Genome Engineering Using Adeno-associated Virus: Basic and Clinical Research Applications

Thomas Gaj[1], Benjamin E Epstein[2] and David V Schaffer[1,2,3,4]

[1]Department of Chemical and Biomolecular Engineering, University of California, Berkeley, California, USA; [2]Department of Bioengineering, University of California, Berkeley, California, USA; [3]Department of Cell and Molecular Biology, University of California, Berkeley, California, USA; [4]Helen Wills Neuroscience Institute, University of California, Berkeley, California, USA

In addition to their broad potential for therapeutic gene delivery, adeno-associated virus (AAV) vectors possess the innate ability to stimulate homologous recombination in mammalian cells at high efficiencies. This process—referred to as AAV-mediated gene targeting—has enabled the introduction of a diverse array of genomic modifications both *in vitro* and *in vivo*. With the recent emergence of targeted nucleases, AAV-mediated genome engineering is poised for clinical translation. Here, we review key properties of AAV vectors that underscore its unique utility in genome editing. We highlight the broad range of genome engineering applications facilitated by this technology and discuss the strong potential for unifying AAV with targeted nucleases for next-generation gene therapy.

Received 7 June 2015; accepted 21 July 2015; advance online publication 29 September 2015. doi:10.1038/mt.2015.151

INTRODUCTION

Rapid advances in next-generation sequencing technologies are driving genome-wide association studies[1] and facilitating the annotation of previously unclassified genomic elements.[2] The resulting wealth of information is offering researchers unprecedented insights into the molecular basis of human disease. With the emergence of highly versatile genome editing tools,[3] including zinc-finger nucleases (ZFNs), TALE nucleases (TALENs), and CRISPR/Cas9, investigators are now positioned to capitalize on this information and develop new therapies that raise the possibility of correcting the underlying genetic causes of disease. However, successfully translating these concepts toward the clinic will require the development of methods and delivery vehicles that will both facilitate and enhance the ability of these tools to correct disease-associated mutations.

Adeno-associated virus (AAV) has emerged as a highly promising gene delivery vector due to its low immunogenicity and ability to mediate persistent gene expression in nondividing cells.[4] The potential of AAV is evidenced by its efficacy in a number of clinical trials, including those for hemophilia B,[5,6] Leber's congenital amaurosis type II,[7,8] choroideremia,[9] and lipoprotein lipase deficiency,[10,11] the last of which gained regulatory approval in the European Union in 2012. In addition to its growing therapeutic potential as a vehicle for gene delivery, AAV can serve as a donor template for homologous recombination (HR), the process by which two highly similar DNA sequences undergo strand exchange. Indeed, the AAV vector genome is endowed with the surprising ability to stimulate HR[12] at efficiencies that exceed conventional plasmid donor systems or other viral vectors. To date, this method, referred to as AAV-mediated gene targeting, has enabled the introduction of a broad range of genomic modifications into mammalian cells both *in vitro*[13] and *in vivo*.[14,15] Coupled with targeted nucleases—which can further enhance the efficiency of HR via induction of targeted DNA double-strand breaks (DSBs)[16]—AAV technology is now poised to accelerate both basic research and clinical applications of genome engineering.

Here, we review the key features of AAV that make it uniquely suited for genome engineering. We highlight the diverse range of genomic modifications enabled with this technology, as well as the prospects and potential for unifying AAV with targeted nucleases for human gene therapy.

AAV BIOLOGY

AAVs are nonenveloped, single-stranded DNA viruses that replicate only in the presence of helper virus, such as adenovirus or herpes simplex virus. The AAV viral genome is ~4.7 kilobases (kb) in length and contains two inverted terminal repeats that flank two genes, *rep* and *cap*, which encode proteins that facilitate viral replication and capsid assembly, respectively. By utilizing alternative splicing and start codons, the *rep* gene can be translated into four overlapping proteins (Rep78/Rep68 and Rep50/Rep42) that are essential for viral replication, integration, transcriptional regulation, and assembly. The *cap* gene can be translated into three structural proteins (VP1, 2, and 3) that self-assemble into a ~26 nm diameter icosahedral particle and, through use of an alternative reading frame, the assembly-activating protein (AAP),[17] which can assist in capsid formation (**Figure 1**). The inverted terminal repeats, which contain palindromic sequences that form an internal T-shaped hairpin structure with specific binding sites for Rep proteins,[18] are the only *cis* elements required for viral packaging. As a result, *rep* and *cap* can function in *trans* to support

Correspondence: David V Schaffer, Department of Chemical and Biomolecular Engineering, University of California, Berkeley, California, USA. E-mail: schaffer@berkeley.edu

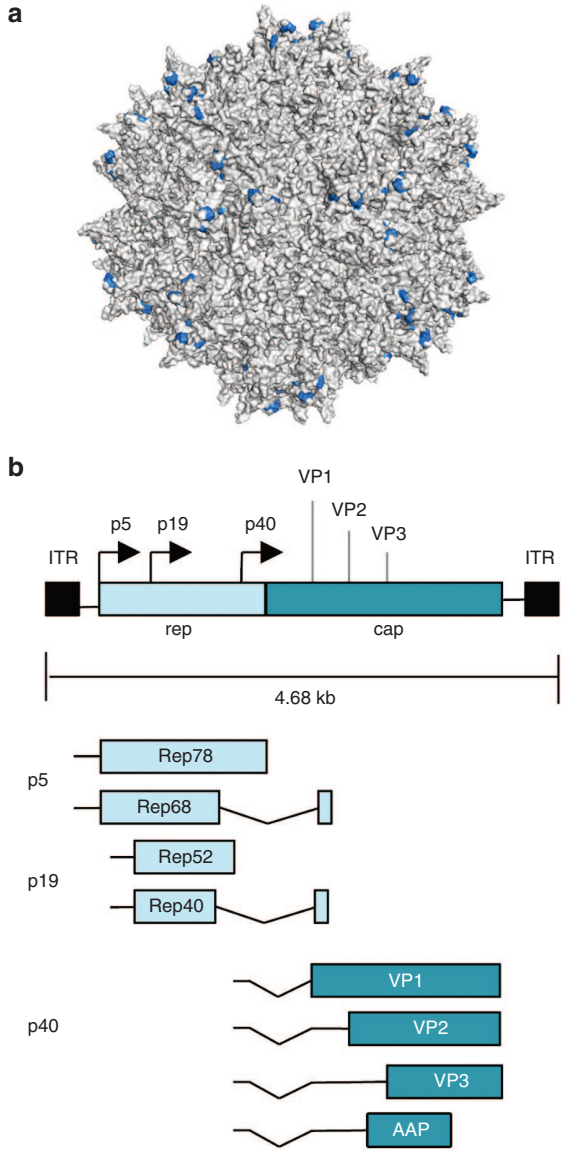

Figure 1 Adeno-associated virus (AAV) structure and genome organization. (**a**) Surface representation of the AAV2 capsid structure. The residues important for heparin binding, Arg 484, Arg 487, Lys 532, Arg 585, and Arg 588,[106] are colored blue (PDB ID: 1LP3).[107] (**b**) Structure of the wild-type AAV genome. Rep78 and Rep68 are expressed from the p5 promoter, and Rep52 and Rep40 are expressed from the p19 promoter. VP1, 2, 3, and the assembly-activating protein (AAP) are translated from the p40 transcript encoded by the cap gene. Solid black boxes indicate the inverted terminal repeats (ITRs).

virion assembly and production of vectors that deliver recombinant genetic payloads.

AAV infects cells by attaching to specific primary cell-surface receptors, such as heparin sulfate proteoglycans for AAV2[19] or sialic acid for AAV5,[20] and then to a secondary receptor that mediates endocytic uptake.[21] This choice of primary and secondary receptors strongly contributes to viral tropism. Once internalized, AAV traffics through the endocytic pathway, escapes the endosome with the aid of a phospholipase domain in the capsid, and transports to the nucleus where the viral genome is released and converted from single-stranded to double-stranded DNA in large part by host DNA polymerases.[22,23] The majority of these genomes then form concatemers that persist extrachromosomally as linear episomes within nondividing cells. In the presence of helper virus, wild-type AAV initiates a productive viral infection, while in the absence of helper, AAV can establish latency in the human genome through Rep-mediated integration.

NONHOMOLOGOUS INTEGRATION OF AAV INTO THE HUMAN GENOME

Wild-type AAV viruses encoding the *rep* gene integrate within a defined region of human chromosome 19, termed AAVS1 (**Figure 2**). Although AAVS1 contains no large regions of homology with the viral genome, up to 70% of AAV integration events occur within this site,[24,25] primarily through a nonhomologous deletion–substitution recombination mechanism.[26] While many details about this process remain unknown, studies using recombinant AAV vectors have indicated that this mechanism favors single-stranded over self-complementary genomes.[27] AAV integration into AAVS1 is mediated entirely by Rep proteins,[18,28] which recognize specific Rep-binding elements in the vector inverted terminal repeats[29] and an adjacent 138-bp *cis* integration efficiency element. Co-delivery of *rep*-deficient, nonviral AAV-derived plasmid[30–32] with Rep78/Rep68 facilitates site-specific integration into the AAVS1 locus. The efficiency of this process, however, is exceedingly low (~0.1%), and overexpression of the Rep proteins is associated with numerous undesirable side effects, including apoptosis.[33]

AAV vectors without the *rep* gene can also integrate at random chromosomal sites via non-HR[34,35] at efficiencies near 0.1% (**Figure 2**).[36] Numerous studies have mapped the integration preferences of AAV in multiple cell lines and tissues, yielding insights into some of the factors that drive this process. In particular, insertions have been found to predominantly occur within regions associated with genomic instability,[37–39] including segmental duplications, noncoding satellite DNA, palindromic sequences, and ribosomal RNA-encoding DNA repeats. Vector integration has also been frequently observed within CpG islands.[37,38] Other studies have reported that AAV insertions can occur within actively transcribed genes.[40,41] While these findings raise concern about the possibility of insertional mutagenesis[42] or aberrant gene activation or inactivation, the wealth of data collected to date indicates that AAV vectors are safe.[6,43]

The elucidation of "hot spots" where AAV integration events can occur, albeit at very low frequencies, provides insight into the potential mechanism for integration. These sites tend to have a highly repetitive nature and are thus dynamic and likely undergo routine recombination,[44] leading to transient DNA DSBs that have the potential to attach to the AAV vector genome via nonhomologous end joining.[45,46] Indeed, the frequency of AAV integration has been shown to be dramatically increased by DSB induction,[46] indicating that breaks are a critical factor driving nonhomologous integration.

AAV-MEDIATED GENE TARGETING

While highly useful for modifying certain cell types, such as mouse embryonic stem cells,[47] HR between exogenous DNA and a chromosomal locus has been limited in many other cell types, routinely displaying frequencies of 10^{-6} or lower. AAV

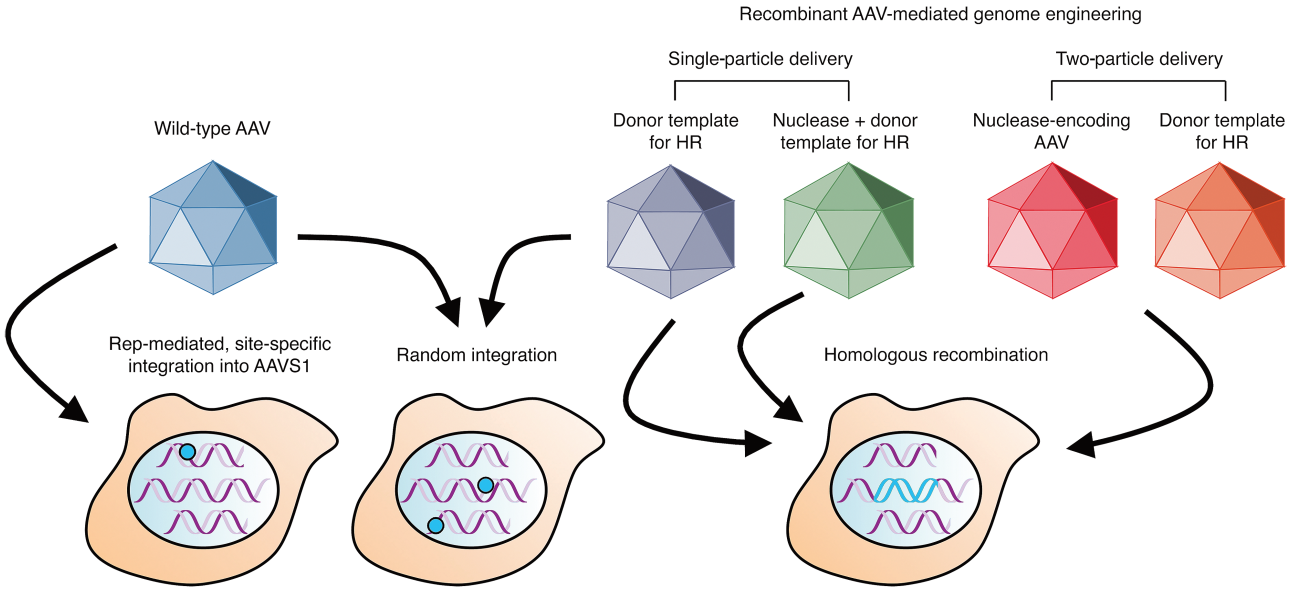

Figure 2 AAV integration into the human genome. Wild-type AAV vectors encoding the *rep* gene can facilitate AAV integration into a region of human chromosome 19 termed AAVS1, denoted by a blue circle. Wild-type and recombinant AAV vectors can also integrate into random chromosomal sites via nonhomologous end joining (denoted by blue circles). When the AAV vector genome is modified to contain a genomic sequence homologous to a specific chromosomal site, homologous recombination (HR) between the AAV vector and target site can occur (denoted by blue DNA fragment). Co-delivering a targeted nuclease within the same or a separate AAV particle can further enhance the frequency of HR.

vectors offer means to overcome this limitation (**Figure 2**). In 1998, Russell and Hirata[12] reported that an AAV viral genome possesses the innate ability to stimulate HR by 1,000-fold compared to conventional donor systems or other viral vectors.[48] In some cases, up to 1% of treated cells have been reported to undergo recombination.[49] Gene targeting by AAV, which proceeds through the canonical HR pathway,[50] can be achieved by introducing genome sequences homologous to a specific chromosomal region into the AAV vector (**Figure 3**). By altering the DNA sequence between these homology arms, a broad range of modifications can be introduced in a targeted manner by natural recombination with the target locus, including single-base substitutions,[51,52] and site-specific integration.[49] The efficiency and specificity afforded by AAV-mediated gene targeting has streamlined the creation of isogenic cell lines,[53–56] enabling the *in vitro* study of human disease. Impressively, AAV-mediated gene targeting has proven effective *in vivo*, enabling proof-of-principle correction of a mutant *lac*Z transgene in the ROSA26 locus,[14] and therapeutically relevant correction of the β-glucuronidase gene (whose loss of function causes mucopolysaccharidosis type VII[14]) at efficiencies of 1–2 corrected hepatocytes per 10^4 cells in a mouse model of the disease. Frequencies up to 10^{-3} have also been reported for repair of the fumarylacetoacetate hydrolase (Fah) gene in a mouse model of hereditary tyrosinemia.[57] Transgene integration into ribosomal DNA has also been demonstrated *in vivo*[58,59] by using AAV vectors containing homology to ribosomal DNA, which has shown increased chromosomal integration in comparison to other genomic targets. More recently, AAV was used to mediate integration of the F9 gene into the albumin locus,[60] leading to integration in ~0.5% of all albumin alleles in hepatocytes and amelioration of hemophilia B symptoms in a mouse model of the disease.

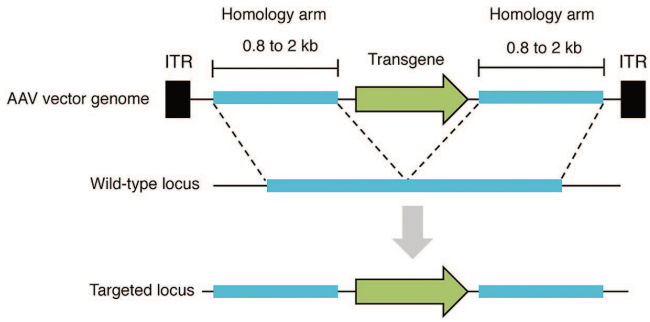

Figure 3 Overview of AAV-mediated gene targeting. AAV vectors containing DNA sequences homologous to a specific chromosomal site can be recombined with the matching genomic locus. By modifying the DNA sequence between the homology arms, targeted modifications (i.e., transgenes, single-base substitutions) can be introduced into the host genome. Dashes denote homologous regions of DNA. Black boxes indicate inverted terminal repeats (ITRs).

AAV-mediated integration can also facilitate human disease modeling, as evidenced by the *in vivo* insertion of enhancer/promoter elements to convert normal hepatocytes to hepatocellular carcinoma cells in the liver of mice.[61] In another example of AAV-mediated disease modeling, knockout and knockin modifications were introduced within the cystic fibrosis transmembrane conductance regulator gene in pig fetal fibroblasts,[62] enabling the generation of a model of cystic fibrosis in newborn pigs by somatic cell nuclear transfer.[15]

In addition to animal models, AAV has also shown success facilitating modifications *ex vivo* for potential regenerative medicine applications. In mesenchymal stem cells from patients with brittle bone disorder osteogenesis imperfecta, AAV was able to inactivate up to 90% of all dominant-negative COL1A1 mutant alleles,[13] and in primary keratinocytes derived from persons with

recessive inherited junctional epidermolysis bullosa, ~1% of mutant LAMA3 genes were corrected.[63] Additionally, repair of the human HPRT1 and HMGA1 genes in human embryonic stem cells was achieved.[64,65] AAV has also been used to reengineer the human leukocyte antigen,[66] indicating potential to combat graft-versus-host disease and generate universal cells from allogeneic donors.

The efficiency of AAV-mediated gene targeting depends both on the multiplicity of infection (MOI)[46] and the length of homology arms inserted into the AAV vectors,[67] though recombination frequencies with AAV genomes containing fourfold shorter homology arms than a conventional system can still be >100-fold higher. Indeed, gene targeting has been observed with vectors containing as little as 200 bases of homology on one arm,[67] although increased homology leads to higher targeting rates. Central positioning of the desired mutations within the viral genome has also been shown to increase modification rates.[67] In addition, the chromosomal position of the genomic target site has been reported to influence integration.[68] In particular, a recent analysis of over 2,000 targeted sites revealed that AAV vector integration is biased toward target sites located within transcriptional units,[69] with a preference for loci embedded within chromosomal genes being transcribed in the opposite direction of the gene-targeting event. The mechanism behind the enhanced efficiency of AAV-mediated gene targeting requires further elucidation but likely involves the linear single-stranded nature of the AAV viral genome,[67] which may simulate DNA damage and induce the cellular DNA repair pathway.

COMBINING AAV WITH TARGETED NUCLEASES

Although AAV-mediated gene targeting is highly versatile and can mediate the introduction of a variety of modifications, its *in vivo* efficiency is too low for most clinical applications, and to date, positive results been largely focused on the liver.[14,57,60] However, as with conventional plasmid DNA repair templates,[16,70,71] the induction of DSBs can increase the frequency of AAV-mediated gene targeting by up to 100-fold in human cells.[72,73] In recent years, a number of highly flexible tools, including ZFNs[74]—which are currently in clinical trials[75]—TALENs,[76] and CRISPR/Cas9[77] have emerged and endowed investigators with the ability to induce DSBs at user-specified genomic loci. Both ZFNs[78–81] and CRISPR/Cas9,[82] in particular, have been combined with AAV to induce targeted gene disruptions and chromosomal deletions in cell culture (**Figure 2**), including, for example, the inactivation of essential hepatitis B viral factors.[81] ZFNs have also been employed with AAV to enhance site-specific integration in a number of settings, and when combined with vector engineering and optimization approaches[4] can be extended to difficult-to-transduce targets such as human embryonic stem cells.[80] Indeed, evolved AAV variants with enhanced gene delivery capabilities are able to considerably increase the frequency of gene targeting.[80]

Combining targeted nucleases with AAV raises the possibility of therapeutic *in vivo* genome editing. As proof-of-principle, systemic delivery of AAV vectors encoding ZFNs targeting a defective copy of the F9 gene and its repair template have led to permanent correction and increased levels of F9 production in a murine model of hemophilia B,[83] indicating the potential for combining these technologies to correct inherited genetic disorders. More recently, the CRISPR/Cas9 system,[84] which enables highly efficient RNA-guided genome editing in the absence of protein engineering,[85,86] was combined with AAV to facilitate *in vivo* gene disruption in the brain and liver. Specifically, the *Streptococcus pyogenes* Cas9 (SpCas9) and its single guide RNA (sgRNA) were used to disrupt expression of multiple genes in mouse brain after stereotactic injection,[87] and more recently a Cas9 ortholog from *Staphylococcus aureus* (SaCas9) was used to modify the PCSK9 gene in the mouse liver,[88] leading to a ~95% decrease in Pcsk9 protein levels and a ~40% reduction in total cholesterol 1 week after systemic injection (CRISPR-mediated modification of the PCSK9 gene has also been achieved *in vivo* via adenoviral delivery[89]). Notably, unlike with SpCas9, only a single AAV particle was required to deliver SaCas9 and its sgRNA, as this variant is ~25% smaller than SpCas9. Indeed, the limited carrying capacity of AAV vectors (~4.7 kb) has impeded single particle delivery of SpCas9 (~4.2 kb) and its sgRNA for *in vivo* genome-editing applications, although a split-intein-mediated SpCas9 trans-splicing system has been developed to help address this challenge.[90]

In addition, AAV-mediated delivery of zinc-finger- and TALE-based transcription factors has enabled repression of mutant huntingtin protein in a mouse model of the disease,[91] and optogenetic control of gene expression in the mouse brain,[92] respectively, indicating the potential of combining AAV with a number of different tools capable of controlling gene expression.

ENHANCING AAV-MEDIATED GENOME ENGINEERING

Despite its ability to facilitate gene targeting *in vitro* and *in vivo*, numerous hurdles must still be overcome in order for AAV to reach its full potential for therapeutic genome engineering. Naturally occurring AAV serotypes are capable of infecting some cell types and tissues; however, evasion of neutralizing antibodies, improved biodistribution, tissue penetration, targeted delivery and, in particular, increased efficiency are needed in order to fully unlock the potential of therapeutic AAV gene delivery. Directed evolution can be harnessed to engineer tailored vectors with enhanced properties that can overcome such barriers.[4,93] These AAV vectors can be paired with targeted nucleases to mediate both cell- and gene-specific modifications.[80,94]

Additionally, off-target DSBs introduced by targeted nucleases could pose a serious challenge for AAV-mediated gene targeting, as these breaks can lead to undesired integration of the viral genome. The development of improved genome editing tools with refined specificity or enhanced nicking activity[95,96] could help mitigate these effects. The use of self-inactivating AAV vectors containing pseudocleavage sites embedded within the vector genome could also be used as means to prevent vector integration.[97] In addition, because the off-target activity of nucleases depends on their concentration within a cell, co-delivery of an AAV vector encoding a repair template[98] along with nuclease-encoding mRNA[99] or purified nuclease protein[100] could be utilized to reduce off-target effects. Since nucleases delivered into cells as mRNA or protein undergo rapid clearance, such an approach would limit the amount of time the cell is exposed to the risk of

off-target nuclease activity, yet still enable synergistic viral vector–mediated gene targeting. However, improvements in mRNA and protein delivery are likely needed for efficient *in vivo* implementation of this strategy.

Furthermore, while AAV vectors themselves are highly recombinogenic, the cellular proteins that facilitate HDR are expressed primarily during S phase of the cell cycle.[101] Thus, inducing efficient targeted correction in many therapeutically relevant cell types, including neurons, retinal cells, and cardiomyocytes, is challenging both *in vitro* and *in vivo*. One potential approach for overcoming this hurdle *in vitro* is tailoring culture conditions to increase the proportion of mitotic cells, thereby enabling more efficient genomic modifications.[102] Analogously, synchronizing cells in S phase can also help increase gene targeting *in vitro*.[103] In addition, small molecule inhibitors of the nonhomologous end joining pathway have been shown to enhance AAV-mediated gene targeting *in vivo*[104] and should be compatible with nuclease-induced DSBs. Transient cell cycle arrest has also been shown to enhance nuclease-assisted AAV-mediated gene targeting *in vitro*.[105] Nevertheless, there remains an urgent need for strategies that maximize HR in postmitotic cells *in vivo*, with one possibility being co-delivery of key components of the HR repair pathway within the AAV vector, though multiple particles may be needed to facilitate this.

CONCLUSIONS AND FUTURE DIRECTIONS

AAV vectors are highly promising therapeutic gene delivery vehicles that also offer the potential to facilitate and enhance many clinical applications of genome engineering. In order for AAV-mediated genome engineering to reach its full therapeutic potential, however, many important questions and challenges must be addressed. Chief among them: will AAV-mediated genome modifications be efficient enough to confer a therapeutic benefit? Questions also center on the long-term safety of delivering genome-modifying cargo to cells. In addition to the potential for off-target mutations and insertional mutagenesis, concerns over host immune responses to the AAV vector and the genome editing cargo must also be addressed. Furthermore, due to the size constraints associated with AAV vectors, it also remains unknown whether dual-particle delivery, which may be necessary for certain Cas9-based applications, can support sufficient levels of modification to yield a therapeutic effect. These challenges notwithstanding, AAV-mediated genome engineering is poised to usher in a new era of medicine that promises to convert genomic data to personalized and targeted therapies.

ACKNOWLEDGMENTS

This work was supported by the National Institutes of Health (NIH) grant R01EY022975. Molecular graphics were generated using PyMol (http://pymol.org).

REFERENCES

1. The Wellcome Trust Case Control Consortium (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447**: 661–678.
2. The ENCODE Project Consortium (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**: 57–74.
3. Gaj, T, Gersbach, CA and Barbas, CF 3rd (2013). ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol* **31**: 397–405.
4. Kotterman, MA and Schaffer, DV (2014). Engineering adeno-associated viruses for clinical gene therapy. *Nat Rev Genet* **15**: 445–451.
5. Nathwani, AC, Tuddenham, EG, Rangarajan, S, Rosales, C, McIntosh, J, Linch, DC *et al.* (2011). Adenovirus-associated virus vector-mediated gene transfer in hemophilia B. *N Engl J Med* **365**: 2357–2365.
6. Nathwani, AC, Reiss, UM, Tuddenham, EG, Rosales, C, Chowdary, P, McIntosh, J *et al.* (2014). Long-term safety and efficacy of factor IX gene therapy in hemophilia B. *N Engl J Med* **371**: 1994–2004.
7. Bennett, J, Ashtari, M, Wellman, J, Marshall, KA, Cyckowski, LL, Chung, DC *et al.* (2012). AAV2 gene therapy readministration in three adults with congenital blindness. *Sci Transl Med* **4**: 120ra15.
8. Bainbridge, JW, Smith, AJ, Barker, SS, Robbie, S, Henderson, R, Balaggan, K *et al.* (2008). Effect of gene therapy on visual function in Leber's congenital amaurosis. *N Engl J Med* **358**: 2231–2239.
9. MacLaren, RE, Groppe, M, Barnard, AR, Cottriall, CL, Tolmachova, T, Seymour, L *et al.* (2014). Retinal gene therapy in patients with choroideremia: initial findings from a phase ½ clinical trial. *Lancet* **383**: 1129–1137.
10. Stroes, ES, Nierman, MC, Meulenberg, JJ, Franssen, R, Twisk, J, Henny, CP *et al.* (2008). Intramuscular administration of AAV1-lipoprotein lipase S447X lowers triglycerides in lipoprotein lipase-deficient patients. *Arterioscler Thromb Vasc Biol* **28**: 2303–2304.
11. Carpentier, AC, Frisch, F, Labbé, SM, Gagnon, R, de Wal, J, Greentree, S *et al.* (2012). Effect of alipogene tiparvovec (AAV1-LPL(S447X)) on postprandial chylomicron metabolism in lipoprotein lipase-deficient patients. *J Clin Endocrinol Metab* **97**: 1635–1644.
12. Russell, DW and Hirata, RK (1998). Human gene targeting by viral vectors. *Nat Genet* **18**: 325–330.
13. Chamberlain, JR, Schwarze, U, Wang, PR, Hirata, RK, Hankenson, KD, Pace, JM *et al.* (2004). Gene targeting in stem cells from individuals with osteogenesis imperfecta. *Science* **303**: 1198–1201.
14. Miller, DG, Wang, PR, Petek, LM, Hirata, RK, Sands, MS and Russell, DW (2006). Gene targeting *in vivo* by adeno-associated virus vectors. *Nat Biotechnol* **24**: 1022–1026.
15. Rogers, CS, Stoltz, DA, Meyerholz, DK, Ostedgaard, LS, Rokhlina, T, Taft, PJ *et al.* (2008). Disruption of the CFTR gene produces a model of cystic fibrosis in newborn pigs. *Science* **321**: 1837–1841.
16. Rouet, P, Smih, F and Jasin, M (1994). Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells. *Proc Natl Acad Sci USA* **91**: 6064–6068.
17. Sonntag, F, Schmidt, K and Kleinschmidt, JA (2010). A viral assembly factor promotes AAV2 capsid formation in the nucleolus. *Proc Natl Acad Sci USA* **107**: 10220–10225.
18. Weitzman, MD, Kyöstiö, SR, Kotin, RM and Owens, RA (1994). Adeno-associated virus (AAV) Rep proteins mediate complex formation between AAV DNA and its integration site in human DNA. *Proc Natl Acad Sci USA* **91**: 5808–5812.
19. Summerford, C and Samulski, RJ (1998). Membrane-associated heparan sulfate proteoglycan is a receptor for adeno-associated virus type 2 virions. *J Virol* **72**: 1438–1445.
20. Walters, RW, Yi, SM, Keshavjee, S, Brown, KE, Welsh, MJ, Chiorini, JA *et al.* (2001). Binding of adeno-associated virus type 5 to 2,3-linked sialic acid is required for gene transfer. *J Biol Chem* **276**: 20610–20616.
21. Bartlett, JS, Wilcher, R and Samulski, RJ (2000). Infectious entry pathway of adeno-associated virus and adeno-associated virus vectors. *J Virol* **74**: 2777–2785.
22. Fisher, KJ, Gao, GP, Weitzman, MD, DeMatteo, R, Burda, JF and Wilson, JM (1996). Transduction with recombinant adeno-associated virus for gene therapy is limited by leading-strand synthesis. *J Virol* **70**: 520–532.
23. Ferrari, FK, Samulski, T, Shenk, T and Samulski, RJ (1996). Second-strand synthesis is a rate-limiting step for efficient transduction by recombinant adeno-associated virus vectors. *J Virol* **70**: 3227–3234.
24. Samulski, RJ, Zhu, X, Xiao, X, Brook, JD, Housman, DE, Epstein, N *et al.* (1991). Targeted integration of adeno-associated virus (AAV) into human chromosome 19. *EMBO J* **10**: 3941–3950.
25. Kotin, RM, Siniscalco, M, Samulski, RJ, Zhu, XD, Hunter, L, Laughlin, CA *et al.* (1990). Site-specific integration by adeno-associated virus. *Proc Natl Acad Sci USA* **87**: 2211–2215.
26. Dyall, J and Berns, KI (1998). Site-specific integration of adeno-associated virus into an episome with the target locus via a deletion-substitution mechanism. *J Virol* **72**: 6195–6198.
27. Daya, S, Cortez, N and Berns, KI (2009). Adeno-associated virus site-specific integration is mediated by proteins of the nonhomologous end-joining pathway. *J Virol* **83**: 11655–11664.
28. Young, SM Jr, McCarty, DM, Degtyareva, N and Samulski, RJ (2000). Roles of adeno-associated virus Rep protein and human chromosome 19 in site-specific recombination. *J Virol* **74**: 3953–3966.
29. Yang, CC, Xiao, X, Zhu, X, Ansardi, DC, Epstein, ND, Frey, MR *et al.* (1997). Cellular recombination pathways and viral terminal repeat hairpin structures are sufficient for adeno-associated virus integration *in vivo* and *in vitro*. *J Virol* **71**: 9231–9247.
30. Surosky, RT, Urabe, M, Godwin, SG, McQuiston, SA, Kurtzman, GJ, Ozawa, K *et al.* (1997). Adeno-associated virus Rep proteins target DNA sequences to a unique locus in the human genome. *J Virol* **71**: 7951–7959.
31. Pieroni, L, Fipaldini, C, Monciotti, A, Cimini, D, Sgura, A, Fattori, E *et al.* (1998). Targeted integration of adeno-associated virus-derived plasmids in transfected human cells. *Virology* **249**: 249–259.
32. Philpott, NJ, Gomos, J and Falck-Pedersen, E (2004). Transgene expression after rep-mediated site-specific integration into chromosome 19. *Hum Gene Ther* **15**: 47–61.
33. Schmidt, M, Afione, S and Kotin, RM (2000). Adeno-associated virus type 2 Rep78 induces apoptosis through caspase activation independently of p53. *J Virol* **74**: 9441–9450.
34. Ponnazhagan, S, Erikson, D, Kearns, WG, Zhou, SZ, Nahreini, P, Wang, XS *et al.* (1997). Lack of site-specific integration of the recombinant adeno-associated virus 2 genomes in human cells. *Hum Gene Ther* **8**: 275–284.
35. Samulski, RJ, Chang, LS and Shenk, T (1989). Helper-free stocks of recombinant adeno-associated viruses: normal integration does not require viral gene expression. *J Virol* **63**: 3822–3828.

36. McCarty, DM, Young, SM Jr and Samulski, RJ (2004). Integration of adeno-associated virus (AAV) and recombinant AAV vectors. *Annu Rev Genet* **38**: 819–845.
37. Miller, DG, Trobridge, GD, Petek, LM, Jacobs, MA, Kaul, R and Russell, DW (2005). Large-scale analysis of adeno-associated virus vector integration sites in normal human cells. *J Virol* **79**: 11434–11442.
38. Nakai, H, Wu, X, Fuess, S, Storm, TA, Munroe, D, Montini, E *et al.* (2005). Large-scale molecular characterization of adeno-associated virus vector integration in mouse liver. *J Virol* **79**: 3606–3614.
39. Inagaki, K, Lewis, SM, Wu, X, Ma, C, Munroe, DJ, Fuess, S *et al.* (2007). DNA palindromes with a modest arm length of greater, similar 20 base pairs are a significant target for recombinant adeno-associated virus vector integration in the liver, muscles, and heart in mice. *J Virol* **81**: 11290–11303.
40. Nakai, H, Montini, E, Fuess, S, Storm, TA, Grompe, M and Kay, MA (2003). AAV serotype 2 vectors preferentially integrate into active genes in mice. *Nat Genet* **34**: 297–302.
41. Janovitz, T, Klein, IA, Oliveira, T, Mukherjee, P, Nussenzweig, MC, Sadelain, M *et al.* (2013). High-throughput sequencing reveals principles of adeno-associated virus serotype 2 integration. *J Virol* **87**: 8559–8568.
42. Donsante, A, Miller, DG, Li, Y, Vogler, C, Brunt, EM, Russell, DW *et al.* (2007). AAV vector integration sites in mouse hepatocellular carcinoma. *Science* **317**: 477.
43. Gaudet, D, Méthot, J, Déry, S, Brisson, D, Essiembre, C, Tremblay, G *et al.* (2013). Efficacy and long-term safety of alipogene tiparvovec (AAV1-LPLS447X) gene therapy for lipoprotein lipase deficiency: an open-label trial. *Gene Ther* **20**: 361–369.
44. Shaw, CJ and Lupski, JR (2004). Implications of human genome architecture for rearrangement-based disorders: the genomic basis of disease. *Hum Mol Genet* **13 Spec No 1**: R57–R64.
45. Iizumi, S, Kurosawa, A, So, S, Ishii, Y, Chikaraishi, Y, Ishii, A *et al.* (2008). Impact of non-homologous end-joining deficiency on random and targeted DNA integration: implications for gene targeting. *Nucleic Acids Res* **36**: 6333–6342.
46. Miller, DG, Petek, LM and Russell, DW (2004). Adeno-associated virus vectors integrate at chromosome breakage sites. *Nat Genet* **36**: 767–773.
47. Thomas, KR and Capecchi, MR (1987). Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. *Cell* **51**: 503–512.
48. Ellis, J and Bernstein, A (1989). Gene targeting with retroviral vectors: recombination by gene conversion into regions of nonhomology. *Mol Cell Biol* **9**: 1621–1627.
49. Hirata, R, Chamberlain, J, Dong, R and Russell, DW (2002). Targeted transgene insertion into human chromosomes by adeno-associated virus vectors. *Nat Biotechnol* **20**: 735–738.
50. Vasileva, A, Linden, RM and Jessberger, R (2006). Homologous recombination is required for AAV-mediated gene targeting. *Nucleic Acids Res* **34**: 3345–3360.
51. Inoue, N, Dong, R, Hirata, RK and Russell, DW (2001). Introduction of single base substitutions at homologous chromosomal sequences by adeno-associated virus vectors. *Mol Ther* **3**: 526–530.
52. Inoue, N, Hirata, RK and Russell, DW (1999). High-fidelity correction of mutations at multiple chromosomal positions by adeno-associated virus vectors. *J Virol* **73**: 7376–7380.
53. Topaloglu, O, Hurley, PJ, Yildirim, O, Civin, CI and Bunz, F (2005). Improved methods for the generation of human gene knockout and knockin cell lines. *Nucleic Acids Res* **33**: e158.
54. Kohli, M, Rago, C, Lengauer, C, Kinzler, KW and Vogelstein, B (2004). Facile methods for generating human somatic cell gene knockouts using recombinant adeno-associated viruses. *Nucleic Acids Res* **32**: e3.
55. Khan, IF, Hirata, RK and Russell, DW (2011). AAV-mediated gene targeting methods for human cells. *Nat Protoc* **6**: 482–501.
56. Rago, C, Vogelstein, B and Bunz, F (2007). Genetic knockouts and knockins in human somatic cells. *Nat Protoc* **2**: 2734–2746.
57. Paulk, NK, Wursthorn, K, Wang, Z, Finegold, MJ, Kay, MA and Grompe, M (2010). Adeno-associated virus gene repair corrects a mouse model of hereditary tyrosinemia *in vivo*. *Hepatology* **51**: 1200–1208.
58. Lisowski, L, Lau, A, Wang, Z, Zhang, Y, Zhang, F, Grompe, M *et al.* (2012). Ribosomal DNA integrating rAAV-rDNA vectors allow for stable transgene expression. *Mol Ther* **20**: 1912–1923.
59. Wang, Z, Lisowski, L, Finegold, MJ, Nakai, H, Kay, MA and Grompe, M (2012). AAV vectors containing rDNA homology display increased chromosomal integration and transgene persistence. *Mol Ther* **20**: 1902–1911.
60. Barzel, A, Paulk, NK, Shi, Y, Huang, Y, Chu, K, Zhang, F *et al.* (2015). Promoterless gene targeting without nucleases ameliorates haemophilia B in mice. *Nature* **517**: 360–364.
61. Wang, PR, Xu, M, Toffanin, S, Li, Y, Llovet, JM and Russell, DW (2012). Induction of hepatocellular carcinoma by *in vivo* gene targeting. *Proc Natl Acad Sci USA* **109**: 11264–11269.
62. Rogers, CS, Hao, Y, Rokhlina, T, Samuel, M, Stoltz, DA, Li, Y *et al.* (2008). Production of CFTR-null and CFTR-DeltaF508 heterozygous pigs by adeno-associated virus-mediated gene targeting and somatic cell nuclear transfer. *J Clin Invest* **118**: 1571–1577.
63. Melo, SP, Lisowski, L, Bashkirova, E, Zhen, HH, Chu, K, Keene, DR *et al.* (2014). Somatic correction of junctional epidermolysis bullosa by a highly recombinogenic AAV variant. *Mol Ther* **22**: 725–733.
64. Khan, IF, Hirata, RK, Wang, PR, Li, Y, Kho, J, Nelson, A *et al.* (2010). Engineering of human pluripotent stem cells by AAV-mediated gene targeting. *Mol Ther* **18**: 1192–1199.
65. Mitsui, K, Suzuki, K, Aizawa, E, Kawase, E, Suemori, H, Nakatsuji, N *et al.* (2009). Gene targeting in human pluripotent stem cells with adeno-associated virus vectors. *Biochem Biophys Res Commun* **388**: 711–717.
66. Riolobos, L, Hirata, RK, Turtle, CJ, Wang, PR, Gornalusse, GG, Zavajlevski, M *et al.* (2013). HLA engineering of human pluripotent stem cells. *Mol Ther* **21**: 1232–1241.
67. Hirata, RK and Russell, DW (2000). Design and packaging of adeno-associated virus gene targeting vectors. *J Virol* **74**: 4612–4620.
68. Cornea, AM and Russell, DW (2010). Chromosomal position effects on AAV-mediated gene targeting. *Nucleic Acids Res* **38**: 3582–3594.
69. Deyle, DR, Hansen, RS, Cornea, AM, Li, LB, Burt, AA, Alexander, IE *et al.* (2014). A genome-wide map of adeno-associated virus-mediated human gene targeting. *Nat Struct Mol Biol* **21**: 969–975.
70. Bibikova, M, Carroll, D, Segal, DJ, Trautman, JK, Smith, J, Kim, YG *et al.* (2001). Stimulation of homologous recombination through targeted cleavage by chimeric nucleases. *Mol Cell Biol* **21**: 289–297.
71. Porteus, MH and Baltimore, D (2003). Chimeric nucleases stimulate gene targeting in human cells. *Science* **300**: 763.
72. Miller, DG, Petek, LM and Russell, DW (2003). Human gene targeting by adeno-associated virus vectors is enhanced by DNA double-strand breaks. *Mol Cell Biol* **23**: 3550–3557.
73. Porteus, MH, Cathomen, T, Weitzman, MD and Baltimore, D (2003). Efficient gene targeting mediated by adeno-associated virus and DNA double-strand breaks. *Mol Cell Biol* **23**: 3558–3565.
74. Urnov, FD, Rebar, EJ, Holmes, MC, Zhang, HS and Gregory, PD (2010). Genome editing with engineered zinc finger nucleases. *Nat Rev Genet* **11**: 636–646.
75. Tebas, P, Stein, D, Tang, WW, Frank, I, Wang, SQ, 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.
76. Joung, JK and Sander, JD (2013). TALENs: a widely applicable technology for targeted genome editing. *Nat Rev Mol Cell Biol* **14**: 49–55.
77. Doudna, JA and Charpentier, E (2014). Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science* **346**: 1258096.
78. Händel, EM, Gellhaus, K, Khan, K, Bednarski, C, Cornu, TI, Müller-Lerch, F *et al.* (2012). Versatile and efficient genome editing in human cells by combining zinc-finger nucleases with adeno-associated viral vectors. *Hum Gene Ther* **23**: 321–329.
79. Ellis, BL, Hirsch, ML, Porter, SN, Samulski, RJ and Porteus, MH (2013). Zinc-finger nuclease-mediated gene correction using single AAV vector transduction and enhancement by Food and Drug Administration-approved drugs. *Gene Ther* **20**: 35–42.
80. Asuri, P, Bartel, MA, Vazin, T, Jang, JH, Wong, TB and Schaffer, DV (2012). Directed evolution of adeno-associated virus for enhanced gene delivery and gene targeting in human pluripotent stem cells. *Mol Ther* **20**: 329–338.
81. Weber, ND, Stone, D, Sedlak, RH, De Silva Feelixge, HS, Roychoudhury, P, Schiffer, JT *et al.* (2014). AAV-mediated delivery of zinc finger nucleases targeting hepatitis B virus inhibits active replication. *PLoS One* **9**: e97579.
82. Senís, E, Fatouros, C, Große, S, Wiedtke, E, Niopek, D, Mueller, AK *et al.* (2014). CRISPR/Cas9-mediated genome engineering: an adeno-associated viral (AAV) vector toolbox. *Biotechnol J* **9**: 1402–1412.
83. Li, H, Haurigot, V, Doyon, Y, Li, T, Wong, SY, Bhagwat, AS *et al.* (2011). *In vivo* genome editing restores haemostasis in a mouse model of haemophilia. *Nature* **475**: 217–221.
84. Jinek, M, Chylinski, K, Fonfara, I, Hauer, M, Doudna, JA and Charpentier, E (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**: 816–821.
85. Cong, L, Ran, FA, Cox, D, Lin, S, Barretto, R, Habib, N *et al.* (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**: 819–823.
86. Mali, P, Yang, L, Esvelt, KM, Aach, J, Guell, M, DiCarlo, JE *et al.* (2013). RNA-guided human genome engineering via Cas9. *Science* **339**: 823–826.
87. Swiech, L, Heidenreich, M, Banerjee, A, Habib, N, Li, Y, Trombetta, J *et al.* (2015). *In vivo* interrogation of gene function in the mammalian brain using CRISPR-Cas9. *Nat Biotechnol* **33**: 102–106.
88. Ran, FA, Cong, L, Yan, WX, Scott, DA, Gootenberg, JS, Kriz, AJ *et al.* (2015). *In vivo* genome editing using *Staphylococcus aureus* Cas9. *Nature* **520**: 186–191.
89. Ding, Q, Strong, A, Patel, KM, Ng, SL, Gosis, BS, Regan, SN *et al.* (2014). Permanent alteration of PCSK9 with *in vivo* CRISPR-Cas9 genome editing. *Circ Res* **115**: 488–492.
90. Truong, DJ, Kühner, K, Kühn, R, Werfel, S, Engelhardt, S, Wurst, W *et al.* (2015). Development of an intein-mediated split-Cas9 system for gene therapy. *Nucleic Acids Res* **43**: 6450–6458.
91. Garriga-Canut, M, Agustín-Pavón, C, Herrmann, F, Sánchez, A, Dierssen, M, Fillat, C *et al.* (2012). Synthetic zinc finger repressors reduce mutant huntingtin expression in the brain of R6/2 mice. *Proc Natl Acad Sci USA* **109**: E3136–E3145.
92. Konermann, S, Brigham, MD, Trevino, AE, Hsu, PD, Heidenreich, M, Cong, L *et al.* (2013). Optical control of mammalian endogenous transcription and epigenetic states. *Nature* **500**: 472–476.
93. Dalkara, D, Byrne, LC, Klimczak, RR, Visel, M, Yin, L, Merigan, WH *et al.* (2013). In vivo-directed evolution of a new adeno-associated virus for therapeutic outer retinal gene delivery from the vitreous. *Sci Transl Med* **5**: 189ra76.
94. Jang, JH, Koerber, JT, Kim, JS, Asuri, P, Vazin, T, Bartel, M *et al.* (2011). An evolved adeno-associated viral variant enhances gene delivery and gene targeting in neural stem cells. *Mol Ther* **19**: 667–675.
95. Wang, J, Friedman, G, Doyon, Y, Wang, NS, Li, CJ, Miller, JC *et al.* (2012). Targeted gene addition to a predetermined site in the human genome using a ZFN-based nicking enzyme. *Genome Res* **22**: 1316–1326.
96. Ran, FA, Hsu, PD, Lin, CY, Gootenberg, JS, Konermann, S, Trevino, AE *et al.* (2013). Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* **154**: 1380–1389.
97. Moore, R, Spinhirne, A, Lai, MJ, Preisser, S, Li, Y, Kang, T *et al.* (2015). CRISPR-based self-cleaving mechanism for controllable gene delivery in human cells. *Nucleic Acids Res* **43**: 1297–1303.
98. Kaulich, M, Lee, YJ, Lönn, P, Springer, AD, Meade, BR and Dowdy, SF (2015). Efficient CRISPR-rAAV engineering of endogenous genes to study protein function by allele-specific RNAi. *Nucleic Acids Res* **43**: e45.
99. Mahiny, AJ, Dewerth, A, Mays, LE, Alkhaled, M, Mothes, B, Malaeksefat, E *et al.* (2015). *In vivo* genome editing using nuclease-encoding mRNA corrects SP-B deficiency. *Nat Biotechnol* **33**: 584–586.
100. Gaj, T, Guo, J, Kato, Y, Sirk, SJ and Barbas, CF 3rd (2012). Targeted gene knockout by direct delivery of zinc-finger nuclease proteins. *Nat Methods* **9**: 805–807.

101. Wong, EA and Capecchi, MR (1987). Homologous recombination between coinjected DNA sequences peaks in early to mid-S phase. *Mol Cell Biol* **7**: 2294–2295.
102. Genovese, P, Schiroli, G, Escobar, G, Di Tomaso, T, Firrito, C, Calabria, A *et al.* (2014). Targeted genome editing in human repopulating haematopoietic stem cells. *Nature* **510**: 235–240.
103. Liu, X, Yan, Z, Luo, M, Zak, R, Li, Z, Driskell, RR *et al.* (2004). Targeted correction of single-base-pair mutations with adeno-associated virus vectors under nonselective conditions. *J Virol* **78**: 4165–4175.
104. Paulk, NK, Loza, LM, Finegold, MJ and Grompe, M (2012). AAV-mediated gene targeting is significantly enhanced by transient inhibition of nonhomologous end joining or the proteasome in vivo. *Hum Gene Ther* **23**: 658–665.
105. Rahman, SH, Bobis-Wozowicz, S, Chatterjee, D, Gellhaus, K, Pars, K, Heilbronn, R *et al.* (2013). The nontoxic cell cycle modulator indirubin augments transduction of adeno-associated viral vectors and zinc-finger nuclease-mediated gene targeting. *Hum Gene Ther* **24**: 67–77.
106. Opie, SR, Warrington, KH Jr, Agbandje-McKenna, M, Zolotukhin, S and Muzyczka, N (2003). Identification of amino acid residues in the capsid proteins of adeno-associated virus type 2 that contribute to heparan sulfate proteoglycan binding. *J Virol* **77**: 6995–7006.
107. Xie, Q, Bu, W, Bhatia, S, Hare, J, Somasundaram, T, Azzi, A *et al.* (2002). The atomic structure of adeno-associated virus (AAV-2), a vector for human gene therapy. *Proc Natl Acad Sci USA* **99**: 10405–10410.