P # 859

TransACT enhances detection and characterization of translocation events from high-throughput sequencing data at base-pair resolution for gene editing products

Timothy D. Collingsworth, Kit Cummins, Michael A. Pettiglio, Nipul Patel, Caroline McGowan, Juliana Xavier-Ferrucio, Ruijia Wang, Shu Wang, Michelle Lin, John Lydeard, Gary Ge, and Tirtha Chakraborty

Vor Bio, Cambridge, MA, USA

INTRODUCTION

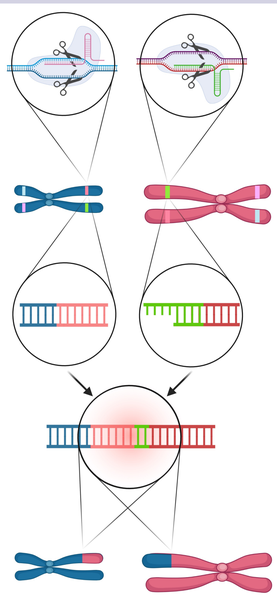
Introducing translocations as an unintended result of gene editing may have irreversible, long-term consequences for patients. Regulators are very sensitive to this risk, as seen with recent clinical holds across the industry. Deep characterization and understanding of translocation events is key to mitigating these risks and safety concerns.

- Removal of surface antigens by editing the genome of hematopoietic stem and progenitor cells (HSPCs) in allogeneic transplants is a novel approach to enable post-transplant targeted therapies.
- Tumor antigen heterogeneity is one of the hurdles in treatment; use of multi-specific (multiplex) therapies will provide greater efficacy.
- Multiplex CRISPR-Cas-based gene editing has broad therapeutic applications but poses a risk of translocation, diminishing or nullifying the therapy by precipitating additional disorders.
- The development of computational tools to support and inform translocation assays is necessary to effectively de-risk gene editing efforts and maximize patient safety.

OBJECTIVE

- Here, we introduce TransACT, a toolkit to assist in the discovery, validation, and quantification of translocation events in edited cells by analyzing unidirectional and targeted next-generation sequencing (NGS) assays.
- We integrate and enhance existing tools to characterize translocation in a portable, scalable, and reproducible manner.

Fig. 1. Translocation from Editing



The occurrence of translocation increases after the introduction of double-strand breaks (DSBs) during genome editing.

- Robust analysis is necessary to detect and validate on-on, on-off, and off-off target fusions.

OVERVIEW

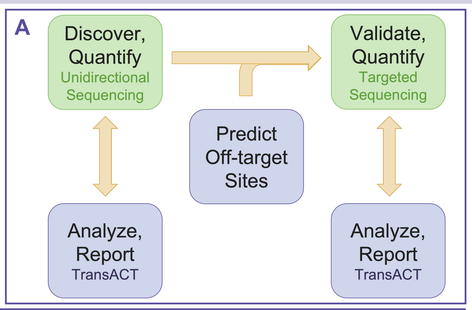
Fig. 2. Toolkit Summary and Workflow Overview

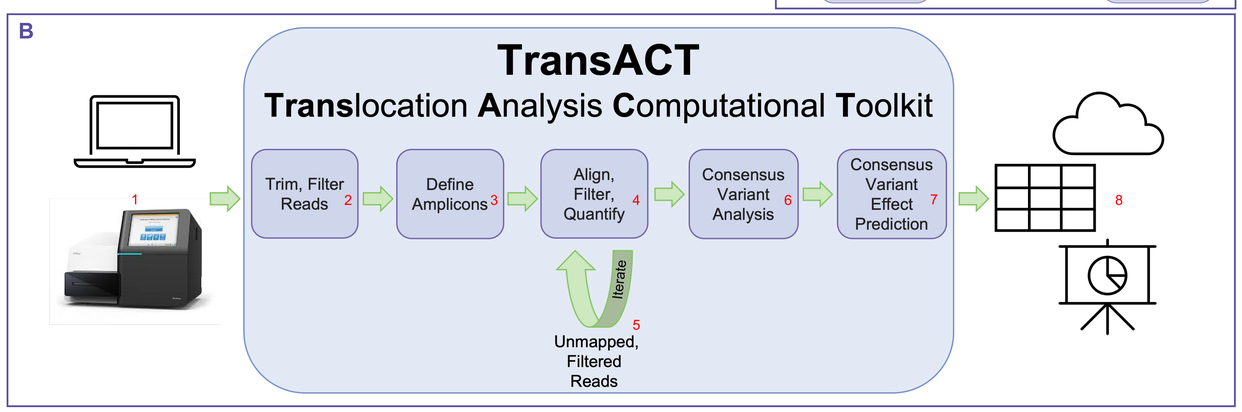
TransACT helps discover, validate and quantify translocation events (**A**).

- Discovery of translocation events using unidirectional sequencing is combined with predicted off-target sites to nominate targets for the design of targeted amplicon sequencing (TAS).

TransACT is an automated sequencer-to-results pipeline (**B**).

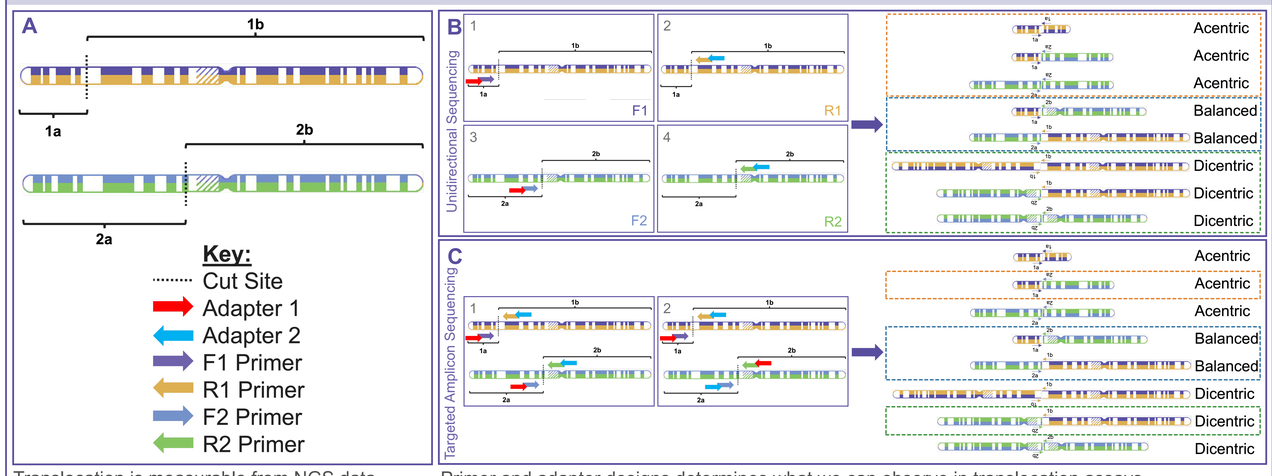
- Experimental design and necessary details are pulled from an electronic laboratory notebook (ELN) and paired with read data from the sequencer (1).
- Reads are first trimmed and filtered to ensure high quality input (2) then, using guide design details, hypothetical translocation sequences (amplicons) are constructed (3, see Fig. 3).
- Reads are then aligned to the amplicon junction, filtered to reduce false positives (see Fig. 4A), and quantified (4). Unmapped and filtered reads are then used in the next iteration, speeding up the process (5).
- Aligned reads are then processed to identify consensus variants (6, see Fig. 4B and Fig. 5) and variant effect(s) are predicted to aid in further de-risking (7).
- Metrics are recorded at every step, stored, and utilized to generate visualizations (8, see Fig. 5).





METHODS

Fig. 3. Translocation Definitions, Assay Organization, and Observable Translocation Species



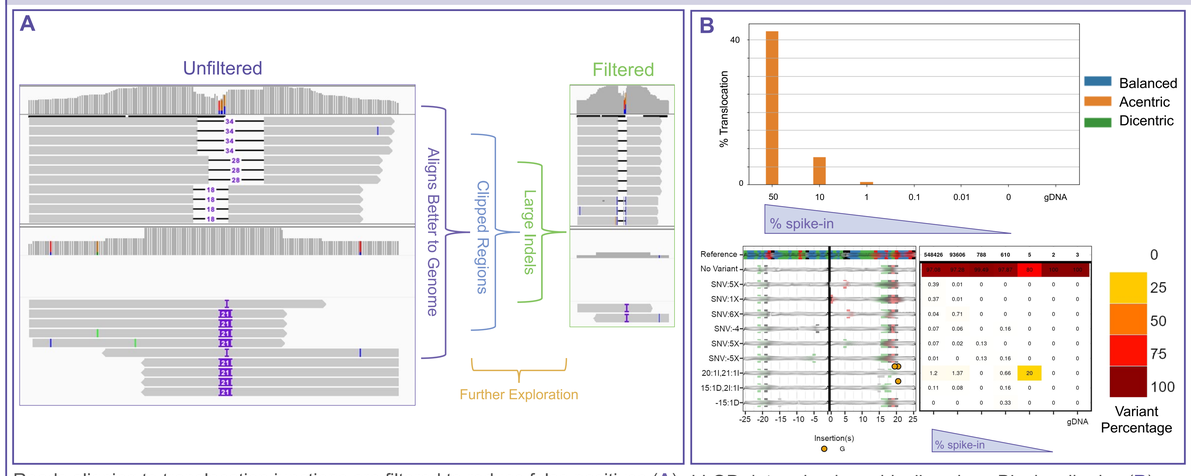
Translocation is measurable from NGS data.

- We focus on potential translocation events resulting from two cuts (dashed lines), each on a different chromosome (ChrA/ChrB): forward primers (F1/F2) extend the top strand, reverse primers (R1/R2) extend the bottom strand; adapter (1/2) allow sequencing of specific primer pairs (**A**).

Primer and adapter designs determines what we can observe in translocation assays.

- Four experiments each with a single primer allows detection and quantification of all 8 translocation species in unidirectional sequencing experiments (**B**).
- In contrast, two primer/adapter pairs allow validation of 4 translocation species in TAS experiments: there are four species that are unobservable in current TAS experiments (**C**).
- Combining these approaches allows us to effectively detect and validate translocation events.

Fig 4. Alignment Filtering and Lower Limit of Detection (LLOD)



Reads aligning to translocation junctions are filtered to reduce false positives (**A**).

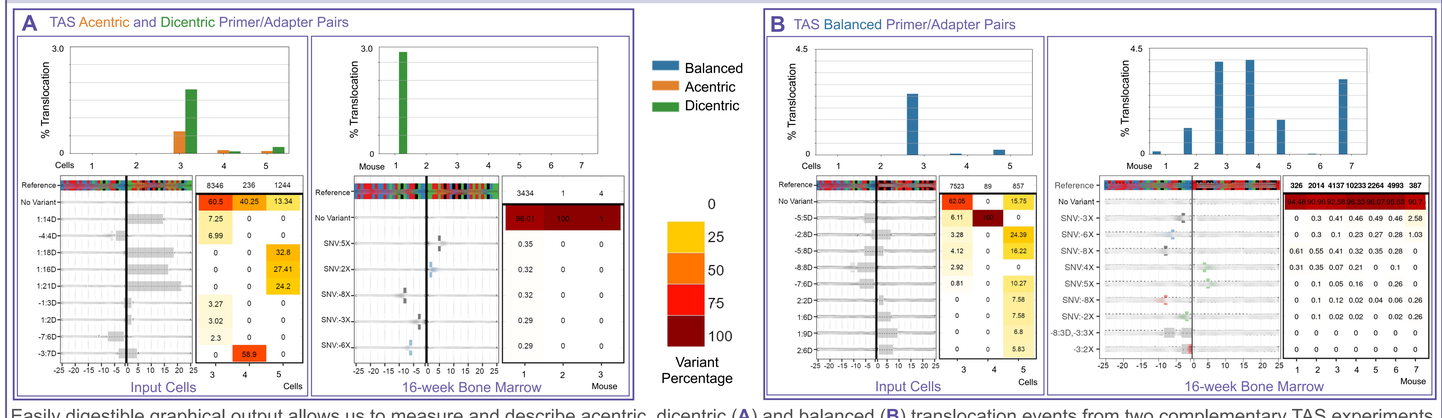
- Reads that align better to the whole genome than the translocation amplicon (AS, MAPQ).
- Alignments with large soft-/hard-clipped regions (> 25bp).
- Alignments with large indels (>30% read length).
- As a result, we increase signal and decrease noise to ensure confidence in observed events.

LLOD determined empirically using gBlock spike-ins (**B**).

- Decreasing ratio of acentric translocation gBlock to WT gBlock (50% to 0%, from left to right plus a gDNA control) is successfully detected and quantified using TransACT and indicates a LLOD of ~0.1-1%.
- Consensus variant sequence analysis corroborates presence of translocation gBlock (i.e., "No Variant") at expected frequencies.

RESULTS

Fig. 5. TransACT Graphical Output: *in vivo* Studies



Easily digestible graphical output allows us to measure and describe acentric, dicentric (**A**) and balanced (**B**) translocation events from two complementary TAS experiments.

- In each separate TAS experiment, the incidence of translocation as well as consensus variants and their frequency is measured in input cells (left) and mice samples taken 16 weeks post-infusion (right).
- While multiple indel variants dominate the recently edited cells, predicted translocation events ('No Variant') and single nucleotide variants (SNVs, likely sequencing errors) persist after 16 weeks in mice.

SUMMARY

TransACT effectively supports the discovery, validation, and quantification of translocation events in NGS data to assess the pre-clinical and clinical safety of gene editing drug products.

- Handles detection of translocation in TAS as well as unidirectional NGS data from edited samples using multiple different enzymes.
- Standard output of summary statistics, visualizations, and explicit translocation species nomenclature facilitate insight into the incidence and type of translocation.
- False positive filtering and translocation alignment visualization ensures confidence in detected translocation events.
- Consensus variant analysis and variant effect prediction allows tracking of problematic translocation species throughout therapy.
- Portable, reproducible, and scalable assay analysis speeds up de-risking and encourages iterative exploration.

References

- Samuelson, C., Radtke, S., Zhu, H. et al. Multiplex CRISPR/Cas9 genome editing in hematopoietic stem cells for fetal hemoglobin reinduction generates chromosomal translocations. Mol Ther 507-523 (2021).
- Clement, K., Rees, H., Canver, M.C. et al. CRISPResso2 provides accurate and rapid genome editing sequence analysis. Nat Biotechnol 37, 224-226 (2019).
- Lindsay, H., Burger, A., Biyong, B. et al. CrispRVariants charts the mutation spectrum of genome engineering experiments. Nat Biotechnol 34, 701–702 (2016).

Disclosures

All authors listed above are current employees of Vor Biopharma.

Acknowledgements

We would like to thank the computational biology, research, technical operations, and lab operations groups at Vor Bio. Fig.1 was generated using BioRender.com.

Presented

Intelligent Systems for Molecular Biology
July 10-14, 2022, Madison, Wisconsin, USA
Session A