

# Section 1

Genome editing

# Section 1 – Genome editing

## Overview

This first section examines the concept of genome editing and its origins in biological research. Genome editing is set in the context of the range of techniques that the life sciences have afforded to allow deliberate influence over organisms and biological materials. It is characterised by its level of action (nucleotide sequences and epigenetic marks), the precision with which it may be targeted, and its controllability.

To elucidate the mode of action of genome editing techniques, the role of DNA (and RNA) in organisms is described and the concepts of 'gene', 'genome' and 'epigenome' are discussed. The difficulties involved in defining these, and the different registers in which they are presented (molecular, informational, functional, genealogical, etc.) are noted. The reproductive re-assortment of DNA and sources of errors are described, as they may lead to variation and, in some cases, to disease. The complexity of associating genetic variation with phenotypic outcomes is noted.

Current techniques of genome editing are described in the context of prior art. Zinc Finger Nucleases (ZFNs), Transcription Activator-Like Effector Nucleases (TALENs) and the CRISPR-Cas9 system make double-stranded breaks in DNA that cells repair using inbuilt pathways. The section outlines how this may be harnessed to 'knock out' genes through non-homologous end-joining, or to insert or remove specific DNA sequences through homology directed repair. The relative advantages and limitations of the different existing approaches, including epigenome editing with Cas9 derivatives, are discussed and some areas of current uncertainty and continuing research are identified. Given the current pace of development it is anticipated that genome editing techniques will continue to be refined and new techniques emerge. Hence, the emphasis in this report will be on what can be achieved using genome editing techniques, rather than the techniques themselves.

## The concept of genome editing

- 1.1 People have long sought and used scientific knowledge to improve the conditions of human life. From breeding crops and the domestication of livestock to modern health care, the biological sciences underpin the possibility of human beings exercising ever greater levels of control over the biosphere, including their environment, the other living beings with which they share it and their own bodies. Contemporary molecular biology affords a particularly powerful set of tools that form the basis of a range of technologies in fields as diverse as medicine, agriculture, industrial production, and environmental management. What we will refer to as 'genome editing' is the practice of making targeted interventions at the molecular level of DNA or RNA function, deliberately to alter the structural or functional characteristics of biological entities. These entities include complex living organisms, such as humans and animals, tissues and cells in culture, and plants, bacteria and viruses. Characteristics of many kinds, from the colour or number of blooms in flowering plants, to some disease traits in animals and plants, can be altered, though the extent to which, and ease with which, such alterations can be made is highly variable.
- 1.2 Targeted alterations may be accomplished in different ways, including through the use of new and emerging techniques such as the CRISPR-Cas9 system described below. In the future, they may be accomplished in ways that have not yet been described or even envisaged. Nevertheless, although recent advances have meant that genome editing has become highly effective in many research contexts (depending on the system used, the conditions of use, and the model organism) there remains some variation in how fully the underlying aims (of deliberate alteration of biological characteristics) have so far been realised.
- 1.3 Throughout this report we refer to 'genome editing' rather than 'gene editing' (although the latter term is also in use) because the concept of genome editing is not limited to genes. For our purposes, 'genome editing' also includes making alterations to non-coding regions of genomes and to epigenomes (in order to modify whether all or part of the genome is active or silent, and to 'tune' the level of activity). Genome editing clearly shares features with techniques of 'genetic engineering' that have been developed and used over the last forty years (in plant breeding, for example) as well as with more recent micromanipulation techniques for cell reconstruction (for

example, ‘cloning’ and mitochondrial donation).<sup>1</sup> By what principles or according to what criteria these different biological interventions should be delineated, and what moral significance should attach to those distinctions, are among the important questions addressed in this review.

## Gene, genome and epigenome

- 1.4 There is no generally agreed definition of the term ‘genome’. On any understanding, however, genomes comprise the chemical deoxyribonucleic acid (DNA) or, in the case of some viruses, the related chemical ribonucleic acid (RNA).<sup>2</sup> DNA is found in almost all cells of living organisms; it plays a crucial role in their development and functioning, and is centrally involved in the transmission of their properties between generations. DNA is often a very long molecule, a polymer, consisting of a sequence of four different sub-units, called nucleotides, arranged in a particular order. This order, or sequence, largely determines the important biological roles of the molecule. DNA comes in a double strand, forming the iconic double helix structure.<sup>3</sup> The double strand can separate, each strand becoming a template for a new second strand, the process that enables identical sequences of DNA to be replicated when cells divide. Genome lengths range from a few thousand nucleotides in the case of bacteria and viruses to several billions in the case of mammals. Some genomes are larger still. The human genome comprises 3.23 billion nucleotides; the wheat genome is about five times this size.
- 1.5 The term ‘genome’ may be used either to refer to the particular sequence of nucleotides in an organism or in a specific kind of organism, or to the material object that they partly constitute.<sup>4</sup> In the latter sense many genomes consist of a set of *chromosomes*, in which the DNA helix is tightly wound around proteins called histones. Modifications to the histones, or alternatively the attachment of additional chemical parts to the nucleotides, is often crucial in determining which parts of the genome are activated or suppressed. These modifications may be passed on from one generation to the next: heritable changes that are not based on changes to the nucleotide sequences are referred to as ‘epigenetic’ changes. The genome includes *genes*, regions that can direct the production of specific proteins or parts of proteins.<sup>5</sup> Proteins are the molecules that make up most biological structures, and also that direct many chemical processes; they are sometimes thought of as the executive molecules in an organism. There are also regions of the genome that help to control which genes are active in the organism at certain times or in certain conditions. Finally, there are regions that appear not to be functional at all, though the extent of this phenomenon is hotly debated.<sup>6</sup> The term ‘genome’ is often used to refer only to the nucleic acid in the cell nucleus, a membrane-enclosed compartment inside the cell<sup>7</sup>; however, at a cellular

<sup>1</sup> The Council’s 2012 report, *Novel techniques for the prevention of mitochondrial DNA disorders: an ethical review* (available at: [http://nuffieldbioethics.org/wp-content/uploads/2014/06/Novel\\_techniques\\_for\\_the\\_prevention\\_of\\_mitochondrial\\_DNA\\_disorders\\_compressed.pdf](http://nuffieldbioethics.org/wp-content/uploads/2014/06/Novel_techniques_for_the_prevention_of_mitochondrial_DNA_disorders_compressed.pdf)), noted the similarities between different biological techniques capable of modifying the human germ line and advised that a fuller discussion of such techniques would benefit the wider policy debate.

<sup>2</sup> The genome of RNA viruses – those that use RNA instead of DNA as their genetic material – usually comprises a single strand of RNA, although some RNA viruses are double-stranded. It is a matter of debate whether viruses should be described as living organisms.

<sup>3</sup> Some DNA viruses contain a single-stranded (rather than double-stranded) DNA molecule as their genetic material.

<sup>4</sup> For the functional importance of the genome as a material object, see Bustin M and Misteli T (2016) Nongenetic functions of the genome *Science* **352**(6286): aad6933, doi: 10.1126/science.aad6933.

<sup>5</sup> There are also non-coding genes, i.e. reasonably discrete functional units that are transcribed (into non-coding RNA) in a tissue-specific fashion, like a protein-coding gene, but which do not apparently encode proteins. See: Yang JX, Rastetter RH and Wilhelm D (2016) Non-coding RNAs: an introduction, in *Non-coding RNA and the reproductive system*, Wilhelm D and Bernard P (Editors) (Dordrecht: Springer Netherlands), pp13-32.

<sup>6</sup> The ENCODE (Encyclopedia of DNA elements) Project is an international research collaboration that aims to build a ‘parts list’ of the functional elements of the human genome, and other genomes (<https://www.encodeproject.org>). There has been some controversy over whether the epigenomic marks characterised by ENCODE can be deemed to be functional. See, for example, the critique in Graur D, Zheng Y, Price N, *et al.* (2013) On the immortality of television sets: “function” in the human genome according to the evolution-free gospel of ENCODE *Genome Biology and Evolution* **5**(3): 578-90, available at: <http://gbe.oxfordjournals.org/content/5/3/578>.

<sup>7</sup> This does not apply to all cells: many single-celled organisms, bacteria and archaea, lack a distinct nucleus. In addition, certain cell types – mature red blood cells in mammals, which have the single specialised function of transporting oxygen in haemoglobin – do not contain a nucleus.

or organismal level it may also encompass nucleic acids found in mitochondria, small organelles that provide energy to the cells of most higher organisms, and plastids, found in plants and algae.

- 1.6 The genome is often described as a code, because those parts of the genome that guide the production of proteins do so by virtue of a precise correlation between nucleotide triplets and amino acids, the chemicals that make up proteins. The process by which proteins are produced involves two stages: first a DNA sequence is ‘transcribed’ into an RNA molecule, a ‘messenger RNA’, which is subsequently ‘translated’ into part of a protein using an intermediate molecule, a ‘transfer RNA’. Various changes and rearrangements of the messenger RNA may occur between transcription and translation, so there is no simple correlation between genes and proteins. Proteins typically depend for their production on many genes and a gene can be involved in the production of many proteins. Regions of the genome can be regulated by proteins that cause an associated region to be active, producing an RNA transcript, or silent, so that no RNA transcript is produced.<sup>8</sup>
- 1.7 Genomes are passed from one generation to the next when organisms reproduce. Sexual reproduction shuffles parental genomes so that offspring receive a new genome that is a unique combination of the two. In sexually reproducing organisms such as mammals, chromosomes come in pairs that are very similar but, importantly, not identical. Each parent contributes one copy of each chromosome to their offspring. Though the particular sequence of DNA that comprises an individual’s genome is thereby inherited from the previous generation, the genome is subject to alteration by a number of causes. Within every cell of the organism, whenever a cell grows and divides, it copies its DNA so that each ‘daughter’ cell has the same genetic code. However, errors in DNA replication occur and, if these are not corrected, mutations may be incorporated. If cell death does not ensue, cells with mutated DNA may be propagated and may lead to pathological states (for example, cancers in humans and animals). DNA can also be damaged by radiation and toxic chemicals, again leading to the incorporation of mutations. Differences may also be introduced through infection: some viruses insert their DNA into their host’s DNA (as in the case of human papilloma virus infection that may lead to cervical cancer). As noted above, epigenetic modifications to the genome, which may also be induced by environmental factors such as diet or stress, may thus reflect the developmental history of the organism. Epigenetic modifications do not affect the sequence of nucleotides in the genome but are a central aspect of how the genome functions. In the case of organisms that reproduce sexually, epigenetic changes can result in genes being expressed in a parent-of-origin-specific manner, a phenomenon known as genomic imprinting.<sup>9</sup>
- 1.8 There are many small variations between any two genome sequences within the same species. Genome sequence variations within protein coding sequences or in regulatory sequences may have specific effects on the ostensible characteristics (the ‘phenotype’) of an organism and its biological function. Early research, before DNA was identified as the genetic material, was entirely concerned with identifying these differences and their mode of transmission across generations. Such work continues, often under the rubric of ‘Mendelian’ genetics, after Gregor Mendel, the Austrian monk who pioneered this kind of investigation. Well-known genetic diseases, such as haemophilia or cystic fibrosis, are caused by a single variation in a specific gene and are sometimes referred to as monogenic or Mendelian.<sup>10</sup> All biological traits, including common diseases such as cancer and coronary heart disease, reflect a complex interaction of multiple

---

<sup>8</sup> Note that the very numerous messenger RNA molecules, regulatory RNAs, and RNAs that may be transcribed but lack an evident biological function, are nucleic acids but are not considered part of the genome. Unlike the genome, which is relatively stable, they are rapidly changing constituents of the cell. They are sometimes referred to collectively as the ‘transcriptome’.

<sup>9</sup> Peters J (2014) The role of genomic imprinting in biology and disease: an expanding view *Nature Reviews Genetics* **15**(8): 517-30.

<sup>10</sup> Even monogenic inheritance may, rarely, be more complex than this presentation suggests due to the effects of genetic variants (modifiers) elsewhere in the genome of some individuals. See Chen R, Shi L, Hakenberg J, *et al.* (2016) Analysis of 589,306 genomes identifies individuals resilient to severe Mendelian childhood diseases *Nature Biotechnology* **34**(5): 531-38.

genetic and environmental factors.<sup>11</sup> Proteins (which are encoded by genes) have structural and (as enzymes) catalytic roles, and perform a vast array of functions, orchestrating the activities of other important molecules in the cells. They perform specific activities such as metabolising glucose, responding to hormones, transporting chemicals such as oxygen (haemoglobin), and protecting against infection (antibodies). Most of these functions are causally upstream of the phenotype, and a given protein may contribute to multiple phenotypic effects, a phenomenon known as 'pleiotropy'. This can occur because the function of a particular protein may vary according to when and where it is produced in the organism. As already noted, furthermore, most of the genome (about 98%) does not code for proteins at all. For these reasons, it is often extremely difficult to correlate variations in genes with specific phenotypes, and many variations in phenotype have no determinate association with genetic characteristics. It is important to stress this in order to contradict the belief, sometimes known as 'genetic determinism', that all differences in physical traits, or even in higher order capacities or behaviours, are directly determined by variations at the level of the genome.

## Techniques of genome editing

- 1.9 Genomes are naturally susceptible to alteration and errors occur every time a cell copies its DNA. If these errors are not corrected by the cell, cancer or some other pathology may arise, or they may confer a competitive advantage, becoming the basis for natural selection. In addition, genomes may be altered by infection (for example, by retroviruses) and ionising radiation (for example, in the case of radiotherapy, X-rays and ultraviolet light), which disrupt DNA at locations that may be difficult or impossible to predict. However, these mechanisms are not targeted and so would not be regarded as genome editing.<sup>12</sup>

## Recombinant DNA technology

- 1.10 With the arrival of molecular biology in the second half of the twentieth century, it became possible to alter genomes in controlled ways. In particular, this development was enabled by a new recombinant DNA technology that allowed the cutting and then splicing together of DNA molecules. This was developed first in bacteria and their viruses, and subsequently applied to multi-cellular organisms, including plants and vertebrates.<sup>13</sup> The first 'transgenic' mice (mice containing DNA from another species) were produced in the mid 1970s. Transgenesis became a powerful biological research tool, although its major limitation was that it only allowed genes to be added, and offered no control over where the added genes would be inserted into the genome.
- 1.11 In 1989 a way was found to introduce directed alterations into the genomes of embryonic stem cells (ES cells) from which entire mice could be generated.<sup>14</sup> ES cells, derived from the inner cell mass of the early embryo, retain the pluripotency of those embryonic cells, meaning that they have the potential to develop into many distinct types of cell in the body. Genetically modified ES cells can, therefore, be re-introduced into the embryo and will contribute to multiple tissues of that individual, including germ cells. This results in germ cells carrying genetic changes that can be used to generate whole animals. Crucially, ES cells with the desired targeted genetic modifications

<sup>11</sup> This manner of speaking about interactions between genetic and environmental factors passes over many complexities, events that mediate between the effects of genes and environments, including details of transcriptional control, alternate splicing or *in vivo* editing of RNA molecules and the chemical modification of proteins, all of which may alter function and the properties of networks in which these molecules operate. For further discussion, see Burian RM (2004) Molecular epigenesis, molecular pleiotropy, and molecular gene definitions *History and Philosophy of the Life Sciences* **26**(1): 59-80.

<sup>12</sup> For estimates on the mutation rate in humans, see: Callaway E (2015) DNA mutation clock proves tough to set *Nature* **519**(7542): 139-40; Scally A (2016) The mutation rate in human evolution and demographic inference *Current Opinion in Genetics & Development* **41**: 36-43.

<sup>13</sup> Berg P and Mertz JE (2010) Personal reflections on the origins and emergence of recombinant DNA technology *Genetics* **184**(1): 9-17.

<sup>14</sup> See Capecchi MR (2005) Gene targeting in mice: functional analysis of the mammalian genome for the twenty-first century *Nature Reviews Genetics* **6**(6): 507-12. Work leading to this breakthrough was recognised in the 2007 Nobel Prize for Physiology or Medicine, which Mario Capecchi shared with Oliver Smithies and Martin Evans.

can be selected from a vast background of non-targeted cells so the relatively inefficient process of changing the ES cell genome is not a significant obstacle. The gene targeting method initially used to alter mouse ES cells also led to advances in other vertebrates as well as non-vertebrates and plants. However, progress has often been technically challenging and, accordingly, some developments have not occurred until recently. In the case of mammals, for example, ES cells have not been obtained for most species and, even in mice, where the technology is relatively refined, it is time-consuming, expensive, variable, often highly inefficient, and requires a special skill set.<sup>15</sup>

### **Engineered endonucleases: ZFNs and TALENs**

- 1.12 These limitations justified a continued search for alternative gene targeting technologies that bore fruit with the first reports, in 2005, of zinc finger nucleases (ZFNs) and, in 2010, of transcription activator-like effector nucleases (TALENs). ZFNs and TALENs are proteins that work in a conceptually similar manner, containing one module that can be engineered to recognise a specific DNA sequence and guide a second, attached module to cut the DNA. ZFNs and TALENs are derived, respectively, from mammalian transcription factors (proteins in mammalian cells that bind to DNA and cause a gene to become active) and the plant pathogen, *Xanthomonas* sp. Although their protein frameworks differ, ZFNs and TALENs each contain a set of 'fingers' that can be designed to recognise a selected DNA sequence with a high degree of specificity. These fingers, which range in length from three or four nucleotides in ZFNs to around ten in TALENs, are physically attached to an enzyme that cuts one strand of DNA; ZFNs and TALENs each work in pairs to produce a double-strand break (a break at opposite points in the two entwined strands of the DNA molecule).
- 1.13 The consequences of double-strand genome breaks are potentially lethal to living cells and are rapidly repaired by cells using one of two principal pathways that are conserved in plants and animals. In one pathway, the DNA ends produced by the break are re-joined by the cell's repair machinery in a sequence-independent manner (i.e. regardless of the sequence at each end). This is known as non-homologous end-joining (NHEJ). NHEJ does not necessarily restore the original sequence as it (and similar pathways) produces an insertion or deletion (an 'indel'), usually of a small number of nucleotides, in a way that cannot be controlled at present. The other major pathway, homology-directed repair (HDR), is DNA sequence dependent and uses an additional matching piece of DNA to provide template information that allows the double-strand break to be repaired correctly. HDR can also be used to add or remove a prescribed DNA sequence at the site of the double-strand break in a manner that can be controlled. The balance between the employment of NHEJ and HDR repair pathways by a cell in particular contexts is not well understood and is an active area of research.<sup>16</sup>
- 1.14 The role of ZFNs and TALENs is therefore to produce a targeted double-strand break in the genome, which the cellular machinery then repairs. The requirement for two engineered ZFN or TALEN proteins for every target is advantageous because it increases the specificity by decreasing the likelihood that the break will be made at an unintended point in the genome that has a similar sequence to the one the 'fingers' are designed to recognise (an 'off-target effect'). The disadvantage is that it requires considerable effort to design, synthesise and optimise a pair of proteins for every editing procedure.

### **CRISPR-Cas9**

- 1.15 In 2012, it was discovered that a system of defence against viral attack found in the bacterium *Streptococcus pyogenes* could be adapted as a programmable system for genome editing.<sup>17</sup> The

<sup>15</sup> Skarnes WC (2015) Is mouse embryonic stem cell technology obsolete? *Genome Biology* **16**: 109.

<sup>16</sup> See, for example, Paquet D, Kwart D, Chen A, *et al.* (2016) Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9 *Nature* **533**(7601): 125-9.

<sup>17</sup> Jinek M, Chylinski K, Fonfara I, *et al.* (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity *Science* **337**(6096): 816-21.

system comprises two elements. The first is generically termed ‘clustered regularly interspaced short palindromic repeat’ (CRISPR) RNA; the second is ‘CRISPR-associated protein 9’ (Cas9), an endonuclease. The prototypical CRISPR components from *S.pyogenes* comprise two types of RNA molecule that scientists combined into one, called a single guide RNA (sgRNA) or guide RNA (gRNA).<sup>18</sup> In genome editing, sgRNA pairs with its predetermined genomic DNA target to form an heteroduplex (so-called because the RNA and DNA that pair together are different types of molecule). One region of the sgRNA matches with its exclusive DNA target site, giving the system specificity, while another region binds to the Cas9 protein. This guides the Cas9 to make a double-strand cut at the target site. Targeted, double-strand genome breaks made by sgRNA-Cas9 are repaired by the same ubiquitous NHEJ or HDR inbuilt cellular repair pathways operating for breaks made by ZFNs and TALENs. In 2013, the CRISPR-Cas9 system was shown to edit mammalian genomes with a high efficiency.<sup>19</sup>

- 1.16 CRISPR-Cas9 has several advantages when compared to its genome engineering forerunners. Its specificity to the target is secured by the way it exploits nucleic acid base-pairing – a feature that also underlies the fidelity of DNA replication and transcription in the animal and plant kingdoms (and consequently makes all DNA molecules, whatever their origin, amenable to editing in this way). It is so much more efficient (in terms of successes per attempt) that for the first time multiplex mammalian genome editing (editing several different genome sites in one procedure) has been achieved. The components are trivial to produce: sgRNA is only approximately one hundred nucleotides in length and can be synthesized with commercially available kits. The system functions with a universal Cas9 protein framework that dispenses with the need to design a different protein for each DNA target.
- 1.17 The comparatively short length of DNA coding for sgRNA-Cas9 renders it amenable to delivery by viruses, some of which are well-characterised in research and clinical contexts, but in which there are often strict limits on the size of additional inserted DNA that allow efficient virus assembly. Relatively short Cas9 species orthologues (Cas9s from different species that are different but perform a similar function), like the one from *Staphylococcus aureus*, help to meet size constraints, and others from *Neisseria meningitidis* and *Francisella novicida* (Cpf1) may offer distinctive advantages regarding size, target selection and target specificity. The current picture is one in which an already efficient system is undergoing continued refinement.
- 1.18 CRISPR-Cas9 works by causing a targeted DNA break but it is possible to replace the DNA cutting activity of Cas9 with other activities. For example, the DNA cutting activity can be replaced with DNA methylating or histone modifying activities. This means that an altered Cas9 (nuclease-dead Cas9) can, instead of making a double-strand break, perform an epigenetic modification at a prescribed site on the genome. This allows it to switch selected genes off or on without altering their sequence. Avoiding the need to alter genomic DNA sequences in this way may have advantages in contexts where the aim is to control gene expression without introducing heritable alterations to DNA.<sup>20</sup>
- 1.19 Given the rapidity of advances with CRISPR-Cas9, it is reasonable to ask what limitations it has and what it promises from a technical standpoint. One concern with the CRISPR-Cas9 system is the potential for off-target effects (editing at sites in the genome other than those intended). These concerns arose originally from analyses of editing in cell populations but single-cell analyses have subsequently suggested that these initial studies exaggerated the lack of specificity. Moreover, engineering of Cas9 protein and sgRNA frameworks have increased specificity further, so that

<sup>18</sup> Ibid. The terms ‘sgRNA’ and ‘gRNA’ are often used synonymously and are used in all editing procedures. Accordingly, we will use sgRNA when referring to editing constructs, and to CRISPR-Cas9 when referring to the system generically.

<sup>19</sup> Cong L, Ann Ran F, Cox D, *et al.* (2013) Multiplex genome engineering using CRISPR/Cas systems *Science* **339**(6121): 819-23; Mali P, Yang L, Esvelt KM, *et al.* (2013) RNA-guided human genome engineering via Cas9 *Science* **339**(6121): 823-6.

<sup>20</sup> Thakore PI, Black JB, Hilton IB, *et al.* (2016) Editing the epigenome: technologies for programmable transcription and epigenetic modulation *Nature Methods* **13**(2):127-37.

now experiments have been performed in which no off-target cutting has been found, even when it is searched for by whole-genome sequence analysis.<sup>21</sup> Another complicating factor is that 'mosaicism' has been commonly observed in mice generated by the injection of CRISPR-Cas9 reagents into single-cell mouse zygotes. Mosaicism describes the situation in which not all cells of an individual are genetically identical but, instead, cells harbouring distinct mutations co-exist in the same organism. This implies such individuals may transmit several distinct mutations to the next generation. Strategies to reduce or eliminate mosaicism are being developed.<sup>22</sup>

- 1.20 So young a technology has, nevertheless, yet to be fully delineated. There may, for example, be classes of genomic DNA sequence that are refractory to CRISPR-Cas9. A very significant limitation to the practical deployment of the technique is the state of knowledge of gene function: CRISPR-Cas9 cannot be used to introduce or eliminate traits until its users know which regions of the genome to edit. Overcoming this obstacle requires considerable advances in the domain of genetics, although this is something to which genome editing can itself contribute as a powerful technique in laboratory research.<sup>23</sup> However, if genome editing is to prove practically valuable in the way that crop breeding, livestock domestication and biomedicine have done to date, it will be equally important, and arguably much more difficult, to demonstrate that the phenotypic modifications that may be achieved in the laboratory can be achieved in the field, the barn, and the clinic, and, equally importantly, to ensure that they can be introduced safely, ethically and acceptably in these contexts.

---

<sup>21</sup> Kleinstiver BP, Pattanayak V, Prew MS, *et al.* (2016) High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects *Nature* **529**(7587): 490-5; Slaymaker IM, Gao L, Zetsche B, *et al.* (2016) Rationally engineered Cas9 nucleases with improved specificity *Science* **351**(6268): 84-8.

<sup>22</sup> Singh P, Schimenti JC and Bolcun-Filas E (2015) A mouse geneticist's practical guide to CRISPR applications *Genetics* **199**(1): 1-15.

<sup>23</sup> Genome editing can, for example, relatively easily 'knock out' a gene of interest by introducing a mutation (indel) that would inactivate the gene, allowing researchers to make inferences about the gene function from subsequently observed differences in the modified organism. However, further research would be necessary to establish that an effect that is reproducible in experimental conditions is also reproducible outside those conditions (in a 'natural' or 'wild' environment).