Gene editing and genotoxicity: targeting the off-targets

Health & Medicine

Introduction

Therapeutic approaches relying on the genetic engineering of cells for the treatment of hereditary diseases has long been a promising strategy to overcome the shortcomings of conventional drug therapies. The principle of these gene therapies is to counteract, correct, or replace a malfunctioning gene within cells that are most severely affected by the caused condition. However, any process affecting DNA integrity or causing DNA or chromosomal damage bears the risk of genotoxicity (Bohne and Cathomen, 2008).

While viral vectors utilized for gene addition-based strategies showed encouraging initial results (Anderson, 1990 ; Rosenberg et al., 1990 ; Gaspar et al., 2004), subsequent trials targeting hematopoietic stem and progenitor cells (HSPCs) exposed the risk of therapy-related toxicities, particularly insertional activation of proto-oncogenes leading to malignant cell transformation. Indeed, the activation of MDS1-EVI1 and LMO2 oncogenes caused by the integration of the gamma retroviral vector led to clonal skewing and development of malignancies in patients enrolled in several gene therapy clinical trials. (Hacein-Bey-Abina et al., 2003 ; Raper et al., 2003 ; Ott et al., 2006 ; Cattoglio et al., 2007 ; Schwarzwaelder et al., 2007 ; Howe et al., 2008 ; Metais and Dunbar, 2008 ; Stein et al., 2010 ; Zhou et al., 2016). These issues have been partly addressed through development of next generation vectors including in particular lentiviruses (Naldini et al., 1996 ; Aiuti et al., 2013 ; Hacein-Bey-Abina et al., 2014 ; Kohn et al., 2020)

and adeno-associated virus (AAV) (Nathwani et al., 2011) lowering, but not eliminating, the risk of insertional mutagenesis and immunogenicity.

Besides the risk of insertional mutagenesis and immunogenicity (Cavazzana-Calvo et al., 2010), viral vectors have additional drawbacks, including their inability to address dominant mutations and their potential influence on the host cell's gene expression (Maeder and Gersbach, 2016). Further attempts to address these issues have been made, for example by using chimeric proteins to retarget lentiviral integration to sites with reduced transcriptional activity (Gijsbers et al., 2010 ; Vranckx et al., 2016).

Many of these limitations can however be overcome by gene therapy approaches that rely on genome editing techniques which enable more precise, targeted genomic modifications to restore wild-type sequences, while preserving the temporal and tissue-specific control of the afflicted gene, or to specifically knock out genes.

Initially the four main families of nucleases–meganucleases (Chevalier et al., 2001 ; Epinat et al., 2003), zinc finger nucleases (ZFNs) (Urnov et al., 2005), transcription activator-like effector nucleases (TALENs) (Bogdanove and Voytas, 2011), and Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)-associated nucleases (Cas) (Jinek et al., 2012) were used to induce targeted DNA double-strand breaks (DSBs). Meganucleases, ZFNs, and TALENS are tethered toward specific DNA sequences by means of DNA-binding protein domains while the CRISPR-Cas system is based on a nuclease protein guided by an RNA molecule complementary to the targeted DNA sequence (gRNA) via Watson-Crick base pairing. The introduction of DSBs

activates one of the two main endogenous cellular repair pathways, including the error prone non-homologous end joining (NHEJ) and the homology directed repair (HDR) pathways. Some repurposed derivatives of the engineered nucleases, in particular of Cas, have been developed to fulfill different tasks. Nickases, which are Cas9 proteins with only one functional nuclease domain (Sapranauskas et al., 2011 ; Jinek et al., 2012), are used to induce DNA single-stranded breaks (SSBs), combined with base editors (BEs) fused to a cytidine or an adenine deaminase to induce precise transition mutations (Komor et al., 2016 , 2017 ; Gaudelli et al., 2017 ; Kurt et al., 2020). In addition, the prime editing strategy uses a nickase fused with a reverse transcriptase complexed with a prime editing guide RNA (pegRNA) to mediate targeted insertions of few bases, deletions, and base conversions (Anzalone et al., 2019).

Similar to viral vectors, genome editing techniques have been rapidly adopted and they have proven suitable for clinical application in various fields. So far, seven patients have been infused with CRISPR-Cas9 modified autologous CD34+ HSPCs for the treatment of beta-hemoglobinopaties, showing encouraging results; among those, two patients affected by beta thalassemia are transfusion independent after 5 and 15 months after infusion while one patient affected by sickle cell disease is free of vaso-occlusive crises at 9 months after treatment (NCT03655678; NCT03745287). CRISPR-Cas9 or TALENs have also been applied to engineer patient or Universal Chimeric Antigen Receptor (CAR)T lymphocytes for improved antitumor immunity (Qasim et al., 2017 ; Stadtmauer et al., 2020) (NCT02735083; NCT02808442; NCT02746952; NCT03081715;

NCT02793856; NCT04244656; NCT04035434). Moreover, TALENs have been used to create allogenic CS1, CD123, or CD22-specific CAR-T cells (NCT04142619; NCT03190278; NCT04150497) and ZFNs helped to engineer T cells with a C-C motif chemokine receptor 5 knockout to induce resistance to HIV infection (Tebas et al., 2014) (NCT03617198). In addition, AAV vectors together with ZFN-mediated genome editing were applied for the insertion of a correct copy of the α-L-iduronidase gene for subjects with attenuated Mucopolysaccharidosis type I (MPS I) (NCT02702115).

While genome editing techniques address certain limitations and reduce particular risks of genotoxicity in viral vector-based gene therapy, they entail new complications. Off-target activity, the induction of DNA modifications at unintended sites, is a concern with all designer nucleases (Fu et al., 2013 ; Koo et al., 2015). Such off-target activity can potentially lead to point mutations, deletions, insertions or inversions. Besides off-targeting due to a sequence similarity to the targeted site, there can also be collateral cleavage activity. This has been observed for CRISPR-Cas12a, which, upon RNA-guided on-target DNA binding, non-specifically cleaves single-stranded DNA molecules (Chen et al., 2018). While high fidelity variants of the Cas9 protein have successfully been developed to reduce off-target activity (Kleinstiver et al., 2016 ; Vakulskas et al., 2018) they still bear the risk of inducing on-target damage after cleavage in the form of large deletions spanning several kilobases or translocations (Kosicki et al., 2018 ; Connelly and Pruett-Miller, 2019 ; Turchiano et al., 2020). As those large deletions can bring relatively distant elements close together, they could have a genotoxic potential similar to the insertional mutagenesis caused by viral

vectors. A key prerequisite for the clinical application of genome editing tools is the monitoring of their safety before, during and after the administration of the treatment. However, while gene therapy and genome editing are advancing at a rapid pace, the application of appropriate assays to evaluate unintended genomic effects suffers from a lack of standardized methods and guidelines (Corrigan-Curay et al., 2015). A multitude of techniques have been developed in the recent years to detect small insertions and deletions (Indels), potential off-target DNA breaks, translocations, or viral integration sites but the lack of standardized analyses that allow an absolute quantification of those modifications makes a direct comparison among these tools cumbersome. The aim of this review is to give an overview of the drawbacks and benefits of the currently available tools to assess the safety of gene editing applications and of the parameters that need to be taken into account for a correct safety assessment of a gene therapy approach.

Biased Detection Methods

A major step for the successful use of designer nucleases is the choice of the target site and the according nuclease or gRNA design following criteria of editing efficiency and specificity. Potential off-targets can be predicted *in-silico* (Grau et al., 2013 ; Bae et al., 2014 ; Cradick et al., 2014 ; Montague et al., 2014 ; Concordet and Haeussler, 2018) or identified on cultured cells (in cellula) or *in vitro* assays for DSB detection. The list of potential off-target sites identified must be subsequently verified in the target cell type/tissue alongside the detection of on-target cleavage. There are three common types of approach to quantify indels formation at the selected sites, all of which rely on polymerase chain reactions (PCR) performed on genomic DNA

treated with the designer nuclease of choice. After denaturation and rehybridization of the PCR products, hetero-duplex DNA containing a mutated and a wild-type strand can form, which can then be cleaved by mismatch-detection nucleases such as Surveyor nuclease or T7 endonuclease I (Mashal et al., 1995 ; Qiu et al., 2004), enabling the quantification of the cleavage products by electrophoresis. While cost-effective and simple, a drawback of those techniques is that single nucleotide polymorphisms are poorly recognized. Quantification of Indels in the PCR products can also be determined by methods such as Tracking of Indels by Decomposition (TIDE) or Inference of CRISPR Edits (ICE) analyses, which compare the Sanger sequence chromatograms of an untreated control against a treated sample at the intended editing site (Brinkman et al., 2014 ; Hsiau et al., 2019). Moreover, Indels can also be directly quantified using deep sequencing of the PCR products (Pinello et al., 2016). As all of these approaches are based on PCR amplicons of around 200–700 base pairs (bp) from the potential target loci, they all suffer from the same shortcoming, that is the missed detection of larger deletions or other aberrations that could encompass at least one of the PCR primer binding sites. While it is commonly accepted that most of the indels fall in a size spectrum of under 50 bp (Koike-Yusa et al., 2014 ; van Overbeek et al., 2016), it has been shown that genome editing can also lead to large deletions of several kilobases (kb) (Kosicki et al., 2018 ; Chakrabarti et al., 2019 ; Turchiano et al., 2020). Moreover, even in the case of deletions being amplified/detected by PCR, a minimal sequence length of sufficient quality, required for the alignment with either a control sequence or a reference genome, might not be reached.

Indel assessment by deep sequencing requires additional consideration. Artifactual sequencing errors produce a background signal that is usually filtered out by setting an arbitrary threshold to define the relevant modified loci. This kind of analysis can produce either false-positives or false-negatives since every amplicon can present higher or lower background levels, respectively. Statistically more robust approaches are needed, but they can be laborious particularly when a multitude of targets are investigated. Performing the analysis on a large set of replicates and untreated controls in order to compare the mean editing frequencies using a t-test would be the preferable procedure (Zeng et al., 2020). Alternatively a two-sample test for equality of proportions or a Fisher's exact test can be performed to detect differences between the mutation rates of edited and untreated samples. In order for this approach to be robust, a high number of reads is desirable. Moreover, when different assays are employed, these statistical tests can account for the variability of the NGS measurements to better define the null hypothesis or introduce a false discovery rate correction (e. g., Benjamini–Hochberg; Kuscu et al., 2014 ; Turchiano et al., 2020). This would be a more traceable practice compared to indel values subtractions between the treated and the relative untreated samples (Cameron et al., 2017) or to simply assuming a background noise level (Yang et al., 2014 ; Kim et al., 2016 ; Kim and Kim, 2018). This approach allows compensation for sequencing errors, especially for challenging regions with repetitive elements, which might produce a considerable amount of indel-like reads and that might even require a visual inspection to evaluate potential sequencing or alignment artifacts (Zeng et al., 2020).

Alternatively, oligo integration analysis rather than indel detection allows dramatic reduction in background noise in the untreated control of about 100 fold, enabling off-target detection at <0. 001% rates (Tsai et al., 2017). It is worth noting that the reliability of this method is dependent on the oligo integration efficiency in the cell type of interest, generally higher in cell lines as compared to primary cells; moreover this technique cannot be applied to samples that are meant to be infused into patients, since this would imply integrations of non-therapeutically relevant exogenous sequences into the genome.

Unbiased Detection Methods

DSB Detection

While *in silico* off-target prediction is a fast and cheap option, it suffers from high false-positive rates as it is mostly based on the similarity of a sequence to the target site and does not consider differences due to genetic variants; moreover, it has a limited sensitivity in the detection of bona fide off-targets (Tsai et al., 2015 ; Kim et al., 2019b). This bias can be overcome with the use of *in vitro* methods that are based upon the incubation of purified genomic DNA with the designer nuclease of choice. The DSBs induced by the nuclease are then detected in various ways, either by the circularization of the created DNA fragments in CIRCLE-seq (circularization for *in vitro* reporting of cleavage effects by sequencing) or CHANGE-seq (circularization for high-throughput analysis of nuclease genome-wide effects by sequencing) (Tsai et al., 2017 ; Lazzarotto et al., 2020), the ligation of adapters in SITE-seq (selective enrichment and identification of adapter-tagged DNA ends by sequencing) (Cameron et al., 2017) or End-seq (DNA

end sequencing) (Canela et al., 2016), or deep sequencing and identification of identical 5' DNA fragments in Digenome-seq (*in vitro* Cas9-digested whole-genome sequencing) (Kim et al., 2015). While being sensitive, a common drawback of these approaches is a tendency to overestimate the number of sites that are actually modified in cells (Cho et al., 2014), as the influence of the chromatin structure in determining the DNA accessibility is widely disregarded (Kim and Kim, 2018). Moreover, the impact of the nuclease concentration inside the cell (Wu et al., 2014) and of the delivery method on the cleavage footprint are not considered by *in vitro* assays (Kim et al., 2014 ; Cameron et al., 2017). Those *in vitro* techniques are usually returning the highest number of sites, but their relative validation rates disregard at least half of them in the best case scenario. The *in cellula* derived deep sequencing validation deserves some additional considerations: (1) it is usually performed only on the top performing sites disregarding the ones close to the cutoff thresholds; (2) it cannot be sensitive enough to detect rare indel events; (3) some DSBs can be perfectly repaired without creating any mutation and therefore could be missed during the validation process.

A more representative assessment can hence be expected from methods where designer nucleases are applied directly in cellula. In GUIDE-seq (genome-wide, unbiased identification of DSBs Enabled by sequencing; Tsai et al., 2015), IDLV (integrative-deficient lentiviral vectors) capture (Gabriel et al., 2011) and ITR-seq (Inverted Terminal Repeat sequencing; Breton et al., 2020) the DSBs are marked by insertion of exogenous sequences, which are subsequently exploited as specific primer binding sites and then

amplified via linker mediated PCR. Instead techniques like BLESS (direct *in situ* breaks labeling, enrichment on streptavidin and next-generation sequencing; Crosetto et al., 2013 ; Ran et al., 2015), its variant DSB Capture (Lensing et al., 2016) or BLISS (Breaks Labeling *in situ* and Sequencing; Yan et al., 2017 , 2019) are based on *in situ* processing of the DNA at the open DSB ends and ligation of biotinylated adapters or adaptors for *in vitro* transcription. For DSB-seq high molecular weight genomic DNA is isolated from treated cells and the DNA ends are 3'-end tailed with biotinylated nucleotides by terminal deoxynucleotidyl transferase (TdT) before sonication, capturing, and sequencing (Baranello et al., 2014). An alternative approach based on Chromatin immunoprecipitation sequencing (ChIP-seq) targets the phosphorylated histone variant H2A. X or other repair factors that are recruited to cleaved sites (Iacovoni et al., 2010). However, as those factors can spread several kb around DSBs, an identification of the cleavage sites at nucleotide resolution is difficult. In DISCOVER-seq (discovery of *in situ* Cas off-targets and verification by sequencing), detection of the MRE11 subunit of the MRN complex binding by ChIP-seq returns a more specific and sensitive information filtered by a custom algorithm that retains cleaved sites followed by the protospacer-adjacent motif (PAM) and the putative protospacer binding site (Wienert et al., 2019 , 2020). It is worth mentioning here that unbiased DSB discovery is also performed by some of the techniques described in section " Translocation and Other Chromosomal Aberration Detection."

SSB/BE Detection

Compared to the variety of assays for DSB detection, methods to monitor SSBs induced by designer nickases and/or base editors are less abundant. This type of gene editors is generally thought to be less harmful than designer nucleases generating DSBs (Hu et al., 2016 ; Bothmer et al., 2017) but thorough and more specific analyses could report a higher genome and RNA mutational rate with some BEs (Rees et al., 2019 ; Xin et al., 2019).

Detection techniques designed for and tested on samples treated with designer nickases linked to BEs are EndoV-seq (Endonuclease V-based sequencing; Liang et al., 2019) and Digenome-seq (Kim et al., 2017 , 2019a , 2020) (Supplementary Table 2). Both techniques rely on an *in vitro* nicking, base modification and subsequent DNA end repairing in order to obtain a particular pattern after whole genome sequencing (WGS). This approach and its respective bioinformatic analysis help to filter out the majority of the natural occurring DNA nicks, but requires a good WGS coverage (> 30 ×), which makes these techniques expensive and only applicable to studies in a pre-clinical phase.

The validation rate for these techniques can vary greatly compared to DSB detection methods. Since SSBs cannot be directly revealed by indel quantification, base editing frequencies at potential off-target sites can be measured by NGS or, in alternative, DSBs induced by an active designer nuclease can be employed as surrogates.

Other techniques are showing great potential but they have not been tested on designer editors/nickases. Among those, SSB-seq (Baranello et al., 2014

), SSiNGLe (single-strand break mapping at nucleotide genome level) (Cao et al., 2019), GLOE-seq (genome-wide ligation of 30-OH ends followed by sequencing) (Sriramachandran et al., 2020), and Nick-seq (Cao et al., 2020) are able to return data from an in cellula approach.

The " Prime editors" strategy instead presents a new challenge for this kind of techniques since its nickase activity is coupled with a reverse transcriptase that could potentially introduce indels at off-target sites or cause retrotransposon activations and integrations of random reverse transcribed RNA sequences into the genomic DNA (Anzalone et al., 2019). A recent work using Digenome-seq (Kim et al., 2020) exploited the aspecific capacity of the dCas9-H840A protein, utilized in the PE, to cleave also the non-targeted strand, resulting in a characteristic signature after WGS and enabling the use of an analysis compatible with the Digenome-seq bioinformatic pipeline. The authors showed that not all the off-target sites detected by Digenome-seq and validated for the presence of indels are prime-edited, confirming the importance of the pegRNA specific priming activity. In support of this, a different work recently showed the presence of unexpected large deletions after prime editing in mice embryos, mainly ascribed to the dCas9-H840A activity (Aida et al., 2020). However, this proposed strategy do not allow the detection of all the possible mutations that this system may induce *in cellula* and therefore it is not completely exhaustive.

Translocation and Other Chromosomal Aberration Detection
Off-target mutations and insertional mutagenesis are considered to be harmful because they can perturb the expression of nearby genetic elements

by means of different mechanisms (McCormack and Rabbitts, 2004). While this dysregulation is usually localized, rare or innocuous, major concerns derive from general genomic instability and the several chromosomal aberrations that we may or may not detect after editing. Increasing evidences are showing how those gross chromosomal aberrations generate after designer nucleases activity (Weinstock et al., 2008 ; Kosicki et al., 2018 ; Turchiano et al., 2020). Oncongenic translocations have been reproduced *in vivo* in lung tissues of mouse models by the simultaneous introduction of two DSBs, confirming this major concern (Blasco et al., 2014). Large deletions, loss of heterozygosity, large inversions, or translocations may also impact the 3D genomic organization and cause dysregulations of entire topological associated domains (median size ~880 kb) usually organized to be transcriptionally active or repressed (Dixon et al., 2012 ; Bonev and Cavalli, 2016).

Recent studies also observed on-target related chromosomal aberrations with formation of micronuclei and chromosome bridges leading to copy number variation, telomeric portion loss, and chromotripsis (Cullot et al., 2019 ; Leibowitz et al., 2020) rising further concerns for the safety of designer nucleases in clinic.

The portfolio of translocation detection techniques developed to recognize those mutations with increasing sensitivity comprises TC-seq (translocation capture sequencing; Klein et al., 2011), UDiTaS (UniDirectional Targeted Sequencing; Giannoukos et al., 2018), AMP-seq (anchored multiplexed PCR sequencing; Zheng et al., 2014), LAM-HTGTS (linear amplification-mediated

high-throughput genome-wide sequencing; Frock et al., 2015), and CAST-seq (chromosomal aberration analysis by single targeted LM-PCR sequencing; Turchiano et al., 2020).

All of these methods are based on nested PCRs with primers binding between a known target site and fused unknown sites marked by an adapter. NGS sequencing is then used to identify the fusion partners. Differences among these techniques include in particular the adapter attachment via tagmentation (UDiTas), bridge adaptor ligation (LAM-HTGTS), or dsDNA ligation (CAST-seq and AMP-seq). Besides, the amount of input DNA and the bioinformatic pipeline can differ substantially for these techniques as shown in Supplementary Table 3 . For designer nucleases, the validation of translocation sites can be performed by looking for off-target cleavage at the fused sites by deep sequencing. On the other hand, translocations can also be directly validated through PCR or ddPCR with specific primers recognizing the two fusion partners (Bak et al., 2017). As translocations themselves exhibit a distinguishable element in the form of the specific point of fusion, the quantification of individual events becomes easier even without the addition of unique molecular identifiers (UMIs) barcodes.

Technical Challenges

All of these techniques have potential drawbacks in their methodology or bioinformatic analysis as summarized in Supplementary Tables 1 – 3 (excluding techniques not optimized to describe designer nucleases activity). Besides the above mentioned biological biases, the distinction between *in vitro* and *in cellula* assays is essential due to their expected difference in

sensitivity and hence in the number of returned off-target sites; for *in vitro* techniques, the introduction of an arbitrary threshold (e. g., the amount of reads per element) to select a smaller subset of sites could be an option, even though it may introduce a significant bias in the off-target detection. In a clinical setting, this bias could be reduced by further validating some of the sites that have been discarded by the arbitrary threshold, as for example those that are found close to oncogenic elements within a window of 50–100 kb of distance.

Another aspect that is crucial from the perspective of clinical application is the required amount of DNA or cell input, since attainable patient samples are limited. Some techniques might not be suitable for the analysis of the most clinically relevant samples. The introduction of additional elements (IDLV, DNA oligonucleotides) into the cell makes approaches like GUIDE-seq and IDLV capture unsuitable for performing the off-target screening directly on the cells intended for the treatment. When using these techniques, off-target detection must be performed in surrogate cell lines hence the cleavage footprint might not be accurate for a particular patient or the particular treatment due to diverging sequences, chromatin state, or DNA accessibility.

While the qualitative description of the off-target sites is an important information, their cleavage frequencies, and their reliable ranking can be equally of value, especially when monitoring the clonal expansion of modified cells in patients. Barcode sequences can be introduced at the cleavage site or upon adapter ligation prior to amplification steps in order to

quantify individual events in an unambiguous manner. On the other hand, a semiquantitative/quantitative information might still be retrieved, without barcoding, by calculating the relative reads amount of a certain mutation over the total amount of reads (Crosetto et al., 2013 ; Wienert et al., 2019) or by utilizing other unique molecular signatures such as the linker ligation point and the translocation fusion point (Zheng et al., 2014 ; Frock et al., 2015 ; Turchiano et al., 2020).

Not least the bioinformatic pipelines are of major importance due to the potential biases they can introduce or remove. Sequences filtering process is mainly borne by the reads/alignments quality and the reads amount counted in a defined genomic region. In order to avoid false positive results, the comparison with an untreated control can be beneficial, as it limits the biases coming from sequence misalignment or indexing hopping phenomena (Kircher et al., 2012) in multiplex NGS. A problematic practice is filtering out sites that do not reach a defined degree of homology to the on-target site. This kind of filtering, sometimes arbitrarily defined, can have a particular impact in case of differences between the patient's and the reference genome. Additionally, unspecific cleavage phenomena such as collateral activity of Cas12a (Li et al., 2018) would never be observed with this kind of filtering. This approach can be difficult to be applied when using other heterodimeric designer nucleases such as TALENs or ZFNs, especially if homodimers or other unintended dimers orientations and distances that can lead to cleavage are considered.

The parameters that may be used to define the potency of an assay could be the sensitivity and the accuracy. For a well-founded sensitivity assessment a known amount of potentially detectable events is ideally present within a sample or where a specific detected event can be quantified by other means and tested in a dilution series. Estimating the sensitivity based on measurements of indels for example suffers from the aforementioned uncertainty of the NGS analysis itself. It is also worth noting that the sensitivity of a technique depends also on the experimental conditions, and could be directly proportional to the numbers of cells treated, the amount of input material and the sequencing depth. The accuracy of this kind of technique relies on the ability to detect off-targeted sites with a minimal error rate or bias and is dependent by the amount of validated false positive and false negative sites. The validation process usually relies on deep sequencing of the inquired genomic regions to discern the false positive sites with all the drawbacks beforehand described. Supplementary Tables 1 – 3 reports the validation rates for the described techniques; however we do not have an objective and complete overview of all the mutations induced by the genome editing tools therefore calculating the false negative rate parameter is impossible with the current state of the art.

In this scenario, we also have to include the possibility of a designer nuclease having off-target activity in some genome widespread repetitive elements. Genomic instability events in these regions would be worrisome as they can be abundant, difficult to align and might be filtered out by the relative bioinformatic pipelines. In this case, use of unmasked reference genomes and a tolerant alignment algorithm together with the comparison

with the untreated control would help mitigating the bias and finding a balance with the accuracy rate.

Conclusions and Outlooks

Gene therapy based on integrating viral vector evolved in the last decades together with techniques and analyses that can at least in part evaluate safety (Modlich et al., 2006 ; Montini et al., 2006 ; Zhou et al., 2016 ; Biasco, 2017). In the same way, the different gene editing strategies are shaping novel reagents, techniques and strategies to improve their safety and efficiency for clinical application (Miller et al., 2007 ; Kleinstiver et al., 2016 ; Casini et al., 2018 ; Gao et al., 2018 ; Vakulskas et al., 2018 ; Rai et al., 2020). Further studies are required to understand and analyse the genotoxicity of new therapeutic strategies, and compare it with existing technologies. So far, the genotoxic effects of retroviral vectors employed in gene therapy approaches have been linked to insertional mutagenesis events mediated by viral enhancers (Bohne and Cathomen, 2008 ; Hacein-Bey-Abina et al., 2014), while designer nucleases act differently and may be more detrimental in regards to cell viability and genome integrity (Schiroli et al., 2019 ; Leibowitz et al., 2020). Delivery methods were also shown to impact differently on the mutational capacity of designer nucleases and the scientific community is moving toward a hit-and-run approach, utilizing ribonucleoproteins, or mRNA, that ensures a quick clearance of the exogenous nuclease and a more specific activity (Hendel et al., 2015). A balanced discussion should also consider the different impact of mutations in stem or in differentiated cells, with the latter likely to bear a lower risk of genotoxicity due to their shorter lifespan.

In light of these observations we can derive a new definition for genotoxicity which can be described as the property of an agent able to alter the genetic function within a cell causing unwanted mutations/effects, which may lead to functional impairment (e. g., cancer, therapy impairment, differentiation impairment).

CRISPR-Cas technology has largely democratized and accelerated the gene editing field but there are not yet standard techniques that can evaluate all the possible mutations induced directly or indirectly by the editing procedure. Recent publications revealed that off-targeting is responsible only for a minor portion of mutations characterized in edited cells, while there are on-target related mutations that justify careful evaluation (Kosicki et al., 2018 ; Connelly and Pruett-Miller, 2019 ; Cullot et al., 2019 ; Turchiano et al., 2020).

Hence, a combination of an *in cellula* technique for the discovery of off-targets sites (DISCOVER-seq, BLISS) and one for all other chromosomal aberrations (CAST-seq, HTGTS) would likely detect most of the unexpected mutations without the need to modify the current ex-vivo clinical procedures. *In vitro* techniques can describe a worst-case-scenario of off-target editing in a pre-clinical setting but require in addition a thorough validation via deep sequencing to exclude the abundant false positive sites that can be returned by the analysis, even when using base editors (see Supplementary Tables 2 , 3). In clinical settings, where hundreds of millions of cells need to be edited, the deep sequencing indel detection threshold of 0. 01–0. 1% may not be sufficient to detect the actual off-target activity. CAST-seq shows an

increased sensitivity reaching 0. 006% (1 mutation out of 15, 000 genome haplotypes) when compared with the absolute ddPCR quantification capacity, and hopefully also DSB detection techniques may be improved in the near future to achieve or lower that threshold in therapeutic settings.

In this review, we have highlighted the currently available techniques to detect DSBs, SSBs, translocations, and other chromosomal aberrations and the methods to quantify cleavage of designer nucleases. Overall, the amount of input DNA, the reliable quantification of events, an unbiased bioinformatic pipeline, the traceable sensitivity and the validation rate assessment are critical to evaluate the suitability of a technique.

The gene therapy field is moving fast; new molecular strategies are being proposed or are now under investigation in order to expand the applications and improve the editing efficiency. Prime editing (Anzalone et al., 2019), for example, could entail new potential. Alternatively, a new site specific and scareless integrative strategy could be developed soon by harnessing the transposase activity (Voigt et al., 2012 ; Klompe et al., 2019 ; Kovac et al., 2020), making another giant leap forward in the field. Advanced methods to assess genotoxicity of such technologies must be devised and will hopefully incorporate additional sensitivity and capacity to quantify all genetic modifications introduced by current and next-generation gene editing platforms.

Author Contributions

GB, AC, AT, and GT contributed to the final version of the manuscript. All authors contributed to the article and approved the submitted version.

Funding

Conflict of Interest

GT and GB have filed a patent application on CAST-seq. AT is on the Scientific Advisory Board of Orchard Therapeutics and Rocket Pharmaceuticals.

The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Supplementary Material

The Supplementary Material for this article can be found online at:

https://www. frontiersin. org/articles/10. 3389/. 2020. 613252/full#supplementary-material

References

Aida, T., Wilde, J. J., Yang, L., Hou, Y., Li, M., Xu, D., et al. (2020). Prime editing primarily induces undesired outcomes in mice. *bioRxiv* 2020. 2008. 2006. 239723. doi: 10. 1101/2020. 08. 06. 239723

CrossRef Full Text | Google Scholar

Aiuti, A., Biasco, L., Scaramuzza, S., Ferrua, F., Cicalese, M. P., Baricordi, C., et al. (2013). Lentiviral hematopoietic stem cell gene therapy in patients with Wiskott-Aldrich syndrome. *Science* 341: 1233151. doi: 10. 1126/science. 1233151

PubMed Abstract | CrossRef Full Text | Google Scholar

Anderson, W. F. (1990). September 14, 1990: the beginning. *Hum. Gene Ther.* 1, 371–372. doi: 10. 1089/hum. 1990. 1. 4-371

PubMed Abstract | CrossRef Full Text | Google Scholar

Anzalone, A. V., Randolph, P. B., Davis, J. R., Sousa, A. A., Koblan, L. W., Levy, J. M., et al. (2019). Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* 576, 149–157. doi: 10. 1038/s41586-019-1711-4

CrossRef Full Text | Google Scholar

Bae, S., Park, J., and Kim, J. S. (2014). Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics* 30, 1473–1475. doi: 10. 1093/bioinformatics/btu048

PubMed Abstract | CrossRef Full Text | Google Scholar

Bak, R. O., Dever, D. P., Reinisch, A., Cruz Hernandez, D., Majeti, R., and Porteus, M. H. (2017). Multiplexed genetic engineering of human

hematopoietic stem and progenitor cells using CRISPR/Cas9 and *AAV* 6. *Elife* 6. doi: 10. 7554/eLife. 27873. 025AQ

PubMed Abstract | CrossRef Full Text | Google Scholar

Baranello, L., Kouzine, F., Wojtowicz, D., Cui, K., Przytycka, T. M., Zhao, K., et al. (2014). DNA break mapping reveals topoisomerase II activity genome-wide. *Int. J. Mol. Sci.* 15, 13111–13122. doi: 10. 3390/ijms150713111

PubMed Abstract | CrossRef Full Text | Google Scholar

Biasco, L. (2017). Integration site analysis in gene therapy patients: expectations and reality. *Hum. Gene. Ther.* 28, 1122–1129. doi: 10. 1089/hum. 2017. 183

PubMed Abstract | CrossRef Full Text | Google Scholar

Blasco, R. B., Karaca, E., Ambrogio, C., Cheong, T. C., Karayol, E., Minero, V. G., et al. (2014). Simple and rapid *in vivo* generation of chromosomal rearrangements using CRISPR/Cas9 technology. *Cell Rep.* 9, 1219–1227. doi: 10. 1016/j. celrep. 2014. 10. 051

PubMed Abstract | CrossRef Full Text | Google Scholar

Bogdanove, A. J., and Voytas, D. F. (2011). TAL effectors: customizable proteins for DNA targeting. *Science* 333, 1843–1846. doi: 10. 1126/science. 1204094

PubMed Abstract | CrossRef Full Text | Google Scholar

Bohne, J., and Cathomen, T. (2008). Genotoxicity in gene therapy: an account of vector integration and designer nucleases. *Curr. Opin. Mol. Ther.* 10, 214–223.

PubMed Abstract | Google Scholar

Bonev, B., and Cavalli, G. (2016). Organization and function of the 3D genome. *Nat. Rev. Genet.* 17, 661–678. doi: 10. 1038/nrg. 2016. 112

CrossRef Full Text | Google Scholar

Bothmer, A., Phadke, T., Barrera, L. A., Margulies, C. M., Lee, C. S., Buquicchio, F., et al. (2017). Characterization of the interplay between DNA repair and CRISPR/Cas9-induced DNA lesions at an endogenous locus. *Nat. Commun.* 8: 13905. doi: 10. 1038/ncomms13905

PubMed Abstract | CrossRef Full Text | Google Scholar

Breton, C., Clark, P. M., Wang, L., Greig, J. A., and Wilson, J. M. (2020). ITR-seq, a next-generation sequencing assay, identifies genome-wide DNA editing sites *in vivo* following adeno-associated viral vector-mediated genome editing. *BMC Genomics* 21: 239. doi: 10. 1186/s12864-020-6655-4

CrossRef Full Text | Google Scholar

Brinkman, E. K., Chen, T., Amendola, M., and van Steensel, B. (2014). Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic. Acids Res.* 42: e168. doi: 10. 1093/nar/gku936

PubMed Abstract | CrossRef Full Text | Google Scholar

Cameron, P., Fuller, C. K., Donohoue, P. D., Jones, B. N., Thompson, M. S., Carter, M. M., et al. (2017). Mapping the genomic landscape of CRISPR-Cas9 cleavage. *Nat. Methods* 14, 600–606. doi: 10. 1038/nmeth. 4284

PubMed Abstract | CrossRef Full Text | Google Scholar

Canela, A., Sridharan, S., Sciascia, N., Tubbs, A., Meltzer, P., Sleckman, B. P., et al. (2016). DNA breaks and end resection measured genome-wide by end sequencing. *Mol. Cell* 63, 898–911. doi: 10. 1016/j. molcel. 2016. 06. 034

PubMed Abstract | CrossRef Full Text | Google Scholar

Cao, B., Wu, X., Zhou, J., Wu, H., Liu, L., Zhang, Q., et al. (2020). Nick-seq for single-nucleotide resolution genomic maps of DNA modifications and damage. *Nucleic Acids Res.* 48, 6715–6725. doi: 10. 1093/nar/gkaa473

PubMed Abstract | CrossRef Full Text | Google Scholar

Cao, H., Salazar-Garcia, L., Gao, F., Wahlestedt, T., Wu, C. L., Han, X., et al. (2019). Novel approach reveals genomic landscapes of single-strand DNA breaks with nucleotide resolution in human cells. *Nat. Commun.* 10: 5799. doi: 10. 1038/s41467-019-13602-7

PubMed Abstract | CrossRef Full Text | Google Scholar

Casini, A., Olivieri, M., Petris, G., Montagna, C., Reginato, G., Maule, G., et al. (2018). A highly specific SpCas9 variant is identified by *in vivo* screening in yeast. *Nat. Biotechnol.* 36, 265–271. doi: 10. 1038/nbt. 4066

PubMed Abstract | CrossRef Full Text | Google Scholar

Cattoglio, C., Facchini, G., Sartori, D., Antonelli, A., Miccio, A., Cassani, B., et al. (2007). Hot spots of retroviral integration in human CD34+ hematopoietic cells. *Blood* 110, 1770–1778. doi: 10. 1182/blood-2007-01-068759

PubMed Abstract | CrossRef Full Text | Google Scholar

Cavazzana-Calvo, M., Payen, E., Negre, O., Wang, G., Hehir, K., Fusil, F., et al. (2010). Transfusion independence and HMGA2 activation after gene therapy of human beta-thalassaemia. *Nature* 467, 318–322. doi: 10. 1038/nature09328

PubMed Abstract | CrossRef Full Text | Google Scholar

Chakrabarti, A. M., Henser-Brownhill, T., Monserrat, J., Poetsch, A. R., Luscombe, N. M., and Scaffidi, P. (2019). Target-specific precision of CRISPR-mediated genome editing. *Mol. Cell* 73, 699–713. e696. doi: 10. 1016/j. molcel. 2018. 11. 031

PubMed Abstract | CrossRef Full Text | Google Scholar

Chen, J. S., Ma, E., Harrington, L. B., Da Costa, M., Tian, X., Palefsky, J. M., et al. (2018). CRISPR-Cas12a target binding unleashes indiscriminate single-stranded DNase activity. *Science* 360, 436–439. doi: 10. 1126/science. aar6245

PubMed Abstract | CrossRef Full Text | Google Scholar

Chevalier, B. S., Monnat, R. J. Jr., and Stoddard, B. L. (2001). The homing endonuclease I-CreI uses three metals, one of which is shared between the two active sites. *Nat. Struct. Biol.* 8, 312–316. doi: 10. 1038/86181

PubMed Abstract | CrossRef Full Text | Google Scholar

Cho, S. W., Kim, S., Kim, Y., Kweon, J., Kim, H. S., Bae, S., et al. (2014). Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Res.* 24, 132–141. doi: 10. 1101/gr. 162339. 113

PubMed Abstract | CrossRef Full Text | Google Scholar

Concordet, J. P., and Haeussler, M. (2018). CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. *Nucleic. Acids Res.* 46, W242–W245. doi: 10. 1093/nar/gky354

PubMed Abstract | CrossRef Full Text | Google Scholar

Connelly, J. P., and Pruett-Miller, S. M. (2019). CRIS. py: a versatile and high-throughput analysis program for CRISPR-based genome editing. *Sci. Rep.* 9: 4194. doi: 10. 1038/s41598-019-40896-w

PubMed Abstract | CrossRef Full Text | Google Scholar

Corrigan-Curay, J., O'Reilly, M., Kohn, D. B., Cannon, P. M., Bao, G., Bushman, F. D., et al. (2015). Genome editing technologies: defining a path to clinic. *Mol. Ther.* 23, 796–806. doi: 10. 1038/mt. 2015. 54

PubMed Abstract | CrossRef Full Text | Google Scholar

Cradick, T. J., Qiu, P., Lee, C. M., Fine, E. J., and Bao, G. (2014). COSMID: a web-based tool for identifying and validating CRISPR/Cas off-target sites. *Mol. Ther. Nucleic Acids* 3: e214. doi: 10. 1038/mtna. 2014. 64

PubMed Abstract | CrossRef Full Text | Google Scholar

Crosetto, N., Mitra, A., Silva, M. J., Bienko, M., Dojer, N., Wang, Q., et al. (2013). Nucleotide-resolution DNA double-strand break mapping by next-generation sequencing. *Nat. Methods* 10, 361–365. doi: 10. 1038/nmeth. 2408

PubMed Abstract | CrossRef Full Text | Google Scholar

Cullot, G., Boutin, J., Toutain, J., Prat, F., Pennamen, P., Rooryck, C., et al. (2019). CRISPR-Cas9 genome editing induces megabase-scale chromosomal truncations. *Nat. Commun.* 10: 1136. doi: 10. 1038/s41467-019-09006-2

PubMed Abstract | CrossRef Full Text | Google Scholar

Dixon, J. R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., et al. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 485, 376–380. doi: 10. 1038/nature11082

PubMed Abstract | CrossRef Full Text | Google Scholar

Epinat, J. C., Arnould, S., Chames, P., Rochaix, P., Desfontaines, D., Puzin, C., et al. (2003). A novel engineered meganuclease induces homologous recombination in yeast and mammalian cells. *Nucleic Acids Res.* 31, 2952–2962. doi: 10. 1093/nar/gkg375

PubMed Abstract | CrossRef Full Text | Google Scholar

Frock, R. L., Hu, J., Meyers, R. M., Ho, Y. J., Kii, E., and Alt, F. W. (2015). Genome-wide detection of DNA double-stranded breaks induced by engineered nucleases. *Nat. Biotechnol.* 33, 179–186. doi: 10. 1038/nbt. 3101

PubMed Abstract | CrossRef Full Text | Google Scholar

Fu, Y., Foden, J. A., Khayter, C., Maeder, M. L., Reyon, D., Joung, J. K., et al. (2013). High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol.* 31, 822–826. doi: 10. 1038/nbt. 2623

PubMed Abstract | CrossRef Full Text | Google Scholar

Gabriel, R., Lombardo, A., Arens, A., Miller, J. C., Genovese, P., Kaeppel, C., et al. (2011). An unbiased genome-wide analysis of zinc-finger nuclease specificity. *Nat. Biotechnol.* 29, 816–823. doi: 10. 1038/nbt. 1948

PubMed Abstract | CrossRef Full Text | Google Scholar

Gao, X., Tao, Y., Lamas, V., Huang, M., Yeh, W. H., Pan, B., et al. (2018). Treatment of autosomal dominant hearing loss by *in vivo* delivery of genome editing agents. *Nature* 553, 217–221. doi: 10. 1038/nature25164

PubMed Abstract | CrossRef Full Text | Google Scholar

Gaspar, H. B., Parsley, K. L., Howe, S., King, D., Gilmour, K. C., Sinclair, J., et al. (2004). Gene therapy of X-linked severe combined immunodeficiency by

use of a pseudotyped gammaretroviral vector. *Lancet* 364, 2181–2187. doi: 10. 1016/S0140-6736(04)17590-9

PubMed Abstract | CrossRef Full Text | Google Scholar

Gaudelli, N. M., Komor, A. C., Rees, H. A., Packer, M. S., Badran, A. H., Bryson, D. I., et al. (2017). Programmable base editing of A * T to G * C in genomic DNA without DNA cleavage. *Nature* 551, 464–471. doi: 10. 1038/nature24644

PubMed Abstract | CrossRef Full Text | Google Scholar

Giannoukos, G., Ciulla, D. M., Marco, E., Abdulkerim, H. S., Barrera, L. A., Bothmer, A., et al. (2018). UDiTaS, a genome editing detection method for indels and genome rearrangements. *BMC Genomics* 19: 212. doi: 10. 1186/s12864-018-4561-9

PubMed Abstract | CrossRef Full Text | Google Scholar

Gijsbers, R., Ronen, K., Vets, S., Malani, N., De Rijck, J., McNeely, M., et al. (2010). LEDGF hybrids efficiently retarget lentiviral integration into heterochromatin. *Mol. Ther.* 18, 552–560. doi: 10. 1038/mt. 2010. 36

PubMed Abstract | CrossRef Full Text | Google Scholar

Grau, J., Boch, J., and Posch, S. (2013). TALENoffer: genome-wide TALEN off-target prediction. *Bioinformatics* 29, 2931–2932. doi: 10. 1093/bioinformatics/btt501

PubMed Abstract | CrossRef Full Text | Google Scholar

Hacein-Bey-Abina, S., Pai, S. Y., Gaspar, H. B., Armant, M., Berry, C. C., Blanche, S., et al. (2014). A modified gamma-retrovirus vector for X-linked severe combined immunodeficiency. *N. Engl. J. Med.* 371, 1407–1417. doi: 10. 1056/NEJMoa1404588

PubMed Abstract | CrossRef Full Text | Google Scholar

Hacein-Bey-Abina, S., Von Kalle, C., Schmidt, M., McCormack, M. P., Wulffraat, N., Leboulch, P., et al. (2003). LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. *Science* 302, 415–419. doi: 10. 1126/science. 1088547

CrossRef Full Text | Google Scholar

Hendel, A., Bak, R. O., Clark, J. T., Kennedy, A. B., Ryan, D. E., Roy, S., et al. (2015). Chemically modified guide RNAs enhance CRISPR-Cas genome editing in human primary cells. *Nat. Biotechnol.* 33, 985–989. doi: 10. 1038/nbt. 3290

PubMed Abstract | CrossRef Full Text | Google Scholar

Howe, S. J., Mansour, M. R., Schwarzwaelder, K., Bartholomae, C., Hubank, M., Kempski, H., et al. (2008). Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients. *J. Clin. Invest.* 118, 3143–3150. doi: 10. 1172/JCI35798

PubMed Abstract | CrossRef Full Text | Google Scholar

Hsiau, T., Conant, D., Rossi, N., Maures, T., Waite, K., Yang, J., et al. (2019). Inference of CRISPR edits from Sanger Trace Data. *bioRxiv* 251082. doi: 10. 1101/251082

CrossRef Full Text | Google Scholar

Hu, J., Meyers, R. M., Dong, J., Panchakshari, R. A., Alt, F. W., and Frock, R. L. (2016). Detecting DNA double-stranded breaks in mammalian genomes by linear amplification-mediated high-throughput genome-wide translocation sequencing. *Nat. Protoc.* 11, 853–871. doi: 10. 1038/nprot. 2016. 043

PubMed Abstract | CrossRef Full Text | Google Scholar

Iacovoni, J. S., Caron, P., Lassadi, I., Nicolas, E., Massip, L., Trouche, D., et al. (2010). High-resolution profiling of gammaH2AX around DNA double strand breaks in the mammalian genome. *EMBO J.* 29, 1446–1457. doi: 10. 1038/emboj. 2010. 38

PubMed Abstract | CrossRef Full Text | Google Scholar

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816–821. doi: 10. 1126/science. 1225829

PubMed Abstract | CrossRef Full Text | Google Scholar

Kim, D., Bae, S., Park, J., Kim, E., Kim, S., Yu, H. R., et al. (2015). Digenome-seq: genome-wide profiling of CRISPR-Cas9 off-target effects in human cells. *Nat. Methods* 12, 237–243, 231 p following 243. doi: 10. 1038/nmeth. 3284

PubMed Abstract | CrossRef Full Text | Google Scholar

Kim, D., Kim, D. E., Lee, G., Cho, S. I., and Kim, J. S. (2019a). Genome-wide target specificity of CRISPR RNA-guided adenine base editors. *Nat. Biotechnol.* 37, 430–435. doi: 10. 1038/s41587-019-0050-1

PubMed Abstract | CrossRef Full Text | Google Scholar

Kim, D., and Kim, J. S. (2018). DIG-seq: a genome-wide CRISPR off-target profiling method using chromatin DNA. *Genome Res.* 28, 1894–1900. doi: 10. 1101/gr. 236620. 118

PubMed Abstract | CrossRef Full Text | Google Scholar

Kim, D., Kim, S., Kim, S., Park, J., and Kim, J. S. (2016). Genome-wide target specificities of CRISPR-Cas9 nucleases revealed by multiplex Digenome-seq. *Genome Res.* 26, 406–415. doi: 10. 1101/gr. 199588. 115

PubMed Abstract | CrossRef Full Text | Google Scholar

Kim, D., Lim, K., Kim, S. T., Yoon, S. H., Kim, K., Ryu, S. M., et al. (2017). Genome-wide target specificities of CRISPR RNA-guided programmable deaminases. *Nat. Biotechnol.* 35, 475–480. doi: 10. 1038/nbt. 3852

CrossRef Full Text | Google Scholar

Kim, D., Luk, K., Wolfe, S. A., and Kim, J.-S. (2019b). Evaluating and enhancing target specificity of gene-editing nucleases and deaminases. *Ann. Rev. Biochem.* 88, 191–220. doi: 10. 1146/annurev-biochem-013118-111730

PubMed Abstract | CrossRef Full Text | Google Scholar

Kim, D. Y., Moon, S. B., Ko, J. H., Kim, Y. S., and Kim, D. (2020). Unbiased investigation of specificities of prime editing systems in human cells. *Nucleic Acids Res.* 48, 10576–10589. doi: 10. 1093/nar/gkaa764

PubMed Abstract | CrossRef Full Text | Google Scholar

Kim, S., Kim, D., Cho, S. W., Kim, J., and Kim, J. S. (2014). Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Res.* 24, 1012–1019. doi: 10. 1101/gr. 171322. 113

PubMed Abstract | CrossRef Full Text | Google Scholar

Kircher, M., Sawyer, S., and Meyer, M. (2012). Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. *Nucleic Acids Res.* 40: e3. doi: 10. 1093/nar/gkr771

PubMed Abstract | CrossRef Full Text | Google Scholar

Klein, I. A., Resch, W., Jankovic, M., Oliveira, T., Yamane, A., Nakahashi, H., et al. (2011). Translocation-capture sequencing reveals the extent and nature of chromosomal rearrangements in B lymphocytes. *Cell* 147, 95–106. doi: 10. 1016/j. cell. 2011. 07. 048

PubMed Abstract | CrossRef Full Text | Google Scholar

Kleinstiver, B. P., Pattanayak, V., Prew, M. S., Tsai, S. Q., Nguyen, N. T., Zheng, Z., et al. (2016). High-fidelity CRISPR-Cas9 nucleases with no

detectable genome-wide off-target effects. *Nature* 529, 490–495. doi: 10. 1038/nature16526

PubMed Abstract | CrossRef Full Text | Google Scholar

Klompe, S. E., Vo, P. L. H., Halpin-Healy, T. S., and Sternberg, S. H. (2019). Transposon-encoded CRISPR-Cas systems direct RNA-guided DNA integration. *Nature* 571, 219–225. doi: 10. 1038/s41586-019-1323-z

PubMed Abstract | CrossRef Full Text | Google Scholar

Kohn, D. B., Booth, C., Kang, E. M., Pai, S. Y., Shaw, K. L., Santilli, G., et al. (2020). Lentiviral gene therapy for X-linked chronic granulomatous disease. *Nat. Med.* 26, 200–206. doi: 10. 1038/s41591-019-0735-5

PubMed Abstract | CrossRef Full Text | Google Scholar

Koike-Yusa, H., Li, Y., Tan, E. P., Velasco-Herrera Mdel, C., and Yusa, K. (2014). Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. *Nat. Biotechnol.* 32, 267–273. doi: 10. 1038/nbt. 2800

PubMed Abstract | CrossRef Full Text | Google Scholar

Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A., and Liu, D. R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 533, 420–424. doi: 10. 1038/nature17946

PubMed Abstract | CrossRef Full Text | Google Scholar

Komor, A. C., Zhao, K. T., Packer, M. S., Gaudelli, N. M., Waterbury, A. L., Koblan, L. W., et al. (2017). Improved base excision repair inhibition and bacteriophage Mu Gam protein yields C: G-to-T: A base editors with higher efficiency and product purity. *Sci. Adv.* 3: eaao4774. doi: 10. 1126/sciadv. aao4774

PubMed Abstract | CrossRef Full Text | Google Scholar

Koo, T., Lee, J., and Kim, J. S. (2015). Measuring and reducing off-target activities of programmable nucleases including CRISPR-Cas9. *Mol. Cells* 38, 475–481. doi: 10. 14348/molcells. 2015. 0103

PubMed Abstract | CrossRef Full Text | Google Scholar

Kosicki, M., Tomberg, K., and Bradley, A. (2018). Repair of double-strand breaks induced by CRISPR-Cas9 leads to large deletions and complex rearrangements. *Nat. Biotechnol.* 36, 765–771. doi: 10. 1038/nbt. 4192

CrossRef Full Text | Google Scholar

Kovac, A., Miskey, C., Menzel, M., Grueso, E., Gogol-Doring, A., and Ivics, Z. (2020). RNA-guided retargeting of sleeping beauty transposition in human cells. *Elife* 9: 53868. doi: 10. 7554/eLife. 53868

PubMed Abstract | CrossRef Full Text | Google Scholar

Kurt, I. C., Zhou, R., Iyer, S., Garcia, S. P., Miller, B. R., Langner, L. M., et al. (2020). CRISPR C-to-G base editors for inducing targeted DNA transversions in human cells. *Nat. Biotechnol* . doi: 10. 1038/s41587-020-0609-x

PubMed Abstract | CrossRef Full Text | Google Scholar

Kuscu, C., Arslan, S., Singh, R., Thorpe, J., and Adli, M. (2014). Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. *Nat. Biotechnol.* 32, 677–683. doi: 10. 1038/nbt. 2916

PubMed Abstract | CrossRef Full Text | Google Scholar

Lazzarotto, C. R., Malinin, N. L., Li, Y., Zhang, R., Yang, Y., Lee, G., et al. (2020). CHANGE-seq reveals genetic and epigenetic effects on CRISPR-Cas9 genome-wide activity. *Nat. Biotechnol* . doi: 10. 1038/s41587-020-0555-7

PubMed Abstract | CrossRef Full Text | Google Scholar

Leibowitz, M. L., Papathanasiou, S., Doerfler, P. A., Blaine, L. J., Yao, Y., Zhang, C.-Z., et al. (2020). Chromothripsis as an on-target consequence of CRISPR-Cas9 genome editing. *bioRxiv* . 2020. 2007. 2013. 200998. doi: 10. 1101/2020. 07. 13. 200998

CrossRef Full Text | Google Scholar

Lensing, S. V., Marsico, G., Hansel-Hertsch, R., Lam, E. Y., Tannahill, D., and Balasubramanian, S. (2016). DSBCapture: in situ capture and sequencing of DNA breaks. *Nat. Methods* 13, 855–857. doi: 10. 1038/nmeth. 3960

PubMed Abstract | CrossRef Full Text | Google Scholar

Li, S. Y., Cheng, Q. X., Liu, J. K., Nie, X. Q., Zhao, G. P., and Wang, J. (2018). CRISPR-Cas12a has both cis- and trans-cleavage activities on single-stranded DNA. *Cell Res.* 28, 491–493. doi: 10. 1038/s41422-018-0022-x

PubMed Abstract | CrossRef Full Text | Google Scholar

Liang, P., Xie, X., Zhi, S., Sun, H., Zhang, X., Chen, Y., et al. (2019). Genome-wide profiling of adenine base editor specificity by EndoV-seq. *Nat. Commun.* 10: 67. doi: 10. 1038/s41467-018-07988-z

PubMed Abstract | CrossRef Full Text | Google Scholar

Maeder, M. L., and Gersbach, C. A. (2016). Genome-editing technologies for gene and cell therapy. *Mol. Ther.* 24, 430–446. doi: 10. 1038/mt. 2016. 10

PubMed Abstract | CrossRef Full Text | Google Scholar

Mashal, R. D., Koontz, J., and Sklar, J. (1995). Detection of mutations by cleavage of DNA heteroduplexes with bacteriophage resolvases. *Nat. Genet.* 9, 177–183. doi: 10. 1038/ng0295-177

PubMed Abstract | CrossRef Full Text | Google Scholar

McCormack, M. P., and Rabbitts, T. H. (2004). Activation of the T-cell oncogene LMO2 after gene therapy for X-linked severe combined immunodeficiency. *N. Engl. J. Med.* 350, 913–922. doi: 10. 1056/NEJMra032207

PubMed Abstract | CrossRef Full Text | Google Scholar

Metais, J. Y., and Dunbar, C. E. (2008). The MDS1-EVI1 gene complex as a retrovirus integration site: impact on behavior of hematopoietic cells and implications for gene therapy. *Mol. Ther.* 16, 439–449. doi: 10. 1038/sj. mt. 6300372

PubMed Abstract | CrossRef Full Text | Google Scholar

Miller, J. C., Holmes, M. C., Wang, J., Guschin, D. Y., Lee, Y. L., Rupniewski, I., et al. (2007). An improved zinc-finger nuclease architecture for highly specific genome editing. *Nat. Biotechnol.* 25, 778–785. doi: 10. 1038/nbt1319

PubMed Abstract | CrossRef Full Text | Google Scholar

Modlich, U., Bohne, J., Schmidt, M., von Kalle, C., Knoss, S., Schambach, A., et al. (2006). Cell-culture assays reveal the importance of retroviral vector design for insertional genotoxicity. *Blood* 108, 2545–2553. doi: 10. 1182/blood-2005-08-024976

PubMed Abstract | CrossRef Full Text | Google Scholar

Montague, T. G., Cruz, J. M., Gagnon, J. A., Church, G. M., and Valen, E. (2014). CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing. *Nucleic Acids Res.* 42, W401–407. doi: 10. 1093/nar/gku410

PubMed Abstract | CrossRef Full Text | Google Scholar

Montini, E., Cesana, D., Schmidt, M., Sanvito, F., Ponzoni, M., Bartholomae, C., et al. (2006). Hematopoietic stem cell gene transfer in a tumor-prone mouse model uncovers low genotoxicity of lentiviral vector integration. *Nat. Biotechnol.* 24, 687–696. doi: 10. 1038/nbt1216

PubMed Abstract | CrossRef Full Text | Google Scholar

Naldini, L., Blomer, U., Gage, F. H., Trono, D., and Verma, I. M. (1996). Efficient transfer, integration, and sustained long-term expression of the

transgene in adult rat brains injected with a lentiviral vector. *Proc. Natl. Acad. Sci. U. S. A.* 93, 11382–11388. doi: 10. 1073/pnas. 93. 21. 11382

PubMed Abstract | CrossRef Full Text | Google Scholar

Nathwani, A. C., Tuddenham, E. G., Rangarajan, S., Rosales, C., McIntosh, J., Linch, D. C., et al. (2011). Adenovirus-associated virus vector-mediated gene transfer in hemophilia B. *N. Engl. J. Med.* 365, 2357–2365. doi: 10. 1056/NEJMoa1108046

PubMed Abstract | CrossRef Full Text | Google Scholar

Ott, M. G., Schmidt, M., Schwarzwaelder, K., Stein, S., Siler, U., Koehl, U., et al. (2006). Correction of X-linked chronic granulomatous disease by gene therapy, augmented by insertional activation of MDS1-EVI1, PRDM16 or SETBP1. *Nat. Med.* 12, 401–409. doi: 10. 1038/nm1393

CrossRef Full Text | Google Scholar

Pinello, L., Canver, M. C., Hoban, M. D., Orkin, S. H., Kohn, D. B., Bauer, D. E., et al. (2016). Analyzing CRISPR genome-editing experiments with CRISPResso. *Nat. Biotechnol.* 34, 695–697. doi: 10. 1038/nbt. 3583

PubMed Abstract | CrossRef Full Text | Google Scholar

Qasim, W., Zhan, H., Samarasinghe, S., Adams, S., Amrolia, P., Stafford, S., et al. (2017). Molecular remission of infant B-ALL after infusion of universal TALEN gene-edited CAR T cells. *Sci. Transl. Med.* 9. doi: 10. 1126/scitranslmed. aaj2013

CrossRef Full Text | Google Scholar

Qiu, P., Shandilya, H., D'Alessio, J. M., O'Connor, K., Durocher, J., and Gerard, G. F. (2004). Mutation detection using surveyor nuclease. *Biotechniques* 36, 702–707. doi: 10. 2144/04364PF01

PubMed Abstract | CrossRef Full Text | Google Scholar

Rai, R., Romito, M., Rivers, E., Turchiano, G., Blattner, G., Vetharoy, W., et al. (2020). Targeted gene correction of human hematopoietic stem cells for the treatment of Wiskott–Aldrich syndrome. *Nat. Commun.* 11: 4034. doi: 10. 1038/s41467-020-17626-2

PubMed Abstract | CrossRef Full Text | Google Scholar

Ran, F. A., Cong, L., Yan, W. X., Scott, D. A., Gootenberg, J. S., Kriz, A. J., et al. (2015). *In vivo* genome editing using *Staphylococcus aureus* Cas9. *Nature* 520, 186–191. doi: 10. 1038/nature14299

PubMed Abstract | CrossRef Full Text | Google Scholar

Raper, S. E., Chirmule, N., Lee, F. S., Wivel, N. A., Bagg, A., Gao, G. P., et al. (2003). Fatal systemic inflammatory response syndrome in a ornithine transcarbamylase deficient patient following adenoviral gene transfer. *Mol. Genet. Metab.* 80, 148–158. doi: 10. 1016/j. ymgme. 2003. 08. 016

PubMed Abstract | CrossRef Full Text | Google Scholar

Rees, H. A., Wilson, C., Doman, J. L., and Liu, D. R. (2019). Analysis and minimization of cellular RNA editing by DNA adenine base editors. *Sci. Adv.* 5: eaax5717. doi: 10. 1126/sciadv. aax5717

PubMed Abstract | CrossRef Full Text | Google Scholar

Rosenberg, S. A., Aebersold, P., Cornetta, K., Kasid, A., Morgan, R. A., Moen, R., et al. (1990). Gene transfer into humans–immunotherapy of patients with advanced melanoma, using tumor-infiltrating lymphocytes modified by retroviral gene transduction. *N. Engl. J. Med.* 323, 570–578. doi: 10. 1056/NEJM199008303230904

PubMed Abstract | CrossRef Full Text | Google Scholar

Sapranauskas, R., Gasiunas, G., Fremaux, C., Barrangou, R., Horvath, P., and Siksnys, V. (2011). The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli* . *Nucleic Acids Res.* 39, 9275–9282. doi: 10. 1093/nar/gkr606

PubMed Abstract | CrossRef Full Text | Google Scholar

Schiroli, G., Conti, A., Ferrari, S., Della Volpe, L., Jacob, A., Albano, L., et al. (2019). Precise gene editing preserves hematopoietic stem cell function following transient p53-mediated DNA damage response. *Cell Stem Cell* 24, 551–565. e558. doi: 10. 1016/j. stem. 2019. 02. 019

CrossRef Full Text | Google Scholar

Schwarzwaelder, K., Howe, S. J., Schmidt, M., Brugman, M. H., Deichmann, A., Glimm, H., et al. (2007). Gammaretrovirus-mediated correction of SCID-X1 is associated with skewed vector integration site distribution *in vivo* . *J. Clin. Invest.* 117, 2241–2249. doi: 10. 1172/JCI31661

PubMed Abstract | CrossRef Full Text | Google Scholar

Sriramachandran, A. M., Petrosino, G., Mendez-Lago, M., Schafer, A. J., Batista-Nascimento, L. S., Zilio, N., et al. (2020). Genome-wide nucleotide-resolution mapping of DNA replication patterns, single-strand breaks, and lesions by GLOE-seq. *Mol. Cell* 78, 975–985. e977. doi: 10. 1016/j. molcel. 2020. 03. 027

PubMed Abstract | CrossRef Full Text | Google Scholar

Stadtmauer, E. A., Fraietta, J. A., Davis, M. M., Cohen, A. D., Weber, K. L., Lancaster, E., et al. (2020). CRISPR-engineered T cells in patients with refractory cancer. *Science* 367. doi: 10. 1126/science. aba7365

PubMed Abstract | CrossRef Full Text | Google Scholar

Stein, S., Ott, M. G., Schultze-Strasser, S., Jauch, A., Burwinkel, B., Kinner, A., et al. (2010). Genomic instability and myelodysplasia with monosomy 7 consequent to EVI1 activation after gene therapy for chronic granulomatous disease. *Nat. Med.* 16, 198–204. doi: 10. 1038/nm. 2088

PubMed Abstract | CrossRef Full Text | Google Scholar

Tebas, P., Stein, D., Tang, W. W., Frank, I., Wang, S. Q., Lee, G., et al. (2014). Gene editing of CCR5 in autologous CD4 T cells of persons infected with HIV. *N. Engl. J. Med.* 370, 901–910. doi: 10. 1056/NEJMoa1300662

PubMed Abstract | CrossRef Full Text | Google Scholar

Tsai, S. Q., Nguyen, N. T., Malagon-Lopez, J., Topkar, V. V., Aryee, M. J., and Joung, J. K. (2017). CIRCLE-seq: a highly sensitive *in vitro* screen for genome-wide CRISPR-Cas9 nuclease off-targets. *Nat. Methods* 14, 607–614. doi: 10. 1038/nmeth. 4278

PubMed Abstract | CrossRef Full Text | Google Scholar

Tsai, S. Q., Zheng, Z., Nguyen, N. T., Liebers, M., Topkar, V. V., Thapar, V., et al. (2015). GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat. Biotechnol.* 33, 187–197. doi: 10. 1038/nbt. 3117

PubMed Abstract | CrossRef Full Text | Google Scholar

Turchiano, G., Andrieux, G., Blattner, G., Pennucci, V., Klermund, J., Monaco, G., et al. (2020). Quantitative evaluation of chromosomal rearrangements in primary gene-edited human stem cells by preclinical CAST-seq. *Cell Stem Cell* . doi: 10. 2139/ssrn. 3565007

CrossRef Full Text | Google Scholar

Urnov, F. D., Miller, J. C., Lee, Y. L., Beausejour, C. M., Rock, J. M., Augustus, S., et al. (2005). Highly efficient endogenous human gene correction using

designed zinc-finger nucleases. *Nature* 435, 646–651. doi: 10. 1038/nature03556

PubMed Abstract | CrossRef Full Text | Google Scholar

Vakulskas, C. A., Dever, D. P., Rettig, G. R., Turk, R., Jacobi, A. M., Collingwood, M. A., et al. (2018). A high-fidelity Cas9 mutant delivered as a ribonucleoprotein complex enables efficient gene editing in human hematopoietic stem and progenitor cells. *Nat. Med.* 24, 1216–1224. doi: 10. 1038/s41591-018-0137-0

PubMed Abstract | CrossRef Full Text | Google Scholar

van Overbeek, M., Capurso, D., Carter, M. M., Thompson, M. S., Frias, E., Russ, C., et al. (2016). DNA Repair profiling reveals nonrandom outcomes at Cas9-mediated breaks. *Mol. Cell* 63, 633–646. doi: 10. 1016/j. molcel. 2016. 06. 037

PubMed Abstract | CrossRef Full Text | Google Scholar

Voigt, K., Gogol-Doring, A., Miskey, C., Chen, W., Cathomen, T., Izsvak, Z., et al. (2012). Retargeting sleeping beauty transposon insertions by engineered zinc finger DNA-binding domains. *Mol. Ther.* 20, 1852–1862. doi: 10. 1038/mt. 2012. 126

PubMed Abstract | CrossRef Full Text | Google Scholar

Vranckx, L. S., Demeulemeester, J., Debyser, Z., and Gijsbers, R. (2016). Towards a safer, more randomized lentiviral vector integration profile

exploring artificial LEDGF chimeras. *PLoS One* 11: e0164167. doi: 10. 1371/journal. pone. 0164167

PubMed Abstract | CrossRef Full Text | Google Scholar

Weinstock, D. M., Brunet, E., and Jasin, M. (2008). Induction of chromosomal translocations in mouse and human cells using site-specific endonucleases. *J. Natl. Cancer Inst. Monogr.* 39, 20–24. doi: 10. 1093/jncimonographs/lgn009

PubMed Abstract | CrossRef Full Text | Google Scholar

Wienert, B., Wyman, S. K., Richardson, C. D., Yeh, C. D., Akcakaya, P., Porritt, M. J., et al. (2019). Unbiased detection of CRISPR off-targets in vivo using DISCOVER-seq. *Science* 364, 286–289. doi: 10. 1101/469635

PubMed Abstract | CrossRef Full Text | Google Scholar

Wienert, B., Wyman, S. K., Yeh, C. D., Conklin, B. R., and Corn, J. E. (2020). CRISPR off-target detection with DISCOVER-seq. *Nat. Protoc.* 15, 1775–1799. doi: 10. 1038/s41596-020-0309-5

PubMed Abstract | CrossRef Full Text | Google Scholar

Wu, X., Scott, D. A., Kriz, A. J., Chiu, A. C., Hsu, P. D., Dadon, D. B., et al. (2014). Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nat. Biotechnol.* 32, 670–676. doi: 10. 1038/nbt. 2889

PubMed Abstract | CrossRef Full Text | Google Scholar

Xin, H., Wan, T., and Ping, Y. (2019). Off-targeting of base editors: BE3 but not ABE induces substantial off-target single nucleotide variants. *Signal. Transduct. Target Ther.* 4: 9. doi: 10. 1038/s41392-019-0044-y

PubMed Abstract | CrossRef Full Text | Google Scholar

Yan, W., Mirzazadeh, R., Garnerone, S., Scott, D., Schneider, M., Kallas, T., et al. (2019). Breaks labeling *in situ* and sequencing (BLISS). *Res. Square* . doi: 10. 21203/rs. 2. 1448/v2

PubMed Abstract | CrossRef Full Text | Google Scholar

Yan, W. X., Mirzazadeh, R., Garnerone, S., Scott, D., Schneider, M. W., Kallas, T., et al. (2017). BLISS is a versatile and quantitative method for genome-wide profiling of DNA double-strand breaks. *Nat. Commun.* 8: 15058. doi: 10. 1038/ncomms15058

PubMed Abstract | CrossRef Full Text | Google Scholar

Yang, L., Grishin, D., Wang, G., Aach, J., Zhang, C. Z., Chari, R., et al. (2014). Targeted and genome-wide sequencing reveal single nucleotide variations impacting specificity of Cas9 in human stem cells. *Nat. Commun.* 5: 5507. doi: 10. 1038/ncomms6507

PubMed Abstract | CrossRef Full Text | Google Scholar

Zeng, J., Wu, Y., Ren, C., Bonanno, J., Shen, A. H., Shea, D., et al. (2020). Therapeutic base editing of human hematopoietic stem cells. *Nat. Med.* 26, 535–541. doi: 10. 1038/s41591-020-0790-y

PubMed Abstract | CrossRef Full Text | Google Scholar

Zheng, Z., Liebers, M., Zhelyazkova, B., Cao, Y., Panditi, D., Lynch, K. D., et al. (2014). Anchored multiplex PCR for targeted next-generation sequencing. *Nat. Med.* 20, 1479–1484. doi: 10. 1038/nm. 3729

PubMed Abstract | CrossRef Full Text | Google Scholar

Zhou, S., Fatima, S., Ma, Z., Wang, Y. D., Lu, T., Janke, L. J., et al. (2016). Evaluating the safety of retroviral vectors based on insertional oncogene activation and blocked differentiation in cultured thymocytes. *Mol. Ther.* 24, 1090–1099. doi: 10. 1038/mt. 2016. 55

PubMed Abstract | CrossRef Full Text | Google Scholar