Genome-wide Mapping of CRISPR-Cas Off-Target Cleavage Sites via GUIDE-Seq

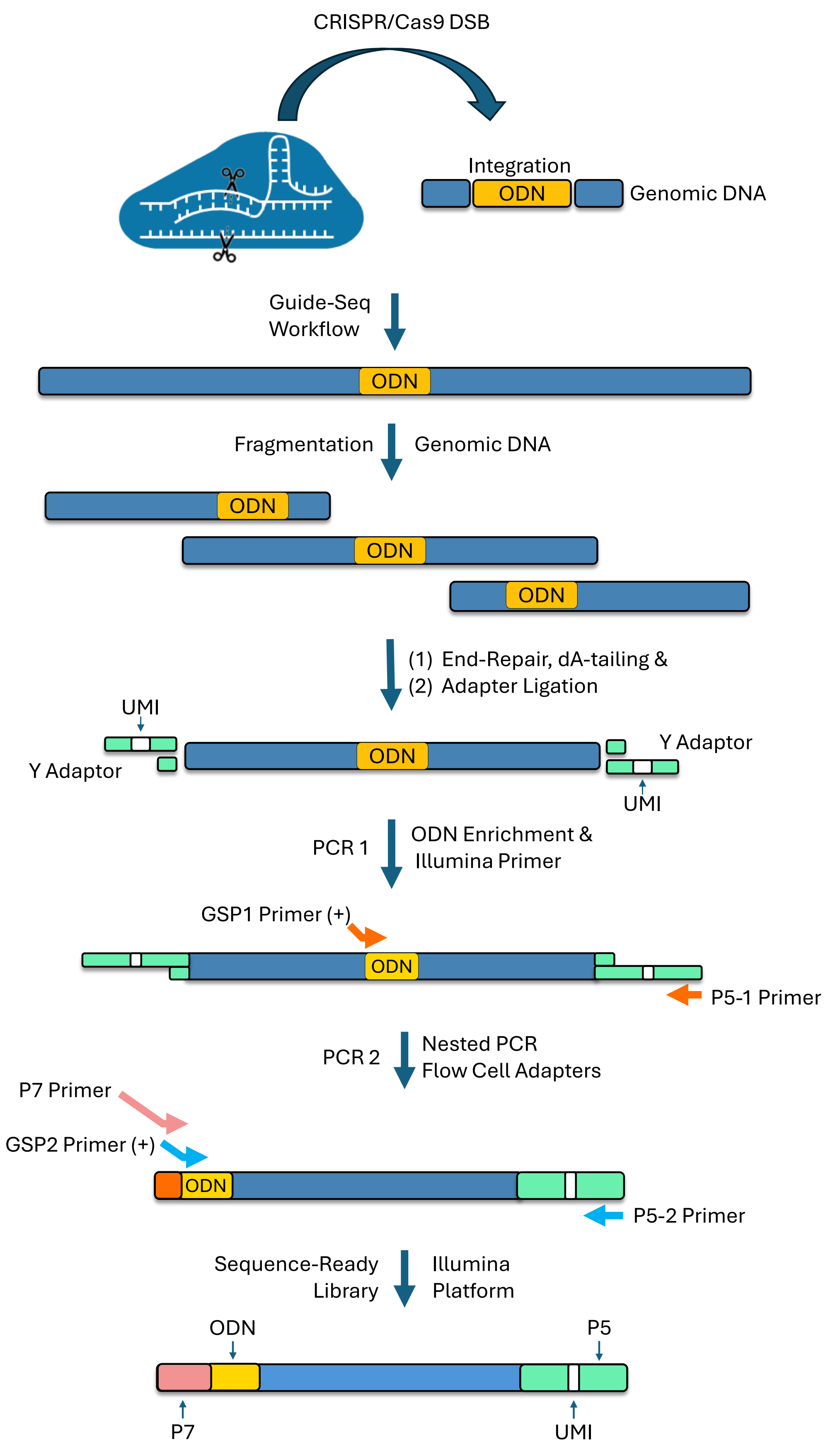
Jimmy Goncalves, Aldrin Lugena, David Zhou, Emma Cross, Xuening Huang

INTRODUCTION

Genome editing has transformed molecular biology by enabling precise genetic manipulation. Yet off-target double-strand breaks (DSBs) remain a concern—particularly in therapies, where the FDA prioritizes genome-wide specificity for patient safety. Edits in functional or regulatory regions may disrupt tumor suppressors, activate oncogenes, or alter essential genes—posing serious risks. GUIDE-Seq (Genome-wide Unbiased Identification of DSBs Enabled by Sequencing) enables sensitive, high-throughput detection of on- and off-target DSBs in living cells, capturing the real genomic context shaped by chromatin, repair, and epigenetic state. The method relies on the integration of blunt-ended double-stranded oligodeoxynucleotides (dsODNs) at DSB sites via non-homologous end joining (NHEJ), serving as molecular tags for enrichment and sequencing. Unlike in vitro assays that overpredict without cellular context, GUIDE-Seq offers a physiologically relevant, genome-wide map of CRISPR-Cas activity for research and safety evaluation.

GuideSeq Workflow

gRNA Screening & Ranking	• Screen gRNAs for editing efficiency and viability. • Validate on-target activity (indel or HDR) prior to GUIDE-Seq.
Cas9 + gRNA Delivery Optimization	• Optimized nucleofection with viability assessment to confirm delivery and minimize cytotoxicity.
dsODN Tag Delivery	• Integration-ready dsODNs (34–46 bp) with end-protection chemistry.
dsODN Integration Efficiency Testing	• Use NdeI digestion or amplicon sequencing to quantify insertion rate.
Control Arm Design & Background Correction	• Include dsODN-only, untreated, and matched cell type controls.
GUIDE-seq Library Prep & Analysis	• Shearing, UMI tagging, sequencing, and modified pipeline analysis.



After DSB induction, blunt-ended dsODNs are inserted at break sites via NHEJ, tagging genomic loci. Genomic DNA is extracted, fragmented, end-repaired, and ligated to adapters with a unique molecular identifier (UMI) and P5 binding site. The UMI is a random tag used for deduplication and error correction during analysis. The P5 sequence enables Illumina flow cell binding for cluster generation and sequencing.

PCR1 amplifies dsODN-tagged fragments using a P5-specific primer (P5-1) and a gene-specific primer (GSP1) targeting the dsODN tag. PCR2 uses nested GSP2 and P5-2 primers, which act as scaffolds for further amplification by a P7 primer homologous to GSP2, completing the Illumina-compatible library. The resulting library enables genome-wide identification of on- and off-target sites via Illumina sequencing.

Why Guide-Seq?

Gold Standard for In Vivo Off-Target Detection

GUIDE-Seq stands out among off-target detection methods for its ability to capture DSBs in living cells using an integration-based approach that reflects the genomic context of CRISPR-Cas activity. *In silico* tools like Cas-OFFinder, CRISPOR, and CRISPRme predict off-target sites but lack experimental validation. In vitro methods such as Digenome-seq, CIRCLE-seq, and SITE-seq are sensitive but may not reflect repair dynamics in cells. Integration-based methods like GUIDE-Seq, iGUIDE-Seq, INDUCE-seq, and DISCOVER-seq offer a more relevant view by tagging DSBs through endogenous repair. Among these, GUIDE-Seq remains widely used, balancing coverage, specificity, and sensitivity for both discovery and regulatory use.

Case Study: Avance Guide-Seq Product Results

Target Site	Replicate 1 (Unique UMIs)	Replicate 2 (Unique UMIs)	Replicate 3 (Unique UMIs)	%INDEL (Amplicon-Seq)
On-target	1,953	2,227	2,045	78.03
Off-target A	ND	ND	3	0.456
Off-target B	14	15	7	0.159
Off-target C	7	7	7	0.118
Off-target D	4	5	ND	0.094
Off-target E	3	7	ND	0.022
Off-target F	2	ND	4	0.02
Off-target G	ND	ND	ND	0.02
Off-target H	20	16	23	0.0189
Off-target I	ND	2	ND	0.009

GUIDE-Seq detection of on- & off-target editing events in a test article processed in triplicate. The table reports the number of unique molecular identifiers (UMIs) corresponding to dsODN integration at each genomic site, including an on-target site and a panel of previously characterized off-target sites. Although all three replicates were derived from the same sample, variability in detection was observed, particularly at low-frequency off-target sites, which is consistent with the stochastic nature of rare editing events. Indel frequencies (%INDEL) were calculated from previously performed amplicon sequencing as %INDEL = indel reads / total reads (i.e., wild-type + indel reads). This calculation was performed for both edited and unedited samples, and the reported %INDEL values are background-corrected by subtracting the %INDEL of the unedited control from that of the edited sample.

Case study: Avance-modified GUIDE-SEQ analysis PIPELINE

	GUIDE-Seq Pipeline	Replicate 1	Replicate 2	Replicate 3
Total Detected Sites	Original	454	452	507
	Modified (Filtered)	263	291	319
Nominated Off-Target Sites	Original	11	14	12
	Modified (Filtered)	10	11	12
Verified Off-Target Sites	Original	6	6	5
	Modified (Filtered)	5	6	5

Comparison of original and modified GUIDE-Seq pipelines for detecting and verifying off-target editing sites. Libraries were analyzed using either the original pipeline or an Avance-modified version that uses longer dsODNs to filter mispriming artifacts. The original workflow, which relies on 34 bp dsODNs, is prone to PCR mispriming due to short sequence length, leading to background reads from non-specific amplification in PCR1. The modified pipeline, using iGUIDE's longer 46 bp dsODNs with an added 12 bp sequence (6 bp added to both ends), serves as a filter to eliminate misprimed reads and improve specificity. Total detected sites decreased significantly with the modified pipeline, reflecting removal of spurious events. However, the number of nominated and verified off-targets remained consistent between the two pipelines, indicating that biologically relevant signals were retained after filtering.

CONCLUSIONS

GUIDE-Seq enabled successful genome-wide profiling of CRISPR-Cas nuclease activity, accurately identifying both on-target and off-target double-strand breaks (DSBs). Detection of low-frequency off-target sites showed variability, which may be improved by increasing replicate number. To enhance DSB calling accuracy, the analysis pipeline was refined with a custom filtering step adapter from iGuide-Seq that removes misprimed PCR artifacts. Avance supports sample processing and analysis with both short and long dsODNs using a unified workflow. Following GUIDE-Seq, nominated off-target sites and *in* silico predicted sites can be verified by rhAmpSeq or amplicon sequencing, supporting downstream assay development and validation.

ACKNOWLEDGEMENTS
Fade Gong, Xihong Xie, Hadir Madrigal, Peyton Frantom, Apurva Vansadia, Tina He, Aaron Smith, Jacob Boggs, James Spurlin

www.avancebio.com
(832) 968-8698
9770 West Little York Road, Houston TX 77040