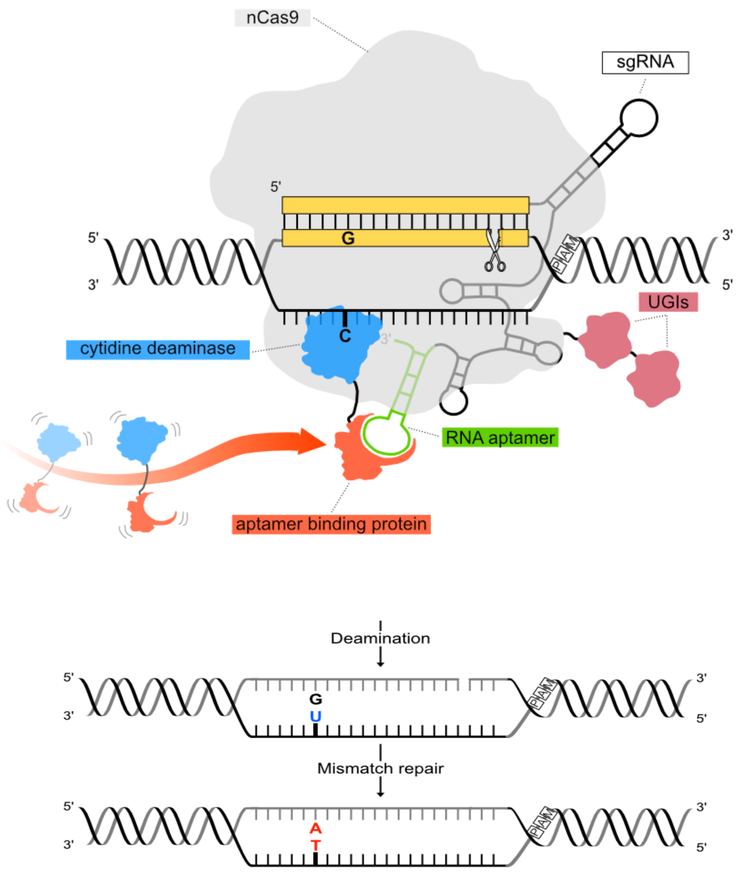
Development and validation of customized guide RNA design and efficiency prediction tools for the Pin-point™ base editing platform.

Paul Russell, Robert Blassberg, Alexis Duringer, John Chambers, Giabao Tonthat, Glynn Martin, Immacolata Porreca, Pablo Perez-Duran.

1 Abstract

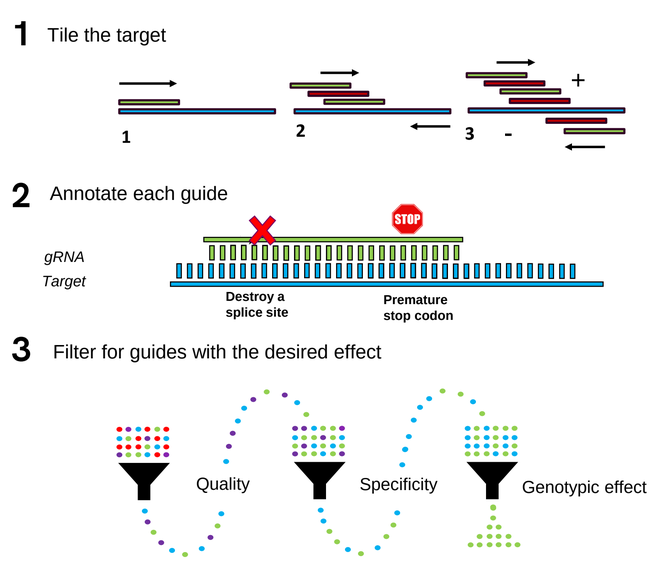
Base editors are a class of promising next generation genome editing technologies with the potential to both precisely correct disease-causing genetic variants and to safely knockout multiple gene targets simultaneously. In one configuration, the Pin-point base editing platform is a modular assembly of DNA binding Cas and DNA modifying deaminase components associated via an aptamer encoded in the sequence-targeting guide RNA (gRNA). A major challenge in the application of base editors in general is accurately in silico predicting efficiency and specificity of editing at target sequences for a given combination of Cas and deaminase. The modularity of the Pin-point base editing system allows the creation of a large number of configurations, that can vary in their PAM specificity, sequence editing preference and editing efficiency. To facilitate and accelerate the development of applications based on the Pin-point platform, we created a custom tool to design gRNAs to target the gene of interest and to install base conversions, including those that would either introduce premature STOP codons or destroy splice sites to knockout the target gene. In addition, we performed a massive parallel cell-based screen to analyse the editing activity of two different Pin-point base editor configurations with gRNAs targeting thousands of target sequences. We used the data obtained from the screen to construct models of the observed editing outcomes for each configuration. We applied these models to rank gRNAs designed to generate functional knockout of multiple clinically-relevant gene targets, including CIITA and PCSK9. After analysing the correlation of the in silico prediction with the cell-based performance of the gRNAs, we confirmed that the model predictions correlate with the observed editing efficiency for the Pin-point base editing platform. The combination of the custom gRNA design tool and the predictive model led to the identification of a novel, highly efficient gRNA able to knockout PCSK9 by disrupting a splice site, and we confirmed the predicted performance of other gRNA designs previously reported in the literature. Our gRNA design rules were informed using our broad cell-based performance dataset, creating reliable custom tools to prioritize gRNAs and select those with high editing efficiency.

2 The Pin-point base editing platform



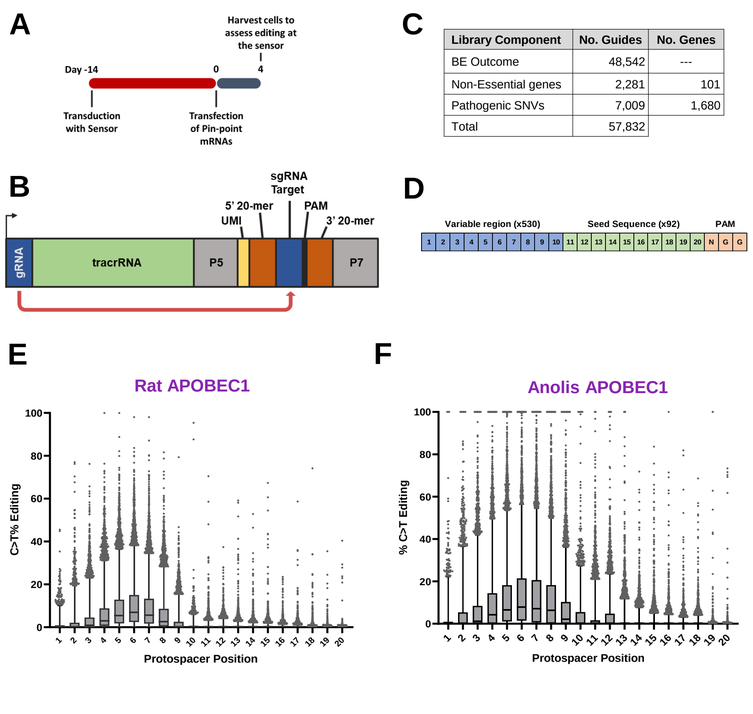
The Pin-point base editing platform allows precise genome modification by single nucleotide conversion. The platform is agnostic to the sequence targeting Cas, DNA modifying deaminase, and deaminase recruiting aptamer modules, and therefore provides the flexibility to assemble the most appropriate combination of elements for a given application.
In one possible configuration of the platform, a Cas9 nickase (nCas9) is guided to the DNA target site via a gRNA with an aptameric region engineered into the scaffold. The aptamer recruits a deaminase via fusion with an aptamer binding protein. When a cytidine deaminase is recruited, conversion of a cytidine to thymidine in the target sequence is achieved. The combination of nCas9, an aptamer binding protein fused to a deaminase, and an aptameric gRNA efficiently base edit a DNA target of interest.

3 Pin-point custom gRNA design tool



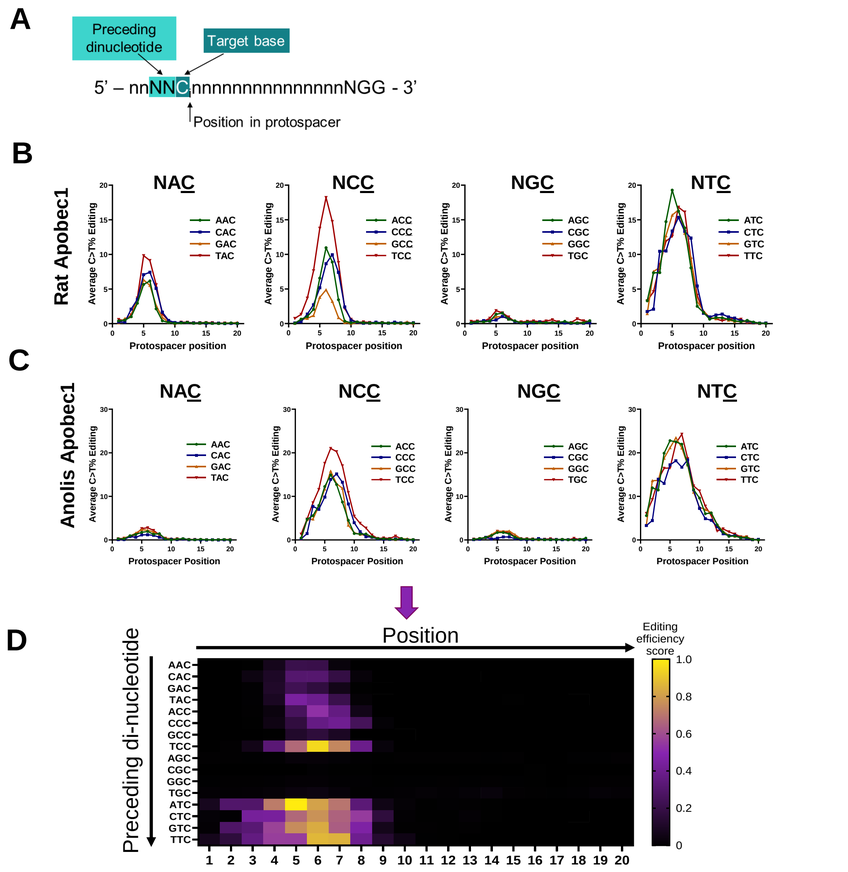
1)The design tool tiles the target by walking its sequence to produce a list of every possible 20-mer (or alternative length) subsequence. Subsequences that are missing a PAM or would produce an invalid guide are removed (red). This process is performed for both genomic strands.
2) Guides are annotated and labeled using genomic records from NCBI and Ensembl. This includes noting the genotypic effects each guide may produce against its target such as introducing a premature stop codon or disrupting a splice acceptor or donor site or introducing an aminoacidic change. Additional annotations included at this step are specificity score based on human genome scan, melting temperature and the presence of poly-nucleotide repeats.
3) Final guide candidates can be produced from the original list of tiled guides by selecting for genotypic effect, specificity, and quality. By starting with a comprehensive tiling approach, extensively annotating each guide, and removing guides based on multitude of factors, we ensure that the best guides are produced for any given target(s).

4 Editing profiles of base editor configurations



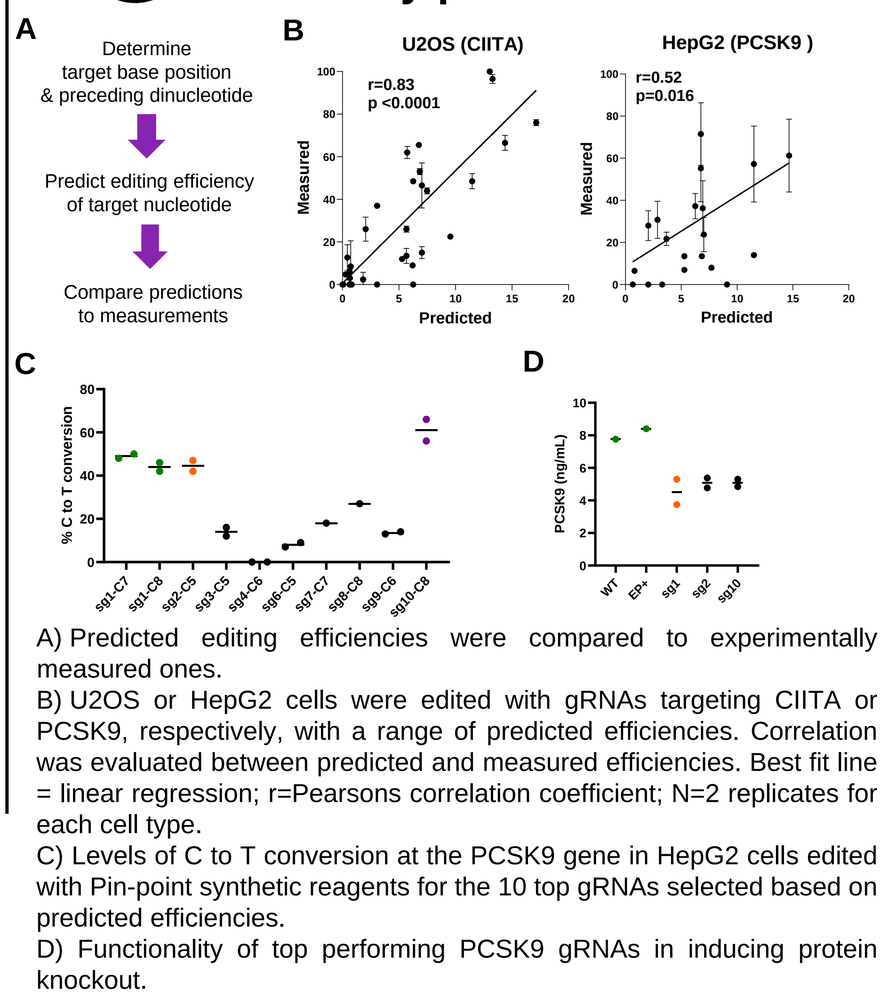
A) Workflow of a pooled screen approach to dissect the base editing behaviour of the Pin-point base editing platform. The screen was performed in HEK293 and in U2OS cells.
B) Design of the sensor molecule, comprising of a gRNA and the corresponding target sequence. The molecule also includes the Illumina P5 & P7 sequences to facilitate NGS sequencing post screen.
C) Description of the gRNA library, comprising ~58K gRNAs. The BE Outcome consists of synthetic sequences that captures all the possible cytosine positions (up to 4 C per gRNA) within the 2-9 nt editing window (shown in D). The library also includes gRNAs that would introduce nonsense mutations into non-essential genes and ~7K gRNAs against known pathogenic SNVs.
E-F) C to T editing percentages per position for all gRNAs of the library define the editing window of two configurations of the Pin-point base editing platform. E) SpCas9 nickase and rat APOBEC1 deaminase. F) SpCas9 nickase and anolis APOBEC1 deaminase.

5 Pin-point cytidine base editor predictive model



A) Dinucleotide context is defined by the identity of the target base, and the identity of the two bases immediately preceding the target.
B-C) Analysis of editing efficiency across different sequence contexts for rat (B) and anolis (C) APOBEC1 highlight the differential impact of the target base position and preceding dinucleotide. Average editing profiles for each of 16 preceding dinucleotides.
D) Lookup table of predicted editing efficiency as a function of the position in the protospacer and the identity of the preceding dinucleotide.

6 Validating gRNA editing efficiency predictions



A) Predicted editing efficiencies were compared to experimentally measured ones.
B) U2OS or HepG2 cells were edited with gRNAs targeting CIITA or PCSK9, respectively, with a range of predicted efficiencies. Correlation was evaluated between predicted and measured efficiencies. Best fit line = linear regression; r=Pearsons correlation coefficient; N=2 replicates for each cell type.
C) Levels of C to T conversion at the PCSK9 gene in HepG2 cells edited with Pin-point synthetic reagents for the 10 top gRNAs selected based on predicted efficiencies.
D) Functionality of top performing PCSK9 gRNAs in inducing protein knockout.

7 Summary

- A custom tool streamlines the design of gRNAs for experimental validation for the Pin-point platform.
- A massive parallel gRNA screen defines the editing behaviour of specific configurations of the Pin-point platform.
- Editing efficiency is a function of position of the target base in the protospacer and preceding dinucleotide identity.
- Predicted editing efficiencies allow to identify the most functional gRNAs.

8 References

Collantes et al. The CRISPR Journal, DOI:10.1089/crispr.2020.003
Porreca et al. Mol Ther, DOI: 10.1016/j.ymthe.2024.06.033

Revvity, 77 4th Avenue, Waltham, MA USA www.revvity.com