1600

High fidelity C-to-T editing with TALE base editors

Louisa Mayer[1], Maria Feola[1], Diane Tkach[1], Aminah Lawal[1], Sylvain Pulicani[2], Robert Hong[1], J Lomibao[1], Alex Boyne[1], Aymeric Duclert[2], J Valton[2], Philippe Duchateau[2], Alexandre Juillerat[1]

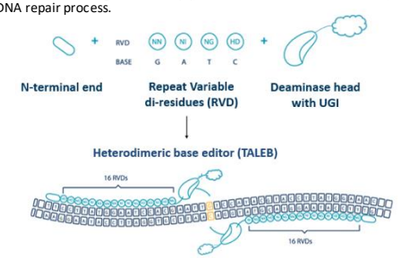
[1] Cellectis Inc, New York, NY; [2] Cellectis SA, Paris, France

Abstract

TALE base editors (TALEB) are fusions of a transcription activator-like effector domain (TALE), split-DddA deaminase halves, and a uracil glycosylase inhibitor (UGI). The C-to-T class of TALEB edits double strand DNA by converting a cytosine (C) to a thymine (T) via the formation of a uracil intermediate. We recently developed and applied a strategy that allowed the comprehensive characterization of C-to-T conversion efficiencies within the target editing window. This method also takes advantage of a highly precise and efficient TALEN®-mediated knock-in of ssODN in primary T cells to assess how the composition and spacer variations of target sequences affect TALEB activity/efficiency. Additionally, we highlighted that the composition of bases surrounding the target bases (TC) may strongly influence the editing efficiencies. We also demonstrated that different TALEB scaffolds could be used to relax target sequence limitations (to increase the targetable sequence) or be used to decrease or eliminate bystander editing within the editing window (to increase specificity). Overall, these findings allow for the fine-tuning of TALEB for a desired gene editing outcome. We then applied a range of different techniques to assess characteristics of nuclear genome editing. First, we focused on on-target editing and then explored the possibility and risk associated with genome-wide TALE-dependent/independent binding and editing. By using an experimental model relevant for therapeutic application with TALEB mRNA vectorization in primary cells (e.g. PBMCs and HSCs), we demonstrated that targeted binding of a single TALEB arm does not lead to detectable editing (detection limit: 0.1-0.2%). Finally, we further applied hybrid capture assays to test for off-target editing, in particular, within regions of the genome that were previously highlighted in cell lines. We demonstrated, in our relevant experimental setup with primary T cells, that editing at these sites was not detectable (detection limit: 0.2%). Altogether, the results of this study have enhanced our control and use of TALEB and allow for the design of extremely efficient (high editing frequencies and edit purity) and specific (absence of TALE-independent off sites and very limited, if not absent, DSB generation) TALEB, compatible with the development of future therapeutic applications.

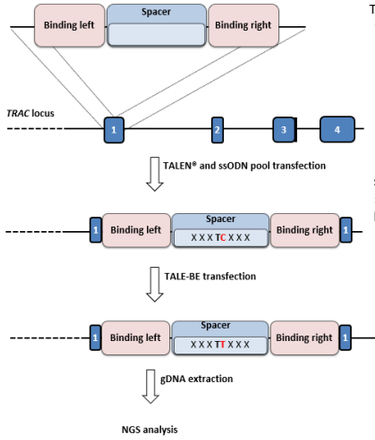
#1 Background

Base editing is a technology that leads to the introduction of point mutations in defined loci of a targeted DNA sequence. Base editors create mutations by deamination of the targeted bases (C), which are then converted into T, during the DNA repair process.



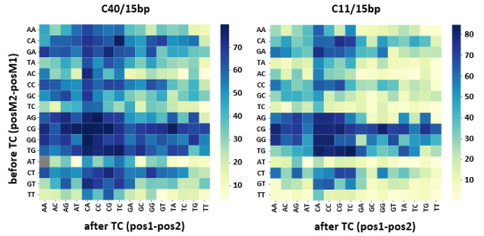
Schematic representation of a TALEB.

#2 Development of a high throughput screening platform for TALEB



To allow medium-to-high throughput screening of TALEB *in cellulo* in a defined genetic environment, we developed a cell-based assay to test the impact of various parameters such as (i) the architecture, (ii) the spacer length (sequence separating the two TALE binding site) and (iii) the sequence composition surrounding the targeted TC, on TALEB editing efficiency.

#3 Defining TALEB architecture and the impact of TC context on editing efficiencies

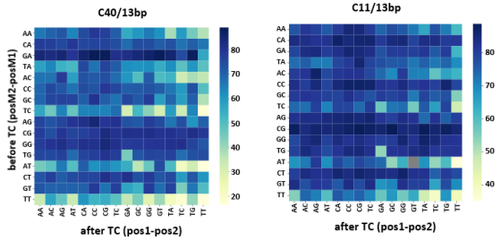


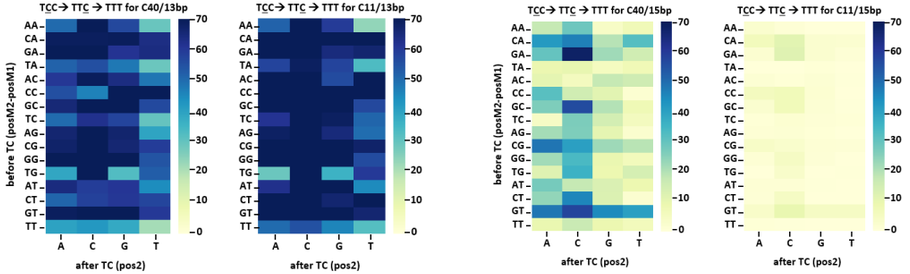
In a **15bp** spacer collection, the C11 architecture editing is more stringent than a C40 architecture against a suboptimal context. The C40 and C11 architectures show the following nucleotide preferences:

posM2: A = T<< G < C
posM1: T = C < A << G
pos1: T< G < A << C
pos2: T< C < G < A

For a **13bp** spacer, the C40 vs C11 architectures seem to have less importance for editing efficiency, which shows higher overall editing. The 13 bp spacer collection shows a similar preference to the 15bp:

posM2: T < A < G < C
posM1: T < C < A < G
pos1: T< G < A < C
pos2: T< A = C < G



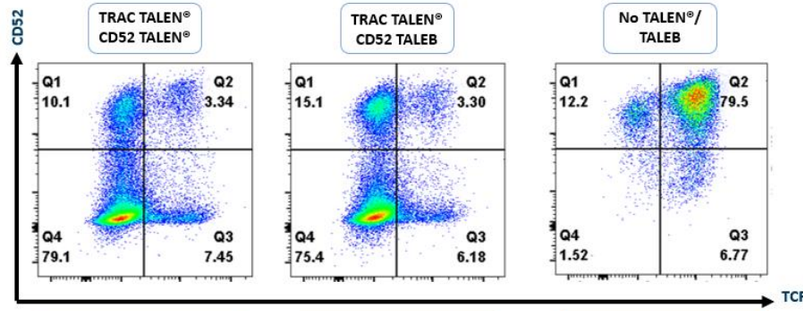


C-to-T conversion (%) of the pos1 C was measured in a 13 bp spacer context. Both the C11 and C40 architecture showed a permissive profile with high rates of editing (C11: 70.96 +/- 14.53, C40: 60.78 +/- 13.74; median +/- stddev). Comparison of the editing results showed small differences in mutation rates on the pos1 C, almost independent of context requirements.

In the 15 bp spacer context, editing results showed a clear difference in mutation rates on the pos1 C, with the C40 architecture showing more permissive activity overall, with less dependence on context requirements than the C11 scaffold. The 15bp spacer in both architectures showed more restrictive editing (C11: 1.09 +/- 2.68, C40: 17.21 +/- 15.41, median +/- stddev). The use of the C11 architecture with the 15 bp spacer nearly abolished editing in most contexts, revealing the possibility to prevent edits on stretches of multiple cytosines.

#4 TALEB do not generate DSBs

No translocations were detected between a bait DSB (nuclease) and a TALEB editing site. Translocations between the two targeted loci, detected by multiplexed amplicon sequencing, were only observed in the TALEN®/TALEN® treated sample (479 reads out of 224,406 for the TALEN®/TALEN® sample; 0 reads out of 144,323 for the TALEN®/TALEB sample, N=1, 1 single T-cell donor).



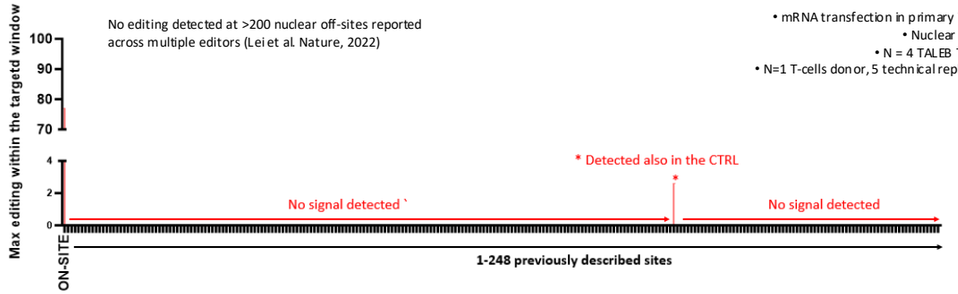
Translocation rate	10^{-4} to $2x10^{-2}$	Not detected	Not detected

Poirot et al. Cancer Research., 2015

#5 TALEB do not edit previously described nuclear off-sites

No editing detected at >200 nuclear off-sites reported across multiple editors (Lei et al. Nature, 2022)

- mRNA transfection in primary T-cells
- Nuclear target
- N = 4 TALEB Tested
- N=1 T-cells donor, 5 technical replicates



#6 Conclusions

This data demonstrates that:

- High throughput screening of TALEB allowed identification of optimal combination of TALEB architecture and positioning on the target to maximize the desired editing outcome
- In primary T-cells, when vectorized as mRNA,
  - TALEB do not show generation of DSBs
  - TALEB do not show editing at previously reported off-sites