

# What rheumatologists need to know about CRISPR/Cas9

Gary J. Gibson and Maozhou Yang

**Abstract** | CRISPR/Cas9 genome editing technology has taken the research world by storm since its use in eukaryotes was first proposed in 2012. Publications describing advances in technology and new applications have continued at an unrelenting pace since that time. In this Review, we discuss the application of CRISPR/Cas9 for creating gene mutations — the application that initiated the current avalanche of interest — and new developments that have largely answered initial concerns about its specificity and ability to introduce new gene sequences. We discuss the new, diverse and rapidly growing adaptations of the CRISPR/Cas9 technique that enable activation, repression, multiplexing and gene screening. These developments have enabled researchers to create sophisticated tools for dissecting the function and inter-relatedness of genes, as well as noncoding regions of the genome, and to identify gene networks and noncoding regions that promote disease or confer disease susceptibility. These approaches are beginning to be used to interrogate complex and multilayered biological systems and to produce complex animal models of disease. CRISPR/Cas9 technology has enabled the application of new therapeutic approaches to treating disease in animal models, some of which are beginning to be seen in the first human clinical trials. We discuss the direct application of these techniques to rheumatic diseases, which are currently limited but are sure to increase rapidly in the near future.

## CRISPR

Segments of prokaryotic DNA containing short repetitions of DNA sequences, which are interrupted by so-called spacer DNA derived from past invaders. CRISPR serves as the bacterial adaptive immune system that protects against invading genetic materials.

Rheumatic diseases represent a diverse family of complex diseases. Many are associated with disorders of the immune system and most have a strong genetic predisposition. Their causes are generally unknown. Advances in gene editing technologies offer the ability to define the underlying biology of these diseases and provide rational targets for new drug development. Initial gene editing techniques used zinc finger nucleases or transcriptional activator-like nucleases (TALEN) to provide the necessary double-strand DNA breaks. Both nucleases are modular proteins that can be engineered to fit a desired sequence and thus create double-strand breaks at a defined DNA sequence.

The gene editing technology, clustered regularly interspersed short palindromic repeats (CRISPR) associated protein 9 (CRISPR/Cas9), is the latest addition to this molecular tool box and has seen a spectacular level of research interest since the 2012 publication of data suggesting its potential as a gene editing tool for eukaryotes<sup>1</sup>. The Cas9 nuclease uses short RNAs to target the desired DNA sequence, avoiding the laborious and expensive protein engineering necessary with the two previously described techniques. This landmark paper<sup>1</sup> was soon followed by a succession of papers establishing and adapting the technique for gene editing in eukaryotes<sup>2–8</sup>.

The rapid pace of publication has continued to the present time, establishing CRISPR/Cas9 as a broadly versatile and technically simple technique for gene editing. This technique will undoubtedly transform the way we conduct basic biology research in the future.

In this Review, we briefly describe the developments in CRISPR/Cas9 technology that have improved its specificity, greatly simplified the creation of complex animal models of disease and enabled more efficient insertion of DNA sequences, the repression and activation of multiple genes in a single cell, the control of disease-carrying vectors and the identification of multiple genes, gene pathways and gene interactions essential for specific phenotypic changes and disease pathologies. Such applications could have a profound effect on our understanding of the biology underlying rheumatic diseases, and could lead to the identification of new therapeutic targets and the possibility of radical new treatment strategies.

## Brief history of CRISPR gene editing

The characteristic CRISPR palindromic repeats were first recognized as an interesting feature in *Escherichia coli* in 1987 (REF. 9) but were not further investigated until much later. With an expanding number of sequenced

Henry Ford Hospital, Bone and Joint Center, 6135 Woodward Avenue, Detroit, Michigan 48202, USA.

Correspondence to G.J.G. GGIBSON1@bjc.hfh.edu

doi:10.1038/nrrheum.2017.6  
Published online 9 Feb 2017

## Key points

- Advances in CRISPR technology have provided the capacity to precisely identify and define the function of genes and noncoding regulatory elements associated with disease development and susceptibility
- CRISPR technology has made the generation of mouse models of disease much quicker and less expensive than traditional approaches, and has facilitated the development of much-needed larger animal models of disease
- CRISPR technology has enabled the generation of gene drives, whereby genetic changes propagate rapidly through a species, providing the potential to eliminate disease vectors and thus vector-borne diseases such as malaria
- Successful treatment of mouse models of human diseases suggests that CRISPR technology can be applied to treat human diseases in the future
- CRISPR technology has the ability to facilitate a breakthrough in our understanding of the more common and complex human diseases, including rheumatic diseases
- The potential of CRISPR/Cas9 technology in the development of new treatment strategies is confidently expected to have a major effect on the practice of rheumatology

### CRISPR/Cas9

RNA-guided gene editing platform based on the Cas9–gRNA ribonucleoprotein complex. The two-component complex can mediate gRNA-programmed recognition of specific DNA sequences and create a site-specific double-strand cleavage of the targeted DNA.

### Cas9

An RNA-guided DNA endonuclease enzyme that is the universal component of the RNA-guided CRISPR/Cas9 gene editing machinery. Cas9 by itself is inactive; upon binding to the gRNA scaffold, Cas9 goes through conformational changes that initiate its target recognition, binding and cleavage activity.

### Spacer

The spacer sequence refers to the 5' end, ~20 nucleotide variable sequence of the targeting gRNA construct. The spacer contains a targeting sequence that matches a region of DNA substrate and guides Cas9 nuclease activity.

### Protospacer

The protospacer sequence refers to the targeted site on the DNA substrate. The nucleotide sequences of the spacer and the corresponding protospacer are identical.

genomes available, the prevalence of CRISPR palindromic sequences was recognized in a wide range of archaea and bacteria and the CRISPR name was coined<sup>10</sup>. CRISPR loci vary between bacterial and archaeal species but have the characteristic feature of a sequence of short, repetitive, partially palindromic sequences separated by equally short (30 to 40 base pairs) spacer sequences. The CRISPR loci are flanked by a variable number of CRISPR associated (Cas) genes<sup>11</sup>. Recognition that a spacer sequence in *E. coli* was homologous with that of an *E. coli*-infecting virus (also known as a bacteriophage or phage) lead to the realization that CRISPR constituted an adaptive immune system for bacteria<sup>12</sup>.

Unsurprisingly, bacteria and archaea have a variety of viral defence systems; they are outnumbered ten to one by phages<sup>13</sup>. The mechanisms of innate immunity in bacteria and archaea have been known for a long time, and include induction of cell death, expression of restriction enzymes and the presence of a sticky cell coat<sup>14</sup>. However, the cellular machinery and complex interactions that constitute an adaptive immune response had been considered a property of eukaryotes. The identification of CRISPR loci in nearly all species of archaea and most species of bacteria established that these organisms also have an adaptive immune system to defend against viral invasion (reviewed elsewhere<sup>11,13</sup>). The CRISPR system performs two of the fundamental functions of an immune system. First, this system enables the acquisition of an immunological memory. Upon attack by a phage, some of an organism's Cas genes are activated to cut DNA sequences, termed protospacers, from the attacking phage and integrate them into their own CRISPR array, thus creating an immunological memory of their viral attackers<sup>11,13</sup>. To avoid targeting its own DNA, a specific sequence termed the protospacer adjacent motif (PAM), which is rarely present in the host genome, is selected by the host organism. The PAM sequence varies according to the CRISPR/Cas type. The second immune system function provided by CRISPR is destruction of invading viruses. CRISPR RNA (crRNA) is generated from the stored CRISPR sequences in response to viral infection and then used to guide an

attack on the invader<sup>11,13</sup>. The crRNAs are initially transcribed as long transcripts, which are then cleaved by endogenous RNase or specific Cas proteins to make smaller crRNAs, which direct Cas nucleases to cleave both DNA strands of the invader (FIG. 1).

The wide variety of CRISPR modules present in bacteria and archaea are divided into two classes, five types and 16 subtypes primarily on the basis of the number and type of Cas genes involved<sup>15</sup>. Type II CRISPR has been developed for gene editing in eukaryotes. Cas9, a large protein with two nuclease active sites, one that cleaves the target strand and another that cleaves the non-complementary strand<sup>18</sup>, is the nuclease used in type II CRISPR and requires an additional small RNA, *transactivating crRNA* (tracrRNA) for target recognition and cleavage.

Jinek *et al.*<sup>1</sup> first demonstrated that CRISPR/Cas9, in combination with crRNA and tracrRNA, could be used to specifically target DNA cleavage *in vitro*. In addition, they demonstrated that crRNA and tracrRNA could be combined to create a single guide RNA (gRNA) to direct sequence-specific Cas9 double-stranded DNA cleavage, and suggested that this technique might represent a simple, programmable RNA method that could be used for genome targeting and genome editing in eukaryotes (FIG. 2). This paper opened the flood gates for CRISPR/Cas9 directed genome editing. Within months, several papers were published that described the application of CRISPR/Cas9 gene editing in mammalian cells and made substantial improvements to the technique<sup>2–8</sup>. The nucleotide sequence of *cas9* was reconstructed by codon optimization and inclusion of nuclear localization signals to optimize nuclear expression in mammalian cells<sup>3</sup>. The efficiency of gRNA was substantially improved by restoring the critical 3' hairpin structure, and the capacity of the system to edit several genomic sites using multiple gRNA sequences encoded in a single construct was demonstrated<sup>2–8</sup>.

The CRISPR system in bacteria and archaea has limited application to rheumatology; however, it is the basis for understanding the mechanisms of action and components of CRISPR/Cas9 that have been adapted to genome editing in eukaryotes.

## Improvements in CRISPR/Cas9 technology

### Off-target mutations

DNA cleavage can occur with the CRISPR/Cas9 system even if there is imperfect complementarity between the gRNA and target DNA, particularly if the mismatches are in the 5' region of the target sequence. In the original report by Jinek *et al.*<sup>1</sup>, gRNA was shown to tolerate up to five mismatches. This finding raised serious concerns for the use of CRISPR/Cas9 in genome editing, particularly for *in vivo* editing. Fortunately, methods to minimize and potentially eliminate off-target mutations have been developed over the past few years. Freely available software developed by measuring off-target cleavage of thousands of gRNAs has enhanced gRNA design, enabling the elimination of promiscuous gRNAs, minimizing off-target cleavage and maximizing effectiveness<sup>16,17</sup>. Additional strategies that reduce off-target cleavage include reducing the size of the gRNA target site

**Protospacer adjacent motif (PAM)**

A three base pair DNA sequence immediately following the protospacer or the DNA sequence targeted by the Cas9/gRNA ribonuclease. The canonical PAM sequence for CRISPR/Cas9 gene editing machinery is 5'-NGG-3'.

**Guide RNA (gRNA)**

gRNA, also known as short guide RNA (sgRNA) is a short synthetic RNA sequence consisting of a scaffold structure and a programmable ~20 nucleotide spacer at the 5' end. The ~80 nucleotide RNA scaffold structure is essential for mediating both Cas9 protein binding and activation. The unique spacer sequence dictates the DNA target site to be recognized and cleaved by Cas9 protein.

**Non-homologous end joining (NHEJ)**

A cellular pathway that repairs double-strand breaks in DNA. NHEJ is active throughout the cell cycle and requires no repair template. NHEJ is frequently imprecise and the repair process can generate an open reading frame shift with insertions, deletions or mutations at the site of double-strand breaks. The inaccurate nature of the NHEJ repair process forms the basis of the CRISPR/Cas knockout strategy.

**Homology-directed repair (HDR)**

The HDR pathway (also known as homologous recombination), involving a homologous template (either a sister chromatid or an exogenous DNA template), repairs double-strand DNA breaks accurately according to the template. The template or donor DNA consists of left and right arms identical to sequences flanking the double-strand break. Between the arms, any DNA sequence or marker can be inserted and HDR will force the additional genetic material to be knocked in to the particular locus. HDR is usually believed to be active only during S and G2 phases of the cell cycle.

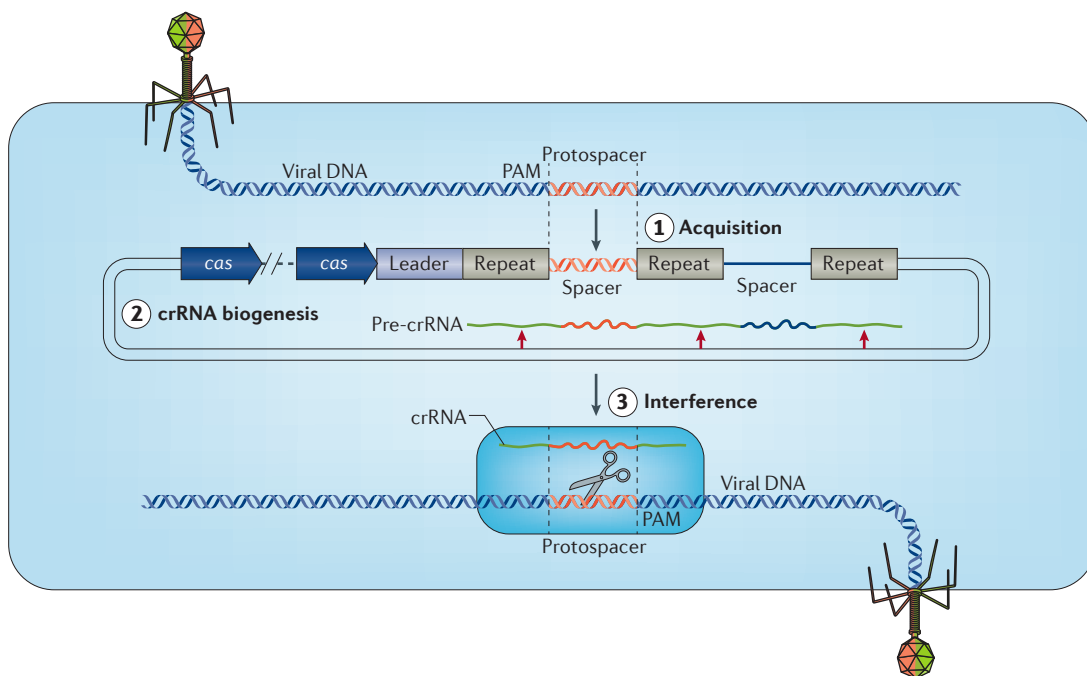


Figure 1 | **Adaptive immune system of bacteria and archaea.** CRISPR-associated (Cas) genes (*cas*) (blue arrows) encode proteins required for new spacer sequence acquisition, CRISPR RNA (crRNA) biogenesis and target interference. Step 1. Acquisition. Unique sequences (protospacers) are acquired from invading viruses and inserted into the host genome, separated by partially palindromic repeats (Repeat). Adjacent to protospacers are short sequences called protospacer adjacent motifs (PAMs). Step 2. crRNA biogenesis. In response to viral invasion, long CRISPR transcripts (pre-crRNA) are processed into short crRNAs that guide Cas proteins to invading DNA through complementary base-pairing. Step 3. Target interference. Cas nucleases initiate double-strand breaks in the DNA at the target site. Permission obtained from Annual Reviews © Sorek, R. *et al. Annu. Rev. Biochem.* 82, 237–266 (2013).

from 20 nucleotides to 17–18 nucleotides<sup>18</sup>. These shortened gRNAs seem to have the same efficiency as full-length gRNA in directing DNA targeting and cleavage, but show decreased off-target cleavage and increased sensitivity to gRNA:DNA mismatches<sup>18</sup>. One of several molecular modifications of Cas9 has been to mutate one of the nuclease active sites such that the enzyme cleaves only one of the DNA strands (termed a Cas9 nickase). Using Cas9 nickase and two paired and appropriately offset gRNAs, larger DNA double-strand breaks were created and the specificity improved in some sites by 50-fold to 1,000-fold<sup>19</sup>.

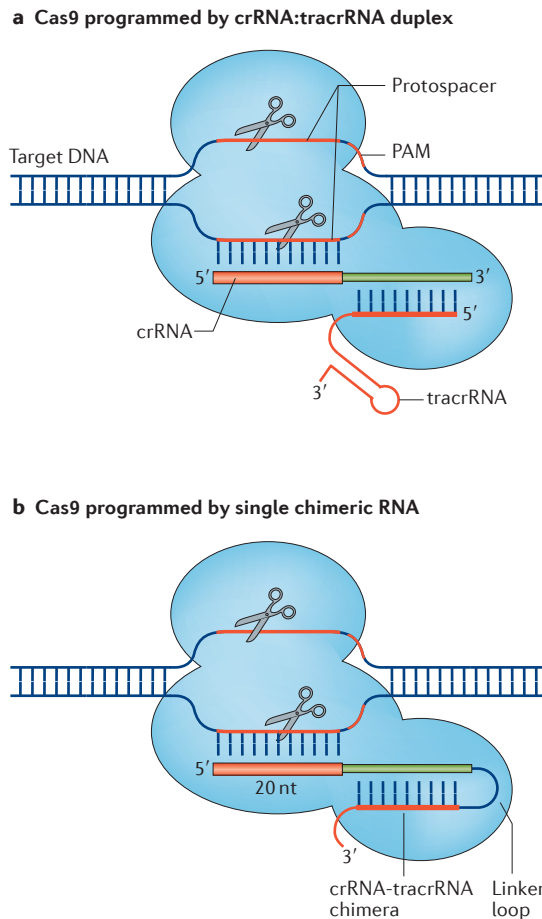
A promising new approach based on the detailed ultrastructure of the Cas9 protein–gRNA complex bound to target DNA has been described<sup>20,21</sup>. Modifying regions of Cas9 to reduce the strength of interaction with DNA was reasoned to increase the reliance on the gRNA:DNA interaction and thus increase the sensitivity to gRNA:DNA mismatches. Kleinstiver *et al.*<sup>20</sup> engineered a version of Cas9 that targets sites of hydrogen bonding to the phosphate backbone of the target DNA strand. After systematically comparing multiple mutations and combinations of mutations, they demonstrated that one variant with four substitutions, in which alanine was substituted for charged amino acids, functioned as a high fidelity Cas9 (Cas9HF) with on-target activity similar to the wild-type, but with largely undetectable off-target mutations<sup>20</sup>. Slaymaker *et al.*<sup>21</sup> used a similar

rational engineering approach, but targeted the interaction between Cas9 and the complementary DNA strand by neutralizing three positively charged amino acid residues to generate a Cas9 mutant with similarly enhanced specificity (eCas9). Both studies used broadly specific, sensitive methods for detecting off-target mutations and examined multiple cell systems. Although these improvements go some way towards quashing concerns over off-target mutations, confidence in their use will only come with the continued application of high specificity Cas9 nucleases in a wide variety of systems, as off-target activity has been shown to be highly dependent on the cell type and culture system studied<sup>22</sup>.

**Homology directed repair**

The generation of knockout mutations using CRISPR/Cas9 is exceptionally efficient, primarily because of the high efficiency of DNA cleavage and high error rate of non-homologous end joining (NHEJ). Double-strand breaks in the host DNA are highly cytotoxic lesions that are efficiently repaired by NHEJ, the predominant repair mechanism in eukaryotes. This repair is usually inaccurate and frequent insertions and/or deletions (indels) occur<sup>23</sup>. The consequent frameshift mutations generate premature stop codons, which result in loss of function of the target gene.

By contrast, homology directed repair (HDR) is typically highly inefficient<sup>24</sup>. HDR makes use of homologous recombination to intentionally generate precise and



**Figure 2 | Cas9 targeting using crRNA (CRISPR RNA)–tracrRNA (transactivating crRNA) or a single guide RNA chimera.** **a** | In type II CRISPR/Cas9 systems, Cas9 is guided by a two-RNA structure formed by tracrRNA and crRNA to cleave targeted double-stranded DNA (dsDNA). **b** | A chimeric RNA generated by fusing crRNA to tracrRNA via a linker loop is able to target and cleave dsDNA. Permission obtained from Science © Jinek, M. *et al. Science* **337**, 816–821 (2012). Nt, nucleotide

specific alterations to the DNA sequence (FIG. 3). HDR is essential for many gene editing applications, particularly correction of genetic mutations. The double-strand DNA break generated by Cas9 can boost the HDR pathway by several orders of magnitude<sup>25</sup>. However, efficiency still remains low, meaning that very large numbers of cells are required for successful insertion of the target sequence. Several approaches for enhancing HDR activity, including use of inhibitors of the NHEJ pathway<sup>26</sup> and use of longer homology arms for the donor DNA<sup>27</sup>, have been described<sup>26,28,29</sup>.

A promising new approach based on detailed investigation of the interaction of Cas9 with its target DNA substrate provides substantial improvements in the efficiency of HDR<sup>29</sup>. The study demonstrated that Cas9 binds tightly to its DNA substrate for at least 5 hours after DNA cleavage<sup>29</sup>, blocking access for donor DNA templates. However, one end of the cleaved DNA, the PAM-distal, non-target strand, was free of protein

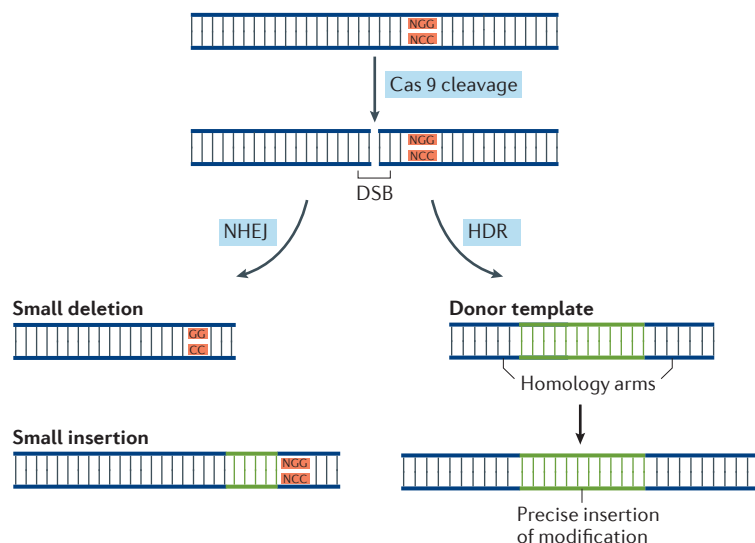
interaction and could anneal exogenous DNA. By targeting this free DNA strand and optimizing donor DNA orientation, polarity and length, the researchers achieved a 60% frequency of HDR<sup>29</sup>. This approach still needs to be corroborated in a variety of systems, but the increase in efficiency suggests that homologous recombination-based gene targeting should be amenable to routine laboratory manipulation. If similar widespread improvements in efficiency are achieved, this technique will substantially advance the prospects of CRISPR/Cas9-based therapeutic gene editing.

### Gene repression and activation

The CRISPR/Cas9 system also offers the capability to selectively switch genes on or off without manipulating their sequence. Several groups have demonstrated that Cas9 can be mutated in both nuclease domains to generate a nuclease-deactivated Cas9 (dCas9)<sup>30,31</sup>. dCas9 can be converted into a programmable gene activator or repressor via fusion with protein regulators while maintaining its ability to strongly bind specific DNA sites via target-directed gRNAs.

Systems that employ multiple activator proteins fused to dCas9 achieve consistently high levels of gene activation, ranging from 10-fold to 1,000-fold across multiple cell types and species<sup>32,33</sup>. Examples of such fusion proteins include the following: VPR, which is a fusion of multiple synergistic activators, VP64 (an engineered tetramer of the herpes simplex VP16 transcriptional activator domain), transcription factor p65 and the Epstein–Barr virus replication and transcription activator (Rta); scaffolds, such as the SunTag array that binds multiple VP64 activator domains; and the synergistic activator mediator, a modified gRNA that contains binding sites for RNA-binding proteins fused with transcription activators<sup>33–35</sup> (FIG. 4). A histone demethylase and a histone acetyltransferase have also been fused to dCas9 to specifically suppress or activate gene enhancers or promoters<sup>36,37</sup>. These systems offer another approach to modifying gene expression, as well as helping to decipher site-specific epigenetic modifications and the role of histone methylation and acetylation in cellular function.

Fusion of dCas9 with transcription repressors (such as the Krüppel-associated box (KRAB) domain) has been effective in the generation of CRISPR-mediated gene interference (CRISPRi) (FIG. 4). When localized to DNA, KRAB recruits a protein complex that initiates chromatin remodelling, methylation and deacetylation<sup>30,38</sup>. CRISPRi is similar to RNA interference (RNAi), a process whereby specific RNA molecules bind to mRNA, initiate its breakdown and thus inhibit gene expression. RNAi technology has been available for >15 years and is actively investigated as a tool for inhibiting specific gene expression both in the laboratory and as gene therapy. CRISPRi differs from and has some advantages over RNAi in that it primarily affects the process of transcription rather than affecting the levels of mature mRNA in a cell. CRISPRi is based on the Watson–Crick base-pairing model of gRNA binding to DNA and offers the same technical simplicity and broad versatility as CRISPR/Cas9.



**Figure 3 | Endogenous repair of double-strand DNA breaks by non-homologous end joining (NHEJ) or homology directed repair (HDR).** Double-strand DNA breaks (DSB) generated by Cas9 in the absence of a donor template are repaired by NHEJ, which is frequently inaccurate, resulting in the insertion of small deletions or insertions. Inclusion of a donor template can result in HDR, leading to the introduction of desired mutations or modifications through homologous recombination.

CRISPRi can also target noncoding sequences, including noncoding RNAs. In addition, CRISPRi can efficiently knock down gene expression by 90–99%; however, the effects of CRISPRi are highly dependent on the target sites of gRNAs and differ across genes, suggesting that chromatin structure and the presence of regulatory elements can limit gene knockdown<sup>31</sup>.

### Multiplexing

One of the exciting prospects for the use of CRISPR technology is the capacity to simultaneously activate, repress and knock out multiple specific genes in a single cell. This capacity has been made more accessible by the surprising observation that short gRNAs (14 nucleotides) bind strongly to their target sequence while simultaneously inhibiting the nuclease activity of Cas9 and preventing DNA cleavage<sup>32</sup>. These 14 nucleotide gRNAs activate genes to the same degree as 20 nucleotide gRNAs when combined with dCas9 fused to the gene activator VPR. When active Cas9 fused with the gene activator VPR (Cas9–VPR) was targeted with 20 nucleotide gRNA, the DNA cleavage activity (and associated gene mutation and, usually, gene deletion) observed was similar to that seen with wild-type Cas9, but when Cas9–VPR was targeted with 14 nucleotide gRNA, gene activation occurred. Kiani *et al.*<sup>32</sup> demonstrated that gene knockout and activation of target genes could be performed simultaneously in single cells by transfection with Cas9–VPR and targeting with 14 nucleotide (activation) or 20 nucleotide (knockout) gRNAs<sup>32</sup>.

Our understanding of the cell-specific regulation of gene expression is rudimentary. The ability of CRISPR systems to activate, inhibit and knock out multiple genes in a single cell creates the capacity to decipher complex gene networks such as the multi-layered immune

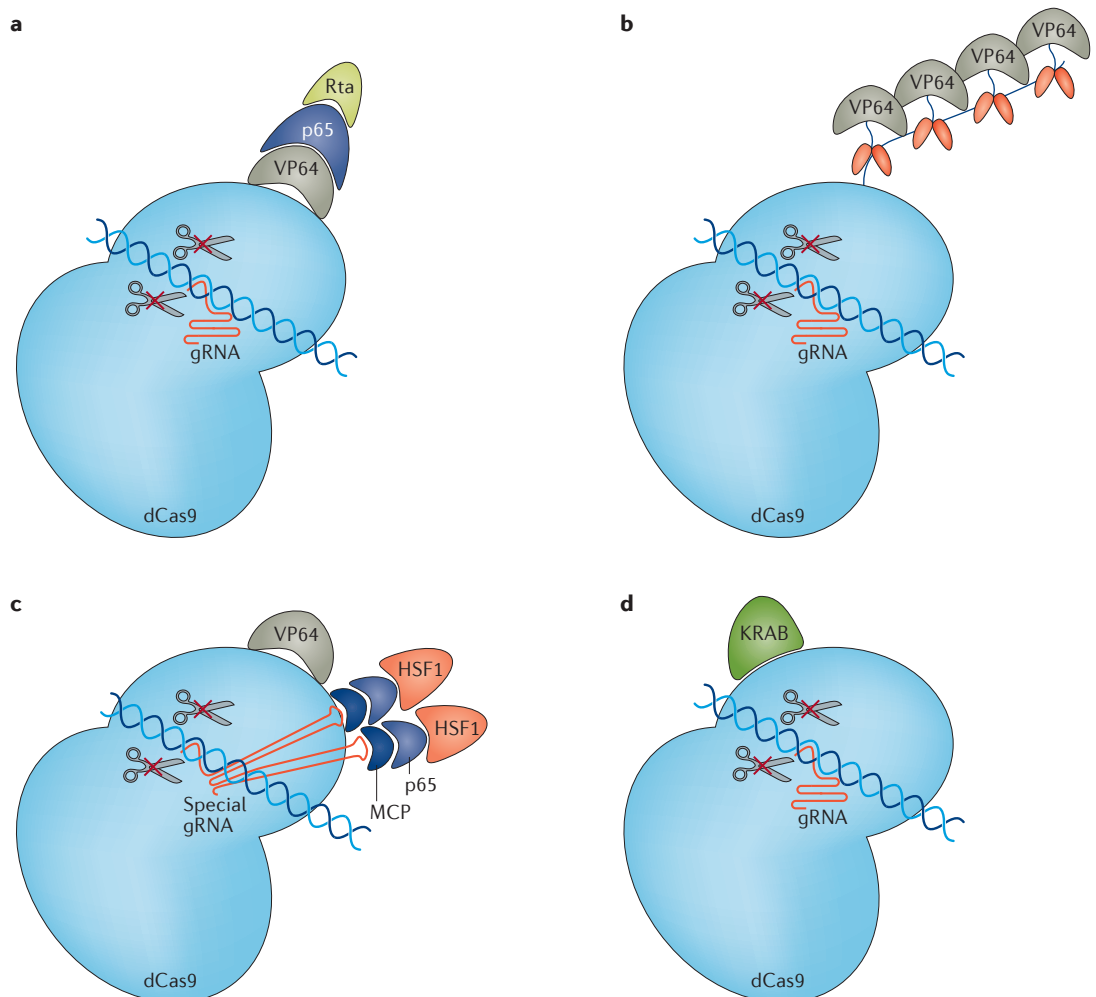
systems associated with rheumatic diseases, the vast majority of which are multigenic. The capacity offered by dCas9 and multiplexed CRISPR systems to activate and repress multiple genes simultaneously should begin to break down barriers to understanding and, eventually, the manipulation of rheumatic diseases.

### Applications of CRISPR/Cas9

#### Gene editing of mammalian cells

The simplest and, to date, most common application of the CRISPR/Cas system is in the generation of cell lines with complete and permanent loss of function of target genes. This application requires only transfection with plasmids containing *cas9* and the gRNA targeting the desired gene. For many cell types (particularly chondrocytes), transfection can be inefficient, partly due to the large size of *cas9* (4,104 bp)<sup>22</sup>. The creation of permanent cell lines expressing *cas9* in safe genomic loci overcomes the problems associated with transfection and creates a versatile and useful tool that can be used to create an almost unlimited range of cell lines with targeted mutations. A permanent chondrocyte cell line has been established that expresses Cas9 (rat chondrosarcoma Cas9, or RCS Cas9)<sup>39</sup>. Subsequent editing of a target locus requires only transfection of this cell line with a specific gRNA. Transfection with gRNA complementary to the third exon of the aggrecan gene resulted in indel mutations in >80% of transfected cells<sup>39</sup>. Most indels were predicted to generate premature stop codons and, because the target site was near the 5' end of the gene, resulted in loss of expression of the aggrecan core protein. Several genes have now been knocked out in RCS Cas9 cells by transfection with specific gRNAs, including *Has2* (REF 40), *Col6a1*, *Col6a2*, *Col6a3*, *Inpp1* and *Kank1* as well as miR-140 (all G.J.G., unpublished data). The associated loss of cell-surface hyaluronan in cells lacking hyaluronan synthase 2 (*Has2*) resulted in loss of the pericellular proteoglycan matrix and helped to define the role of hyaluronan in retention of pericellular matrix<sup>40</sup>.

RCS Cas9 or similar cell lines are expected to facilitate the investigation of the complex interactions regulating chondrocyte function, differentiation, homeostasis and the role of disease-associated genetic traits in cartilage degeneration. One example is the a disintegrin and metalloproteinase with thrombospondin motifs (ADAMTS) proteinase family, which has a central role in cartilage matrix turnover and cell–matrix interactions, although the complex structure and membrane association of ADAMTS proteinases has made identification of their precise role challenging. The generation of cell lines with targeted mutations using CRISPR/Cas9 from cells such as the RCS Cas9 cells could provide an excellent system to identify the specific, detailed function of ADAMTS proteinases and their role in cartilage pathology<sup>41</sup>. A wide variety of genes and noncoding regions that regulate cell function and pathologic changes associated with rheumatic diseases are likely to be identified in the near future using CRISPR/Cas9 technology (BOX 1). The availability of cell lines relevant to rheumatology research, like RCS Cas9, will be vital for verifying the role of these elements and more precisely defining their function.



**Figure 4 | Engineering nuclease-deactivated Cas9 (dCas9) for gene activation and repression.** Several strategies can be used to generate gene activators: **a** | fusion of dCas9 with three activator domains, namely the herpes simplex activation domain (VP64), transcription factor p65 and the Epstein–Barr virus replication and transcription activator (Rta) (together known as VPR); **b** | an array of small peptide epitopes fused to dCas9 to recruit multiple copies of single chain variable fragment fused to VP64 (Sun Tag); or **c** | a modified guide RNA (gRNA) encoding extra loop structures that bind to the MS2 coat protein (MCP) that is in turn fused to p65 and heat shock factor 1 (HSF1). **d** | Similar strategies are employed for the generation of gene repressors and include fusion of dCas9 with the repressor domain Krüppel-associated box (KRAB). Modified with permission from NPG © Chavez, A. *et al. Nat. Methods* **13**, 563–567 (2016).

**Analysis of gene regulatory elements**

Noncoding *cis*-regulatory elements are estimated to make up as much as 10% of the human genome, and evidence suggests that 75% of polymorphisms associated with heritable diseases occur within these sequences<sup>42</sup>. The precise identification of such regulatory elements and our understanding of their role in regulating development-specific, cell-specific and disease-specific expression is very limited. New methods that analyse the entire genome, such as ChIP-seq and DNase-seq, provide extensive identification of distal regulatory elements within a specific cell type<sup>43</sup>. The ability to delete large regions of the genome and prepare daughter cell lines with specific gene edits, multiple edits or gene deletions using CRISPR/Cas9 technology has provided a mechanism to precisely identify regulatory elements and decipher the function of individual enhancer regions and their interactions. For example,

researchers have identified a new class of regulatory elements with temporary enhancer activity that was lost after a few cell divisions and contributed to the complex temporal regulation of cell-specific gene expression<sup>44</sup>. The role of distal enhancer elements that regulate expression of one of the critical matrix degrading enzymes, MMP13 is also beginning to be defined<sup>45</sup>. This study<sup>45</sup> demonstrated the interplay between the transcription factors RUNX2 and C/EBPβ and the vitamin D receptor in regulation of osteoblastic differentiation. Similar studies in chondrocyte cell lines are expected to define the genome environment and interactions that regulate MMP13 expression during cartilage homeostasis, growth and degradation. More broadly, application of this approach will enable the identification of regulatory elements and the definition of their function and interaction in any system or cell type, including those most relevant to rheumatologists.

### Box 1 | Application of CRISPR/Cas9 in rheumatology research

Examples of published applications of the CRISPR/Cas9 technology related to rheumatology research are relatively few, owing to the fact that the technology is only a few years old. A large proportion of the studies employing CRISPR/Cas9 are proof-of-principal for new modifications, developments and adaptations of the CRISPR toolbox.

A chondrocyte cell line expressing Cas9 has been developed that enables quick gene-editing by simple transfection with gRNAs targeting the gene of interest<sup>39</sup>. These cells have been used to demonstrate the role of aggrecan in isolating chondrosarcoma cells from the host immune system<sup>39</sup> and further defining the role of cell-surface hyaluronan<sup>40</sup>.

CRISPR/Cas9 technology has begun to define the interplay of genes and noncoding regions regulating the expression of the central matrix degrading proteinase MMP13 in bone cells<sup>45</sup>.

The use of CRISPR/Cas9 gene screening has identified the genes present in mosquitos and humans that are necessary for the replication of families of viruses<sup>53,54</sup>; together with the development of CRISPR/Cas9 gene drives<sup>82,84</sup>, these findings offer the potential to eliminate insect vectors of diseases, including those with rheumatologic significance, such as Lyme disease.

Gene screening using CRISPR/Cas9 libraries has provided a new understanding of pathways of the innate immune system and identified several critical regulatory genes not previously recognized<sup>55,56,61</sup>. These studies have the potential to provide new targets for the treatment of various rheumatic and associated diseases, including systemic lupus erythematosus<sup>58</sup>, crystal-induced arthritides<sup>59</sup> rheumatoid arthritis and inflammatory bowel disease<sup>60</sup>.

The potential and scope of CRISPR/Cas9 technology is enormous. These few examples of applications to rheumatic diseases will rapidly increase in the near future as new applications that are either underway or proposed in laboratories around the world are published.

Polymorphisms identified by genome-wide association studies (GWAS) are commonly in the form of single nucleotide polymorphisms (SNPs)<sup>42</sup>. The DNA regions typically contain multiple closely spaced SNPs that are co-inherited and thus unable to be distinguished by classical genetic association studies<sup>42,43</sup>. However, CRISPR/Cas9 technology can specifically distinguish genetic variants associated with disease traits from bystander variants and enable researchers to start to decipher their causal role<sup>42,43</sup>. An excellent example of the power of this technology was demonstrated in a study of the role of SNP variants in Parkinson disease<sup>46</sup>. The accumulation of  $\alpha$ -synuclein in the brains of patients with Parkinson disease has been associated with pathophysiology. Using CRISPR/Cas9 genetic modification of multiple SNPs, the authors of this study<sup>46</sup> showed that a specific SNP variant increased expression of  $\alpha$ -synuclein by reducing the binding of transcription inhibitory factors.

The capacity to identify and modify SNPs has received a further major boost with publication of a method that the authors term 'RNA guide tuning'<sup>47</sup>. This technique identifies gRNAs that are able to distinguish target sites differing by only a single base. The authors predict that CRISPR/Cas9 technology using these gRNAs will enable identification of the causal role of thousands of disease-associated SNPs, including those that have been reported for rheumatic diseases<sup>48</sup>. The authors also speculate that in the future these tuned gRNAs could provide a means to disable disease alleles delineated by SNPs, in order to treat the associated disease<sup>47</sup>.

#### CRISPR/Cas9 genomic screen

CRISPR/Cas9-mediated genome-wide knockout screen system. In this platform, a gRNA library targeting most genes in the genome with multiple sites per gene is cloned into lentiviral vectors and delivered as a pool into target cells that express Cas9. A low multiplicity of infection is used to ensure that each cell will receive no more than one gRNA or viral particle. By proper phenotype selection, gRNAs that are enriched or depleted in cells are determined and, correspondingly, genes that are required for that particular phenotype can be systematically identified.

#### Gene drives

Gene drives are genetic manipulations that enable a gene to force its inheritance to all, rather than half, of its offspring.

### Gene Screening

The human genome project provided an almost complete catalogue of our genes<sup>49</sup>. A CRISPR/Cas9 genomic screen now offers a high-throughput method of assigning functions to these genes. Several groups have synthesized genome-wide libraries of gRNAs that target almost the entire human and mouse genomes multiple times<sup>50–52</sup>. Lentiviral libraries either contain gRNAs alone or gRNAs in addition to Cas9. Libraries with gRNAs alone require stable cell lines expressing Cas9, whereas libraries expressing gRNA and Cas9 can be used with almost any cell line (although these libraries require much higher cell numbers than for gRNAs alone owing to the size of the construct and consequent low viral titre). The technique for gene screening requires transduction of target cells with the gRNA library, growth of infected cells in culture and selection for transduced cells by either antibiotic resistance or fluorescence activated cell sorting<sup>50–52</sup>. Cell populations selected for any target phenotype, for example survival, growth, differentiation or resistance to anticancer agents, are then isolated. The gRNA (identified by barcodes using deep sequencing) present or lost from the cell population is compared with the transfected library. Second generation gRNA libraries typically contain >100,000 gRNAs targeting 17,000–20,000 genes, each gene targeted with 5–10 gRNAs. The libraries have been carefully designed for efficient gene-knockout and minimal crossreactivity.

Genome-wide CRISPR/Cas9 screening has focused largely on cancer, but has also identified genes necessary for viral infection and innate immune system pathways. The host genes necessary for dengue and Zika virus infection<sup>53,54</sup> include endoplasmic reticulum peptidases and oligosaccharide transferases that are expressed in both mosquito and human hosts. Most of the genes identified in a hepatitis C virus (another *Flaviviridae* virus) screen are distinct from those of dengue and Zika and include viral receptors, RNA binding proteins and enzymes associated with the conversion of riboflavin to flavin adenine dinucleotide<sup>54</sup>. These studies emphasize the power of the CRISPR/Cas9 screening approach, and the findings represent much needed new pharmacologic targets for inhibition of *Flaviviridae*. In combination with targeted gene drives (described below), they present the exciting potential to control and possibly eliminate vector-borne viral diseases, including those most relevant to rheumatologists, such as Lyme disease and chikungunya.

Of direct interest to rheumatologists, several research groups have employed CRISPR/Cas9 gRNA libraries to identify innate immune system pathways, including a comprehensive unbiased CRISPR/Cas9 analysis to identify genes controlling the induction of TNF in response to dendritic cell stimulation<sup>55</sup>. In this study, the authors used bone-marrow-derived dendritic cells isolated from transgenic mice expressing Cas9. The cells were infected with a library of gRNAs and monitored for gRNA abundance associated with high TNF expression in response to stimulation with lipopolysaccharide. The authors of this study identified and validated the role of many genes not previously associated with innate immune

circuits and described new pathways associated with endoplasmic reticulum stress and the polymerase associated complex, a regulator of transcription elongation, not previously implicated in inflammatory gene expression<sup>55</sup>. In addition, two papers examining the innate immune pathways in macrophages have provided new understanding of pyroptosis and the inflammasome-mediated immune diseases<sup>56,57</sup> implicated in systemic lupus erythematosus<sup>58</sup>, crystal-induced arthritides<sup>59</sup> rheumatoid arthritis, inflammatory bowel disease and rare hereditary periodic fever syndromes<sup>60</sup>. Although the involvement of caspase activation associated with the inflammasome protein complex and downstream activation of proinflammatory cytokines (including IL-1 $\beta$ ) had been recognized previously, the molecular pathways identified upstream of pyroptosis and the mechanisms inducing cell lysis and cytokine release were previously unknown. The two studies<sup>56,57</sup> identified critical roles for Nek7 kinase upstream of inflammasome activation and caspase cleavage of gasdermin in driving cell lysis and the release of inflammatory cytokines. An additional study using the CRISPR/Cas9 system in human macrophages also revealed an alternative inflammasome pathway for secretion of IL-1 $\beta$  that does not seem to be active in mouse macrophages<sup>61</sup>.

These studies, published since 2015, emphasize the utility of the CRISPR/Cas9 system for unbiased description of critical molecular pathways. This technique will provide rapid expansion of our understanding of complex molecular pathways across diverse biological systems, and might reveal new therapeutic targets for a wide variety of diseases, including rheumatic diseases.

#### **Animal models and xenotransplantation**

CRISPR/Cas9 technology has had an enormous effect on the ability to develop mouse models of disease. The technique makes the generation of genetically engineered mice quicker and cheaper than traditional techniques. Generation of transgenic mice using CRISPR/Cas9 takes ~11 weeks, in contrast to the traditional approach using embryonic stem cells that takes ~1 year<sup>62</sup>. In its simplest form the CRISPR/Cas9 approach involves injection of a single plasmid construct consisting of *cas9* and gRNA genes into fertilized mouse oocytes<sup>63</sup>. Targeted sites in the genome are cleaved and mutation rates resulting from error prone NHEJ are high with ~50% of pups affected<sup>63</sup>. CRISPR/Cas9 also enables the generation of complex models with large deletions, inversions and duplications<sup>64</sup>. The technique enables the generation of mice carrying mutations in multiple genes<sup>65</sup> and the disruption of large topological domains<sup>64</sup> that would be very difficult and time consuming to generate by traditional methods. Many hundreds of genetically engineered mice have been generated using this technique<sup>66</sup>.

As with all areas of CRISPR/Cas9 technology, improvements in generating mouse models of human disease are advancing rapidly. The development of an adult-onset and tissue-specific model of heart disease<sup>67</sup> has opened the door for simple and efficient development of temporally and tissue-specific models of other human diseases. Carroll *et al.*<sup>67</sup> described the generation

of a transgenic mouse line expressing Cas9 exclusively in cardiomyocytes with no overt effects. Delivery of gRNAs targeting the gene encoding cardiac myosin heavy chain 6, *Myh6* using adeno-associated virus (AAV) in adult mice demonstrated high levels of cardiac-specific mutation and cardiac failure<sup>67</sup>. The use of mice with tissue-specific expression of Cas9 overcomes the difficulty of delivering components of the CRISPR/Cas9 complex that are at the packaging limit of many viral delivery systems. In addition, this method enables analysis of the function of any cardiac gene, including those that are embryonic lethal or widely expressed in other tissues, by the simple delivery of specific gRNAs. Mice with expression of Cas9 specifically in other tissues are expected to be developed and will enable analysis of gene function in a wide variety of adult-onset diseases. Mice with cartilage-specific or joint-specific expression of Cas9 would be invaluable in the description of the gene function associated with a wide variety of rheumatic diseases. This idea is made more tantalizing by the flexibility of the CRISPR/Cas9 system to mutate multiple coding and noncoding sites in a genome.

CRISPR/Cas9 technology might overcome our reliance on mouse models of disease, and several models of disease and disease resistance have been generated in other species including goats, cattle, ferrets, fish, monkeys and even elephants<sup>68</sup>. A surprisingly large number of studies have reported the use of CRISPR/Cas9 to engineer mutations in domestic species, particularly in pigs, sheep, cattle and goats<sup>69–73</sup>. Most studies employ somatic cell nuclear transfer. The technique involves editing the desired gene in a somatic cell, usually fibroblasts, and replacing the nucleus in isolated oocytes with the nucleus of the gene-edited somatic cell. Gene editing can also be performed in embryos by direct injection of CRISPR components into the pronucleus or cytoplasm (reviewed elsewhere<sup>74,75</sup>).

One application that has received a lot of research and commercial interest addresses the growing demand for human tissue for transplantation and the chronic shortage of organ donors. For many years scientists proposed the use of pig organs for transplantation with such enthusiasm that several companies were established with this goal in mind. However, this work ran into two major obstacles: namely, endogenous viruses and immune incompatibility. The ability of the CRISPR/Cas9 system to delete multiple copies of a gene in a single cell system has enabled eradication of 60 copies of the family of porcine retroviruses from the pig genome and enabled the deletion or mutation of 20 genes known to trigger a human immune response<sup>69</sup>. These advances suggest porcine cartilage and bone for human transplantation might provide a radically new approach for treatment of end-stage arthritides.

A study examining the role of the Mohawk (Mkx) transcription factor has demonstrated the value in moving from transgenic mice to larger animals, in this case transgenic rats. Studies in mice suggested an important role in tendon development, with the *Mkx* mouse knockout resulting in tendon hypoplasia<sup>76,77</sup>. The rat *Mkx*<sup>-/-</sup> (generated by direct injection of *cas9* and gRNA



into fertilized oocytes) had similar hypoplasia to the mice, but onset was earlier and more severe<sup>78</sup>. Unlike the *Mkx*<sup>-/-</sup> mice, the *Mkx*<sup>-/-</sup> rats developed chondral lesions and heterotopic ossification of the Achilles tendon. As well as providing access to more tissue and the capacity to conduct more cell-based and biochemical experiments, the larger size of the animal, the authors suggested, increased the mechanical stimulation to the tendons which resulted in chondrogenic differentiation and a more severe phenotype<sup>78</sup>. For diseases with skeletal pathology, the transition to larger animal models made feasible by CRISPR/Cas9 technology is expected to provide a quantum leap in our understanding and ability to model human diseases.

### Gene drives

During normal sexual reproduction the copy of a gene inherited from one parent will not spread through a wild population because in each generation there is only a 50% chance of passing it on. However, if the gene is modified so that it causes the gene from the other parent to be modified in the same way, the offspring will always receive the modified gene and the gene will rapidly spread through a wild population, potentially reaching 100% of the population within a few generations. These gene drives (as they have been termed) have been proposed as a way of eliminating disease vectors, controlling invasive species, immunizing threatened species and generating crops with resistance to herbicides<sup>79,80</sup>. Although proposed many years ago little progress had been made in the development of gene drives, primarily because of the difficulty in precisely engineering genomes. The development of the CRISPR/Cas9 system has caused a rapid reversal of that inactivity. Several publications have described model gene drives confined to the laboratory in fruitflies and mosquitoes<sup>81–84</sup>. The systems employed in these studies were similar and comprised a construct with three components: a *cas9* gene, a gRNA targeting the sequence of interest and homology arms enabling Cas9 cleavage of the second allele and expression of the Cas9-gRNA via HDR. Three genes were targeted in *Anopheles gambiae*<sup>84</sup> and *Anopheles stephensi*<sup>82</sup> (the main malaria vectors) with transmission rates of 90–99%. The studies suggest that this approach, in combination with the identification of the genes that are essential for viral replication, could be developed to eliminate malaria from affected regions<sup>82,84</sup>. Similar approaches might be used to target other vector transmitted diseases including Lyme disease, hepatitis C, dengue and Zika virus.

The capacity to wipe out or at least drastically alter entire wild populations clearly has serious ecological concerns. These concerns have received wide attention and have been the subject of a National Academy of Sciences Report<sup>85,86</sup>. The use of gene drives remains controversial and studies to date are not sufficient to allow the release of gene-drive-modified organisms into the environment.

### Treating human disease

A cancer study in which one gene was inserted and another deleted in T cells was the first using CRISPR/Cas9 technology to pass the first committee (NIH Recombinant

DNA Advisory Committee) on its way to clinical trials<sup>87</sup>. An additional similar trial is about to begin in China<sup>88</sup>. These researchers plan to use CRISPR technology to insert a receptor for a protein often expressed in tumours, but not in healthy cells, into patients' T cells, and to delete PD-1, a T cell surface protein that has been shown to dampen cell activity after an immune response. These trials are based on very promising studies in mice<sup>89</sup> that demonstrated tumour regression using this approach.

Several other promising studies in animal models suggest that CRISPR/Cas9 technology will soon be applied to treat diseases that affect humans (TABLE 1). Three concurrent publications provide proof of principle that CRISPR/Cas9 technology can be used to correct Duchenne muscular dystrophy (DMD)<sup>90–92</sup>. DMD is a devastating progressive muscle disease that is typically caused by small mutations in the dystrophin gene, resulting in the generation of a premature stop codon and thus the loss of protein expression. Dystrophin is a very large muscle protein composed of many domains, some of which are dispensable for protein function. Because most mutations affecting patients with DMD occur in these non-essential protein regions it has been proposed that exon skipping strategies would provide effective treatment in the majority of patients. The authors of these three studies investigated a mouse model of the human disease with a nonsense mutation in exon 23 of the dystrophin gene<sup>90–92</sup>. AAV delivery of *cas9* and two gRNAs targeting the 3' and 5' ends of exon 23 caused skipping of this exon and expression of a functional dystrophin protein. Postnatal systemic delivery of AAV vectors also restored dystrophin expression and enhanced muscle function. Dystrophin expression in muscle progenitors was consistent and continued improvement in muscle function was maintained for at least 6 months post-treatment. With continued development to enhance safety and efficacy it is hoped that this technology will realize its promise in treating patients with DMD.

Delivery of the CRISPR components remains one of the challenges for treatment of human disease (CRISPR delivery is reviewed in detail elsewhere<sup>22,93,94</sup>). AAV delivery shows great promise; however, the relatively small packing capacity of AAV presents a problem. The most widely used *cas9* comes from *Streptococcus pyogenes* and is at the limits of AAV's packing capacity. A smaller *cas9* isolated from *Staphylococcus aureus* should overcome these problems, but has not been widely used to date<sup>95</sup>. In addition, the long-term persistent expression of the CRISPR system that results from viral delivery exaggerates the problems of off-target effects. Even very low levels of off-target cleavage could become problematic if expression is maintained for the extended periods that occur with viral delivery. Transient expression can be achieved by use of lipid nanoparticles for delivery (reviewed elsewhere<sup>22</sup>). Nanoparticle delivery of Cas9 with AAV delivery of gRNAs and a repair template has shown success in the treatment of a mouse model of tyrosinaemia, achieving therapeutic correction in 6% of hepatocytes after one treatment<sup>96</sup>. Furthermore, a very encouraging chondroprotective effect has been obtained using intra-articular nanoparticle delivery

Table 1 | Preclinical research using CRISPR/Cas9

| Disease                                   | Protein (gene)  | Platform                  | Delivery   | Refs  |
|---|---|---------------------------|--|-------|
| Metabolic liver disease                   | Ornithine carbamoyltransferase ( <i>Otc</i> )                       | Mouse model               | Intravenous delivery of two AAVs; HDR  | 102   |
| Retinitis pigmentosa                      | X-Linked retinitis pigmentosa GTPase regulator ( <i>RPGR</i> )      | Patient iPSCs             | Transfection; HDR  | 103   |
| Dominant dystrophic epidermolysis bullosa | Collagen alpha-1(VII) chain ( <i>COL7A1</i> )                       | Patient iPSCs             | Transfection; NHEJ   | 104   |
| Hereditary tyrosinemia                    | Fumarylacetoacetase ( <i>Fah</i> )                                  | Mouse model               | Lipid nanoparticle delivery of Cas9; AAV delivery of gRNA and repair template; systemic injection. | 96    |
| Fanconi anaemia                           | Fanconi anemia group I protein ( <i>FANCI</i> )                     | Patient iPSCs             | Cas9 nickase; repair template transfection; hematopoietic differentiation                          | 105   |
| Duchenne muscular dystrophy               | Dystrophin ( <i>Dmd</i> )   | Mouse model               | Exon deletion using paired gRNAs; AAV systemic and muscle delivery                                 | 90–92 |
| Hepatitis B virus                         | Cell surface antigen  | Cell culture, Mice        | Transfection; NHEJ   | 106   |
| Sickle-cell anaemia                       | Haemoglobin ( <i>Hbb</i> )  | Mouse model               | <i>Ex vivo</i> repair of haematopoietic stem cells and transplantation back into mice.             | 107   |
| β-Thalassemia                             | Haemoglobin ( <i>HBB</i> )  | Patient iPSCs             | iPSCs derived from patient fibroblasts; HDR; excisable antibiotic selection                        | 108   |
| Cystic fibrosis                           | Cystic fibrosis transmembrane conductance regulator ( <i>CFTR</i> ) | Patient iPSCs             | Skin iPSCs; HDR; excisable antibiotic selection; differentiation to lung epithelial cells          | 109   |
| Latent HIV infection                      | C–C chemokine receptor type 5 ( <i>CCR5</i> )                       | Haematopoietic stem cells | CCR5 ablation in haematopoietic stem cells   | 110   |

Abbreviations: AAV, adeno-associated virus; gRNA, guide RNA; HDR, homology-directed repair; iPSC, induced pluripotent stem cell; NHEJ, non-homologous end joining.

of NF-κB siRNAs for treatment of a mouse model of post-traumatic osteoarthritis<sup>97</sup>. Importantly, this peptide nanoparticle complex delivered RNA deep within the cartilage, providing a unique vehicle for the treatment of cartilage diseases<sup>97</sup>. Nanoparticle constructs containing Cas9 protein and gRNA (ribonucleoprotein complexes) have been shown to provide very efficient gene manipulation in cells<sup>98</sup>. Although untested to date, the intra-articular delivery of CRISPR ribonucleoprotein complexes using similar peptide nanoparticles offers a unique opportunity to treat arthritic joint diseases.

Genetic diseases, such as DMD, will probably comprise the first focus of the therapeutic use of CRISPR technology. With the rapid rate of improvements and diversification in CRISPR technology and delivery, it is expected that many of the promising prospects for treatment of human genetic disease will reach clinical trials. We optimistically anticipate that as CRISPR/Cas9 technology expands our understanding of more complex diseases and reveals new treatment strategies, our understanding and capacity to treat the more common human diseases, including rheumatic diseases, will also expand.

### Conclusions

CRISPR/Cas9 technology is transforming molecular biology in a way similar to how PCR transformed it >30 years ago. The application of PCR for the analysis of gene expression and function, diagnosis of disease, DNA cloning, phylogeny associations, molecular fingerprinting and much more has made a critical contribution to our capacity to perform analyses and test gene function. Similarly, the ability to quickly and simply edit the genome using rapidly expanding CRISPR/Cas9

technologies across a wide range of systems from cell culture to animal models has already transformed our capacity to address fundamental previously intractable questions of normal gene function and the molecular pathogenesis of diseases. CRISPR/Cas9 technologies are showing enormous promise in untangling the interactions of gene networks responsible for the regulation of specific biologic and disease functions. GWAS have identified many hundreds of variants associated with diseases, including rheumatic diseases<sup>48,99</sup>. Large population studies have identified common variants associated with immune traits<sup>100</sup> and provided detailed analysis of common variants among autoimmune diseases<sup>100,101</sup>. CRISPR/Cas9 technology provides genome-wide high-throughput screening and fine mapping of these regions, including the ability to distinguish causal and bystander SNPs. As the application of these analyses to research questions continues, and advances in CRISPR/Cas9 and other CRISPR technologies are developed, their effect will no doubt be enormous. These advances are anticipated to offer a new and exciting conceptual understanding of the complex regulation of cell function and the molecular pathogenesis of human diseases. Like PCR, the CRISPR technologies will rapidly become accepted as just another critical component of the molecular biology tool kit.

The capacity of CRISPR/Cas9 to repair endogenous genes while preserving the physiological regulation of gene expression offers the potential for human gene therapy. Although the efficiency and accuracy of CRISPR-based gene editing is bound to continue to advance, the primary challenges of gene therapy (probably the greatest of which are the development of safe and efficient delivery systems) remain to be

overcome before extensive application of gene therapy using this technology can be widely employed. The initial PCR technique was very cumbersome and time consuming until heat stable polymerases transformed the technique to the quick, simple, routine procedure

that is used in every laboratory. Although CRISPR technology is already transformative, additional new technologies, such as new gene delivery techniques, might be expected to have a similar additive effects on its applications.

- Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012).
- Mali, P. *et al.* RNA-guided human genome engineering via Cas9. *Science* **339**, 823–826 (2013).
- Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823 (2013).
- Cho, S. W., Kim, S., Kim, J. M. & Kim, J. S. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat. Biotechnol.* **31**, 230–232 (2013).
- Hwang, W. Y. *et al.* Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat. Biotechnol.* **31**, 227–229 (2013).
- Jiang, W., Bikard, D., Cox, D., Zhang, F. & Marraffini, L. A. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat. Biotechnol.* **31**, 233–239 (2013).
- Hsu, P. D. *et al.* DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* **31**, 827–832 (2013).
- Gasiunas, G., Barrangou, R., Horvath, P. & Siksnys, V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc. Natl Acad. Sci. USA* **109**, E2579–E2586 (2012).
- Ishino, Y., Shinagawa, H., Makino, K., Amemura, M. & Nakata, A. Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *J. Bacteriol.* **169**, 5429–5433 (1987).
- Jansen, R., Embden, J. D., Gaastera, W. & Schouls, L. M. Identification of genes that are associated with DNA repeats in prokaryotes. *Mol. Microbiol.* **43**, 1565–1575 (2002).
- Sorek, R., Lawrence, C. M. & Wiedenheft, B. CRISPR-mediated adaptive immune systems in bacteria and archaea. *Annu. Rev. Biochem.* **82**, 237–266 (2013).
- Mojica, F. J., Diez-Villasenor, C., Garcia-Martinez, J. & Soria, E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J. Mol. Evol.* **60**, 174–182 (2005).
- Marraffini, L. A. CRISPR-Cas immunity in prokaryotes. *Nature* **526**, 55–61 (2015).
- Eberl, G. Immunology: close encounters of the second type. *Nature* **464**, 1285–1286 (2010).
- Makarova, K. S. *et al.* An updated evolutionary classification of CRISPR-Cas systems. *Nat. Rev. Microbiol.* **13**, 722–736 (2015).
- Kim, D. *et al.* Digenome-seq: genome-wide profiling of CRISPR-Cas9 off-target effects in human cells. *Nat. Methods* **12**, 237–243 (2015).
- Doench, J. G. *et al.* Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat. Biotechnol.* **34**, 184–191 (2016).
- Fu, Y., Sander, J. D., Reyon, D., Casicio, V. M. & Joung, J. K. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat. Biotechnol.* **32**, 279–284 (2014).
- Ran, F. A. *et al.* Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* **154**, 1380–1389 (2013).
- Kleinstiver, B. P. *et al.* High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* **529**, 490–495 (2016).
- Slaymaker, I. M. *et al.* Rationally engineered Cas9 nucleases with improved specificity. *Science* **351**, 84–88 (2016).
- Gori, J. L. *et al.* Delivery and specificity of CRISPR-Cas9 genome editing technologies for human gene therapy. *Hum. Gene Ther.* **26**, 443–451 (2015).
- 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).
- Hudson, D. F., Morrison, C., Ruchaud, S. & Earnshaw, W. C. Reverse genetics of essential genes in tissue-culture cells: 'dead cells talking'. *Trends Cell Biol.* **12**, 281–287 (2002).
- Rouet, P., Smih, F. & Jasin, M. Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells. *Proc. Natl Acad. Sci. USA* **91**, 6064–6068 (1994).
- Chu, V. T. *et al.* Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. *Nat. Biotechnol.* **33**, 543–548 (2015).
- Wang, B. *et al.* Highly efficient CRISPR/HDR-mediated knock-in for mouse embryonic stem cells and zygotes. *Biotechniques* **59**, 201–208 (2015).
- He, X. *et al.* Knock-in of large reporter genes in human cells via CRISPR/Cas9-induced homology-dependent and independent DNA repair. *Nucleic Acids Res.* **44**, e85 (2016).
- Richardson, C. D., Ray, G. J., DeWitt, M. A., Curie, G. L. & Corn, J. E. Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. *Nat. Biotechnol.* **34**, 339–344 (2016).
- Gilbert, L. A. *et al.* CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* **154**, 442–451 (2013).
- Dominguez, A. A., Lim, W. A. & Qi, L. S. Beyond editing: repurposing CRISPR-Cas9 for precision genome regulation and interrogation. *Nat. Rev. Mol. Cell Biol.* **17**, 5–15 (2016).
- Kiani, S. *et al.* Cas9 gRNA engineering for genome editing, activation and repression. *Nat. Methods* **12**, 1051–1054 (2015).
- Chavez, A. *et al.* Comparison of Cas9 activators in multiple species. *Nat. Methods* **13**, 563–567 (2016).
- Konermann, S. *et al.* Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* **517**, 583–588 (2015).
- Gilbert, L. A. *et al.* Genome-scale CRISPR-mediated control of gene repression and activation. *Cell* **159**, 647–661 (2014).
- Kearns, N. A. *et al.* Functional annotation of native enhancers with a Cas9-histone demethylase fusion. *Nat. Methods* **12**, 401–403 (2015).
- Hilton, I. B. *et al.* Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nat. Biotechnol.* **33**, 510–517 (2015).
- Thakore, P. I. *et al.* Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. *Nat. Methods* **12**, 1143–1149 (2015).
- Yang, M., Zhang, L., Stevens, J. & Gibson, G. CRISPR/Cas9 mediated generation of stable chondrocyte cell lines with targeted gene knockouts: analysis of an aggregon knockout cell line. *Bone* **69**, 118–125 (2014).
- Huang, Y., Askew, E. B., Knudson, C. B. & Knudson, W. CRISPR/Cas9 knockout of HAS2 in rat chondrosarcoma chondrocytes demonstrates the requirement of hyaluronan for aggregan retention. *Matrix Biol.* **56**, 74–94 (2016).
- Dubail, J. & Apte, S. S. Insights on ADAMTS proteases and ADAMTS-like proteins from mammalian genetics. *Matrix Biol.* **44–46**, 24–37 (2015).
- Tak, Y. G. & Farnham, P. J. Making sense of GWAS: using epigenomics and genome engineering to understand the functional relevance of SNPs in non-coding regions of the human genome. *Epigenetics Chromatin* **8**, 57 (2015).
- Yao, L., Berman, B. P. & Farnham, P. J. Demystifying the secret mission of enhancers: linking distal regulatory elements to target genes. *Crit. Rev. Biochem. Mol. Biol.* **50**, 550–573 (2015).
- Diao, Y. *et al.* A new class of temporarily phenotypic enhancers identified by CRISPR/Cas9-mediated genetic screening. *Genome Res.* **26**, 397–405 (2016).
- Meyer, M. B., Benkusky, N. A. & Pike, J. W. Selective distal enhancer control of the *Mmp13* gene identified through clustered regularly interspaced short palindromic repeat (CRISPR) genomic deletions. *J. Biol. Chem.* **290**, 11093–11107 (2015).
- Soldner, F. *et al.* Parkinson-associated risk variant in distal enhancer of alpha-synuclein modulates target gene expression. *Nature* **533**, 95–99 (2016).
- Chavez, A. *et al.* Precise Cas9 targeting enables genomic mutation prevention. Preprint at *bioRxiv* <http://dx.doi.org/10.1101/058974> (2016).
- Kirino, Y. & Remmers, E. F. Genetic architectures of seropositive and seronegative rheumatic diseases. *Nat. Rev. Rheumatol.* **11**, 401–414 (2015).
- Birney, E. & Soranzo, N. Human genomics: the end of the start for population sequencing. *Nature* **526**, 52–53 (2015).
- Wang, T., Wei, J. J., Sabatini, D. M. & Lander, E. S. Genetic screens in human cells using the CRISPR-Cas9 system. *Science* **343**, 80–84 (2014).
- Shalem, O. *et al.* Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* **343**, 84–87 (2014).
- Zhou, Y. *et al.* High-throughput screening of a CRISPR/Cas9 library for functional genomics in human cells. *Nature* **509**, 487–491 (2014).
- Zhang, R. *et al.* A CRISPR screen defines a signal peptide processing pathway required by flaviviruses. *Nature* **535**, 164–168 (2016).
- Marceau, C. D. *et al.* Genetic dissection of Flaviviridae host factors through genome-scale CRISPR screens. *Nature* **535**, 159–163 (2016).
- Parnas, O. *et al.* A genome-wide CRISPR screen in primary immune cells to dissect regulatory networks. *Cell* **162**, 675–686 (2015).
- Schmid-Burgk, J. L. *et al.* A genome-wide CRISPR (clustered regularly interspaced short palindromic repeats) screen identifies NEK7 as an essential component of NLRP3 inflammasome activation. *J. Biol. Chem.* **291**, 103–109 (2016).
- Shi, J. *et al.* Cleavage of GSMDM by inflammatory caspases determines pyroptotic cell death. *Nature* **526**, 660–665 (2015).
- Reeves, W. H. Editorial: systemic lupus erythematosus: death by fire and ICE? *Arthritis Rheumatol.* **66**, 6–9 (2014).
- So, A., Ives, A., Joosten, L. A. & Busso, N. Targeting inflammasomes in rheumatic diseases. *Nat. Rev. Rheumatol.* **9**, 391–399 (2013).
- Saavedra, P. H., Demon, D., Van Gorp, H. & Lamkanfi, M. Protective and detrimental roles of inflammasomes in disease. *Semin. Immunopathol.* **37**, 313–322 (2015).
- Gaidt, M. M. *et al.* Human monocytes engage an alternative inflammasome pathway. *Immunity* **44**, 833–846 (2016).
- Parikh, B. A., Beckman, D. L., Patel, S. J., White, J. M. & Yokoyama, W. M. Detailed phenotypic and molecular analyses of genetically modified mice generated by CRISPR-Cas9-mediated editing. *PLoS ONE* **10**, e0116484 (2015).
- Henaoui-Mejia, J. *et al.* Generation of genetically modified mice using the CRISPR-Cas9 genome-editing system. *Cold Spring Harb. Protoc.* <http://dx.doi.org/10.1101/pdb.prot090704> (2016).
- Lupianez, D. G. *et al.* Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* **161**, 1012–1025 (2015).
- Wang, H. *et al.* One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* **153**, 910–918 (2013).
- Bono, J. M., Olesnicky, E. C. & Matzkin, L. M. Connecting genotypes, phenotypes and fitness: harnessing the power of CRISPR/Cas9 genome editing. *Mol. Ecol.* **24**, 3810–3822 (2015).
- Carroll, K. J. *et al.* A mouse model for adult cardiac-specific gene deletion with CRISPR/Cas9. *Proc. Natl Acad. Sci. USA* **113**, 338–343 (2016).
- Reardon, S. Welcome to the CRISPR zoo. *Nature* **531**, 160–163 (2016).
- Yang, L. *et al.* Genome-wide inactivation of porcine endogenous retroviruses (PERVs). *Science* **350**, 1101–1104 (2015).
- Zhou, X. *et al.* Generation of CRISPR/Cas9-mediated gene-targeted pigs via somatic cell nuclear transfer. *Cell. Mol. Life Sci.* **72**, 1175–1184 (2015).

71. Wu, M. *et al.* Rosa26-targeted sheep gene knock-in via CRISPR-Cas9 system. *Sci. Rep.* **6**, 24360 (2016).
72. Wang, Z. Genome engineering in cattle: recent technological advancements. *Chromosome Res.* **23**, 17–29 (2015).
73. Ni, W. *et al.* Efficient gene knockout in goats using CRISPR/Cas9 system. *PLoS ONE* **9**, e106718 (2014).
74. Whitelaw, C. B., Sheets, T. P., Lillico, S. G. & Telugu, B. P. Engineering large animal models of human disease. *J. Pathol.* **238**, 247–256 (2016).
75. Tu, Z., Yang, W., Yan, S., Guo, X. & Li, X. J. CRISPR/Cas9: a powerful genetic engineering tool for establishing large animal models of neurodegenerative diseases. *Mol. Neurodegener.* **10**, 35 (2015).
76. Yokoyama, S. *et al.* A systems approach reveals that the myogenesis genome network is regulated by the transcriptional repressor RP58. *Dev. Cell* **17**, 836–848 (2009).
77. Shimizu, H. *et al.* The AERO system: a 3D-like approach for recording gene expression patterns in the whole mouse embryo. *PLoS ONE* **8**, e75754 (2013).
78. Suzuki, H. *et al.* Gene targeting of the transcription factor Mohawk in rats causes heterotopic ossification of Achilles tendon via failed tenogenesis. *Proc. Natl Acad. Sci. USA* **113**, 7840–7845 (2016).
79. Esvelt, K. M., Smidler, A. L., Catteruccia, F. & Church, G. M. Concerning RNA-guided gene drives for the alteration of wild populations. *eLife* **3**, e03401 (2014).
80. Champer, J., Buchman, A. & Akbari, O. S. Cheating evolution: engineering gene drives to manipulate the fate of wild populations. *Nat. Rev. Genet.* **17**, 146–159 (2016).
81. Basu, S. *et al.* Silencing of end-joining repair for efficient site-specific gene insertion after TALEN/CRISPR mutagenesis in *Aedes aegypti*. *Proc. Natl Acad. Sci. USA* **112**, 4038–4043 (2015).
82. Gantz, V. M. *et al.* Highly efficient Cas9-mediated gene drive for population modification of the malaria vector mosquito *Anopheles stephensi*. *Proc. Natl Acad. Sci. USA* **112**, E6736–E6743 (2015).
83. Gantz, V. M. & Bier, E. Genome editing. The mutagenic chain reaction: a method for converting heterozygous to homozygous mutations. *Science* **348**, 442–444 (2015).
84. Hammond, A. *et al.* A CRISPR-Cas9 gene drive system targeting female reproduction in the malaria mosquito vector *Anopheles gambiae*. *Nat. Biotechnol.* **34**, 78–83 (2016).
85. Akbari, O. S. *et al.* BIOSAFETY. Safeguarding gene drive experiments in the laboratory. *Science* **349**, 927–929 (2015).
86. National Academies of Sciences, Engineering, and Medicine. *Gene Drives on the Horizon: Advancing Science, Navigating Uncertainty, and Aligning Research with Public Values* (National Academies Press, 2016).
87. Reardon, S. First CRISPR clinical trial gets green light from US panel. *Nature* <http://www.nature.com/news/first-crispr-clinical-trial-gets-green-light-from-us-panel-1.20137> (2016).
88. Cyranoski, D. Chinese scientists to pioneer first human CRISPR trial. *Nature* **535**, 476–477 (2016).
89. Ren, J. *et al.* Multiplex Cripsr/Cas9 genome editing to generate potent universal CART and PD1-deficient cells against leukemia. *Blood* **126**, 4280–4280 (2015).
90. Tabebordbar, M. *et al.* In vivo gene editing in dystrophic mouse muscle and muscle stem cells. *Science* **351**, 407–411 (2016).
91. Nelson, C. E. *et al.* In vivo genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy. *Science* **351**, 403–407 (2016).
92. Long, C. *et al.* Postnatal genome editing partially restores dystrophin expression in a mouse model of muscular dystrophy. *Science* **351**, 400–403 (2016).
93. Juliano, R. L. The delivery of therapeutic oligonucleotides. *Nucleic Acids Res.* **44**, 6518–6548 (2016).
94. Li, L., He, Z. Y., Wei, X. W., Gao, G. P. & Wei, Y. Q. Challenges in CRISPR/CAS9 delivery: potential roles of nonviral vectors. *Hum. Gene Ther.* **26**, 452–462 (2015).
95. Ran, F. A. *et al.* In vivo genome editing using *Staphylococcus aureus* Cas9. *Nature* **520**, 186–191 (2015).
96. Yin, H. *et al.* Therapeutic genome editing by combined viral and non-viral delivery of CRISPR system components in vivo. *Nat. Biotechnol.* **34**, 328–333 (2016).
97. Yan, H. *et al.* Suppression of NF-kappaB activity via nanoparticle-based siRNA delivery alters early cartilage responses to injury. *Proc. Natl Acad. Sci. USA* **113**, E6199–E6208 (2016).
98. Liang, X. *et al.* Rapid and highly efficient mammalian cell engineering via Cas9 protein transfection. *J. Biotechnol.* **208**, 44–53 (2015).
99. Wang, S., Wiley, G. B., Kelly, J. A. & Gaffney, P. M. Disease mechanisms in rheumatology — tools and pathways: defining functional genetic variants in autoimmune diseases. *Arthritis Rheumatol.* **67**, 1–10 (2015).
100. Roederer, M. *et al.* The genetic architecture of the human immune system: a bioresource for autoimmunity and disease pathogenesis. *Cell* **161**, 387–403 (2015).
101. Cotsapas, C. & Hafler, D. A. Immune-mediated disease genetics: the shared basis of pathogenesis. *Trends Immunol.* **34**, 22–26 (2013).
102. Yang, Y. *et al.* A dual AAV system enables the Cas9-mediated correction of a metabolic liver disease in newborn mice. *Nat. Biotechnol.* **34**, 334–338 (2016).
103. Bassuk, A. G., Zheng, A., Li, Y., Tsang, S. H. & Mahajan, V. B. Precision medicine: genetic repair of retinitis pigmentosa in patient-derived stem cells. *Sci. Rep.* **6**, 19969 (2016).
104. Shinkuma, S., Guo, Z. & Christiano, A. M. Site-specific genome editing for correction of induced pluripotent stem cells derived from dominant dystrophic epidermolysis bullosa. *Proc. Natl Acad. Sci. USA* **113**, 5676–5681 (2016).
105. Osborn, M. J., Belanto, J. J., Tolar, J. & Voytas, D. F. Gene editing and its application for hematological diseases. *Int. J. Hematol.* **104**, 18–28 (2016).
106. Zhen, S. *et al.* Harnessing the clustered regularly interspaced short palindromic repeat (CRISPR)/CRISPR-associated Cas9 system to disrupt the hepatitis B virus. *Gene Ther.* **22**, 404–412 (2015).
107. DeWitt, M. A. *et al.* Selection-free genome editing of the sickle mutation in human adult hematopoietic stem/progenitor cells. *Sci. Transl. Med.* **8**, 360ra134 (2016).
108. Xie, F. *et al.* Seamless gene correction of beta-thalassemia mutations in patient-specific iPSCs using CRISPR/Cas9 and piggyBac. *Genome Res.* **24**, 1526–1533 (2014).
109. Firth, A. L. *et al.* Functional gene correction for cystic fibrosis in lung epithelial cells generated from patient iPSCs. *Cell Rep.* **12**, 1385–1390 (2015).
110. Mandal, P. K. *et al.* Efficient ablation of genes in human hematopoietic stem and effector cells using CRISPR/Cas9. *Cell Stem Cell* **15**, 643–652 (2014).

### Author contributions

G.J.G. contributed to researching data for the article, writing, and reviewing and editing the manuscript before submission. M.Y. contributed to discussion of content and writing the article.

### Competing interests statement

The authors declare no competing interests.