Current capabilities for human genome editing: report for the WHO Expert Advisory Committee on Developing Global Standards for Governance and Oversight of Human Genome Editing

Güneş Taylor and Christophe Galichet

April 2021

World Health Organization

WHO/SCI/RFH/2021.06

Suggested citation. Taylor G, Galichet C. Current capabilities for human genome editing: report for the WHO Expert Advisory Committee on Developing Global Standards for Governance and Oversight of Human Genome Editing, April 2021. Geneva: World Health Organization; 2021 (WHO/SCI/RFH/2021.06). Licence: CC BY-NC-SA 3.0 IGO.

Cataloguing-in-Publication (CIP) data. CIP data are available at http://apps.who.int/iris.

Contents

Abbreviations iv
Context 1
1. Introduction to genome editing 1
2. Improving NHEJ and HDR outcomes 3
3. Base and prime editing 4
4. Precision editing of mitochondrial DNA 6
5. Transcriptional modulators and epigenetic editing 6
6. Improving the sensitivity of ways to assess the outcome of genome editing 7
7. Alternative routes for heritable human genome editing using stem cells 8
8. Prenatal (in utero) somatic gene therapy or heritable genome editing 9
9. Conclusion and future scoping 10
References 11

Abbreviations

ABE	Adenine Base Editor
CBE	Cytidine Base Editor
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
dCas9	nuclease-dead Cas9
DSB	Double-Strand Break
GONAD	Genome editing via Oviductal Nucleic Acid Delivery
gRNA	guide RNA
HDR	Homology-Directed Repair
indel	insertion or deletion
MMEJ	Microhomology-Mediated End Joining
mtDNA	mitochondrial DNA
NHEJ	Non-Homologous End Joining
pegRNA	prime editing guide RNA
PGC	Primordial Germ Cell
REDIT	RecT Editor via Designer-Cas9-Initiated Targeting
RS-1	RAD51-stimulatory compound-1
SCGE	Somatic Cell Genome Editing
TALEN	Transcription Activator-Like Effector Nuclease
ZFN	Zinc Finger Nuclease

Context

To develop enduring global standards for governance and oversight of human genome editing, it is necessary both to define the genome editing technologies currently available and to appreciate the trajectory of the research underpinning them. The unprecedented rate of progress in this field necessitates regular revisions to benchmark reports such as that of the International Commission on the Clinical Use of Human Germline Genome Editing in 2020, and as such, this report focuses particularly on recent developments and highlights areas of current research efforts.

At the outset, it must be stressed that despite all that has been learned in the last two decades about human genes, genomes and genetic variation, there are still many gaps in our knowledge. Multiple genes and "environmental" effects can influence the incidence and severity of many human disorders with an underlying genetic cause. Genetic background effects are well known in animal models, and also apply to humans. These can reflect one, several, or many gene variants that can map anywhere within the genome or be closely linked to the main causative "mutant" allele. Many genetic variants identified in humans have an unknown impact on phenotype - some will be in protein coding regions, others in RNA products of genes, and many will be in non-transcribed regions, where they could affect gene expression of single or even multiple genes. Even for diseases that are commonly referred to as monogenic, diseases attributable to the malfunction of a single gene, sound evidence for a causative role of a specific genetic variant would be needed prior to genome editing. This can come from family studies or population genetics, but ideally it requires knowledge of whole genomes, not just of the specific gene in question. Environmental effects can be due to the influence of nutrition in utero or postnatally, gut and other microbiota, exposure to pathogens and other harmful substances such as pharmaceuticals, and so on. It follows that, without further research designed to explore these variables, it might be hard to predict the exact outcome in attempts at human genome editing. This will be true for somatic as well as heritable applications.

1. Introduction to genome editing

Genome editing encompasses a number of techniques to modify DNA within the genome of a targeted cell, thereby altering the information encoded at the target site. These techniques include the use of meganucleases, zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs) and, most recently, CRISPR-Cas9.[1] Each of these systems employs a different repertoire of molecules to alter DNA, and thus has different properties, which need to be carefully considered in any clinical research or applications.

Genome editing technologies traditionally rely upon the generation, and subsequent repair, of double-strand breaks (DSBs) of DNA at user-defined target sites. It is this ability to

[1] CRISPR: clustered regularly interspaced short palindromic repeats.

programme user-defined targets, in combination with improved DNA sequencing methods giving both reference and patient-specific genome sequences, that sets modern genome editing apart from previous efforts to alter genomes. DSBs are predominantly repaired via an error-prone mechanism known as non-homologous end joining (NHEJ), which regularly results in small insertions or deletions (indels). These indels can cause significant changes to the information encoded at the target site, thereby disruptively "editing" the underlying genome. However, under certain circumstances a higher-fidelity mechanism for DNA repair within cells called homology-directed repair (HDR) can be employed**.** HDR occurs with less frequency as it requires the presence of a repair template with high sequence similarity to the regions flanking the target site. Thus, if a donor DNA template carrying a specific edit is introduced together with the genome editing components, cells are able to repair DSBs at target sites via HDR, thus introducing the desired template sequence into the genome.

Cells can also repair DNA DSBs by microhomology-mediated end joining (MMEJ), which makes use of homologies of just a few base pairs and leads to usually short (but sometimes quite extensive) deletions. Another cell-based repair mechanism that can be triggered during HDR uses the DNA sequence from the other chromosome (instead of the DNA template provided), leading to "gene conversion". This can lead to repair of the mutant allele in a heterozygote, or if this happens after the allele has been repaired, in a homozygote. However, it can also spread over long distances and lead to loss of heterozygosity in adjacent regions of the chromosome, thus resulting in large stretches of identical DNA sequences in the chromosome pair. This could have significant consequences: for example, if the individual was heterozygous for a recessive mutation in a gene adjacent to the one being targeted, gene conversion could result in this now becoming homozygous, for either the normal or the disease-causing variant, the latter leading to a new, perhaps unexpected, phenotype.

Of note, all genome editing events occur independently within targeted cells, whether by NHEJ or HDR. Consequently, screening for successful editing outcomes is vitally important, especially in any potential clinical context.

While all genome editing platforms have been used extensively to model and understand human conditions in laboratory animals *(1)*, in recent years, CRISPR-mediated genome editing has become the focus of most efforts to develop potential therapeutic tools for humans. Unlike ZFNs and TALENs, which are proteins engineered to have both a nuclease, usually a FokI domain, and a part made up of individual modules that each recognize specific DNA bases, CRISPR is a two-part molecular system, comprising a CRISPR guide RNA (gRNA) and a Cas9 DNA-cleaving enzyme that interact with each other. Partly because it is simple to generate specific gRNAs, the latter offers superior speed, efficiency and programmability compared to the other genome editing technologies.

Since the publications first demonstrating the power of CRISPR genome editing in bacteria *(2)* and subsequently in mammals *(3, 4)*, the scientific community has spent enormous efforts to further refine its effectiveness and develop a sizable repertoire of other CRISPR-based tools to address specific needs. For example, most pathogenic mutations found in the human genome must be corrected rather than simply disrupted in order to benefit patients, meaning that the rare HDR events are required rather than the disruptive NHEJ. Unfortunately, the

inefficiency of HDR in most cell types and the challenges associated with donor DNA template delivery have limited the therapeutic relevance of HDR to date. Consequently, technical improvements in HDR rate, component delivery and specific single nucleotide alterations have particular clinical significance, and most clinical trials using genome editing focus on cases where gene disruption by indel generation is therapeutic. Despite being the most recent innovation in the genome editing toolkit, CRISPR has already been used in the clinic. In a landmark case in February 2020, two patients with refractory myelomas and one with metastatic sarcoma had their own immune cells edited in vitro to allow these to better recognize and eliminate cancer cells upon reintroduction, with positive initial results reported *(5)*.

Below are summary overviews of current and contemplated uses of genome editing technologies in humans. However, these cannot be all inclusive, given the rapid rate of development in the field. Moreover, methods that would be very challenging to engineer and are likely to raise significant ethical and societal objections, such as gene drives in humans (which are being developed to manage important disease vectors such as mosquitoes and invasive species), are not included.

2. Improving NHEJ and HDR outcomes

As indicated in the Stadtmauer et al. report *(5)*, when in vitro screening of the CRISPR-targeted cells is possible, the error-prone nature of NHEJ can be mitigated by growing colonies (clones) from single cells, screening these and selecting only those for further use that have their genome edited correctly. However, not all tissues or cell types are as amenable to screening as immune cells; thus efforts are being directed to improving NHEJ outcome predictability. Recent work has demonstrated that predictable and precise NHEJ-mediated deletions are possible using paired Cas9 gRNAs to generate adjacent DSBs *(6)*. While off-target prediction precision has also improved, the generation of DSBs at non-target sites and consequently the possibility of other NHEJ events occurring still needs to be considered. Minimizing the time that the Cas9 is active within a cell, for example, by using short-lived versions of the protein, can help reduce off-target events, as can using specific variants of either the nuclease or the gRNAs that favour their activity at the cognate target DNA sequence rather than those that differ by one or two base pairs.

Unlike HDR, even with improvements to outcome prediction, NHEJ-edited alterations are ultimately not user defined. However, HDR is usually favoured less frequently than NHEJ within cells and requires the presence of a donor sequence template carrying the intended replacement information. Advances have been made in the use of chemical inhibitors of key components of the NHEJ and microhomology-mediated end joining DNA damage repair pathways to shift the balance in favour of HDR. The use of SCR7, an inhibitor of DNA ligase IV involved in the NHEJ pathway, permits an increased incidence of HDR-mediated repair in human cell lines *(7)*. Furthermore, introduction of RAD51, an essential protein for HDR, has also been shown to favour HDR occurrence. Provision of RAD51-coated oligonucleotides has also been used to favour integration of repair templates, as has introduction of RS-1 (RAD51-stimulatory compound-1), a compound known to stabilize the association of RAD51

to the DNA, into cells *(8)*. In human fibroblast and induced pluripotent stem cells, direct fusions of Cas9 to a key factor for DSB resection and HDR initiation, CtIP, focuses its activity on the target site, thereby enhancing HDR *(9)*. Competition between NHEJ and HDR is cell cycle dependent, with HDR being restricted to late S and G2 phases *(10)*. Consequently, synchronizing cell cycle phases has also been shown to help increase the rate of HDR *(11)*. It is worth noting that HDR repair only occurs in cells that are dividing (have an active cell cycle), thus this method cannot be used in differentiated cells that no longer divide (they are "post-mitotic"), such as neurons.

Aside from shifting the balance from NHEJ to HDR, the availability of all necessary components at the target site is an important factor for determining high editing efficiency. Accordingly, multiple methods to maximize the presence of all components have been developed. One such method, REDIT,[2] features a modified gRNA containing MS2 stem loop scaffolds, which, upon binding of the MS2 coat protein, bring the DNA donor template to the cleavage site *(12)*. Another approach is the use of streptavidin-tethered Cas9 enzyme, which can be used in combination with repair templates containing biotinylated oligonucleotides, to favour incorporation of the template at the CRISPR on-target site *(13)*. Furthermore, as reported in a 2019 study, two gRNAs in conjunction with a long single-stranded DNA template can dramatically improve large and complex gene edits in early mouse embryos *(14)*. In light of these developments, we predict that combinations of multiple methods to further improve HDR rates will be increasingly employed.

3. Base and prime editing

Many disease-associated human genetic variants are attributable to point mutations affecting single DNA base pairs, such as the transition mutations in the ageing disease progeria or transversion events leading to sickle cell anaemia, or small indels, such as the four base pair insertion underlying the majority of cases of Tay-Sachs disease. New tools allowing modification without DNA cleavage or DSB formation have also been gaining increased attention because of concerns regarding the possibility of large genomic rearrangements subsequent to DSB formation. Base editors were developed to circumvent the issues of random indels from NHEJ and ineffective HDR rates in CRISPR-mediated genomic editing, which hinder their effective use in correcting specific and small mutations. Base editors fuse enzymes with particular DNA-altering properties to Cas9, thereby focusing their action on