Soft genome editing based on CRISPR nickases: it takes one break to tango

Wang, Q.

The manipulation of the human genome with customized genetic information makes it possible to further decipher the basis of biological processes under both physiological and disease states. In the past decades, rapid technological breakthroughs originating from fundamental microbiology research have yielded novel genome engineering tools and principles that greatly facilitate our ability to efficiently modify specific genomic sequences in living cells and organisms. As a corollary, these technologies are also starting to permeate the realm of medicine when applied as a form of "genomic surgery". These genetic therapies aim at tacking the root cause of human pathologies, inherited or acquired, by correcting or modulating the genetic content or expression, respectively, present in target cells, tissues and organs. To this end, delivery vehicles capable of introducing, in an efficient and safe manner, the increasingly sophisticated (epi)genome editing reagents are in demand, especially when considering *in vivo* genetic therapies.

Owing to their robustness, simplicity, and versatility, engineered RNA-guided nucleases (RGNs) built on prokaryotic clustered regularly interspaced short palindromic repeat (CRISPR)–Cas9 systems, consisting of sequence-tailored single guide RNAs (gRNAs) coupled to Cas9 endonucleases, remain amongst the most powerful genome editing tools since their introduction by independent groups in 2013. Commonly, RGN-based genome editing manoeuvres start by the triggering of site-specific chromosomal double-stranded DNA breaks (DSBs) that, upon endogenous DNA repair pathways activation, yield gene knockouts and, in the presence of exogenous donor DNA, gene knock-ins. Generically, the former and latter genome editing outcomes involve the recruitment of non-homologous end joining (NHEJ) and homology-directed repair (HDR) factors, respectively, at the RGN-induced DSBs. Yet, precise genome editing is often hindered due to the multiple-copy character of the vast majority of chromosomal sequences and off-target RGN activities. Additionally, targeted DSBs required for cellular DNA repair activation as well as DSBs resulting from off-target DNA cleavage inevitably produce inaccurate and unpredictable genetic structural variants in the form of small insertions and deletions (indels) and local or genome-wide chromosomal rearrangements, e.g., duplications, large deletions and/or translocations. Moreover, regardless of their specificity, RGNs trigger P53-dependent cell cycle arrest and apoptosis. Indeed, the activation of this DNA damage response (DDR) limits the efficacy of genome editing procedures. This is especially so in the case of DSB-dependent genome editing in regular P53 proficient stem cells that serve as highly relevant substrates for human disease modelling and therapy, e.g., induced pluripotent stem cells (iPSCs) and hematopoietic stem cells, respectively. Equally insidious, DDR activation is known to create selective pressure for the emergence of gain-of-function and loss-of-function gene mutations linked to tumorigenesis.

Hence, the research presented in this thesis is primarily directed to the heightening of the specificity and fidelity of genome editing procedures by investigating and harnessing nicking RGNs and their prime editing derivatives based on prototypic CRISPR–Cas9 systems. Unlike regular Cas9 nucleases, sequence- and strand-specific Cas9 nucleases ("nickases") contain either their RuvC or HNH nuclease domains disabled. Therefore, when compared with intrinsically mutagenic DSBs, single-stranded DNA breaks (SSBs), or nicks, made by such RNA-programmable enzymes, are less disruptive to the genome in that they do not constitute canonical substrates for error-prone DNA repair processes, e.g., NHEJ and microhomology-mediated end joining. Importantly, albeit at low frequencies, SSBs are capable nonetheless of triggering HDR in mammalian cells. This knowledge has laid a foundation for further investigating herein the *in trans* paired nicking (ITPN) concept based on enhancing HDR-mediated genome editing by combining nicking RGNs with nicking-susceptible donor DNA constructs. In addition, the feasibility and utility of deploying adenoviral vector (AdV) technologies for the purpose of prime editing in DSB-sensitive and hard-to-transfect cell types, namely, muscle progenitors, mesenchymal stem cells and iPSCs, is established. Finally, research presented in this thesis further discloses a role for higher-order chromatin conformations on the ultimate efficiency and fidelity of prime editors and base editors, both comprising nicking RGNs fused to effector domains responsible for the installation of specific genomic edits in a DSB- and donor DNA-independent manner.

Chapter 1 serves as an introductory chapter in that it provides a detailed overview about the principles governing the main genome editing strategies and associated effector platforms focusing on those that have entered the clinical trial arena, i.e., zinc-finger nucleases, transcription activator-like effector nucleases and RNA-guided CRISPR-Cas9 nucleases. In addition, **Chapter 1** reviews applications of these genome editing tools and strategies in human stem cells focusing on the use of adenoviral vectors

(AdVs) as delivery vehicles. This chapter further highlights the opportunities offered by high-capacity AdVs (HC-AdVs) in particular for ferrying large donor DNA payloads and DNA-editing fusion constructs, such as those underlying DSB-independent base editing and prime editing processes.

Research presented in **Chapter 2** formally demonstrates that the formation of indels resulting from RGN-induced DSBs at target sites can lead to the loss of fitness by gene-edited cells and reports that simultaneous SSB formation at donor DNA constructs and acceptor chromosomal sequences by nicking RGNs (i.e., ITPN genome editing), can overcome such disruptive genotype-phenotype associations. Moreover, ITPN compared favorably with DNA manipulations involving the exclusive formation of SSBs or DSBs at chromosomal sequences as it yields more frequent and seamless, respectively, HDR-mediated genome editing events in human cells.

A major concern in the genome editing field as a whole, that acquires particular relevance when considering therapeutic gene-editing interventions, regards the activity of RGNs at off-target chromosomal sequences. These unintended off-target activities and associated collateral effects result from the fact that wild-type Cas9 proteins often remain proficient at DNA cleaving even when multiple mismatches and/or bulges exist between gRNA spacer and genomic sequences. This is especially the case when the mismatches and/or bulges locate distally to protospacer adjacent motif (PAM) sites, i.e., sites that constitute the initial engagement points of RGN complexes with DNA. Therefore, the cumulative work described in **Chapter 3** and **Chapter 4** sought to tackle this issue through the assembly, testing and validation of dual and single RGNs with nicking Cas9 variants capable of triggering gene knock-outs and gene knock-ins in an efficient and highly specific manner. In particular, in **Chapter 3**, a systematic assessment of the activities and specificities attained by a representative panel of high-specificity Cas9 nucleases and their corresponding RuvC-disabled Cas9^{D10A} variants, was conducted. Importantly, dual nicking RGNs based on specific Cas9^{D10A} variants were shown to outperform parental dual nicking RGNs and achieve selective cleavage of target sequences with high similarity to off-target sites. Following from these findings, **Chapter 4** further investigates the capability of orthogonal and high-specificity Cas9 proteins in directing gene targeting through homologous recombination (HR) and homology-mediated end joining (HMEJ), and explores the compatibility of the IPTN principle with orthogonal and high-specificity Cas9 nickases.

Prime editing is a recent precision genome editing modality that permits installing any single base-pair substitution and well-defined indels at specific genomic positions requiring to this effect neither DSBs nor donor DNA substrates. However, the large size of prime editing complexes poses substantial production and delivery issues. As discussed in **Chapter 1**, the HC-AdV platform presents a particularly valuable set of features that warrants its exploitation for genome editing purposes, namely (i) efficient transduction of cycling and quiescent cells; (ii) amenability to tropism modifications; (iii) high genetic stability; (iv) strict episomal nature; (v) absence of viral genes, and (vi) vast packaging capacity (i.e., up to 36 kb). In this regard, **Chapter 5** reports the feasibility of tailoring these biological nanoparticles for all-in-one transfer of full-length prime editing components into both transformed and non-transformed cell types. In addition, the positive influence of cellular replication on prime editing activity is disclosed by exploiting the HC-AdV cell cycle independency. Building on these findings, **Chapter 6** outlines the therapeutic potential of HC-AdV delivery of advanced prime editing machineries comprising optimized and multiplexing components. In these HC-AdV-enabled prime editing experiments, Duchenne muscular dystrophy (DMD) was targeted as a disease model. DMD (OMIM #310200) is an X-linked progressive muscle-wasting disorder (incidence: ~1:4700 boys) caused by loss-of-function mutations in the large *DMD* gene (~2.4 Mb) that normally codes for the striated muscle-stabilizing protein dystrophin (427 kDa). Of notice, in-frame *DMD* deletions result in a less acute form of muscular dystrophy, named Becker muscular dystrophy (BMD; OMIM #300376), owing to the formation of internally truncated, yet partially functional, dystrophin molecules. Hence, **Chapter 6** concerns investigations on the resetting of defective *DMD* reading frames in human myogenic cells by using HC-AdV delivery of the aforementioned optimized and multiplexing prime editing components. Finally, research described in **Chapter 6** establishes that combining straightforward HC-AdV transductions with seamless prime editing allows for stacking chromosomal edits in target cell populations through successive delivery rounds.

Similar to prime editors, base editors permit installing specific base pair changes in the genome requiring in the process neither DSB formation nor donor DNA delivery. The main base editing platforms consist of cytidine base editors (CBEs) and adenine base editors (ABEs) with CBEs and ABEs yielding C•G-to-T•A (C→T) and A•T-to-G•C (A→G) transitions, respectively. Importantly, although prime editors and base editors both constitute powerful high-potential tools for genetic therapies, their performance and precision at alternate chromatin states governing cell differentiation and identity, remain ill-defined. To address this knowledge gap, in **Chapter 7**, complementary loss-of-function and gain-of-function cellular systems are implemented to provide in-depth information concerning the efficiency and fidelity attained by using prime editors and base editors at euchromatin versus heterochromatin. The resulting findings inform and help guiding the development, selection and application of these powerful tools in specific cell types and contexts.

Taken together, the research presented in this thesis expands the current knowledge and toolbox underlying genome editing procedures through a comprehensive investigation of fast-developing genome editing systems and strategies in different cellular contexts. In particular, it reveals the feasibility and utility of using regular and high-specificity nicking RGNs for achieving efficient and accurate genetic modification of human cells involving targeted gene knockouts and HDR-mediated gene knock-ins. Moreover, it also establishes the suitability of the HC-AdV platform for the versatile investigation of advanced prime editing systems independently of their size and component numbers, which should facilitate the screening and application of the said systems in basic science and biotechnological settings. Finally, this thesis establishes causal relationships between specific chromatin states and the activities and fidelities attained by base editing and prime editing complexes in human cells, which has consequences for their further development and optimal deployment.