Chapter 3

Dynamics of indel profiles induced by various Cas9 delivery methods

3.1 Introduction

3.1.1 Cas9 mutagenesis of genes

CRISPR/Cas9 has made site-specific mutagenesis highly efficient. Therefore, experimental and therapeutic edits can be performed on populations of cells, even without integration of an exogenous selection cassette. In many cases, good results can be achieved without extensive optimization, especially when cells constitutively express Cas9 or when selection for the desired phenotype is possible. In such circumstances, optimization may not be necessary. However, it may become paramount when phenotypic selection is impossible, when Cas9 has to be delivered into the cells, when the phenotypic effect is small or when cell numbers are limiting (e.g. patient material).

Ideally, the efficiency of a Cas9 experiment should be measured on the level of phenotype. However, it may be expensive, time consuming, the necessary reagents may not be available, the gene product may not be expressed in the edited cells or the phenotypic effect may not be detectable within a reasonable timeframe (e.g. if cells need to be quickly reintroduced into the patient). In this case, the genotype may be a good proxy for the phenotype. This is particularly the case, if the phenotype-genotype relationship has been established in a pilot experiment.

3.1.2 Genotyping of small indels in bulk cell populations

Routine genotyping of bulk cell populations can be performed by methods such as Enzyme Mismatch Cleavage (EMC), Indel Detection by Amplicon Analysis (IDAA), Tracking of Indels by DEcomposition (TIDE) and Next-Generation Sequencing (NGS) (Brinkman et al., 2014; Lonowski et al., 2017; Yang et al., 2015; Yeung et al., 2005). Each of these method requires short-range PCR amplification of the genomic region of interest (<600 bp), followed by various methods of allele separation and detection.

In EMC, colloquially known as the T7 or Surveyor assay, the pool of PCR products is denatured and rehybridized, resulting in formation of homo- and heteroduplexes. The latter are selectively digested by a heteroduplex sensitive endonucleases (e.g. T7EI, CELI or Surveyor) and separated from wild-type sized homoduplexes by agarose gel electrophoresis. Quantification of band intensities determines a simple efficiency score. In IDAA, fluorescently labelled PCR products are resolved at basepair resolution and quantified using a fragment analyzer machine (similar to those used in Sanger sequencing). This yields an indel profile, the frequency of different-sized indels (Fig. 3.1). In TIDE, PCR products from wild-type and mutagenized cells are Sanger sequenced and the resulting traces are computationally deconvoluted into an indel profile. Assessment of templated editing is also possible using a related TIDER procedure. In the NGS approach, the PCR products are sequenced using the Illumina platform. Both

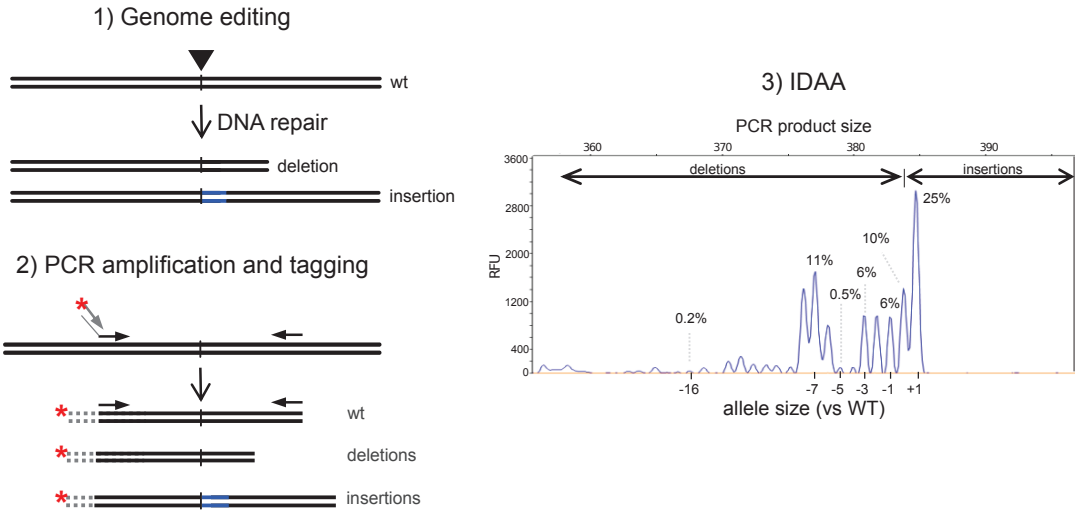


Figure 3.1: IDAA workflow. Cells are mutagenized in a pool, resulting in small deletions or insertions. The region of interest is amplified using a triprimer PCR reaction. One of the two genomic primers contains an overhang that allows annealing and amplification by a universal fluorescently labelled third primer. Fluorescent PCR products are separated at single basepair resolution and quantified using a fragment analysis machine (a capillary DNA sequencer). PCR product size (top x-axis) can be expressed as the allele size versus wild-type (wt, bottom x-axis) and the amount of given product/allele corresponds to the intensity of the peak expressed in relative fluorescent units (RFU, y-axis). Modified from Lonowski et al., 2017.

indel profiles and editing can be investigated using computational approaches (CRISPR Genome Analyzer, CRISPResso, CRISP-R).

EMC is the "quick and dirty" method. It takes less than a day and offers only a crude quantitative measure of efficiency, as it cannot distinguish between in-frame and out-of-frame mutations (both of which can form heteroduplexes that are cleaved by the endonuclease). Furthermore, endonucleases used in the assay are not sensitive to single base changes and heterozygosity at the wild-type locus may produce spurious signal due to formation of "natural" heteroduplexes in unedited samples. IDAA and TIDE resolve alleles by size at single basepair resolution with sensitivity similar to NGS (~0.1%, assuming the usual read quality cut-off). This makes them tools of choice for estimation of out-of-frame mutations, a proxy for functional knockout. When investigating many pools at once, NGS becomes an economically viable option, as many pools can be sequenced in one run for the same price. As the only method to produce sequence level data, it can resolve indels of same size, but different basepair composition. This property also allows it to reliably quantify templated editing. However, NGS is considerably more time-consuming than other methods to execute and to analyze.

3.1.3 Factors influencing Cas9 mutagenesis and genotyping

Genotyping gives an estimate of the overall efficiency of generating mutant alleles. This can be influenced by generic factors, such as transfection efficiency (proportion of cells receiving the reagents) and concentration of reagents within the cell over time. The latter is influenced by transfection multiplicity, activity of the gRNA and Cas9 promoters, whether expression constructs are stably integrated or transiently transfected, etc. If necessary, most of these factors can be optimized in a given cell system. Furthermore, the efficiency varies between different gRNAs, independently of the generic factors. The underlying reason is likely a combination of high and accurate expression of a particular gRNA sequence from a PolIII promoter (if used), the binding and cutting activity

of Cas9 at a given locus as well as how often the given locus reverts to wild-type upon repair.

How mutant genotype translates into gene knock-out will depend on both gene and gRNA-specific factors. Knock-out of non-coding genes is usually achieved by introducing a large deletion with paired gRNAs, as such genes are robust to small indels introduced by single gRNAs. Conversely, a small out-of-frame mutations in the first few exons of a protein-coding gene is often enough to abolish protein expression through nonsense-mediated decay (NMD) of its RNA transcript. However, alternative splicing and transcription start sites may lead to production of functional protein despite such mutations. Activity of many protein domains is also sensitive to in-frame mutations. Composition of mutant genotypes (the "indel profile") is specific to each gRNA (Chakrabarti et al., 2018; Taheri-Ghahfarokhi et al., 2018; van Overbeek et al., 2016). In consequence, two guides targeting the same protein domain with the same overall level of DNA mutation may differ substantially in the proportion of out-of-frame mutations and thus in the level of protein knock-out.

Little is known about the factors influencing the indel profile of a specific gRNA. Microhomologies around the cut site are speculated to contribute to it. Transcription has been reported to increase mutagenic efficiency of Cas9 in a strand-dependent manner, which implies it may also influence the indel profile. However, no general rules have yet been defined. The indel profile is therefore usually established empirically for each gRNA.

Knowing how different factors influence the "indel profile" may help predict the level of phenotypic knock-out. I set out to study two such previously unexplored factors - time of genotyping and Cas9/gRNA delivery method. If genotyping is performed before the indel profile becomes stable, the efficiency may be assessed incorrectly. On the other hand, delaying genotyping may be inconvenient, especially when cells need to be reinjected into the model organism or patient as soon as possible after the procedure. Similarly, if different delivery methods result in similar outcomes, they could be interchanged at convenience. If not, the differences could be exploited to achieve higher phenotypic knock-out. Finally, different Cas9/gRNA delivery methods may have different dynamics of mutagenesis. It may be beneficial to know whether genotyping time needs to be adjusted depending on the method.

3.2 Results

The Cas9 protein and guide RNA may be delivered into cells in a variety of forms (e.g. plasmid DNA, mRNA, protein) and using a variety of methods (e.g. electroporation, lipofection, transduction). Methods used in this study are summarized in Fig. 3.2. I collected cells at multiple timepoints post-delivery and analyzed the indels using IDAA. I chose IDAA, because it offers rapid results and high resolution at a low cost. I performed the experiment in a commonly used HEK293 cell line, known to be amenable to many delivery methods and previously shown to achieve high Cas9-induced mutagenesis rates. A validated, highly efficient gRNA against the *ST6GALNAC1* gene (Hansen and O'Shea, 2015), which is silent in HEK cells, was picked in hope of avoiding knockout-specific proliferative effects. To minimize differences between methods I utilized a plasmid backbone containing both lentiviral and PiggyBac functional elements (Metzakopian et al., 2017).

3.2.1 Transfection and integration dynamics

Methods compared in this study result in either stable (lentivirus, PiggyBac) or transient (RNP, transient plasmid, P&P - protein & plasmid) expression of Cas9 and gRNA. As neither component of the RNP was fluorescent, I did not monitor the transfection efficiency of this method. As the Cas9 plasmid contains no fluorescent marker, I monitored the BFP expression from BFP-gRNA cassette as a proxy for the overall transfection and integration efficiency (Fig. 3.3a). Cas9 was not selected for and therefore the percentage of

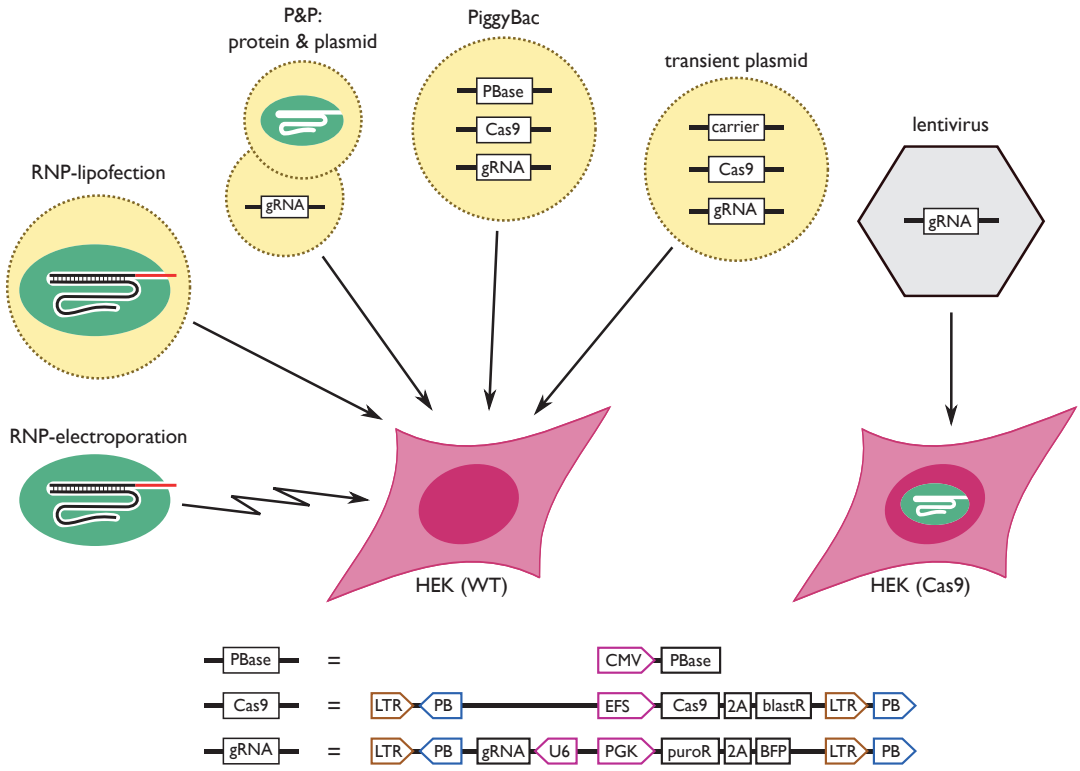


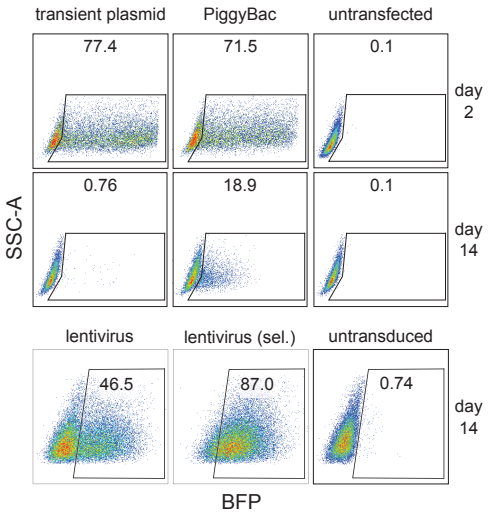
Figure 3.2: Cas9 and gRNA delivery methods. The same gRNA and Cas9 plasmids were used for all experiments. Yellow circle indicates transfection with Lipofectamine 3000. RNP-electro: electroporation of Cas9 protein and synthetic two-part gRNA (crRNA+trRNA) using Neon Transfection System. RNP-lipo: as RNP-electro, using lipofectamine. P&P - protein & plasmid: transfection of Cas9 protein and gRNA-encoding plasmid. A version with addition of carrier plasmid (pBluescript II SK+) was also used (P&P-carrier). PiggyBac: transfection of plasmids encoding Cas9 and gRNA together with PiggyBac transposase (resulting in stable cellular integration). transient plasmid: as PiggyBac, with transposase replaced by the carrier plasmid (pBluescript II SK+). lentivirus: transduction of HEK cells stably expressing Cas9 with gRNA lentivirus. Vector schematics in the bottom part of the figure, LTR = Long Terminal Repeats, lentiviral elements; PB = PiggyBac repeats; blastR and puroR = blasticidin and puromycin resistance cassettes; LTRs, PBs and promoters are marked with colors.

BFP positive cells is an overestimation of overall number of cells containing both Cas9 and gRNA-expressing constructs.

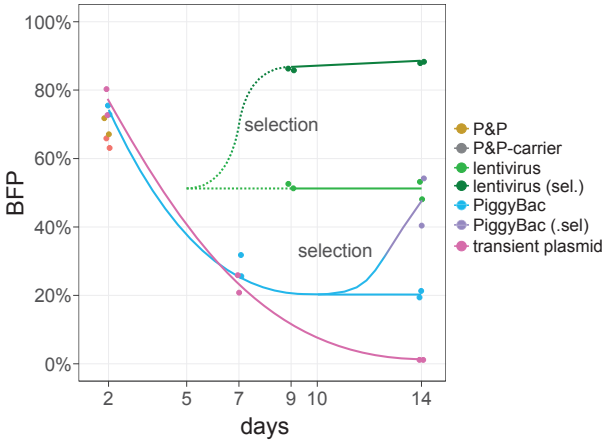
PiggyBac and transient methods resulted in the highest transfection efficiencies on day 2 (>70%), followed by P&P-carrier and P&P (55% and 40%; Fig. 3.3b). As expected, the BFP expression was all but extinguished in the transient plasmid, P&P and P&P-carrier conditions by day 14. The few remaining BFP positive cells may indicate rare cases of stable integration of the plasmids.

In the PiggyBac condition, 20% of the cells (about 1/4 of all transfected cells on day 2) remained BFP positive on day 14, indicating stable transposition of the PiggyBac transposon from the donor DNA into the genome. Short-term selection using puromycin (days 10–14) for the genomically integrated gRNA-BFP construct increased the percentage of positive cells to 44%.

Data from early timepoints in the lentiviral transductions was not collected, but the percentage of BFP positive cells was maintained at approximately 50% between days 9 and 14 post-transduction, indicating stable integration. Selection for the integrated gRNA-BFP pro-virus between days 5 and 14 increased this proportion to 87%. As puromycin normally kills all wild-



(a) Examples of flow cytometry plots. Samples were assessed using Cytoflex machine, except for the lentiviral method, which was assessed using BD Fortessa.



(b) Comparison of transfection levels across methods and timepoints. *Solid line* indicates the most likely path between collected timepoints. *Dotted line* connects day 7, the time when selection of lentiviral condition was started (no data available, efficiency inferred) to day 9 for selected and unselected samples.

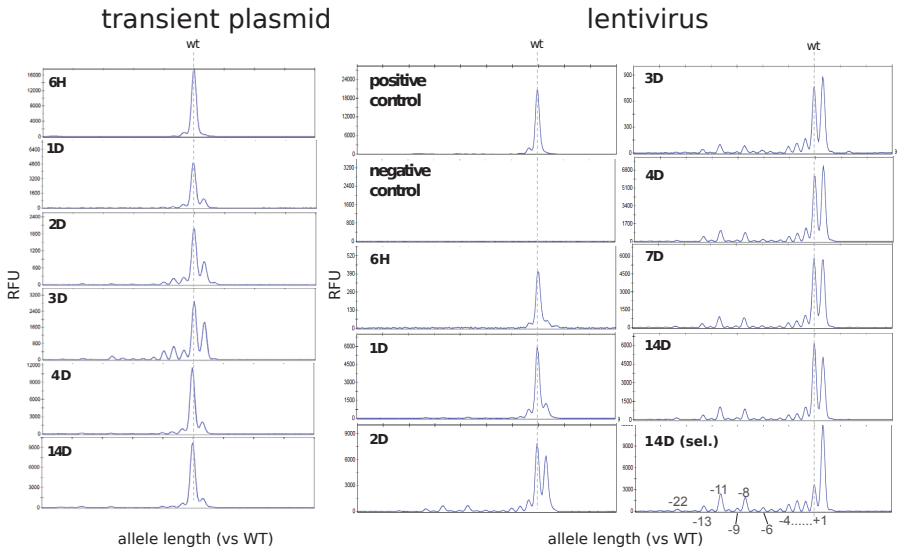
Figure 3.3: Transfection efficiency over time. HEK cells transfected with Cas9 and gRNA were analyzed by flow cytometry at various time points. BFP fluorescence comes from gRNA expressing plasmids and thus likely overestimates the actual percentage of Cas9/gRNA double-positive cells. The exception is lentiviral transduction, where cells constitutively express Cas9. As RNP has no fluorescent component, the transfection efficiency of this methods is unknown. sel.: selected for gRNA construct using puromycin.

type HEK cells (data not shown), I speculate that remaining 13% of cells were resistant, but expressed no or little BFP. Since BFP follows the puromycin resistance gene in the expression cassette (Fig. 3.2), it is possible for a knock-out mutation to occur within BFP without affecting the puromycin resistance (personal communication, Konstantinos Tzelepis).

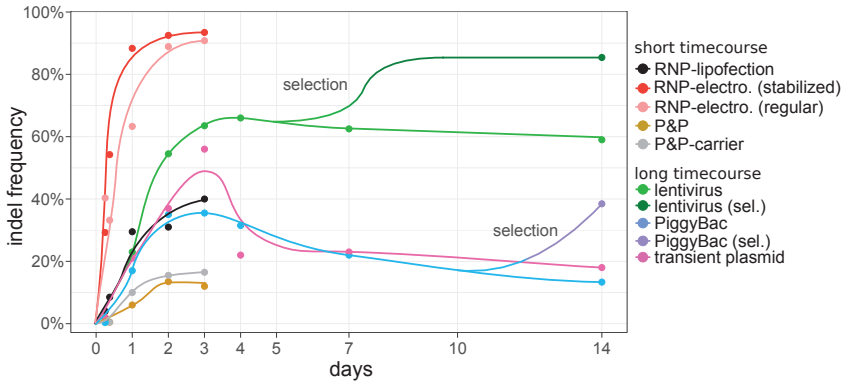
3.2.2 Mutagenesis efficiency over time

I studied the dynamics of indel generation using IDAA. For RNP and P&P methods I only collected samples up to day 3 post-delivery, on the assumption that Cas9 protein is degraded by that time. For other methods, I continued collecting samples until day 14 (examples of IDAA indel profiles in Fig. 3.4a). The mutagenesis efficiency was calculated as a ratio of intensity of the prominent, non wild-type peaks to all the peaks (Fig. 3.4a). As the wild-type size peak may represent rare SNPs or balanced indels in addition to the wild-type allele, this efficiency may be slightly underestimated.

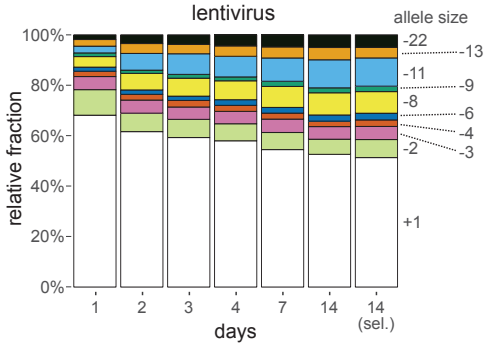
On day 3 post-delivery, RNP-electroporation was the most efficient method (91–93%), followed by lentivirus (64%), transient plasmid (56%), RNP-lipofection, PiggyBac (36%–40%) and P&P/P&P-carrier (12–17%; Fig. 3.4b). The transfection efficiency broadly correlated with the mutagenic efficiency, except for P&P and P&P-carrier, which induced relatively lower levels of indels. For transient plasmid transfection, no smooth curve fit could be found for experimental data, indicating some timepoints were outliers. However, selection for stable PiggyBac integrants between days 10 and 14 using puromycin more than doubled the final percentage of mutagenized alleles (from 17% to 38%). Similarly, selection for provirus integration from day 5 to day 14 increased that percentage from 59% to 86%. Despite stable Cas9 expression and long-term selection for the integration of the gRNA construct 100% mutagenic efficiency was not achieved. This may reflect the



(a) Examples of IDAA plots for lentivirus and transient plasmid conditions. Prominent peaks used for efficiency calculation are indicated in the selected sample. Asterisk indicates truncated product (see Discussion). Representative panels from duplicate experiments are shown. sel.: selected for gRNA construct using puromycin.



(b) Indel frequency over time. Solid line indicates most likely path between collected timepoints. Samples involving transfection of Cas9 protein were only collected up to day 3. N = 1-2.



(c) Indel profiles over time in the lentiviral sample. Prominent peaks, as indicated in Fig. 3.4a, are colored. N = 1-2.

Figure 3.4: Indel frequency and profiles overtime. HEK cells transfected with Cas9 and gRNA were analyzed by IDAA at various timepoints.

inherent limitation of the method, such as inability to distinguish substitutions and "balanced" indels from wild-type and occasional misclassification of the 1 bp insertion as wild-type due to Taq polymerase action (see Discussion). Furthermore, some gRNA-expressing cassettes may have been silenced independent of the puro-BFP cassette before mutagenesis could occur.

Both RNP-electroporation and RNP-lipofection show earliest indels formation at 6h and 9h post-transfection. Most of the methods reached their maximum efficiency on day 2 and 3. Interestingly, transient plasmid and PiggyBac exhibited a significant decrease in efficiency after day 3, while lentivirus plateaued.

To reduce technical variation, I repeated the experiment on a single day using transient plasmid, RNP-lipofection, PiggyBac, P&P and P&P-carrier methods (data not shown). On day 3 post-delivery, I analyzed the cells by flow cytometry and TIDE. All methods using protein Cas9 (RNP, P&P, P&P-carrier) achieved the same mutagenic efficiency of around 12%. Despite P&P-carrier transfection efficiency being slightly higher than with transient plasmid method (82% vs 76%), the mutagenic efficiency was much lower (12% vs 27%). These results indicate that Cas9 protein was equally likely to co-transfect with synthetic gRNA as with gRNA plasmid and that Cas9 protein was the limiting factor for mutagenic efficiency in the experiment.

Both transfection and mutagenic efficiency were higher in transient plasmid than PiggyBac condition on day 3, consistent with a similar difference on day 14 in the experiment presented in Fig. 3.4b. Therefore, higher transfection efficiency likely explains the higher mutagenic efficiency observed at day 14. It is unclear why there was a difference in transfection efficiency in the first place. It may be that the carrier plasmid in transient method increases the plasmid entry rate compared to the transposase plasmid, which it replaces, possibly due to its smaller size and hence higher copy number. Alternatively, the transposase may have effectively reduced the concentration of the gRNA expressed from the transposon, either by direct transcriptional interference or by exposing the transposon plasmid to degradation.

3.2.3 Indel profiles over time

To investigate whether the allelic composition stays stable or fluctuates over time and across methods, I quantified the relative abundances of IDAA indel peaks. As larger deletion indels (e.g. 22, 13, 11 and 8 bp) may correspond to microhomologies that I found around the target site (data not shown), I wondered if they would appear later than the smaller ones. This would be consistent with the observation that NHEJ which usually creates small indels is a faster repair pathways than MMEJ (Mladenov and Iliakis, 2011). In lentiviral condition, which yielded most reliable data, larger indels appear later in the timecourse (Fig. 3.4c). This suggests either some large indels take longer to form or that small indels (such as 1 bp insertion) and substitutions are susceptible to recutting, yielding larger indels upon mutagenic repair.

3.2.4 Dynamic effect of gRNA stability on indel efficiency

It has been proposed that gRNAs protected from cellular exonucleases by chemical modifications may increase the overall indel generation (Hendel et al., 2015). Therefore, I studied the dynamics of indel generation using IDAA in cells electroporated with Cas9 protein and either regular or stabilized gRNAs. Both gRNAs showed similar, rapid indel formation by 6h post-electroporation and demonstrate comparable indel profile and maximum efficiency at the end of the experiment on day 3 (Figs. 3.5). However, the stabilized gRNA reached its maximum efficiency on day 1, faster than the regular gRNA by about 24h. It is known, that gRNA loading is a key regulator of Cas9 enzyme function (Jiang et al., 2016) and I speculate that stabilized gRNA may improve its binding affinity and/or nuclease activity.

3.2.5 Comparison of IDAA and TIDE methods

As both IDAA and TIDE generate indel profiles, I decided to compare the results for selected samples. While the overall profiles were similar, the sensitivity of TIDE differed depending on which primer was used for the Sanger sequencing reaction (Fig. 3.6). As IDAA does not involve a sequencing step, it is not susceptible to this problem.

3.3 Discussion

I have investigated the time dynamics of indel profile generation by different methods of Cas9 and gRNA delivery. I showed that the mutagenesis and the initial transfection efficiency were roughly correlated, except when low protein Cas9 transfection was the likely limiting factor. Explicit measurement of the efficiency of Cas9 delivery, e.g. by using a Cas9-GFP fusion, would help solidify this finding. I have confirmed RNP delivery is characterized by rapid indel induction and found an indication that stabilized gRNAs may further speed up the process, which could potentially be useful in therapeutic setting.

3.3.1 Causes and consequences of mutagenesis efficiency fluctuation

Unexpectedly, when using plasmid transfection methods (but not lentivirus), the proportion of mutagenized cells did not plateau at the maximum level (usually around day 3 or 4), but peaked temporarily around day 3-4 and then decreased. One explanation for this effect may be that alleles, which end up being perfectly repaired by the slower process of homologous recombination remain undetectable for a longer time than the ones repaired mutagenically. Therefore, I observe a temporary enrichment for mutagenic genotypes. Discrepancy between lentivirus and plasmid methods may be explained by higher frequency of stable lentiviral integration. Most cells that are transduced continue expressing the gRNA from an integrated lentiviral cassette and are eventually mutagenized, leading to a plateau. In contrast, many cells in transient plasmid and PiggyBac conditions express the gRNA temporarily from the unintegrated plasmid, lose the construct due to cell division or plasmid degradation, repair the break perfectly and remain wild-type.

Another potential explanation is that cells mutagenized using the plasmid grow slower and end up being outgrown by wild-type cells. This could be due to toxicity associated with the DNA (lentivirus typically infects at low multiplicity, but many copies are delivered by transfection) or with the carrier (lipofectamine). One way to clarify this issue would be to monitor frequency of alleles "in repair" by qPCR and to compare growth rates. Comparing mutagenic efficiency of RNP-lipofectamine and RNP-electroporation beyond day 3 would indicate whether lipofectamine itself causes the effect.

As a consequence of this dynamic, correct estimation of the phenotypic effect from a snapshot genotype becomes difficult. In the "plateau" model, premature genotyping leads to underestimation of the effect, but the window for correct genotyping after reaching the plateau is large. In the "peak" model, premature genotyping may paradoxically better mirror the ultimate phenotypic effect, if only by chance. The decrease, at least in the PiggyBac condition, seems to take a long time and is quite substantial (from around 35% on day 3 to around 15% on day 14). One potential solution to this problem is to genotype at the peak and include the expected decline in phenotypic calculations.

3.3.2 Delayed large indel formation

The indel profile in lentiviral condition showed that larger indels form later (Fig. 3.4c). A recent paper found a similar effect and attributed it to the slower repair kinetics of MMEJ (Brinkman et al., 2018). While this is almost certainly the case early on in the timecourse, completion of MMEJ does not generally take many days, as observed in here (biochemical studies suggest $t_{1/2}$ of 2-20h, Iliakis, 2009; Perrault et al., 2004). Therefore, late formation of larger indels remains unexplained. Recut-

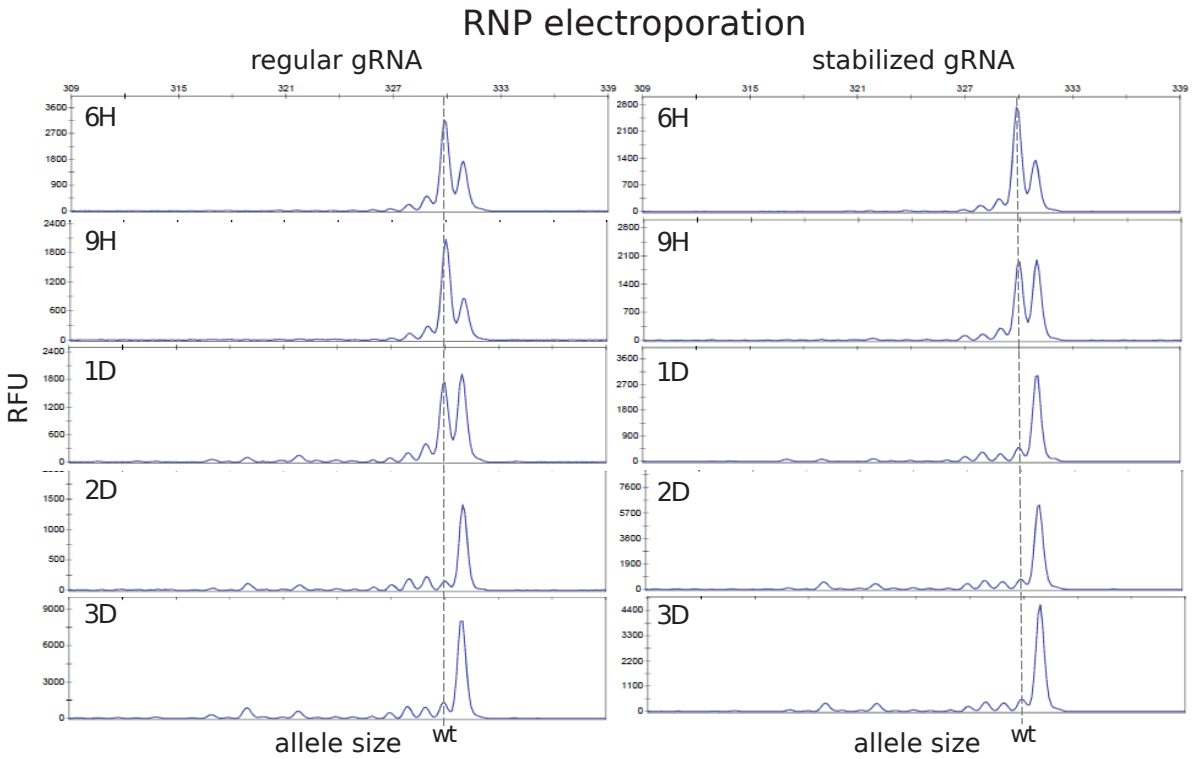


Figure 3.5: Comparison of normal and stabilized synthetic gRNAs. HEK cells were electroporated using Neon transfection system with protein Cas9 and regular or stabilized synthetic gRNAs. Representative panels from duplicate experiments are shown.

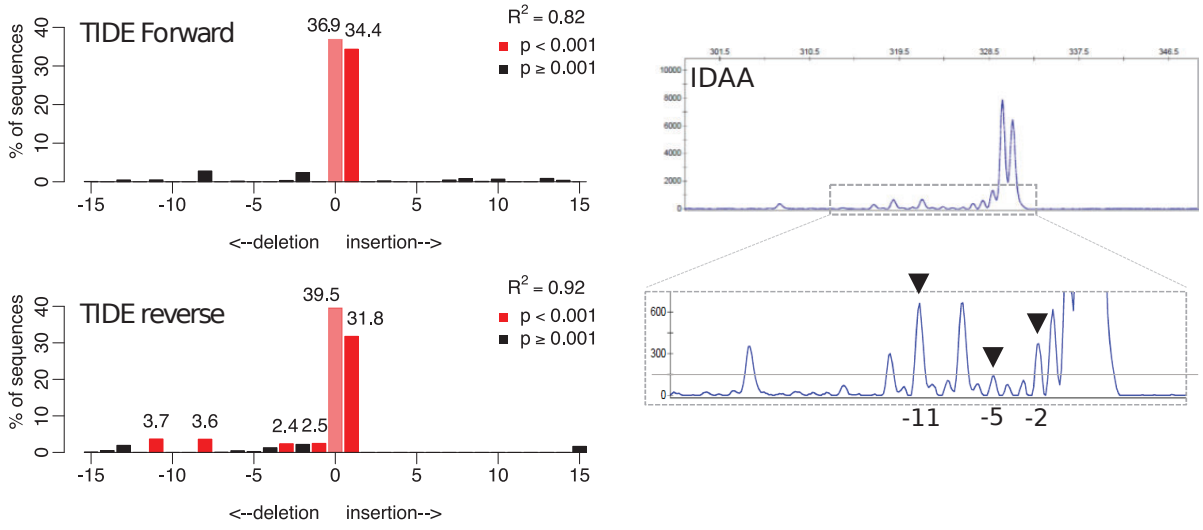


Figure 3.6: Comparison of IDAA and TIDE. The sample shown in Fig. 3.4a, lentivirus 3D, was indel profiled by TIDE analysis using both the forward or reverse Sanger sequencing raw data files. The results were compared to IDAA generated profiles, where a zoom in of the indel peaks is provided as inset for clarity. Recommended detection threshold (150RFU) is indicated by the horizontal line. Selected indels are indicated with black triangles. N = 1.

ting of small indels and substitutions may be the cause. If large indels continue to accumulate after two weeks, it would validate this hypothesis. Conversely, lack of additional accumulation would not falsify this hypothesis as recutting and mutagenic repair of such indels may be slow enough to reach an equilibrium with production of new wild-type alleles by DNA replication. Therefore, investigation of mutagenic dynamic using gRNAs with single mismatches could provide a more direct answer.

3.3.3 Caveats of indel profiling

Genotyping of the other delivery methods did not yield data of sufficient quality to perform a time-course analysis or to compare indel profiles across methods. Two issues caused this. First, the low efficiency of some methods, especially at the early timepoints, made signal detection and quantification difficult. Excess material had to be loaded to detect weak signal from indel alleles, which led to a strong wild-type signal overwhelming adjacent –1 and +1 peaks. Furthermore, I assumed co-linearity between peak intensity and abundance of an allele. However, for low intensity peaks this assumption is likely incorrect. Additional experiments to establish a standard curve and quantify the magnitude of the signal saturation could rectify these issues. The second issue was the presence of a spurious –1 peak. Since it was detected in all control samples (with an intensity of between 5-15% of wild-type peak), I assume it has been created by the use of Taq polymerase, which adds a single 3' adenine to all PCR products with less than 100% efficiency. This precluded accurate quantification of the –1 indel and likely affected quantification of other peaks as well. Usage of high-fidelity non-Taq polymerase would likely remedy this issue.

Finally, this study suffers from lack of replication. Only a single guide was used and experiments were repeated at most twice. More replication would increase confidence in the presented results.