Principles of Virology*, AAV is a parvovirus (also known as a dependovirus) that can establish a latent infection during which its DNA is integrated into the host cell genome in an inactive state. AAV2 integrates into the *PPP1R12C* gene locus on chromosome 19.

Table 1 | **Endogenous genes modified by zinc finger nucleases**

Organism	Gene	ZFN development method*	Refs
Gene disruption			
Fruitflies	*yellow*	Modular assembly	2
	rosy, *brown*	Modular assembly	60
CHO cells	*Dhfr*	Two-finger modules	48
	Dhfr, *Glul*	Two-finger modules	50
	Fut8	Two-finger modules	92
	Bax, *Bak1*	Two-finger modules	49
Zebrafish	*kdr*	Bacterial one-hybrid	36
	golden, *no tail*	Two-finger modules	35
	tfr2, *dat*, *telomerase*, *hif1aa*, *gridlock* (also known as *hey2*)	OPEN	37
	cxcr4a	Modular assembly	93
Human T cells	*CCR5*	Two-finger modules‡	76
Hek293 cells	*CCR5*	Modular assembly	17
Rats	*Rab38*, *IgM*	Two-finger modules	38
	Il2rg	Two-finger modules	39
SupT1 cells	*CXCR4*	Two-finger modules	94
K562 cells, HeLa cells	*PPP1R12C* (the AAVS1 locus), *TP73*, *MAP3K14*, *EP300*, *BTK*, *CARM1*, *GNAI2*, *TSC2*, *RIPK1*, *KDR*, *NR3C1*	Two-finger modules	47
Gene correction			
Fruitflies	*yellow*	Modular assembly	3
	rosy	Modular assembly	60
	coilin, *pask*	Modular assembly	34
K562 cells, human T cells	*IL2RG*	Two-finger modules‡	61
K562 cells	*IL2RG*, *VEGF*, *HOXB13*, *CFTR*	OPEN	62
Tobacco	*SuRA*, *SuRB* (acetolactate synthase genes)	OPEN	63
Arabidopsis thaliana	*ABI4*, *KU80*	Modular assembly	42
	ADH1, *TT4*	OPEN	41
Mouse ES cells	*H3f3b*	Two-finger modules	67
Gene addition			
K562 cells	*IL2RG*	Two-finger modules‡	66
Human ES cells	*IL2RG*, *CCR5*	Two-finger modules‡	68
	PIGA	OPEN	70
	OCT4 (also known as *POU5F1*), *PPP1R12C* (AAVS1 locus), *PITX3*	Two-finger modules	71
Tobacco	Chitinase	Two-finger modules	74
Maize	*lpk1*, *Zein protein 15*	Two-finger modules‡	75
Human tissue culture cells	*PPP1R12C* (AAVS1 locus)	Two-finger modules‡	72
Mouse ES cells	*H3f3b*	Two-finger modules	67

*For details see BOX 2 and the section: 'The zinc finger — a platform for the design of novel DNA binding domains'.
‡The initial active leads assembled from archived modules were further optimized to yield the ZFNs reported in the study.
AAVS1, adeno-associated virus integration site 1; *ABI4*, *ABA-INSENSITIVE 4*; *ADH1*, *ALCOHOL DEHYDROGENASE 1*; *Bak1*, BCL2-antagonist/killer 1; *Bax*, BCL2-associated X protein; *BTK*, Bruton agammaglobulinaemia tyrosine kinase; *CARM1*, coactivator-associated arginine methyltransferase 1; CCR, chemokine (C-C motif) receptor; *CFTR*, cystic fibrosis transmembrane conductance regulator; CHO, Chinese hamster ovary; *dat*, *dopamine transporter* (also known as *slc6a3*); CXCR, chemokine (C-X-C motif) receptor; *Dhfr*, dihydrofolate reductase; *EP300*, E1A binding protein p300; ES, embryonic stem; *Fut8*, fucosyltransferase 8; *Glul*, glutamine synthetase; *GNAI2*, guanine nucleotide binding protein (G protein), alpha inhibiting activity polypeptide 2; *H3f3b*, H3 histone, family 3B; *hif1aa*, *hypoxia-inducible factor 1α*; *HOXB13*, homeobox B13; *IgM*, immunoglobulin M (also known as *Igh6*); *Il2rg*, interleukin-2 receptor-γ; *Ipk1*, *inositol-1,3,4,5,6-pentakisphosphate 2-kinase*; *kdr*, kinase insert domain receptor; *MAP3K14*, mitogen-activated protein kinase kinase kinase 14; *NR3C1*, nuclear receptor subfamily 3, group C, member 1; *pask*, *PAS kinase* (also known as *CG3105*); *PIGA*, phosphatidylinositol glycan anchor biosynthesis, class A; *PITX3*, paired-like homeodomain 3; *PPP1R12C*, protein phosphatase 1, regulatory (inhibitor) subunit 12C; *RIPK1*, receptor (TNFRSF)-interacting serine-threonine kinase 1; *tfr2*, *transferrin receptor 2*; *TP73*, tumour protein p73; *TSC2*, tuberous sclerosis 2; *TT4*, *TRANSPARENT TESTA 4*; *VEGF*, vascular endothelial growth factor; *vegfr2*, *vascular endothelial growth factor receptor type 2*; ZFN, zinc finger nuclease.

Box 2 | **Strategies for developing zinc finger proteins with new sequence specificities**

Several alternatives to modular assembly have been developed for identifying zinc finger proteins (ZFPs) with sufficient affinity and specificity for use in genome engineering. One approach, called the 'OPEN' system[37,62,84], uses bacterial selections to identify finger combinations that will work well together. The method involves two distinct steps. First, multiple, parallel low-stringency selections are performed for binding of randomized fingers to each triplet in the targeted sequence. Mild conditions ensure that the resultant pools retain considerable diversity. Next, fingers from these pools are combinatorially linked and the products are selected at high stringency for binding to the final target. Studies that have used this method are identified by 'OPEN' in TABLE 1.

A second approach for identifying ZFPs with new specificities uses a bacterial selection system that is similar to OPEN but a different strategy for library construction[36]. For each target triplet, a library is assembled that randomizes only a subset of residues at the zinc finger–DNA interface. At the remaining positions, specificity is achieved by the use of residues chosen for their ability to make especially well-understood base contacts. Studies using this method are identified by 'bacterial one-hybrid' in TABLE 1 to reflect the use of this selection system for this method[36].

An alternative path for accommodating finger–finger cooperativity, as well as addressing other limitations, has been to develop more sophisticated versions of the modular assembly strategy. For example, at least one published protocol has incorporated explicit checking for a particularly energetic and well-understood extra-triplet contact during evaluation of prospective designs[85]. Other studies have developed finger design weightings[86] or have sought improved success rates through the use of only a subset of available finger designs, which are chosen for their consistency of function[17].

Another strategy is to use two-finger modules (instead of individual fingers) as the principle unit of DNA recognition[87]. This approach enables optimization of finger junctions within each module for more cooperative and specific base recognition. Moreover, it reduces the number of untested finger–finger junctions in any new ZFP design and therefore the risk of a poor interaction between newly joined fingers. A four-finger ZFP, for example, will contain just one new junction instead of three if assembled from one-finger units. This approach has been used to make zinc finger nucleases (ZFNs) consisting of four, five or six zinc fingers for a range of applications (note that a five-finger ZFN is constructed using one one-finger and two two-finger units). Studies that have used this method are identified by 'two-finger modules' in TABLE 1. Although this approach benefits from both context-dependent DNA recognition and the speed of modular assembly, a limitation is the scale of the initial investment required to develop and characterize the large panel of two-finger units (up to 4,096 for recognition of all 6-bp sequences).

Each of these methods has been successfully used to generate endogenously active ZFNs (TABLE 1). However, they differ substantially from each other in terms of time and cost for reagent development, success rate and 'design density' (the ability to deliver a ZFN dimer that cleaves exactly at, or close to, the ideal location for a given application). A thorough analysis of these considerations is beyond the scope of this Review. Also, several other ZFP development methods have been described but not yet applied for retargeting ZFNs[88–90]. It is likely that many if not all could yield ZFPs of sufficient activity for use in genome modification. However, space constraints prevent a detailed review of these methods.

also proven effective in a range of primary cell types, including purified $CD4^+$ T cells and human ES cells.

More complex types of disruption. The range of mutations that can be created via NHEJ need not be limited to small insertions or deletions at the ZFN target site (FIG. 3). For example, ZFN cleavage at a proximal pair of sites at the same locus deletes the intervening segment[50–53], and other studies have shown that more distal pairs of targets sequences may be cleaved and joined, albeit with reduced efficiency[51,54]. Finally, in both hamster and human cells, supplying the cells with a double-stranded oligonucleotide carrying overhangs complementary to those that the ZFNs generate in the endogenous locus yields chromatids in which the oligonucleotide has precisely ligated into the chromosome[52]; this approach can be used to, for example, add tags to endogenous genes in cells in which HDR is less efficient (see below) or to replace an entire chromosomal stretch with a recombinase recognition site.

With some notable exceptions[45,55], mammalian somatic cell 'genetics' has required quotation marks, as it relied on RNAi — a process that offers an operationally useful transient knockdown of target gene expression but not a *bona fide* genetic modification. As larger numbers of functionally validated gene-specific ZFNs become available, a gene knockout will become the standard for determining gene function.

Homology-based genome editing

The second, mechanistically more complex pathway that can be invoked following a ZFN-induced DSB is called HDR (FIG. 3). Whether spontaneous or induced by the I-*Sce*I homing endonuclease[56,57] or a ZFN[3,58], a DSB is recombinogenic in cells of higher eukaryotes. Homology-based genome editing requires the simultaneous provision of a suitably designed, homology-containing donor DNA molecule along with the locus-specific ZFNs. This enables two related modes of genome editing that are a function of donor DNA design and distinguished by the type of allele being generated.

Gene correction (allele editing). This approach allows the transfer of single-nucleotide changes and short heterologous stretches from an episomal donor to the chromosome following a ZFN-induced DSB; recent experiments support the notion that the endogenous

Homing endonuclease
A 'meganuclease', such as I-*Sce*I, that recognizes and cuts longer sequences (18 bp in the case of I-*Sce*I) than those cut by commonly used restriction enzymes.

Episomal
The formal, technical definition of 'episome', and its distinction from the word 'plasmid' (as proposed by Jacob and Wollman in 1958), is complex. In vernacular use it means 'circular extrachromosomal DNA molecule present inside the nucleus'.

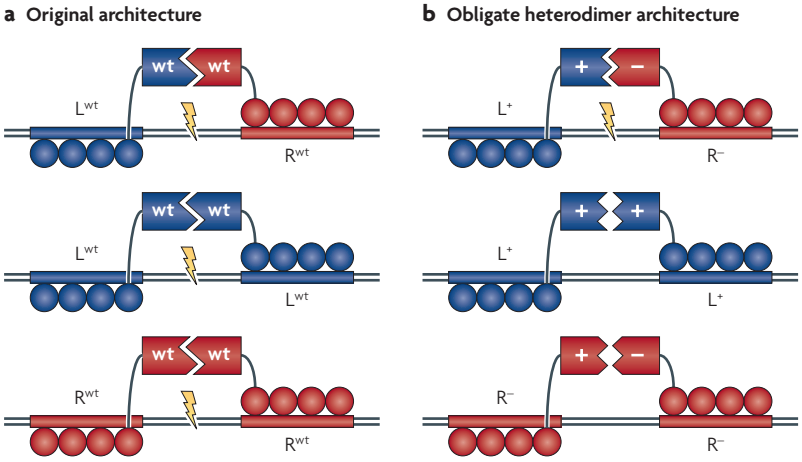

Figure 2 | **Eliminating potential homodimerization.** Sketch of a zinc finger nuclease (ZFN) heterodimer bound to its intended target (top), and two corresponding ZFN homodimers binding to alternative target sites in an architecture that permits (**a**) or prevents (**b**) heterodimerization. If the ZFNs carry *Fok*I domains engineered to function solely as heterodimers (indicated as shapes labelled with plus and minus signs in **b**), binding to the intended target (top) will lead to DNA cleavage (indicated by the lightning symbol), but homodimerization (induced in the example shown by the proximity of two R^- or two L^+ binding sites) will be impeded by the inability of the *Fok*I domains they carry to form a productive dimer. wt, wild type. Adapted, with permission, from REF. 26 © (2007) Macmillan Publishers Ltd. All rights reserved.

repair machinery uses the extrachromosomal, investigator-provided donor as a template for repairing the DSB via the synthesis-dependent strand annealing process[59]. This technique enables the study of gene function and/or the modelling of disease-causing mutations through the creation of a point mutation that is characteristic, for example, of a known disease-predisposing allele or that disables a motif that is thought to be crucial for function. Such point mutations can be efficiently created at a specific position in the target gene. Carroll and colleagues exploited this method at three different genes in *D. melanogaster*[3,60], and up to 90% of the ZFN-treated animals yielded offspring that carried donor-specified alleles of the target gene.

The correction and *de novo* creation of point mutations at a native locus in human cells was first accomplished by designing ZFNs that recognize a mutational hot spot in the interleukin-2 receptor-γ (*IL2RG*) gene[61]. Use of these ZFNs resulted in up to 20% of the endogenous *IL2RG* alleles in the treated pool of K562 cells acquiring a novel, donor-specified restriction fragment length polymorphism (RFLP) in the absence of any selection for the desired event. Of note, 8% of the single-cell-derived clones obtained from this pool showed biallelic modification. This high efficiency allowed the construction of an isogenic panel of cells carrying three different allelic forms of the native gene at the endogenous locus. Comparable gene correction frequency was observed at this locus in human $CD4^+$ T cells[61], and the robustness of this approach is demonstrated by its successful application to three additional human genes in transformed cells[62].

Gene correction using ZFNs has also been achieved in plants, which provides new avenues for agricultural biotechnology. ZFNs were used to specifically edit two paralogous tobacco genes[63]; the ZFN-expressing plasmids were co-delivered to protoplasts along with a linear donor molecule encoding a point mutation that corrects the endogenous gene to a herbicide-resistant form. Correction occurred in 75–96% of all the herbicide-resistant calli. Thus, in organisms as evolutionarily distant as fruitflies, humans and tobacco, engineered ZFNs can create a DSB that facilitates the transfer of defined, small genetic changes from an investigator-provided donor to a native chromosomal locus.

Gene addition in mammalian cells. The same approach allows the transfer of gene-sized heterologous DNA sequences from an episomal or linear extrachromosomal donor to the genome following a ZFN-induced DSB. Early *in vitro* work using a phage DNA polymerase[64] and studies of P-element excision in flies[65] provided the starting point for further efforts to determine whether larger stretches of heterology could be copied to the chromosome following a DSB. This has been demonstrated using ZFNs directed against *IL2RG* in combination with donors carrying homology arms of 750 bp that flank transgenes positioned precisely between the ZFN recognition sites[66]. In this study, ~5% of chromatids acquired transgenes of up to 8 kb in length in the absence of selection for the desired event[66].

ZFN-driven gene addition can now be applied at other loci and mammalian cell types. For example, ZFNs were used to generate an isogenic panel of mouse ES cells carrying a defined series of alleles for an endogenous gene[67]. As demonstrated in this study, the ability to both tag an endogenous gene and make a novel allelic form provides a way to study structure–function relationships at native levels of expression and in isogenic settings.

Gene addition in human ES and iPS cells

The remarkable biology of human stem cells has fuelled interest in their application to modelling of human tissues (diseased and normal) in basic science, drug discovery and regenerative medicine. However, the plasticity of stem cell fate contrasts with their recalcitrance to genetic engineering. Facile methods for the introduction of disease-linked alleles, addition of lineage markers or inducible suicide genes, and/or the correction of genetic defects for cell-based therapy have been lacking.

The initial demonstration of ZFN-driven targeted gene addition to an endogenous locus in human ES cells used an integration-defective lentiviral vector to deliver both the ZFN expression cassette and the donor construct[68]. Gene addition was observed at rates as high as 6% in the absence of selection and resulted in stable gene expression for at least 2 months, both in cultures that retain 'stemness' and following neuronal differentiation. The application of this approach in human mesenchymal stem cells yielded 50% targeted gene addition without selection[69]. Furthermore, ZFNs delivered as plasmid DNA have been used in human ES and iPS cells to efficiently target a drug resistance marker to a specific gene[70]

Protoplasts
Plant cells that lack a cell wall.

Calli
(Sing. callus.) Clusters of undifferentiated plant cells grown on solid medium that, in some species and under specific culture conditions, have the capacity to regenerate a whole plant.

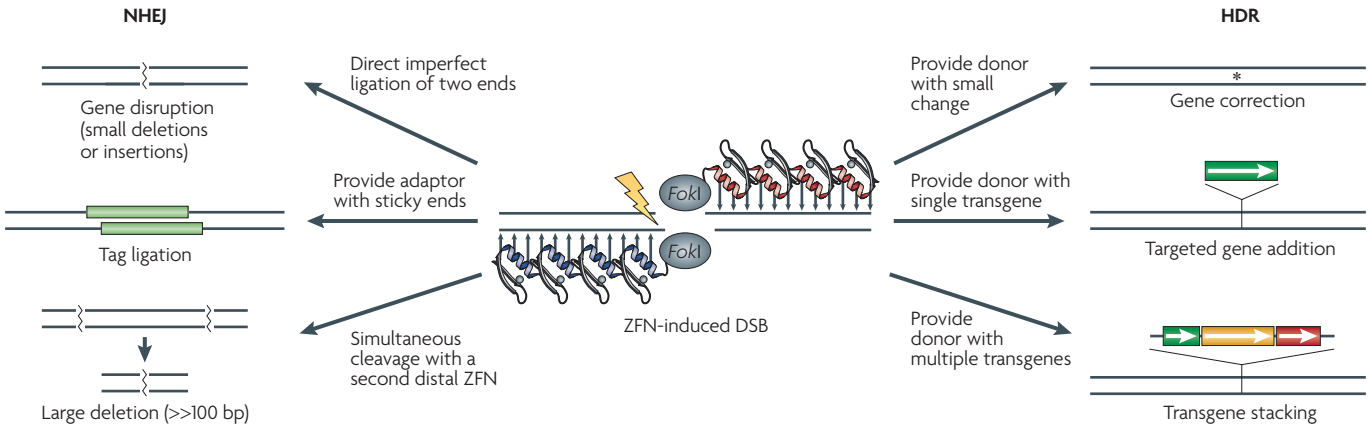

Figure 3 | **Types of genome editing made possible using zinc finger nucleases.** A zinc finger nuclease (ZFN)-induced double-strand break (DSB) allows a range of alleles to be generated at endogenous loci, as specified by the investigator. The diagram shows the different outcomes that can result from the introduction of a site-specific DNA break. A ZFN pair is shown bound to a genomic target site (the two different DNA binding domains are shown in red and blue). The DSB generated by ZFN cleavage induces DNA repair processes that may be influenced by the addition of an investigator-designed donor DNA. As shown on the left, if the break is resolved via non-homologous end joining (NHEJ) (which will occur in the absence of donor DNA), this can lead to the following outcomes (from top to bottom): gene disruption — the two ends can be ligated back together, frequently with loss or gain of genetic information at the site of the break, resulting in small insertions or deletions; tag ligation — if a double-stranded oligonucleotide is provided with overhangs complementary to those left by the ZFNs (an adaptor), it will be ligated into the chromosome, thus producing, for example, a tagged allele; large deletion — two simultaneous DSBs made on the same chromosome can lead to a deletion of the entire intervening stretch. As shown in the panel on the right, if the break is resolved via homology-directed repair (HDR) (which will occur in the presence of donor DNA), this can result in (from top to bottom): gene correction — if the donor specifies solely a single-base-pair change (for example, a restriction fragment length polymorphism encoding a novel allele), this will result in 'gene correction' that subtly edits the endogenous allele; targeted gene addition — if a donor is provided that carries an ORF or a transgene at the position corresponding to the site of the break, its sequence will be transferred to the chromosome via the synthesis-dependent strand annealing pathway; transgene stacking — if the donor carries multiple linked transgenes between the homology arms, they will be transferred into the chromosome via the synthesis-dependent strand annealing pathway, producing essentially 'a stacked trait'.

and to generate novel allelic forms of three endogenous loci[71]. In all of these studies, efficient, specific and stable gene addition was achieved and the cells retained characteristics of pluripotency. Inducible gene cassettes added to the AAVS1 locus retained long-term function, and promoterless ORFs added to the *OCT4* (also known as *POU5F1*) locus correctly reported on its transcriptional status, turning this important marker of stemness into an endogenous reporter of cell fate[71]. ZFN-mediated gene addition has also now been used successfully to introduce short hairpin RNA (shRNA) expression cassettes into the AAVS1 locus in transformed cells and human ES cells; furthermore, the two alleles of the target gene were edited simultaneously to harbour distinct transgenes and thus generate a transheterozygote in a single step[72].

Gene addition in plants. ZFN-driven gene addition to native loci has recently also been achieved in plants: a taxonomic group that historically has proven resistant to targeted gene modification[73]. ZFNs targeting an endogenous endochitinase gene in tobacco were co-delivered to tobacco suspension cells or leaf disk protoplasts along with a donor DNA carrying short homology arms and a herbicide resistance marker[74]. Correct addition of the resistance cassette was observed in ~10% of the cases.

A recent report[75] described the editing of two endogenous loci in *Zea mays* (maize). ZFNs targeted to the gene encoding an enzyme required for the production of phytate were introduced with a donor carrying a herbicide resistance marker, resulting in transgene addition to the ZFN-specified locus (and consequent disruption of the endogenous gene), notably without additional random integration of the donor or of the ZFN-encoding plasmid DNA. ZFN-edited plants were fertile, the transgene was transmitted to the next generation in normal Mendelian fashion and co-segregation of the herbicide-resistant and low-phytate-content phenotype was observed. Site-specific gene addition in a major crop species could be used for 'trait stacking' (FIG. 3), an important goal that involves the creation of plants in which several independent traits are physically linked, thus ensuring their co-segregation throughout the breeding process.

Challenges for gene addition. Two issues associated with all gene addition strategies need to be considered. First, the necessity for co-delivery to the target cell of the donor DNA construct along with the ZFNs can often be limiting and require optimization. DNA, RNA and viral delivery systems have all been used to deliver both

Short hairpin RNA
Small RNAs that form hairpins that can induce sequence-specific silencing in mammalian cells through RNAi.

ZFN and donor molecules, although the ideal method has proven to be dependent on cell type (TABLE 1). A second issue is repair pathway choice or competition — that is, the propensity of a cell to preferentially repair a ZFN-induced DSB using NHEJ rather than HDR (FIG. 3). Elucidation of the crucial factors that determine and/or influence repair pathway selection may ultimately provide fine control over the outcome of DSB repair. Of note, disabling the *Ligase 4* gene in *D. melanogaster* (the product of which is required for NHEJ) strongly biases the resolution of a ZFN-induced break towards HDR[59].

Specificity of genome editing

Nuclease specificity is an important component for the successful application of ZFNs. In general, 'off-target' cleavage can lead to reduced efficiency of on-target modification and cytotoxicity. An off-target cleavage site may confound the interpretation of the intended genome editing event or, worse, lead to an adverse event in a therapeutic setting. To address this concern, one can optimize the reagents being delivered to the cell and analyse cells for any unwanted side effects.

Reagent optimization. High ZFN specificity has been achieved by the combination of two approaches. As discussed in the section on ZFN design, the first of these is the use of two-finger modules to assemble ZFPs with longer (12–18 bp) DNA recognition sites; such long sites are potentially rare even in complex genomes. Note, however, that a longer target site is necessary but not sufficient for improved specificity. Also, the protein and DNA must interact at the correct positions across the entire recognition interface to ensure efficient binding *in vivo*[35,38,61,75,76]. One approach is to analyse an archive of both designed and selected zinc finger modules and use data from such analysis to replace certain residues or even entire α-helices in the ZFP with those that are likely to perform better *in vivo*[61]. An example of such specificity is provided by the paralogue-specific action of the ZFNs designed to target the *inositol 1,3,4,5,6-pentakisphosphate 2-kinase* (*Ipk1*) locus in maize[75] described above.

The second, complementary approach is to require that a DSB can only be induced via a heterodimer of two ZFNs (via the use of obligate heterodimerization domains[26,27]). ZFNs combining both approaches recognize composite sites of 24–36 bp (unique within the genome) and are well-tolerated by primary and transformed mammalian cells[61,68,71,76], zebrafish embryos[35] and rat embryos[38,39].

Monitoring editing specificity. To analyse ZFN action genome-wide, several independent assays have been developed, which can be broadly divided into methods guided by a biochemical determination of the specificity of the two ZFP DNA binding domains that encompass a given ZFN pair and methods that are independent of such prior knowledge.

Determining the specificity for a ZFP DNA binding domain *in vitro* is routinely performed by systematic evolution of ligands by exponential enrichment (SELEX)[77,78], resulting in an experimentally determined consensus binding site. A SELEX protocol has been developed that yields a biologically relevant consensus for naturally occurring C2H2 ZFPs[79]. This information (often displayed as a position weight matrix) can be used to bioinformatically interrogate the genome of interest and generate a rank order of potential off-target sites with highest similarity to the consensus. The genome-wide specificity of ZFN action can thus be studied by direct sequencing of these loci in cells already known to carry the intended modification at the targeted endogenous gene. This analysis has been performed in primary human T cells[76] (reviewed in the next section), zebrafish[35], rats[38], maize[75] and human ES and iPS cells[71]. In each case, a set of such loci (typically 5–15) was genotyped in cells (or organisms) carrying ZFN-induced alleles at the target (intended) locus. Importantly, with one exception (disruption of one off-target site for one ZFN pair in one single-cell-derived clone[71], which serves to validate the SELEX-bioinformatic procedure for off-target identification[71]), all putative off-target sites were found to be wild type; see the next section for a review of the work on editing in T cells[76].

In addition to bioinformatically driven genome-wide analyses, a range of indirect methods for assessing editing fidelity have been exploited. For example, cytological profiling of the epigenetic hallmarks of DSB induction (that is, H2AX and/or tumour suppressor p53-binding protein 1 (TP53BP1) immunostaining) found that one to two DSBs occur above background in cells treated with ZFNs with a 24 bp composite recognition stretch and obligate heterodimer *Fok*I nuclease domains (see above). The number of DSBs resolved to basal levels by day 5 post-ZFN addition[26]. A cytogenetic analysis of ZFN-edited CHO, human ES and human iPS cells to look for gross chromosomal changes showed that all possessed a wild-type karyotype[48,71]. A more refined search for off-target chromosomal changes focused on the potential for ZFNs to increase the rates of random DNA integration in cells. No measurable increase in the rate of random plasmid integration was observed in transformed human cells efficiently edited with 24-bp target ZFNs[66,72]. Southern blot analysis for randomly integrated donor constructs carrying a selectable marker gene in systems as distinct as maize embryogenic suspension cells[75] and human ES cells[71] revealed, in both cases, that the majority of clonal events obtained carry donor-specified selectable markers solely at the intended ZFN target site.

Importantly, although any individual analysis taken in isolation may have limitations of sensitivity or scope, together the breadth of analyses carried out thus far point to a specificity of ZFN action that exceeds what is required for model systems and, as will be discussed below, has proven sufficient for the clinical translation of ZFN-edited T cells in three Phase I clinical trials.

Therapeutic application of ZFNs

Editing of native loci is, in principle, a superior strategy in certain clinical settings. For example, a corrected allele of a disease-causing gene could be curative in several monogenic diseases (for example, replacing a

Systematic evolution of ligands by exponential enrichment
(SELEX). A method for identifying nucleic acid ligands for a chosen 'bait' molecule (typically a protein). In its most general form, the method comprises: incubation of a randomized nucleic acid library with the bait molecule; recovery of the bait, along with any bound nucleic acids; amplification of recovered nucleic acids via PCR; and sequencing to identify binding motifs.

Embryogenic suspension cells
Plant cells, derived from callus tissue or from an embryo, that can be maintained in liquid growth medium and that, under appropriate culture conditions, can be used to regenerate a whole plant.

mutation) while retaining all aspects of endogenous gene control. Alternatively, the knockout of a gene encoding a virus receptor can be shown to eliminate rather than merely reduce infection. This section describes both the progress and the challenges in bringing genome editing to the clinic.

Advances in therapy using ZFNs. Due to its relative simplicity, ZFN-mediated gene disruption (achieved by transient delivery of the ZFNs alone) is the first ZFN-based approach that has been taken to the clinic — specifically for the treatment of glioblastoma[80] (NCT01082926) and HIV[76,81] (NCT00842634 and NCT01044654). In the former case, the glucocorticoid receptor gene is disrupted by ZFNs as part of a T cell-based cancer immunotherapy (autologous T cell therapy for the treatment of prostate cancer was recently approved by the US Food and Drug Administration[82]).

The AIDS trials are based on the fact that HIV infection requires the co-receptors chemokine (C-C motif) receptor type 5 (CCR5) or chemokine (C-X-C motif) receptor type 4 (CXCR4). A naturally occurring human mutation in the *CCR5* gene (*CCR5Δ32*) was shown to confer resistance to the virus without causing detectable pathophysiological effects beyond an increased susceptibility to West Nile virus infection, and reducing or blocking CCR5–HIV interaction is therefore a validated drug target for small molecule inhibitors, small interfering RNA knockdown approaches, antibodies or intrabodies. Unlike these 'knockdown' or blocking strategies (which require persistent exposure to the therapeutic), the potential advantage of a ZFN approach is a fully penetrant and heritable gene knockout (and consequent HIV resistance) that persists for the lifetime of the cell and its progeny. Delivery of ZFNs targeting CCR5 using a recombinant adenoviral vector was recently shown to result in the disruption of *CCR5* in >50% of transduced cells, both in model cell lines and primary human CD4 T cells[76]. In a murine xenotransplantation model, ZFN-modified primary $CD4^+$ T cells preferentially increased in number in the presence of HIV, whereas unmodified cells did not. In the mice that received modified T cells there was a substantial reduction in viral load (>sevenfold reduction) and an overall increase in CD4 counts (>fivefold increase) in the peripheral blood. Moreover, ZFN-modified $CD4^+$ T cells engrafted and functioned normally in response to stimulation, supporting the possibility that these cells may be able to reconstitute immune function in patients with HIV/AIDS via maintenance of an HIV-resistant $CD4^+$ T cell population. Two Phase I clinical trials are in progress that aim to address the clinical and laboratory safety of infusion of *ex vivo* expanded $CD4^+$ T cells treated with ZFNs targeting CCR5 in patients with HIV/AIDS.

Although the $CD4^+$ T cell is crucially important in the prevention of AIDS, elimination of CCR5 in $CD34^+$ HSCs would allow the generation of CCR5-negative cells representing all blood lineages (including macrophages and dendritic cells in addition to T cells), hence potentially protecting most natural targets for HIV infection. A crucial proof of concept study recently documented the successful transplantation of bone marrow from a donor homozygous for the *CCR5Δ32* mutation into an HIV patient with acute myeloid leukaemia. Importantly, the *CCR5Δ32* homozygous cells engrafted and repopulated the peripheral blood and the patient so far remains in a persistent aviraemic state[83]. However, finding sufficient human-leukocyte-antigen-matched *CCR5Δ32* homozygote donors is not feasible given the small percentage of humans who are homozygous for this mutation. In principle, a ZFN approach would allow an autologous $CD34^+$ HSC transplant for any patient. To this end, normal human HSCs carrying a ZFN-induced disruption of *CCR5* were shown to retain stemness, as gauged by full engraftment in immunodeficient mice. Importantly, the T cell progeny derived from these modified HSCs were shown to home to peripheral tissues, including the gut-associated lymphoid tissues, and preferentially expand upon challenge with R5-tropic HIV[81].

Addressing the risk of off-target effects. The inherent risk of a therapeutic approach using a ZFN-induced DSB is the potential for low-frequency off-target cleavage events at undesired locations in the genome. Validating ZFNs for clinical application thus requires more sensitive methods. First, one can exploit the premise that any site of ZFN action is a function of the DNA-binding specificity of the engineered ZFP DNA binding domains. A high-resolution consensus DNA binding preference for each ZFP enables a bioinformatic search for the most similar (and therefore most likely) sites for off-target ZFN action within the target genome. Combining this bioinformatic approach with ultradeep sequencing methods (such as Solexa and 454) has allowed the identification of very rare events and validated the sensitivity of the SELEX-bioinformatics approach for off-target identification. For example, these approaches were implemented in an analysis of the ZFNs that target CCR5, which showed that these ZFNs are highly specific for their intended target site in CCR5 (REF. 76). Specifically, these ZFNs showed a tenfold lower preference for the *CCR2* gene (the closest paralogue to the *CCR5* target in the genome) but, importantly, revealed a very rare off-target event that occurs only once in every 20,000 events and is located in the intron of a gene (actin binding LIM protein family, member 2 (*ABLIM2*)) involved in central nervous system development and maintenance.

In combination with immunostaining for the formation of DSBs and standard cytological karyotyping, the molecular assays above provide the investigator with important additional information with which to assess ZFN safety. However, such assays are valuable only as part of a broad package of safety assessments. Additional assays include a soft-agar transformation assay of control cells exposed to an excess of ZFN, as well as *in vivo* carcinogenicity studies using a sufficient number of the modified target cells (again exposed to excess ZFN) to permit detection of any rare but undesirable events that would lead to cellular transformation.

Autologous
In transplantology, referring to cells or an organ transplanted from an individual to that same individual (often after some *ex vivo* procedure has been performed).

Intrabodies
Antibodies that are directed against intracellular target molecules and expressed within a specific subcellular compartment as directed by localization signals genetically fused to the amino or carboxyl terminus of a given antibody.

Aviraemic
Aviraemia refers to the lack of detectable virus in the circulation of an individual.

Ultradeep sequencing
An umbrella term that refers to several independent, proprietary, high-throughput DNA-sequencing technologies that use massively parallel sequencing-by-synthesis approaches. The new methods allow an increase in generated sequence per run of about two orders of magnitude compared with conventional Sanger sequencing technologies.

Three Phase I clinical trials investigating two different clinical applications of ZFN-modified cells (HIV/AIDS and glioblastoma) are currently in progress. In all cases, given the therapeutic modality (using T cells modified *ex vivo*) the cells can be extensively tested before infusion into the patient. This study is an important milestone towards the broad application of ZFNs for correcting or modifying additional cell types (including human stem cells) for use in cell-based therapy and, ultimately, for editing cells *in vivo*.

Conclusion

The fundamental barriers to using the C2H2 zinc finger motif as a platform for designing DNA binding modules with novel specificities[8] have now been overcome, and several methods are available for the production of novel ZFPs against investigator-specified loci. ZFNs have been shown to allow *bona fide* reverse genetics in diverse model organisms and to enable allele engineering in somatic cell genetics. Furthermore, the current generation of ZFN technology exploits evolutionarily conserved pathways — protein–DNA interactions and DNA repair — that offer the prospect of successfully porting this approach to an even broader range of experimental and applied settings. A potential area of focus could be the development of delivery technologies that will allow an expansion of the settings and organisms that are amenable to ZFN-driven genome editing — for example, in the future cells could be edited within an adult organism.

ZFN-mediated genome editing now offers the ability to carry out sophisticated gene-function studies directly in the model system of interest. For example, the random integration of an additional and overexpressed copy of a marked gene (such as one fused to GFP) can now be replaced with the tagging of the endogenous gene — simultaneously eliminating the risk of a confounding insertional mutagenesis event and potential artefacts driven by variegated and inappropriate expression levels from a cDNA transgene. Thus, the application of current-generation ZFNs removes many of the constraints on experimental design that previously rendered some studies impossible and forced others to an achievable, but less than optimal, surrogate. The addition of ZFN technology to our tool box will perhaps allow the awesome power of genetics to be extended to any eukaryote.

1. Carroll, D. Progress and prospects: zinc-finger nucleases as gene therapy agents. *Gene Ther.* **15**, 1463–1468 (2008).
2. Bibikova, M., Golic, M., Golic, K. G. & Carroll, D. Targeted chromosomal cleavage and mutagenesis in *Drosophila* using zinc-finger nucleases. *Genetics* **161**, 1169–1175 (2002).
3. Bibikova, M., Beumer, K., Trautman, J. K. & Carroll, D. Enhancing gene targeting with designed zinc finger nucleases. *Science* **300**, 764 (2003).
References 2 and 3 are early studies in the field of genome editing that applied ZFNs for the disruption and correction of endogenous genes in *D. melanogaster*.
4. Kim, Y. G., Cha, J. & Chandrasegaran, S. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proc. Natl Acad. Sci. USA* **93**, 1156–1160 (1996).
5. Bibikova, M. *et al.* Stimulation of homologous recombination through targeted cleavage by chimeric nucleases. *Mol. Cell Biol.* **21**, 289–297 (2001).
6. Miller, J., McLachlan, A. D. & Klug, A. Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus* oocytes. *EMBO J.* **4**, 1609–1614 (1985).
This paper describes the discovery of the Cys2-His2 zinc finger — the most common DNA recognition motif in metazoa.
7. Wolfe, S. A., Nekludova, L. & Pabo, C. O. DNA recognition by Cys2His2 zinc finger proteins. *Annu. Rev. Biophys. Biomol. Struct.* **29**, 183–212 (2000).
8. Pavletich, N. P. & Pabo, C. O. Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A. *Science* **252**, 809–817 (1991).
This study describes the first X-ray structure of a Cys2-His2 ZFP–DNA complex.
9. Segal, D. J. *et al.* Evaluation of a modular strategy for the construction of novel polydactyl zinc finger DNA-binding proteins. *Biochemistry* **42**, 2137–2148 (2003).
10. Rebar, E. J. & Pabo, C. O. Zinc finger phage: affinity selection of fingers with new DNA-binding specificities. *Science* **263**, 671–673 (1994).
11. Jamieson, A. C., Kim, S. H. & Wells, J. A. *In vitro* selection of zinc fingers with altered DNA-binding specificity. *Biochemistry* **33**, 5689–5695 (1994).
12. Dreier, B., Beerli, R. R., Segal, D. J., Flippin, J. D. & Barbas, C. F. Development of zinc finger domains for recognition of the 5′-ANN-3′ family of DNA sequences and their use in the construction of artificial transcription factors. *J. Biol. Chem.* **276**, 29466–29478 (2001).
13. Dreier, B. *et al.* Development of zinc finger domains for recognition of the 5′-CNN-3′ family DNA sequences and their use in the construction of artificial transcription factors. *J. Biol. Chem.* **280**, 35588–35597 (2005).
14. Segal, D. J., Dreier, B., Beerli, R. R. & Barbas, C. F. Toward controlling gene expression at will: selection and design of zinc finger domains recognizing each of the 5′-GNN-3′ DNA target sequences. *Proc. Natl Acad. Sci. USA* **96**, 2758–2763 (1999).
15. Choo, Y. & Klug, A. Toward a code for the interactions of zinc fingers with DNA: selection of randomized fingers displayed on phage. *Proc. Natl Acad. Sci. USA* **91**, 11163–11167 (1994).
16. Bae, K. H. *et al.* Human zinc fingers as building blocks in the construction of artificial transcription factors. *Nature Biotech.* **21**, 275–280 (2003).
17. Kim, H. J., Lee, H. J., Kim, H., Cho, S. W. & Kim, J. S. Targeted genome editing in human cells with zinc finger nucleases constructed via modular assembly. *Genome Res.* **19**, 1279–1288 (2009).
18. Segal, D. J., Crotty, J. W., Bhakta, M. S., Barbas, C. F. & Horton, N. C. Structure of Aart, a designed six-finger zinc finger peptide, bound to DNA. *J. Mol. Biol.* **363**, 405–421 (2006).
19. Wolfe, S. A., Grant, R. A., Elrod-Erickson, M. & Pabo, C. O. Beyond the 'recognition code' structures of two Cys2His2 zinc finger/TATA box complexes. *Structure* **9**, 717–723 (2001).
20. Pavletich, N. P. & Pabo, C. O. Crystal structure of a five-finger GLI–DNA complex: new perspectives on zinc fingers. *Science* **261**, 1701–1707 (1993).
21. Fairall, L., Schwabe, J. W., Chapman, L., Finch, J. T. & Rhodes, D. The crystal structure of a two zinc-finger peptide reveals an extension to the rules for zinc-finger/DNA recognition. *Nature* **366**, 483–487 (1993).
22. Houbaviy, H. B., Usheva, A., Shenk, T. & Burley, S. K. Cocrystal structure of YY1 bound to the adeno-associated virus P5 initiator. *Proc. Natl Acad. Sci. USA* **93**, 13577–13582 (1996).
23. Nolte, R. T., Conlin, R. M., Harrison, S. C. & Brown, R. S. Differing roles for zinc fingers in DNA recognition: structure of a six-finger transcription factor IIIA complex. *Proc. Natl Acad. Sci. USA* **95**, 2938–2943 (1998).
24. Ramirez, C. L. *et al.* Unexpected failure rates for modular assembly of engineered zinc fingers. *Nature Methods* **5**, 374–375 (2008).
25. Vanamee, E. S., Santagata, S. & Aggarwal, A. K. FokI requires two specific DNA sites for cleavage. *J. Mol. Biol.* **309**, 69–78 (2001).
26. Miller, J. C. *et al.* An improved zinc-finger nuclease architecture for highly specific genome editing. *Nature Biotech.* **25**, 778–785 (2007).
27. Szczepek, M. *et al.* Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases. *Nature Biotech.* **25**, 786–793 (2007).
28. Shimizu, Y., Bhakta, M. S. & Segal, D. J. Restricted spacer tolerance of a zinc finger nuclease with a six amino acid linker. *Bioorg. Med. Chem. Lett.* **19**, 3970–3972 (2009).
29. Handel, E. M., Alwin, S. & Cathomen, T. Expanding or restricting the target site repertoire of zinc-finger nucleases: the inter-domain linker as a major determinant of target site selectivity. *Mol. Ther.* **17**, 104–111 (2009).
30. Guo, J., Gaj, T. & Barbas, C. F. Directed evolution of an enhanced and highly efficient FokI cleavage domain for zinc finger nucleases. *J. Mol. Biol.* **400**, 96–107 (2010).
31. Moynahan, M. E. & Jasin, M. Mitotic homologous recombination maintains genomic stability and suppresses tumorigenesis. *Nature Rev. Mol. Cell Biol.* **11**, 196–207 (2010).
32. Lieber, M. R. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annu. Rev. Biochem.* **79**, 181–211 (2010).
33. Jackson, S. P. & Bartek, J. The DNA-damage response in human biology and disease. *Nature* **461**, 1071–1078 (2009).
34. Beumer, K. J. *et al.* Efficient gene targeting in *Drosophila* by direct embryo injection with zinc-finger nucleases. *Proc. Natl Acad. Sci. USA* **105**, 19821–19826 (2008).
35. Doyon, Y. *et al.* Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases. *Nature Biotech.* **26**, 702–708 (2008).
36. Meng, X., Noyes, M. B., Zhu, L. J., Lawson, N. D. & Wolfe, S. A. Targeted gene inactivation in zebrafish using engineered zinc-finger nucleases. *Nature Biotech.* **26**, 695–701 (2008).
37. Foley, J. E. *et al.* Rapid mutation of endogenous zebrafish genes using zinc finger nucleases made by Oligomerized Pool ENgineering (OPEN). *PLoS ONE* **4**, e4348 (2009).
38. Geurts, A. M. *et al.* Knockout rats via embryo microinjection of zinc-finger nucleases. *Science* **325**, 433 (2009).
39. Mashimo, T. *et al.* Generation of knockout rats with X-linked severe combined immunodeficiency (X-SCID) using zinc-finger nucleases. *PLoS ONE* **5**, e8870 (2010).
40. Lloyd, A., Plaisier, C. L., Carroll, D. & Drews, G. N. Targeted mutagenesis using zinc-finger nucleases in *Arabidopsis*. *Proc. Natl Acad. Sci. USA* **102**, 2232–2237 (2005).

41. Zhang, F. *et al.* High frequency targeted mutagenesis in *Arabidopsis* thaliana using zinc finger nucleases. *Proc. Natl Acad. Sci. USA* **107**, 12023–12028 (2010).
42. Osakabe, K., Osakabe, Y. & Toki, S. Site-directed mutagenesis in *Arabidopsis* using custom-designed zinc finger nucleases. *Proc. Natl Acad. Sci. USA*, **107**, 12034–12039 (2010).
43. Thomas, K. R., Folger, K. R. & Capecchi, M. R. High frequency targeting of genes to specific sites in the mammalian genome. *Cell* **44**, 419–428 (1986).
44. Kohli, M., Rago, C., Lengauer, C., Kinzler, K. W. & Vogelstein, B. Facile methods for generating human somatic cell gene knockouts using recombinant adeno-associated viruses. *Nucleic Acids Res.* **32**, e3 (2004).
45. Ruis, B. L., Fattah, K. R. & Hendrickson, E. A. The catalytic subunit of DNA-dependent protein kinase regulates proliferation, telomere length, and genomic stability in human somatic cells. *Mol. Cell Biol.* **28**, 6182–6195 (2008).
46. Khan, I. F. *et al.* Engineering of human pluripotent stem cells by AAV-mediated gene targeting. *Mol. Ther.* **18**, 1192–1199 (2010).
47. Doyon, Y., Vo, T., Choi, V. M., Gregory, P. D. & Holmes, M. C. A transient cold shock enhances zinc-finger nuclease-mediated gene disruption. *Nature Methods* **7**, 459–460 (2010).
48. Santiago, Y. *et al.* Targeted gene knockout in mammalian cells using engineered zinc finger nucleases. *Proc. Natl Acad. Sci. USA* **105**, 5809–5814 (2008).
49. Cost, G. J. *et al.* BAK and BAX deletion using zinc-finger nucleases yields apoptosis-resistant CHO cells. *Biotechnol. Bioeng.* **105**, 330–340 (2009).
50. Liu, P. Q. *et al.* Generation of a triple-gene knockout mammalian cell line using engineered zinc-finger nucleases. *Biotechnol. Bioeng.* **106**, 97–105 (2010).
51. Lee, H. J., Kim, E. & Kim, J. S. Targeted chromosomal deletions in human cells using zinc finger nucleases. *Genome Res.* **20**, 81–89 (2010).
52. Orlando, S. *et al.* Zinc-finger nuclease-driven targeted integration into mammalian genomes using donors with limited chromosomal homology. *Nucleic Acids Res.* 8 Jun 2010 (doi:10.1093/nar/gkq512).
53. Petolino, J. F. *et al.* Zinc finger nuclease-mediated transgene deletion. *Plant Mol. Biol.* **73**, 617–628 (2010).
54. Brunet, E. *et al.* Chromosomal translocations induced at specified loci in human stem cells. *Proc. Natl Acad. Sci. USA* **106**, 10620–10625 (2009).
55. Brown, J. P., Wei, W. & Sedivy, J. M. Bypass of senescence after disruption of p21CIP1/WAF1 gene in normal diploid human fibroblasts. *Science* **277**, 831–834 (1997).
56. Rouet, P., Smih, F. & Jasin, M. Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. *Mol. Cell Biol.* **14**, 8096–8106 (1994).
The conclusion drawn by this paper — that the DNA ends generated by a DSB in a mitotically dividing mammalian cell are recombinogenic — provided a starting point for the development of genome editing. See also reference 57.
57. Puchta, H., Dujon, B. & Hohn, B. Homologous recombination in plant cells is enhanced by *in vivo* induction of double strand breaks into DNA by a site-specific endonuclease. *Nucleic Acids Res.* **21**, 5034–5040 (1993).
58. Porteus, M. H. & Baltimore, D. Chimeric nucleases stimulate gene targeting in human cells. *Science* **300**, 763 (2003).
59. Bozas, A., Beumer, K. J., Trautman, J. K. & Carroll, D. Genetic analysis of zinc-finger nuclease-induced gene targeting in *Drosophila*. *Genetics* **182**, 641–651 (2009).
This paper provides the first comprehensive study of the intranuclear requirements for resolving an endogenous DSB by NHEJ rather than HDR.
60. Beumer, K., Bhattacharyya, G., Bibikova, M., Trautman, J. K. & Carroll, D. Efficient gene targeting in *Drosophila* with zinc-finger nucleases. *Genetics* **172**, 2391–2403 (2006).
61. Urnov, F. D. *et al.* Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature* **435**, 646–651 (2005).
62. Maeder, M. L. *et al.* Rapid 'open-source' engineering of customized zinc-finger nucleases for highly efficient gene modification. *Mol. Cell* **31**, 294–301 (2008).
63. Townsend, J. A. *et al.* High-frequency modification of plant genes using engineered zinc-finger nucleases. *Nature* **459**, 442–445 (2009).
64. Formosa, T. & Alberts, B. M. DNA synthesis dependent on genetic recombination: characterization of a reaction catalyzed by purified bacteriophage T4 proteins. *Cell* **47**, 793–806 (1986).
65. Nassif, N., Penney, J., Pal, S., Engels, W. R. & Gloor, G. B. Efficient copying of nonhomologous sequences from ectopic sites via P-element-induced gap repair. *Mol. Cell Biol.* **14**, 1613–1625 (1994).
66. Moehle, E. A. *et al.* Targeted gene addition into a specified location in the human genome using designed zinc finger nucleases. *Proc. Natl Acad. Sci. USA* **104**, 3055–3060 (2007).
67. Goldberg, A. D. *et al.* Distinct factors control histone variant H3.3 localization at specific genomic regions. *Cell* **140**, 678–691 (2010).
68. Lombardo, A. *et al.* Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. *Nature Biotech.* **25**, 1298–1306 (2007).
69. Benabdallah, B. F. *et al.* Targeted gene addition to human mesenchymal stromal cells as a cell-based plasma-soluble protein delivery platform. *Cytotherapy* **12**, 394–399 (2010).
70. Zou, J. et al. Gene targeting of a disease-related gene in human induced pluripotent stem and embryonic stem cells. *Stem Cell* **5**, 97–110 (2009).
71. Hockemeyer, D. *et al.* Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. *Nature Biotech.* **27**, 851–857 (2009).
72. DeKelver, R. C. *et al.* Functional genomics, proteomics, and regulatory DNA analysis in isogenic settings using zinc finger nuclease-driven transgenesis into a safe harbor locus in the human. *Genome Res.* **20**,1133–1142 (2010).
A useful feature of ZFN-driven genome editing is the ability it provides to perform mammalian somatic cell genetics experiments in isogenic settings. This paper provides several examples of such isogenic engineering.
73. Hanin, M. & Paszkowski, J. Plant genome modification by homologous recombination. *Curr. Opin. Plant Biol.* **6**, 157–162 (2003).
74. Cai, C. Q. *et al.* Targeted transgene integration in plant cells using designed zinc finger nucleases. *Plant Mol. Biol.* **69**, 699–709 (2009).
75. Shukla, V. K. *et al.* Precise genome modification in the crop species Zea mays using zinc-finger nucleases. *Nature* **459**, 437–441 (2009).
76. Perez, E. E. *et al.* Establishment of HIV-1 resistance in $CD4^+$ T cells by genome editing using zinc-finger nucleases. *Nature Biotech.* **26**, 808–816 (2008).
This paper provides data underlying the first clinical trial on the genome editing of human cells. See also reference 81.
77. Tuerk, C. & Gold, L. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* **249**, 505–510 (1990).
78. Blackwell, T. K. & Weintraub, H. Differences and similarities in DNA-binding preferences of MyoD and E2A protein complexes revealed by binding site selection. *Science* **250**, 1104–1110 (1990).
79. Phillips, C. M. *et al.* Identification of chromosome sequence motifs that mediate meiotic pairing and synapsis in, *C. elegans*. *Nature Cell Biol.* **11**, 934–942 (2009).
80. Reik, A. et *al.* Zinc finger nucleases targeting the glucocorticoid receptor allow IL-13 zetakine transgenic CTLs to kill glioblastoma cells *in vivo* in the presence of immunosuppressing glucocorticoids. *Mol. Ther.* **16**, S13–S14 (2008).
81. Holt, N. et *al.* Human hematopoietic stem/progenitor cells modified by zinc-finger nuclease targeted to *CCR5* control HIV-1 *in vivo*. *Nature Biotech.* 20 Jul 2010 (doi:10.1038/nbt.1663).
82. Bot., A. The landmark approval of Provenge, what it means to immunology and 'in this issue': the complex relation between vaccines and autoimmunity. *Int. Rev. Immunol.* **29**, 235–238 (2010).
83. Hutter, G. *et al.* Long-term control of HIV by *CCR5* Delta32/Delta32 stem-cell transplantation. *N. Engl. J. Med.* **360**, 692–698 (2009).
84. Hurt, J. A., Thibodeau, S. A., Hirsh, A. S., Pabo, C. O. & Joung, J. K. Highly specific zinc finger proteins obtained by directed domain shuffling and cell-based selection. *Proc. Natl Acad. Sci. USA* **100**, 12271–12276 (2003).
85. Mandell, J. G. & Barbas, C. F. Zinc Finger Tools: custom DNA-binding domains for transcription factors and nucleases. *Nucleic Acids Res.* **34**, W516–W523 (2006).
86. Sander, J. D., Zaback, P., Joung, J. K., Voytas, D. F. & Dobbs, D. An affinity-based scoring scheme for predicting DNA-binding activities of modularly assembled zinc-finger proteins. *Nucleic Acids Res.* **37**, 506–515 (2009).
87. Moore, M., Klug, A. & Choo, Y. Improved DNA binding specificity from polyzinc finger peptides by using strings of two-finger units. *Proc. Natl Acad. Sci. USA* **98**, 1437–1441 (2001).
88. Bartsevich, V. V. & Juliano, R. L. Regulation of the *MDR1* gene by transcriptional repressors selected using peptide combinatorial libraries. *Mol. Pharmacol.* **58**, 1–10 (2000).
89. Isalan, M., Klug, A. & Choo, Y. A rapid, generally applicable method to engineer zinc fingers illustrated by targeting the HIV-1 promoter. *Nature Biotech.* **19**, 656–660 (2001).
90. Greisman, H. A. & Pabo, C. O. A general strategy for selecting high-affinity zinc finger proteins for diverse DNA target sites. *Science* **275**, 657–661 (1997).
91. Rebar, E. J. & Miller, J. C. Design and applications of engineered zinc finger proteins. *Biotech International* 20–23 (April/May 2004).
92. Malphettes, L. *et al.* Highly efficient deletion of *FUT8* in CHO cell lines using zinc-finger nucleases yields cells that produce completely nonfucosylated antibodies. *Biotechnol. Bioeng.* **106**, 774–783 (2010).
93. Siekmann, A. F., Standley, C., Fogarty, K. E., Wolfe, S. A. & Lawson, N. D. Chemokine signaling guides regional patterning of the first embryonic artery. *Genes Dev.* **23**, 2272–2277 (2009).
94. Del Prete, G. Q. *et al.* Derivation and characterization of a simian immunodeficiency virus SIVmac239 variant with tropism for CXCR4. *J. Virol.* **83**, 9911–9922 (2009).

Competing interests statement
The authors declare competing financial interests: see Web version for details.

DATABASES

ClinicalTrials.gov: http://clinicaltrials.gov
NCT00842634 | NCT01044654 | NCT01082926
FlyBase: http://flybase.org
coilin | *Ligase 4*
UniProtKB: http://www.uniprot.org
CCR5 | CXCR4

ALL LINKS ARE ACTIVE IN THE ONLINE PDF