# Frequent, Context-Dependent Effects of Human Genetic Variation on Cas9 Activity Revealed by Population-Scale GUIDE-seq-2 and Deep Combinatorial CHANCE-seq Profiling[1]

Flory et al., *bioRxiv* 2025.02.10.637517

*How GUIDE-seq-2 Unlocked the First Genome-Wide, Large-Scale Study of the Effects of Genetic Variation on Cas9 Off-Target Activity*

**BY THE NUMBERS**

| | | | |
|---|---|---|---|
| **95** Genetically diverse donors studied | **665** GUIDE-seq-2 libraries generated | **3 hrs** library prep time vs. ~1 day for prior method | **4×** reduction in gDNA input required |
| **6** therapeutic gRNA targets profiled | **1,115** Cas9 on/off-target sites identified | **99%** of off-target sites overlap gnomAD variants | **14.1%** of tested variants altered off-target activity |

## CHALLENGE: A Critical Gap in Gene Editing Safety Assessment

Every patient's genome differs from the reference sequence at ~4–5 million positions[2]. When a genetic variant overlaps a Cas9 protospacer or PAM sequence, it can create a new off-target site or dramatically alter activity at an existing one. Prior studies have used off-target scoring algorithms to estimate the theoretical prevalence of variant effects, but these were *in silico* analyses. No direct experimental, population-scale, genome-wide assessment had been carried out as off-target assays have been too resource-intensive at the necessary scale.

# SOLUTION: Tagmentation Transforms GUIDE-seq into a Population-Scale Tool

The research team led by Drs. Yong Cheng and Shengdar Tsai, redesigned the well-known GUIDE-seq assay[3], replacing physical DNA sonication with transposase tagmentation during NGS library construction while maintaining full sensitivity and reproducibility. This single change eliminated six processing steps and cut library preparation time from ~1 day to 3 hours, with fourfold less input DNA required, making the method compatible with the limited material often available from primary cell types.

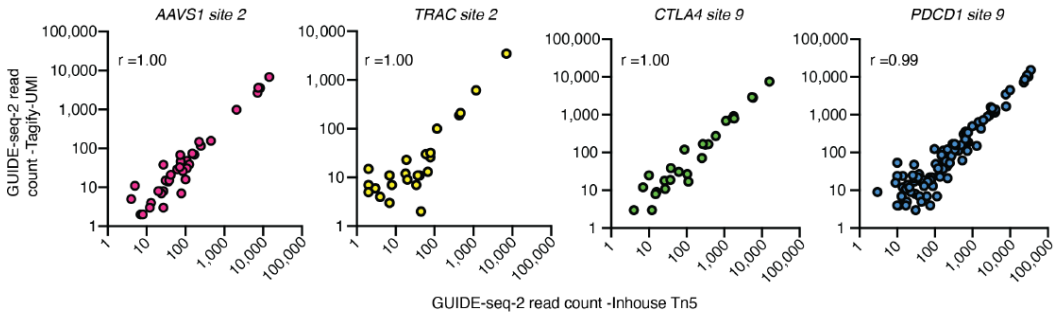
*"Due to a lack of scalable methods to measure genome-wide editing activity in cells from large populations and diverse target libraries, the frequency and extent of these variant effects on editing remain unknown."*

– Flory et al., *bioRxiv* 2025.02.10.637517

### Why Tagmentation Was Key

Tagmentation simultaneously fragments genomic DNA and appends sequencing adapters eliminating end-repair, A-tailing, adapter ligation and the need for shearing equipment. A redesigned single-round PCR strategy further simplified the workflow to 3 hours.

The team validated both in-house transposase preparations and a UMI-loaded Tn5 transposase commercially available from seqWell for the new GUIDE-seq-2 library prep method. Side-by-side comparisons across four therapeutically relevant target sites (AAVS1 site 2, TRAC site 2, CTLA4 site 9, and PDCD1 site 9) demonstrated near-perfect correlation between in-house and seqWell Tagify-UMI-generated libraries (Pearson's r = 0.99–1.00).




**Supplementary figure 1| GUIDE-seq-2 profile correlates well between inhouse Tn5 and commercially available Tn5** Scatterplots of GUIDE-seq-2 read counts (log10) from experiments performed on primary human T cells for 4 target sites. Genomic DNA were tagmented using inhouse Tn5 (x-axis) and seqWell Tagify-UMI (y-axis) before GUIDE-seq-2 PCR. GUIDE-seq-2 library prepared with each method was sequenced individually with NextSeq2000. r = Pearson's correlation coefficient.

Figure source: Flory, A, et.al., *bioRxiv*, 2025.02.10.637517.

## PUBLICATION: What GUIDE-seq-2 Made Possible

The first population-scale, genome-wide cellular off-target profiling study was performed using the GUIDE-seq-2 method for six therapeutically relevant gRNA targets - AAVS1 site 14, CCR5 site 8, CTLA4 site 9, CXCR4 site 8, LAG3 site 9, and TRAC site 1 - in lymphoblastoid cell lines from 95 individuals spanning four genetically diverse populations: African Ancestry in Southwest USA (ASW), Northern Europeans from Utah (CEU), Han Chinese in Beijing (CHB), and Mexican Ancestry in Los Angeles (MXL). The scale of this study—665 libraries, six targets, four populations—would have been logistically and economically unfeasible with the original GUIDE-seq protocol.

**Tagify i5 UMI Adapter-loaded Transposases:**

- Eliminate a multi-step protein purification & oligo annealing workflow
- Provide access to a critical assay reagent for laboratories without a dedicated molecular biology infrastructure
- Ensure lot-to-lot consistency with validated performance specifications to address regulatory expectations for clinical research applications

## RESULTS: Genetic Variation Significantly Alters Cas9 Off-Target Activity

- **Off-Target Sites Frequently Overlap Known Human Variants:** 16.6% of Cas9 off-target sites overlapped a genetic variant in the 95-person cohort; this rises to 99% when considering the full gnomAD database - meaning virtually every off-target site in the human genome has the potential to be affected by patient-specific genetic variation.
- **Variants Can Significantly Alter Cellular Editing Activity:** 14.1% of tested variants showed a significant, at least two-fold change in off-target activity (6.6% increased, 7.5% decreased); effects on GUIDE-seq-2 read counts ranged from an ~5-fold decrease to a 128-fold increase; and four of the increased-activity variants created entirely new ("*de novo*") off-target sites not detected in the reference genome.
- **A Single Nucleotide Difference in a gRNA Can Have Meaningful Impacts:** One T>G variant at the first position of a CCR5-targeting gRNA off-target site led to a 70-fold increase in Cas9 indels, elevating off-target editing from near-zero to approximately 10%.
- **Population-Specific Patterns of Risk:** African-ancestry donors showed the highest variant overlap frequency (13%), highlighting the inadequacy of European-dominated reference databases for diverse patient populations.

## CLINICAL SIGNIFICANCE: Why It Matters for Gene Therapy

Current off-target safety assessments are based on reference-genome sequences in a handful of donors. This study shows that this approach may miss clinically significant risks that are specific to individual patients or ethnic subgroups. For approved therapies like exa-cel (which treats sickle cell disease—a condition disproportionately affecting patients of African ancestry), and for programs currently in clinical trials, these findings argue for incorporating population-aware off-target profiling into therapeutic development.

> *"For exa-cel and similar therapies, a critical question is how genetic variation within patient populations may influence the risk of unwanted off-target editing. As genome editing therapies expand to broader patient groups and different targets, accounting for human genetic diversity will be essential to ensure their consistent safety and efficacy across populations."*
>
> – Flory et al., *bioRxiv* 2025.02.10.637517

### Beyond Cas9: Base Editors, Prime Editors, and Future Modalities

While this study focused on Cas9 nuclease activity, the authors note that similar considerations apply across the broader landscape of genome editing modalities, including base editors and prime editors. As these technologies advance toward clinical application, population-scale off-target analysis will be increasingly important for comprehensive pre-clinical safety characterization.

### ALSO IN THIS STUDY: A Complementary Biochemical Assay – CHANCE-seq

Alongside GUIDE-seq-2, the team also developed CHANCE-seq, an *in vitro* biochemical assay that profiles Cas9 activity across millions of synthetic mismatched sequences. This dataset trained CHANCE-net, a machine learning model that predicts variant effects on off-target activity for genotypes not directly sampled by GUIDE-seq-2, extending the population-scale cellular findings.

**Reference Publication**


1. Flory AR, Lazzarotto CR, Li Y, et al. Frequent, context-dependent effects of human genetic variation on Cas9 activity revealed by population-scale GUIDE-seq-2 and deep combinatorial CHANCE-seq profiling. bioRxiv (2025). doi: 10.1101/2025.02.10.637517
2. 1000 Genomes Project Consortium, et al. A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).
3. Tsai, S. Q. et al. GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat. Biotechnol.* **33**, 187–197 (2015).


**Data & Code Availability**

Sequencing data: NCBI GEO SuperSeries GSE286300 GUIDE-seq-2 code: github.com/tsailabSJ/guideseq
CHANCE-seq code: github.com/tsailabSJ/chanceseq