

UCSF

UC San Francisco Previously Published Works

Title

UNCOVERseq enables sensitive and controlled gene editing off-target nomination across CRISPR-Cas modalities and systems

Permalink

<https://escholarship.org/uc/item/5cg628rf>

Journal

Nature Communications, 17(1)

ISSN

2041-1723

Authors

Kinney, Kyle J

Jia, Kun

Zhang, He

et al.

Publication Date

2026-08-01

DOI

10.1038/s41467-026-74623-7

Copyright Information

This work is made available under the terms of a Creative Commons Attribution-NonCommercial-NoDerivatives License, available at

<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Peer reviewed

UNCOVERseq enables sensitive and controlled gene editing off-target nomination across CRISPR-Cas modalities and systems

Received: 12 September 2025

Accepted: 9 June 2026

Published online: 11 August 2026

 Check for updates

Kyle J. Kinney¹, Kun Jia², He Zhang¹, Ellen Schmaljohn¹, Thomas Osborne¹, Bernice Thommandru¹, Karthik Murugan¹, Andrea Sánchez-Peña¹, Sean West¹, Shengyao Chen¹, Roshani Codipilly¹, Morgan Sturgeon¹, Rolf Turk¹, Matthew S. McNeill¹, Mark Behlke¹, Ashley Jacobi¹, M. Kyle Cromer², Garrett Rettig¹ & Gavin L. Kurgan¹✉

The rapid expansion of CRISPR-Cas gene editing enables new therapeutic strategies but complicates assessment of unintended editing risks due to emerging modalities and unclear analytical standards. We present UNCOVERseq (Unbiased Nomination of CRISPR Off-target Variants using Enhanced RhPCR), an improved in cellulo off-target nomination workflow that sensitively identifies rare off-target events using defined inputs and analytical process controls. Using an inter-method off-target confirmation benchmarking dataset, UNCOVERseq demonstrates high analytical sensitivity (97.6%) and precision (78%), outperforming published nomination methods. We apply UNCOVERseq across 192 guide RNAs and identify six guides spanning a broad specificity range, enabling relative risk assessment across *S. pyogenes* Cas9, high-fidelity variants, and base editors in hematopoietic stem and progenitor cells. We further show that double-strand break nomination sites retain strong rank-order concordance with single-strand break-mediated base editing. Together, these results establish UNCOVERseq as a robust framework for informed off-target risk assessment in translational gene-editing systems.

CRISPR-Cas systems have revolutionized genetic engineering by enabling precise genome modifications. Derived from a bacterial adaptive immune mechanism, CRISPR-Cas9 allows targeted DNA edits in various organisms, enabling vast potential in medicine, agriculture, and biotechnology^{1,2}. However, CRISPR-Cas9 faces challenges, notably off-target effects where unintended genomic regions similar to the target sequence can be edited^{3,4}. This off-target activity is often due to sequence homology between the protospacer region of the gRNA and non-target regions of the genome. These off-target edits can lead to unwanted genetic changes and disrupt the function of essential genes leading to adverse effects on cellular function, viability, development

of secondary diseases, or exacerbate existing conditions that pose risks in therapeutic applications.

Recent advancements in gene editing have driven the parallel development of technologies to nominate off-target editing effects^{5–12}. These technologies broadly fall into three categories: cell-based in cellulo/in situ, computational in silico, and biochemical in vitro assays each with distinct advantages and disadvantages. Cell-based assays like GUIDE-seq are known for high precision due to their ability to accurately mimic the cellular environment^{7,13,14}. While cell-based methods provide nomination in a cell-type-specific and even donor-specific manner, these methods face challenges in representing diverse

¹Integrated DNA Technologies, Coralville, IA, USA. ²Department of Surgery, University of California, San Francisco, CA, USA. ✉e-mail: gkurgan@idtdna.com

genome populations and typically require biologically variable processes such as DNA double-strand breaks (DSBs) and blunt integration of a dsDNA tag through non-homologous end joining (NHEJ) as a mechanism of action. Biochemical in vitro assays are highly sensitive and can represent patient genotypes with diverse editing modalities since they are executed on cell-free genomic DNA, but have low precision and suffer difficulties in replicating the cellular environment, which significantly impacts enzymatic activity^{6,7,13–15}. Computational in silico approaches can rapidly predict all candidate off-target sites with homology to the gRNA in a patient genome/population-specific manner, but result in low precision due to insufficient data to inform predicted enzymatic and biological activity across diverse cell types^{11–14,16}. All methods are setback by a lack empirical data and application of standards to characterize sensitivity against known truth, as it is challenging to ascertain a complete list of true positives for each gRNA in an unbiased manner. To better standardize off-target safety assessment, comprehensive information on assessment methods and guidance for their use is essential.

Most methods for off-target nomination currently utilize next-generation sequencing (NGS) to detect off-targets. The efficiency of library preparation, library complexity (i.e., unique genome equivalents sequenced), read depth, and computational analysis methods are all known limitations affecting performance of NGS methods¹⁷. Some work has endeavored to improve upstream biological and library preparation inefficiencies of in-cellulo methods^{18,19}. Other work has presented evidence of biological optimizations that proxy cell lines promiscuously editing off-targets may represent a superset of off-targets in HSPCs²⁰. However, exploration of other NGS variables commonly benchmarked in parallel fields to drive uniformity in experimental procedures and reporting, are still largely lacking for gene editing off-target nomination^{17,21}. The FDA has recently worked to address this need through draft guidelines for the cell and gene therapy industry, which currently recommends multiple orthogonal methods for off-target nomination²². This guideline for orthogonality ideally aims to mitigate known and unknown weaknesses in nomination, including variability in NGS method execution.

Rapid innovations in gene editing have further complicated the off-target nomination process with the discovery of various editors possessing unique qualities beyond the targeted generation of DSBs with a Type II nuclease. This includes engineered systems generating staggered double-strand breaks (DSBs) in Type V CRISPR-Cas systems or leveraging single-strand DNA breaks (SSBs) coupled to engineered moieties for editing^{23–25}. SSB engineered moieties like adenine base editors (ABE)²⁶, cytosine base editors (CBE)²⁷, and prime editors (PE)²⁸ offer benefits such as decreased frequencies of unintended indels and even the ability to make non-indel DNA changes typically reserved for homology-directed repair (HDR). However, they lack consensus guidance regarding appropriate nomination assays for off-target assessment, and previous work has largely used DSB nomination methods as a proxy^{29,30}. Due to the nature of these editors to generate both indels and SNPs, an orthogonal approach specializing on each event type may be needed to sensitively nominate off-targets.

In this work, we develop an optimized end-to-end in cellulo nomination method and analysis pipeline, UNCOVERseq (Unbiased Nomination of CRISPR Off-target Variants using Enhanced RhPCR), leveraging NHEJ-based dsDNA tag integration for RNase H-dependent PCR (rhPCR)-based library preparation³¹. We demonstrate that using this method with a cell line under promiscuous editing conditions serves as a sensitive proxy for clinically relevant cell types. We establish experimental recommendations for library complexity, coverage depth, and end-to-end process controls that reliably lead to the successful nomination of risk-tiered edited sites with sub-0.01% editing (indel) frequencies observable empirically, and demonstrate that this optimized method outperforms other nomination methods through benchmarking comparisons. By screening 192 gRNAs using

UNCOVERseq, we identify an experimental set of gRNAs spanning a broad range of specificities that are broadly representative of use-cases encountered in the field. We find that off-target effects of base editors are directly, rank-order correlated to DSB editing frequencies, demonstrating UNCOVERseq as an effective technology for nominating meaningful off-targets for both current and next-generation editing modalities. We conclude that, regardless of editor modality, the off-target burden of gRNAs quickly decreases to frequencies not measurable—that is, below the limit of detection—with current NGS methods (0.01% for indels and 0.5% for base editing) as gRNA specificity ratios rise. However, off-targets may still be observable at higher specificities in an editor/gRNA-dependent manner, emphasizing the need to consistently and empirically nominate and interrogate off-targets across the specificity spectrum.

Results

Development and optimization of UNCOVERseq

To create our nomination method, we first started with the original GUIDE-seq protocol and designed an orthogonal dsDNA sequence with sufficient length to perform a modified rhPCR to multiplex primers in close proximity within a single reaction while avoiding primer-dimers (Supplementary Data 1). To streamline the process for preparing the nomination gDNA libraries we additionally converted from a mechanical to enzymatic DNA fragmentation (Fig. 1A). Upon analyzing data, we observed that freely adapted dsDNA tag was allocated an average range of 37–67% of reads, varying across 4 gRNAs (Fig. 1B). This same artifact was also observed with the original GUIDE-seq protocol (Supplementary Fig. 1). To improve usable reads resulting from NGS, we introduced a blocking oligo into the PCR1 preparation designed to target the adapter:dsDNA junction (Fig. 1A). Introduction of this blocker reduced reads belonging to the adapter:dsDNA artifact to an average range of 0.3–0.5%, meaning >99% of reads were now belonging to gDNA:dsDNA junctions (Fig. 1B). Nomination frequencies were found to be conserved for all gRNAs ($R^2 = 0.99$) with and without the blocking oligo (Supplementary Fig. 2).

In parallel to creation of the wet-lab protocol, we created an analysis pipeline with features such as heuristic nomination criteria (Levenshtein distance <7; read-evidence from both sides of a prospective off-target), statistical comparison of treatment:control samples as nomination criteria (FDR < 0.05) and integrated genomic annotations. We then investigated optimizations in the computational pipeline for nomination of gRNAs using a set of 48 gRNAs spread across the *PDCD1*, *LAG3*, *CTLA4*, *NRP1*, *IL2RA*, and *TIGIT* genes. In off-target nomination, off-target loci are generally determined to be trustworthy based on: (1) frequency, (2) reproducibility, and/or (3) similarity to the intended target sequence (gRNA), with Levenshtein distance >6 often being used to disqualify an off-target.

To investigate the effect of alignment method used for determination of an off-target list, we tested existing GUIDE-seq pipeline methods (<https://github.com/aryeelab/guideseq>; commit: 997b892; fuzzy regular expression based; Regex) and historical GUIDE-seq pipeline methods (<https://github.com/aryeelab/guideseq>; tag:v1.0; Smith-Waterman alignment with –100/–100 gap open/extension penalty) as compared to a global implementation of the Needleman-Wunsch algorithm. Investigation of 48 different gRNAs found a significant difference in the number of Levenshtein distance <7 loci nominated using each approach, and that the global Needleman-Wunsch alignment approach yielded a median of 30% and 150% more qualified off-target locations than the current and historical GUIDE-seq analysis approaches (Fig. 1C). To do an end-to-end comparison of the current GUIDE-seq method to our method, we did a complete workflow comparison across the same 48 gRNAs and found that our method nominated a range of 60–4150% more off-targets per gRNA, with an average of 132 off-targets as compared to 23 using the GUIDE-seq protocol (Fig. 1D). We termed the final instantiation of the end-to-

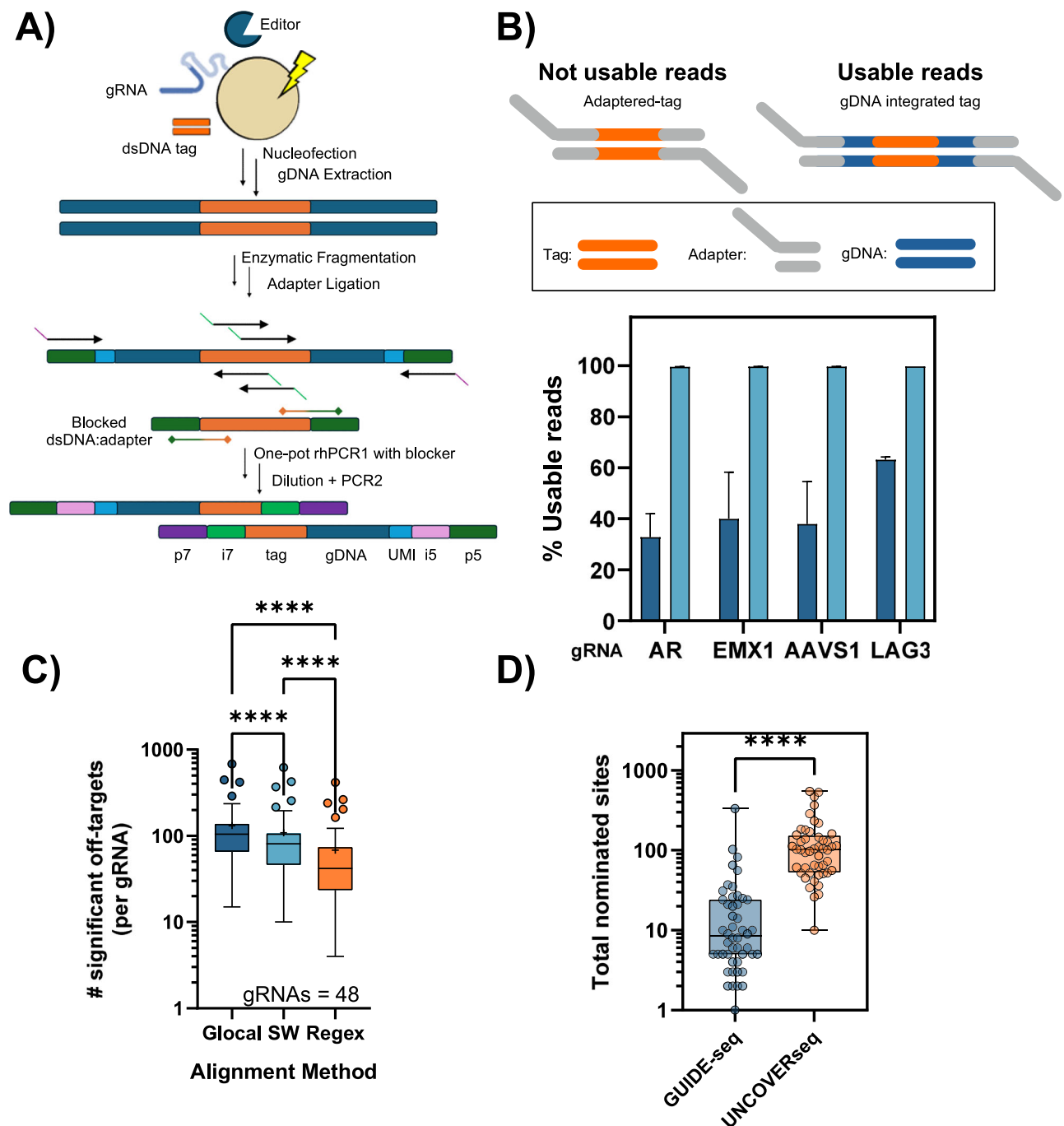


Fig. 1 | Comparison of UNCOVERseq and the GUIDE-seq off-target nomination workflows. **A** An overview of the UNCOVERseq workflow demonstrates that cells with a genomically integrated dsDNA tag have gDNA extracted and amplified with rhPCR in a single reaction, with dsDNA tag:adapter byproducts being blocked by a targeted oligo before being sequenced and analyzed using our described workflow. **B** A depiction is shown regarding what kind of events are targeted by the blocking oligo, with the usable reads (non-dsDNA:adapter reads) measured across nominated off-targets for 4 gRNAs in K562 ($n = 3$ per gRNA) with (light blue) or without (dark blue) the blocking oligo (mean $\pm$ s.d. shown). **C** Tukey box plot comparison of different alignment methods used in publicly available nomination packages was performed to determine differences in Levenshtein distance < 7 sites from a glocal Needleman-Wunsch alignment (Glocal; +2/-1/-10/-1 match/

mismatch/gap open/gap extension; penalize_end_gaps = False), Smith-Waterman (SW; original GUIDE-seq pipeline; +2/-1/-100/-100 match/mismatch/gap open/gap extension) or string-match method (Regex; current GUIDE-seq pipeline). Statistical significance was determined for pairwise comparisons using a two-tailed paired T -test. **D** Box plots (min-max with Q3, median and Q1; all raw data points) show the end-to-end differences in total nominated off-target sites for the current GUIDE-seq method (wet lab protocol and Regex alignment) compared with UNCOVERseq. Data is representative of 48 gRNAs (Targets: *PDCDI*, *LAG3*, *CTLA4*, *NRPI*, *IL2RA*, and *TIGIT*; 8 gRNAs per target). Statistical significance was determined for pairwise comparisons using Wilcoxon matched-pairs signed rank test and for multiple comparisons using a Friedman test with a post-hoc Dunn's test with Bonferroni correction. **** $p < 0.0001$.

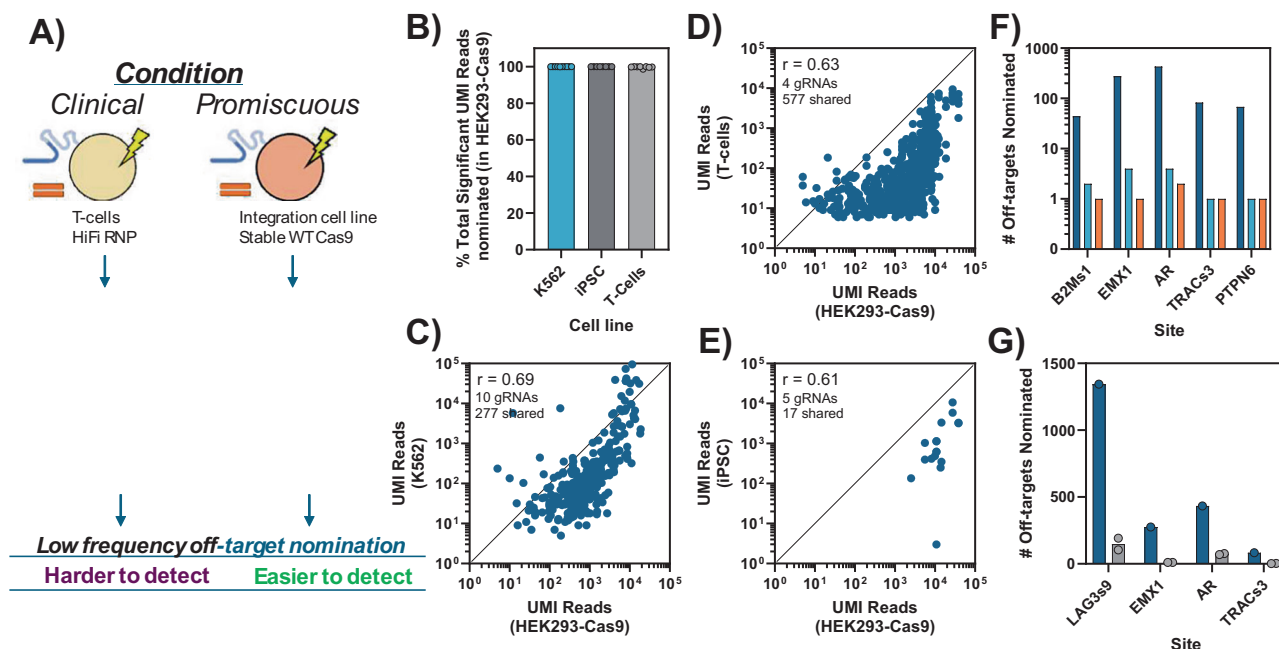


Fig. 2 | Analysis of performance of a promiscuous cell system and the ability to translate to other cellular systems. **A** Example diagram demonstrating promiscuous nomination systems can help sensitively represent off-target lists by increasing the frequency that off-targets are detected. **B** Comparison of total frequencies (represented as the cumulative % significant UMI reads) of nominated sites captured in a single replicate ($n = 1$) of HEK293-Cas9 vs. wildtype Cas9 RNP transfection using three different cell types: K562 ($n = 10$ gRNAs; biological duplicates per gRNA), iPSCs ($n = 5$ gRNAs; one biological replicate per gRNA; wildtype

and HiFi Cas9), and primary T-cells ($n = 4$ gRNAs; 2 biological replicates per gRNA as two different donors) mean $\pm$ s.d. shown. Total UMI-corrected reads, total number of nominated sites, and Spearman correlation are shown for shared nominated sites between HEK293-Cas9 and, **C** K562, **D** T-cells, and **E** iPSCs. Total average number of nominated sites between **F** iPSCs + S.p. Cas9 (light blue), or HiFi Cas9 (orange) and **G** primary T-cells (gray) are displayed in comparison to HEK293-Cas9 (dark blue).

end method as UNCOVERseq (Unbiased Nomination of CRISPR Off-target Variants using Enhanced RhPCR; v1.0).

Promiscuous cell systems as sensitive UNCOVERseq proxy nomination models

To identify ideal biological operating conditions, we explored biological variables with potential workflow impacts on nomination performance. Promiscuous editing conditions are known to increase editing frequencies at off-targets, which we hypothesize should increase the sensitivity of in cellulo methods like UNCOVERseq (Fig. 2A). To test this, we selected 4–10 gRNAs per cell line and nominated off-targets using UNCOVERseq in K562, iPSCs (wildtype *S.p.* Cas9 or HiFi Cas9), primary T-cells, or a promiscuous HEK293 cell line stably expressing *S.p.* Cas9 (HEK293-Cas9). Investigation of the overlap of off-target frequencies between these cell lines and HEK293-Cas9 found an average of 99.7–100% of total UMI-corrected events (corresponding to total frequency) in each cell line could be found in the off-targets of just a single replicate of HEK293-Cas9 (Fig. 2B). Comparison of shared nomination frequencies showed a high rank-order correlation for HEK293-Cas9 nominated off-targets between K562 ($r = 0.69$), iPSC ($r = 0.61$), and primary T-cells ($r = 0.63$), demonstrating that the frequency-based importance of different off-targets for prioritization was still largely conserved (Fig. 2C–E). We also observed that the overall nominated off-targets number for the same gRNAs could vary significantly in different primary cell lines, with low numbers of off-targets nominated in iPSCs, further demonstrating why a promiscuous system can be used to remove this variable (Fig. 2F). Overall nomination in a promiscuous cell line like HEK293-Cas9 was capable of generating an average of 196–1560% more candidate targets per gRNA compared to an efficient primary cell type for nomination, like primary T-cells, supporting that this is a more sensitive model for off-target nomination even in translational contexts (Fig. 2G).

Off-target reproducibility using UNCOVERseq

To assess sample-to-sample reproducibility, we compared biological triplicates of UNCOVERseq in HEK293-Cas9 across four gRNAs. A range of 98.9–99.7% of instances based on frequency were shared between any two biological replicates across gRNAs, indicating high frequency sites are consistently captured with a single replicate (Fig. 3A). Frequency rank order of off-targets was highly conserved ($R^2 = 0.997$) across replicates as well (Fig. 3B). This indicates that both frequencies and sites containing the majority of reads are highly reproducible between UNCOVERseq replicates.

To determine ideal experimental conditions for off-target nomination, we characterized factors affecting reproducibility in a functional context. When making decisions about nominated off-targets, ideally off-targets are prioritized based on: (a) frequency, (b) reproducibility, and (c) genomic impact. To this end, we developed a tiering system based on UNCOVERseq data to prioritize off-targets for confirmation (Tier 1–3) from less important ones (Supplementary Data 2). To assess reproducibility for prioritizing important off-targets (Tier 1–3), we compared biological triplicates to single replicates of the previous 48 gRNAs in HEK293-Cas9. Without biological triplicates, 30–40% of high priority off-targets were not captured or prioritized (Fig. 3C). The average frequency of missed high priority targets increased with gRNA specificity, though this is partially because the denominator (total off-targets) is smaller (Fig. 2C). While high frequency events were reproducible (Fig. 3A), low frequency events lacking full reproducibility or those in important genomic contexts (e.g., exonic) were not appropriately identified without replicates (Fig. 3C).

Replication additionally theoretically allows more appropriate tiering of off-targets. To investigate the effect of tiering on the downstream requirements for confirmation, we compared the number of off-targets nominated for appropriately tiered off-targets of

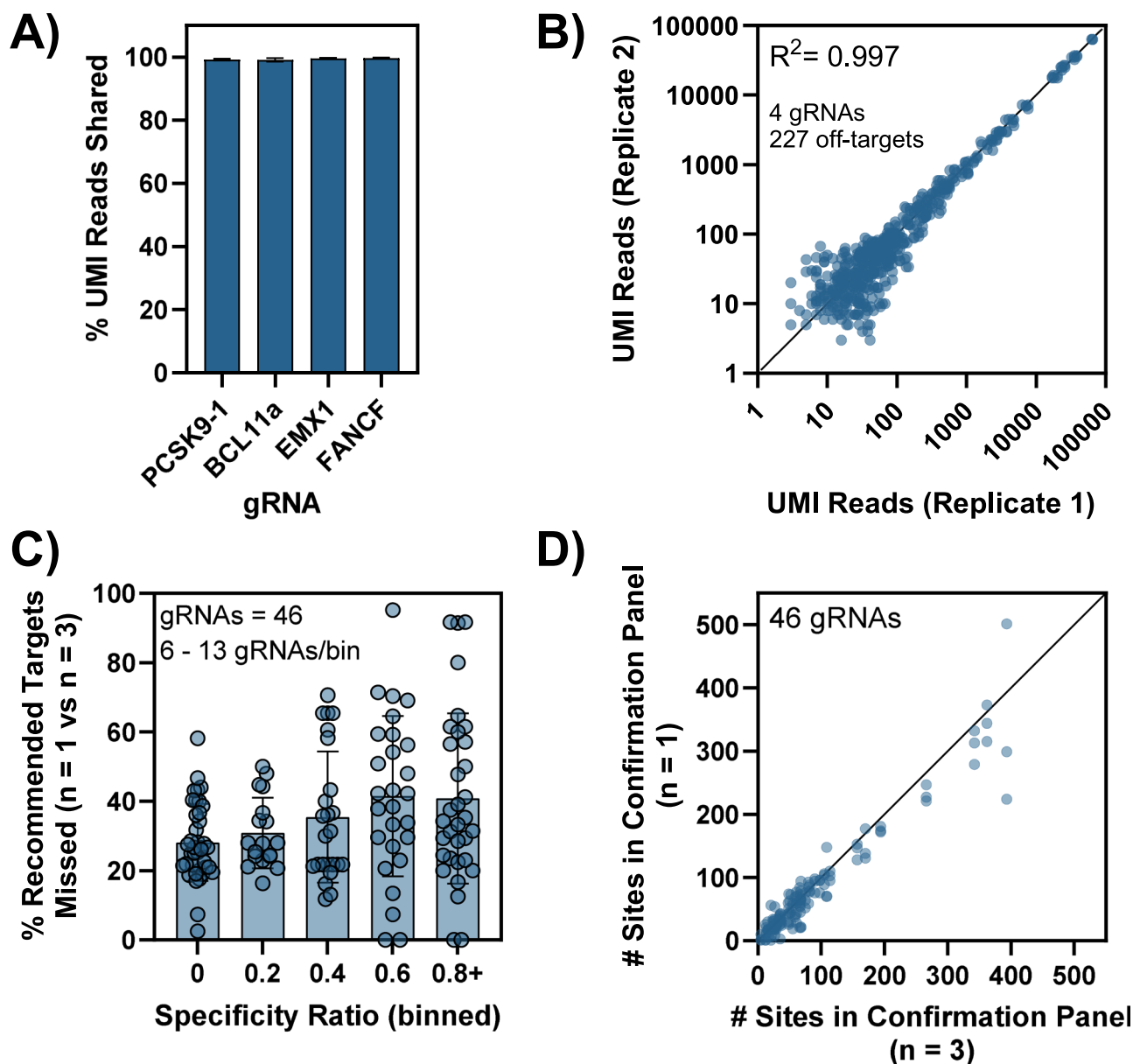


Fig. 3 | UNCOVERseq nomination reproducibility for high frequency and high priority off-targets. To measure the ability to reproducibly nominate off-targets at consistent frequencies, **A** a combinatorial comparison of replicates was performed of targets nominated (by frequency, measured as cumulative UMI reads on overlapping targets/UMI reads on all nominated targets; mean $\pm$ s.d. shown) using UNCOVERseq across 4 gRNAs in HEK293-Cas9 ($n = 3$ biological replicates per gRNA) and **B** compared the target nomination frequency between replicates. To determine the reproducibility to capture high priority off-targets (defined as Tier 1 to

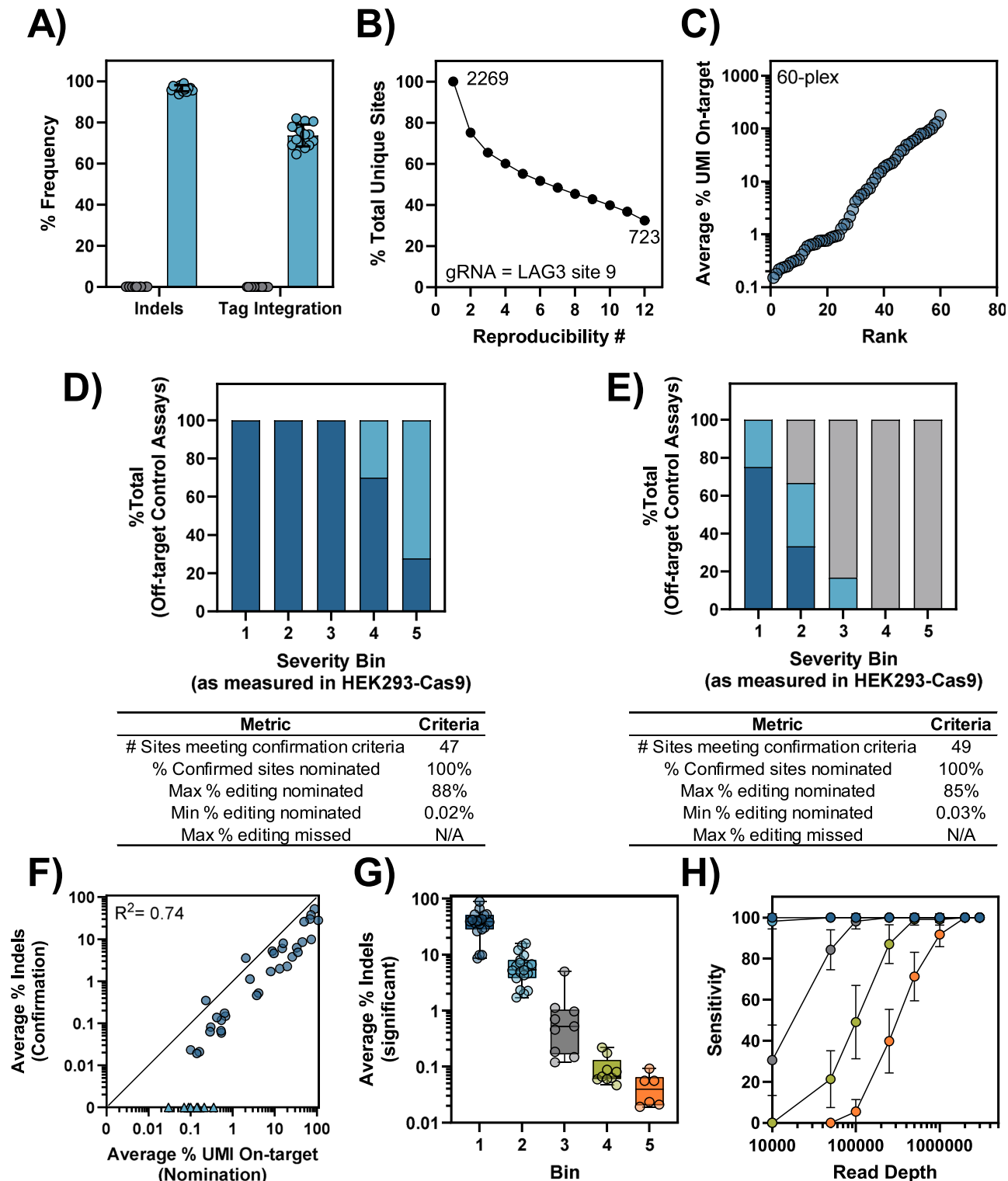
Tier 3 in Supplementary Data 2) 46 gRNAs across a broad specificity ratio (calculated as the sum of on-target reads divided by the sum of on-target and off-target reads) spectrum were nominated in biological triplicate in HEK293-Cas9 and **C** the high priority panel content missed ($n = 138$ comparisons; 3 singleton comparisons per 46 gRNAs; mean $\pm$ s.d. shown) and **D** the total number of off-targets for interrogation are shown for each individual replicate (all nominated sites) compared to the total # of high priority sites (Tier 1 to Tier 3) nominated as a biological triplicate.

biological triplicates (Tier 1–3) versus all off-targets of a single replicate. We observe that even though replicates lead to more off-targets nominated, the number of off-targets for interrogation are roughly equivalent when using tiering criteria (Fig. 3D). This provides evidence that tiering off-targets in replicate nominations leads to more impactful sites being nominated and recommended for interrogation without significantly increasing the number of off-targets recommended.

Determination of UNCOVERseq sensitivity and process controls

To characterize the sensitivity of UNCOVERseq and create routine process controls, we extensively characterized the off-targets of a

promiscuous gRNA for use as a positive control. Using a promiscuous *LAG3*-targeting gRNA (*LAG3* site 9), we nominated off-targets in 12 biological replicates using HEK293-Cas9. All replicates showed high editing (>90%) with tag integration frequencies between 64% and 81% (Fig. 4A). A total of 2269 unique nominated sites were identified, with 723 consistently reproduced across all replicates (Fig. 4B; Supplementary Data 3). The relative rate at which off-targets could not be reproduced between each subsequent replicate rapidly dropped below <5% after 3 replicates and decreased linearly afterwards (Fig. 4B). Furthermore, we observed that as reproducibility decreased, so did the nomination frequency of the site (Supplementary Fig. 3A). Although only off-targets reproduced in all replicates were chosen for



downstream confirmation, the consistent decrease in nomination frequency as reproducibility decreases suggests that putative off-targets that were harder to reproduce due to low editing frequencies and random sampling differences may be lost from high replication requirements.

To subset off-targets into different groups, we developed five Severity Bins based on the nomination frequency of the off-target relative to the identified on-target event (Supplementary Data 4). Five bins were chosen to attempt to capture a dynamic range of frequencies spanning at least 3 orders of magnitude of editing that we

hypothesized could simply represent high off-target editing (Bin 1), moderate-high (Bin 2), moderate (Bin 3), moderate-low (Bin 4) and low (Bin 5) off-target editing. These were as follows: Severity Bin 1, >50%; Severity Bin 2, 10–50%; Severity Bin 3, 1–10%; Severity Bin 4, 0.5–1%; Severity Bin 5, <0.5%. To monitor editing frequencies, we selected 12 sites from each frequency bin (Severity Bin), creating a subset of 60 sites for sequencing and confirmation (Fig. 4C; Supplementary Fig. 3; Supplementary Data 5). Twelve sites were selected per Severity Bin to ensure equal representation of the full dynamic range of nomination frequencies with sub-10% metric resolution (i.e.,

Fig. 4 | Quality control procedures for confirming assay sensitivity and read requirements. To create a positive control for UNCOVERseq process quality control, we extensively characterized a promiscuous *LAG3* (site 9) gRNA in HEK293-Cas9 using 12 biological replicate transfections with paired controls (no gRNA). Following transfection, **A** indel frequency and tag integration frequencies were measured at the *LAG3* on-target site via NGS (mean $\pm$ s.d shown) for each treatment (blue) and control (gray), and **B** the frequency of unique sites relative to the cumulative total reaching each reproducibility frequency is shown. **C** Twelve sites were selected per frequency bin (Severity Bin 5 = <0.49%; Severity Bin 4 = 0.50–0.99%; Severity Bin 3 = 1–9.9%; Severity Bin 2 = 10–49.9%; Severity Bin 1 = >50% UMI reads relative to the on-target) for routine targeted sequencing with average nomination frequency shown. To test application of this, paired confirmation was performed in **D** a highly promiscuous condition (HEK293-Cas9; $n = 3$ biological replicates) and **E** high-specificity condition (K562 nucleofected SpyFi

RNP; $n = 3$ biological replicates) and the status (Nominated + Confirmed = dark blue; Nominated + Not Confirmed = light blue; Not Nominated + Confirmed = orange; Not Nominated + Not Confirmed = gray) of each of the 60 measured sites meeting coverage criteria (>1000 $\times$) recorded. **F** Quantification of HEK293-Cas9 confirmation and nomination frequencies is plotted with final status shown as either Nominated and Confirmed (blue circle), Nominated and Not Confirmed (light blue triangle), or Not Nominated and Confirmed (orange square) with **G** confirmed indel frequencies per bin shown as box plots (min-max with Q3, median and Q1; all raw data points). **H** Downsampling of the HEK293-Cas9 *LAG3* nomination samples ($n = 12$ biological replicates) was performed and sensitivity per Severity Bin (Bin 1 = dark blue, Bin 2 = light blue; Bin 3 = gray; Bin 4 = green; Bin 5 = orange) calculated for recovering all 60 *LAG3* positive control confirmation loci (mean $\pm$ s.d shown).

sensitivity) per bin, while still being able to maintain ultra-deep sequencing on common sequencing instruments. Following selection, it can be seen that the frequency of selected sites for interrogation remains comparable to the distribution of the frequencies of the full 723 off-targets in each frequency bin (Supplementary Fig. 3D).

To test the functionality of this positive control panel in variable conditions, cells were edited with the *LAG3* site 9 gRNA in a promiscuous condition (HEK293-Cas9) and high-fidelity condition (K562 nucleofected with SpyFi RNP), and gDNA sequenced to confirm significant editing detectable across the dynamic range of sites. Sequencing yielded median read depths of 21,174 $\times$ (HEK293-Cas9), and 31,197 $\times$ (K562), with between 80% and 85% of sites above 1000 $\times$ coverage and 63–69% of sites above 10,000 $\times$ coverage (Supplementary Fig. 4). Confirmation of editing for these sites in HEK293-Cas9 showed significant indels ranging from 88 to 0.02%, approaching the limit for detection (0.01%) (Fig. 4D; Fig. 4F). Of the interrogated panel, 100% of sites in Severity Bin 1 to Severity Bin 3 had confirmed editing (Fig. 4D). 72.3% of sites in the <0.5% nomination frequency bin (Severity Bin 5) and 30% in the 0.5–1% nomination frequency bin (Severity Bin 4) could not be confirmed down to 0.01% indels (Fig. 4D). This linear decrease suggests that UNCOVERseq nominates sites with frequencies below 0.01% indels (Fig. 4D). This may be due in part to the higher genomic DNA input in UNCOVERseq (500 ng) as compared to the subsequent confirmation library preparation (100 ng). In high specificity conditions (SpyFi Cas9 + ribonucleoprotein nucleofection), fewer off-targets were edited, but a similar dynamic range of confirmed editing was retained (0.03–85% indels), with 100% of sites with confirmed significant editing successfully nominated (Fig. 4E). This indicates the approach can be used as a process control to ensure high sensitivity with reportable metrics. UNCOVERseq nomination frequencies were highly correlated ($R^2 = 0.74$) with confirmation frequencies (Fig. 4F). The high true positive rate and consistent decrease in confirmation success in bins approaching the detection limit suggest these sites may still represent true positives below detectable frequencies.

Determination of UNCOVERseq input and sequencing requirements

The number of genomes in an amplification reaction and the number of sequencing reads allocated are key limiters for NGS assay performance. To maximize sensitivity, all UNCOVERseq experiments used ~150,000 genome equivalents. While gDNA input could potentially be increased, we rationalize that this amount of gDNA is attainable by most experimental conditions and represents the ability to potentially detect down to ~0.001%, which is below the limit of detection for any currently published confirmation techniques for CRISPR gene editing.

To characterize read depth requirements for reproducible, sensitive off-target nomination using UNCOVERseq, total reads were downsampled from the previous *LAG3* site 9 UNCOVERseq dataset ($n = 12$) to frequencies ranging from 3 million to 10,000 reads per sample. Prior to downsampling, editing across confirmable sites

showed interquartile range (IQR) frequencies for each Severity Bin as follows: Severity Bin 1, 26–48%; Severity Bin 2, 2.6–7.6%; Severity Bin 3, 0.13–0.82%; Severity Bin 4, 0.06–0.13%; and Severity Bin 5, 0.01–0.02% (Fig. 4G). Thus, this dataset represents the full dynamic range of what can currently be detected by confirmation (>0.01% indels). High frequency sites in Severity Bin 1 and Severity Bin 2 were nominated with 100% sensitivity using 50,000 reads per sample (Fig. 4H). To nominate down to 0.01–0.02% indel frequencies with 100% sensitivity, at least 2 million reads per sample were required (Fig. 4H). Given the tendency of HEK293-Cas9 cells to over-nominate off-targets, we recommend performing UNCOVERseq with >500,000 reads per sample to aim for >50% analytical sensitivity at Severity Bin 5, and >2 million reads per sample for maximum sensitivity (~100% analytical sensitivity in all Severity Bins) in assessing candidate off-target sites (Fig. 4H). It is possible that read depth requirements may vary with off-target number. However, by using a promiscuous gRNA (specificity ratio = 0.013) to determine this value, we propose that this represents the number of reads to successfully nominate sites using UNCOVERseq even with gRNAs with very poor specificity.

Comparative nomination analysis of UNCOVERseq to other methods

A comparative analysis of UNCOVERseq to peer-reviewed accounts of other nominations methods was performed to better understand how the sensitivity and nomination frequencies of diverse methods compare to UNCOVERseq. Due to variable operational conditions, precision, and total nomination list sizes reported of different methods, we postulate that sensitivity is most appropriately measured using either off-targets with confirmed editing or off-targets from methods with high precision. Interrogation of the 60 *LAG3* site 9 gRNA off-targets compared to CHANGE-seq⁷ and GUIDE-seq⁷ nominations demonstrated that both methods could nominate the most frequent group of confirmed off-targets (Severity Bin 1) with 91–100% sensitivity, but sensitivity rapidly decreased in the lower frequency off-target bins. CHANGE-seq was demonstrated to have a sensitivity between 66% and 75% for recovering Severity Bin 3–5 off-targets, while GUIDE-seq had a linear decrease from 16 to 0% for these same Severity Bins of confirmable off-targets (Supplementary Fig. 5). Previous analysis of the 60 *LAG3* site 9 off-targets demonstrated that 100% of sites in Severity Bin 1–3 represented confirmed off-targets, demonstrating that published accounts of both GUIDE-seq and CHANGE-seq lack the ability to sensitively detect all off-targets (Fig. 4; Supplementary Fig. 5).

Random sampling of the *LAG3* site 9 dataset with 100% reproducibility showed UNCOVERseq-nominated sites had high precision, with confirmation frequencies correlating to average nomination frequencies down to the limit of detection (Fig. 4). Using this logic, we postulate the full 723 sites in this fully reproducible set are also likely to represent true positives. Investigation of sensitivity and frequencies of previous accounts of CHANGE-seq and GUIDE-seq for this gRNA yielded similar trends in sensitivity per Severity Bin, further supporting

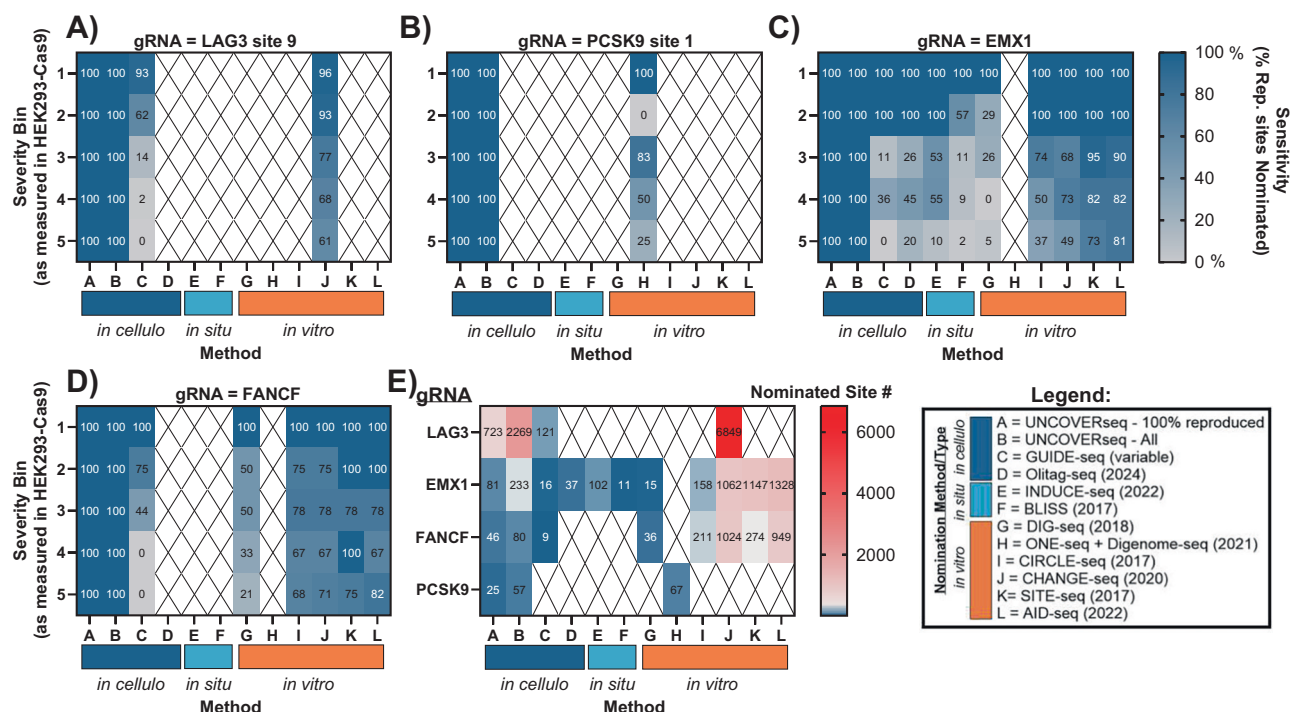


Fig. 5 | Comparative analysis of UNCOVERseq to other nomination technologies. Comparison of the sensitivity of other published accounts of nomination technologies to nominate 100% reproducible UNCOVERseq off-target sites using **A** LAG3 site 9 ($n = 12$ nomination replicates) **B** PCSK9 site 1 ($n = 6$ nomination replicates) **C** EMX1 ($n = 6$ nomination replicates) and **D** FANCF ($n = 6$ nomination replicates). **E** Total number of off-targets nominated per nomination method per gRNA with color gradient shifts at the number of off-targets easily interrogated

using amplicon sequencing (<250 targets; blue/gray) vs those more challenging to interrogate using amplicon sequencing (>250; gray/red). Legend displays the corresponding method to each column in the heatmaps along with the corresponding method type (in cellulo, in situ, in vitro). Rep. sites: Reproducible UNCOVERseq sites between all replicates. UNCOVERseq–All represents all Tier 1 to Tier 3 sites from the experiments.

this (Fig. 5A). This provides additional evidence that reproducible UNCOVERseq sites have high likelihood of being true positives and may serve as an appropriate proxy for measuring sensitivity of different methods.

Using our previous finding that highly reproduced UNCOVERseq off-targets are likely indicative of true positives (Fig. 4B), we compared fully reproduced off-targets from 6 biological replicate UNCOVERseq experiments for three gRNAs (*EMX1*, *FANCF*, and *PCSK9*) to previous accounts of off-targets from GUIDE-seq⁵, INDUCE-seq¹⁰, OliTag-seq¹⁸, CHANGE-seq⁷, CIRCLE-seq⁶, AID-seq³², BLISS³³, ONE-seq³⁴, DIG-seq³⁵, Digenome-seq³⁴, and SITE-seq⁸ nomination methods (Supplementary Data 6). For *PCSK9*, only a combined set of ONE-seq and Digenome-seq sites was available, and nomination sensitivity was found to fluctuate from 0 to 83% through Severity Bins 2–5 (Fig. 5B). For *EMX1*, DIG-seq and BLISS sensitivity decreased to 29% and 57% after Severity Bin 1, respectively, and sharply decreased at lower frequency bins. Cell-based methods GUIDE-seq, OliTag-seq and INDUCE-seq sensitivity dropped after Bin 2 (11–53% sensitivity), ranging from 0 to 55% for off-target nomination frequencies observed after Severity Bin 2 (Fig. 5C). However, nomination frequencies for GUIDE-seq, OliTag-seq and INDUCE-seq correlated well with expected frequencies ($R^2 = 0.52–0.70$), if identified (Supplementary Fig. 5). Remaining biochemical in vitro nomination methods had the highest sensitivity compared to UNCOVERseq, with approximate rank-order as follows: AID-seq/SITE-seq > CHANGE-seq > CIRCLE-seq (Fig. 5C). Sensitivity dropped for all these methods at Severity Bin 3, and ranged from 37 to 95% for Severity Bin 3–5 (Fig. 5C). For *FANCF*, GUIDE-seq and DIG-seq sensitivity quickly decreased below 50% after Severity Bin 2 to a range from 0 to 21% below Severity Bin 3 (Fig. 5D). AID-seq/SITE-seq achieved higher sensitivity compared to CIRCLE-seq/CHANGE-seq, similar to previous observations with *EMX1* (Fig. 5C, D). CHANGE-seq and

CIRCLE-seq sensitivity dropped to a range of 67–78% after Severity Bin 1 (Fig. 5D). SITE-seq and AID-seq sensitivity dropped after Severity Bin 2 and ranged from 75 to 100% (Fig. 5D).

Each method nominated a widely variable number of sites per gRNA, which affects the calculation of analytical sensitivity and precision. The number of nominated sites ranged from 9 for *FANCF* using GUIDE-seq to 6,849 off-targets using *LAG3* site 9 with CHANGE-seq (Fig. 5E). While biochemical in vitro methods like CHANGE-seq, AID-seq, and SITE-seq showed the greatest comparative sensitivity, they also had the largest nominated site list, ranging from 291 to 1328 sites for the *FANCF* and *EMX1* gRNAs shared between most methods (Fig. 5E). Investigation of method normalized nomination frequencies demonstrate that in vitro methods have the poorest correlation to UNCOVERseq derived frequencies ($R^2 = 0.01–0.26$), with AID-seq and SITE-seq having the lowest correlations (Supplementary Fig. 5). Nomination frequencies derived from cell-based in cellulo and in situ methods (GUIDE-seq, OliTag-seq, INDUCE-seq, BLISS) better correlate to UNCOVERseq nomination frequencies ($R^2 = 0.31–0.70$), which previously were shown to correlate to observed confirmation frequencies (Supplementary Fig. 5; Fig. 4F; Fig. 5). These findings demonstrate that UNCOVERseq improves upon the sensitivity of existing in cellulo methods such as GUIDE-seq, in addition to subsequent improved methodologies such as OliTag-seq. Our findings also demonstrate that in vitro methods are not inherently more sensitive than in cellulo methods for discovering true off-targets, and UNCOVERseq nominates likely confirmable off-targets not detected in other methods.

Comparative confirmation analysis of UNCOVERseq to other methods

To characterize the sensitivity and precision of UNCOVERseq more comprehensively relative to other nomination technologies, we

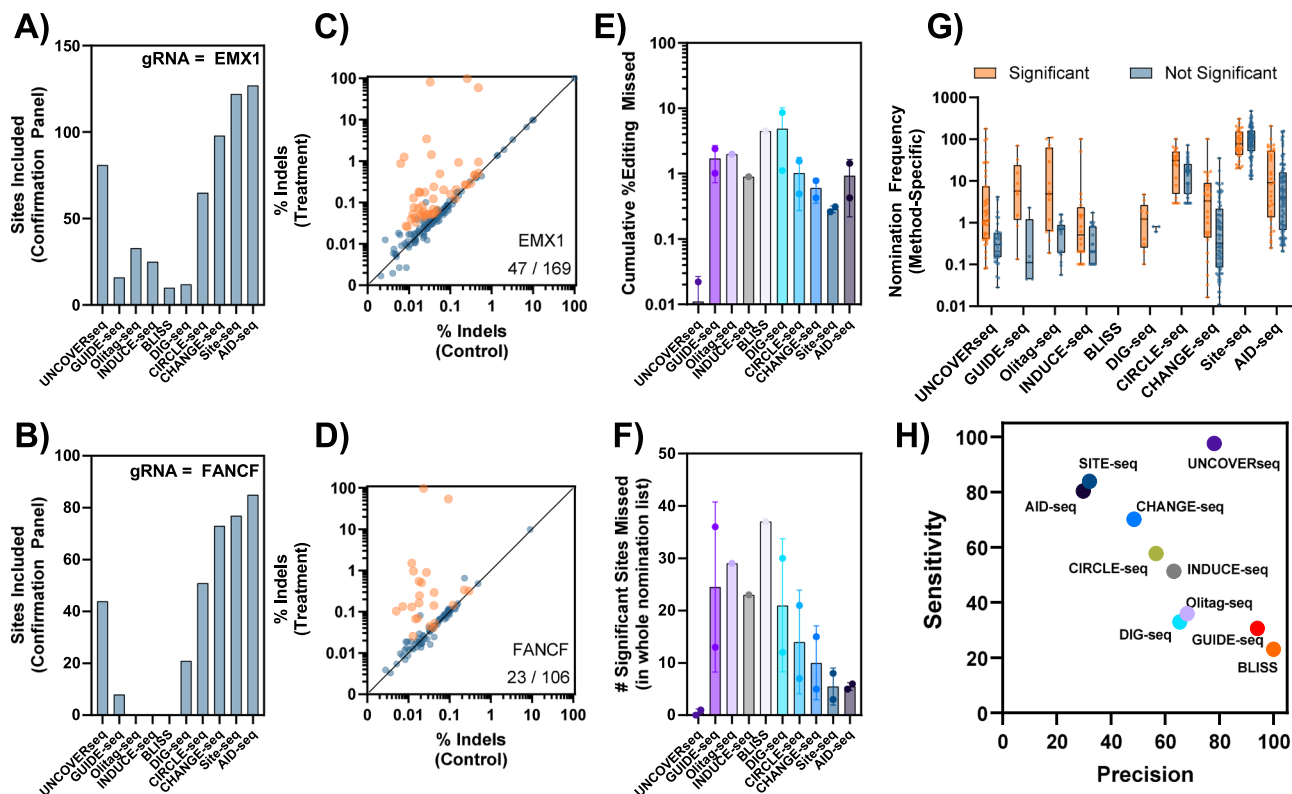


Fig. 6 | Confirmation of editing at off-targets nominated by variable off-target nomination technologies in HEK293-Cas9. **A** To compare the analytical sensitivity and precision of other published accounts of nomination technologies to UNCOVERseq off-target sites were selected across the dynamic range of each nomination technology for two gRNA with the total number of off-targets selected from each nomination methodology quantified for both EMX1 and **B** FANCF. **C** Confirmation of nominated off-targets was performed in HEK293-Cas9 ($n = 3$; paired treatment and control) to ensure true-positives were more easily detected. Editing quantification at assays with $>1000\times$ coverage. Each dot represents a gRNA on/off-target with the average raw frequency of indels of the control (x-axis) and treatment (y-axis) plotted. Blue dots indicate sites with no statistical significance

while orange dots indicate significant sites ($p \text{ adj} < 0.05$) for EMX1 and **D** FANCF. **E** After confirmation, the cumulative frequency of significant off-target events not nominated per method for both gRNAs ($n = 2$) was quantified in addition to **F** the total number of significant off-target sites missed. **G** Method specific nomination frequency was quantified for both confirmed significant and confirmed not significant sites for the full off-target list ($n = 2$ gRNAs). **H** Method-specific analytical sensitivity ($\# \text{ nominated true positives} / \# \text{ confirmed true positives}$) and precision ($\# \text{ confirmed true positives} / \# \text{ confirmed true positives} + \text{false positives}$) was calculated for each method across both gRNAs with each point displaying the average across the two gRNAs considering all 5 Severity Bins as valid for precision calculations.

conducted targeted editing confirmation experiments in the promiscuous HEK293-Cas9 cell system using the two most widely implemented benchmarking gRNA loci in our set, EMX1 and FANCF (Fig. 6). This system was selected because stable Cas9 expression was previously shown to yield highly permissive, low specificity conditions, thereby maximizing the likelihood of correctly detecting confirmable off-target editing (i.e., true positives) that can be edited in biological contexts. To select nomination sites across technologies, nomination frequencies were first method normalized to the on-target value and these frequencies were used to sample sites across the full spectrum of predicted editing probabilities using the previously derived Severity Bins of nomination (Fig. 6A, B; Supplementary Fig. 7). Each nomination method had sites selected from all available Severity Bins per method while prioritizing top frequency sites across methods, for a total of 297 putative off-target loci for EMX1 and 181 putative off-target loci for FANCF (Supplementary Data 7). After intersecting and consolidating nominated sites common to different selections per method, the final confirmation panels comprised of 169 putative off-targets for EMX1 and 106 putative off-targets for FANCF, with variable numbers of sites per nomination methodology (Fig. 6A, B; Supplementary Data 7). Following consolidation of common off-targets across methods, all methods maintained relatively identical nomination frequency distributions to the original selection, with the exception of CIRCLE-seq, CHANGE-seq, SITE-seq, and AID-seq, which had a lower median off-

target nomination frequency following consolidation due to contribution of sites nominated originally from other methods random selections (Supplementary Fig. 7). This indicates that sites interrogated in lower Severity Bins for these in vitro methods were enriched with sites predicted with more precise methods, which may ultimately inflate precision calculations for these methods (Fig. 6; Supplementary Fig. 7; Supplementary Fig. 8). Each nomination method had off-targets tested across all Severity Bins that the respective nomination technology produced targets within, with between 1 to 92 targets tested per nomination Severity Bin per method (Supplementary Fig. 8A, B).

To benchmark the performance of each nomination technology, edited HEK293-Cas9 cells were sequenced and significantly edited sites were compared with the complete nominated site lists generated by each method (Fig. 6E, F). Sequencing of Cas9 edited cells yielded median read depth per assay of 139,208x for EMX1 and 132,620x for FANCF, with more than 95% of assays surpassing 1,000x coverage for both gRNAs (Supplementary Fig. 8). Across all assays, 47 significantly edited off-target sites were identified for EMX1 and 23 for FANCF, with confirmed indel frequencies ranging from 0.03% to 98.27% (Fig. 6C, D; Supplementary Fig. 7). Across the two gRNAs, UNCOVERseq successfully nominated 47 of 47 (100%) of EMX1 confirmed sites and 22 of 23 (95%) of FANCF sites, while missing a single 0.02% editing site (Fig. 6E, F; Supplementary Fig. 8). In contrast, other in cellulo and in situ assays (GUIDE-seq, OliTag-seq, INDUCE-seq, BLISS) each failed to detect

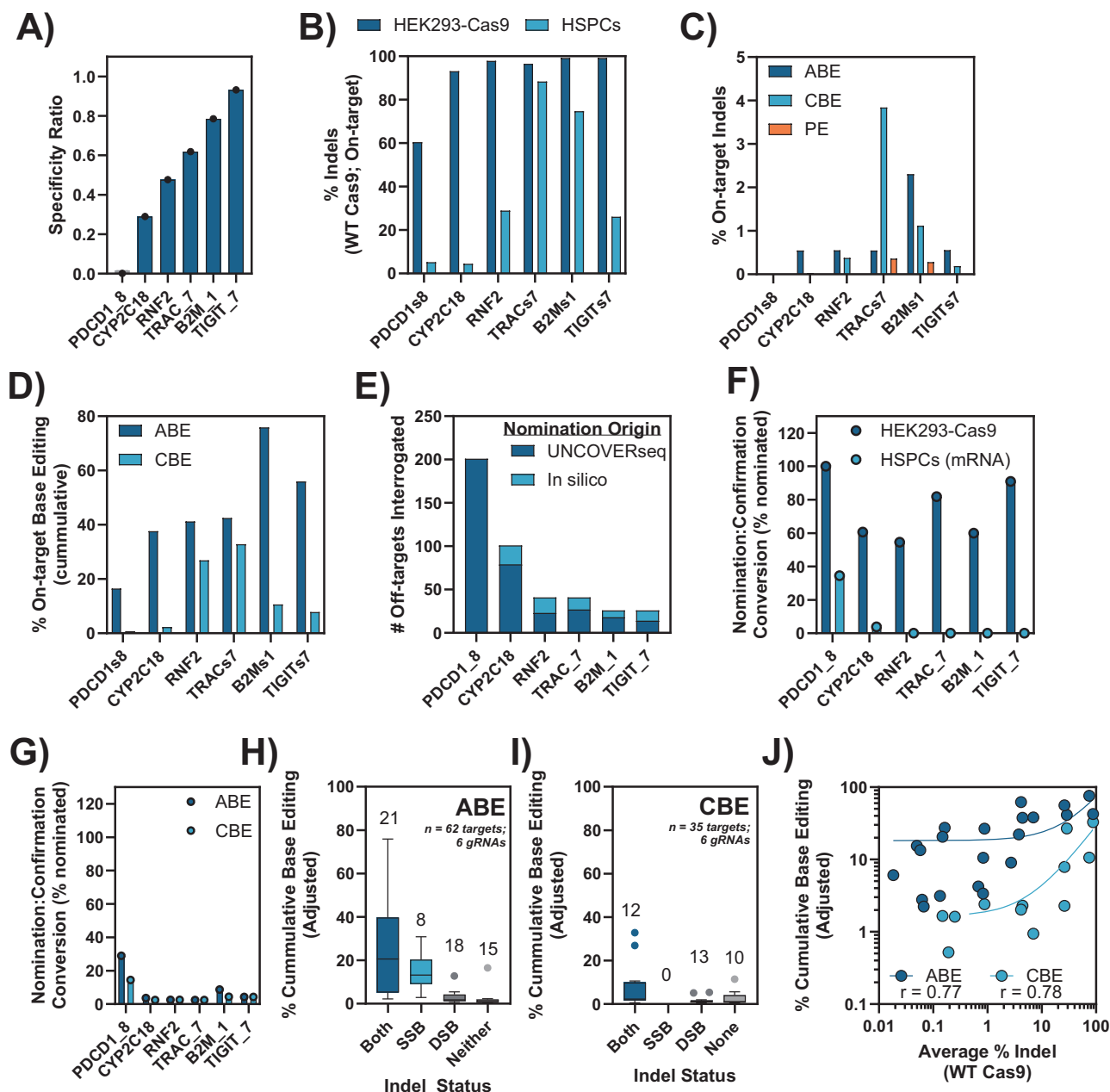


Fig. 7 | On-target and off-target editing in HEK293-Cas9 and HSPCs. **A** Six gRNAs with a single targeted ABE or CBE base were selected with specificity ratios shown. HEK293-Cas9 ($n = 1$) and HSPCs ($n = 3$ donors), with simultaneous delivery of the corresponding HiFi Cas9 mRNA, were delivered each gRNA and had **B** on-target indel editing quantified by NGS. HSPCs were also delivered mRNA with either the *S.p.* Cas9-ABE8 or *S.p.* Cas9-CBE fusion for each gRNA and had **C** indel editing and **D** base editing quantified using NGS. **E** Multiplexed amplicon sequencing (rhAmpSeq) panels were created for each gRNA for off-target quantification based on origin from UNCOVERseq or in silico nomination (in silico) with the number of targets interrogated per gRNA shown. After sequencing/confirmation of off-targets

in all conditions, the frequency that UNCOVERseq nominated sites converted to true positives was measured for **F** *S.p.* Cas9 (HEK293-Cas9; $n = 1$ paired treatment/control) or HiFi Cas9 ($n = 3$ paired biological paired treatment/control) indel editing and **G** *S.p.* Cas9 base editing conditions ($n = 3$ paired biological paired treatment/control). Confirmed base editing sites were categorized by their respective indel confirmation status (DSB = wildtype *S.p.* Cas9; SSB = ABE/CBE nickase) with cumulative base editing as Tukey boxplots for **H** ABE and **I** CBE editing conditions. **J** Frequencies were plotted for all confirmed ABE/CBE sites with both SSB and DSB confirmation with the respective DSB indel frequency and Spearman r calculated. Line indicates linear regression for ABE (dark blue) and CBE (light blue).

between 14 and 37 significantly edited sites per gRNA, with cumulative indel frequencies ranging from 0.90 to 4.51% and single missed editing events as high as 1.25% indel editing (Fig. 6E, Supplementary Fig. 8). Of in vitro methods, DIG-seq was an outlier as the lowest performing nomination methodology, with 14–37 significantly edited off-targets missed per gRNA. DIG-seq also missed cumulative editing frequencies ranging from 1.1 to 8.5% and single editing events as high as 3.4% indel editing (Fig. 6E, F; Supplementary Fig. 8). Comparatively, other in vitro methods (CIRCLE-seq, CHANGE-seq, SITE-seq, AID-seq) performed

better, with missed sites ranging from 7 to 21 per gRNA with missed cumulative editing frequencies ranging from 0.26 to 1.55% (Fig. 6E, F). Among in vitro methods, single missed editing events were as high as 1.25% indel editing, as observed by a missed site by AID-seq (Fig. 6E, F; Supplementary Fig. 8).

Using this data, the empirically derived analytical sensitivity and precision was calculated for each nomination methodology over two gRNAs. Of the nomination methods investigated, UNCOVERseq (97.6%), SITE-seq (83.8%), and AID-seq (80.4%) were the most sensitive

nomination methods (Fig. 6F). Conversely, BLISS (23.1%), DIG-seq (33.0%), and GUIDE-seq (35.9%) were the least sensitive nomination methods (Fig. 6F). Precision of each method also demonstrated that UNCOVERseq maintains high precision (78%) along with other high precision in cellulo and in situ methods such as BLISS (100%), GUIDE-seq (94%) and OliTag-seq (68%) (Fig. 6F). As was observed previously, precision of these methods decreased as the nomination frequency decreased, suggesting that confirmation sensitivity may be a limitation in detecting some of these sites as true positives (Supplementary Fig. 9). The in vitro methods demonstrated the lowest overall precision, with AID-seq (29.7%), SITE-seq (32.1%) and CHANGE-seq (48.4%) performing most poorly. Since it was previously observed that multiple low frequency sites were supplemented to these in vitro methods from nominations of other higher precision methods (and not all in vitro nominated sites were tested), it is also likely that this derived precision value is actually an overestimate of the true value (Supplementary Fig. 7; Supplementary Fig. 9). Each nomination method generally had precision increase as nomination frequency increased, further confirming previous linear correlations between nomination frequency and editing activity (Supplementary Fig. 7C, D; Supplementary Fig. 9). These analyses derive comparative empirical performance metrics between nomination methods and demonstrate markedly higher combined sensitivity and precision of UNCOVERseq in nominating bona fide off-target editing events in comparison to other methods.

Screening gRNAs of variable specificity

To identify the specificity of a broad set of gRNAs for future experimental design, 192 gRNAs were selected and UNCOVERseq performed in HEK293-Cas9 cells. Samples were sequenced to a median of 1.8 million reads, in line with our previous recommendations for maximizing sensitivity (Supplementary Fig. 11A; Supplementary Data 8). Following this, we calculated a specificity ratio for each gRNA, defined as the sum of reads at the on-target locus divided by the sum of reads at the on-target and off-target locations⁷. We rationalized that only a single replicate per gRNA was needed for this experimentation since we were most interested in higher frequency off-targets that would be recovered from a single replicate in the promiscuous system, as previously demonstrated (Fig. 3A). A range of specificity ratios per gRNA were recovered ranging from 0.00063 to 1.0 (Supplementary Fig. 11B). Comparison to previously published CHANGE-seq specificity ratios of overlapping gRNAs demonstrated a high rank-order correlation ($r=0.63$) between specificity ratios derived from both methods, though sometimes large disagreements in specificity ratios for higher specificity gRNAs occurred (Supplementary Fig. 11C). We then subset gRNAs by specificity ratios (binned in increments of 0.2) and obtained a relatively uniform distribution with each bin containing a range between 22 and 47 gRNAs each (Supplementary Fig. 11D, E). We further subset these gRNAs by those that were ABE- and CBE-compatible as defined as having both an “A” in the +4 to +7 positions (5′ to 3′) and a “C” in the +4 to +8 positions (5′ to 3′). This resulted in a less uniform distribution of gRNAs across the specificity spectrum with a range of 8 to 21 gRNAs per specificity ratio bin (Supplementary Fig. 11D, E). From this we selected six gRNAs for further experimentation as representing a continuous range from 0 to 1 supporting all editor modalities: *PDCDI* site 8, *CYP2C18*, *RNF2*, *TRAC* site 7, *B2M* site 1 and *TIGIT* site 7 (Fig. 7A).

Comparative analysis of editors in HEK293-Cas9 and HSPCs (on-target)

Next, we sought to determine the translation of UNCOVERseq off-targets across a broad range of specificities in a translational ex vivo system (HSPCs with mRNA editor nucleofection) across different editing modalities (DSB Cas9, SSB Base Editors, SSB Prime Editors). To do this, we edited HSPCs with one of six gRNAs along with mRNA of either (a) a high fidelity *S.p.* Cas9, variant (HiFi Cas9) (b) *S.p.* Cas9

(D10A) fused to a Cytosine Base Editor (AncBE4max³⁶), (c) *S.p.* Cas9 (D10A) fused to Adenine Base Editor (ABE8e³⁷) and (d), *S.p.* Cas9 (H840A) fused to the PE2 system²⁸ with a pegRNA intended to introduce a single SNP. HEK293-Cas9 was also edited in parallel with the wildtype *S.p.* Cas9 nuclease. Evaluation of on-target *S.p.* Cas9 editing found that editing in HEK293-Cas9 was highly efficient at all sites, ranging from 60.4 to 99.4% indel editing, but with a trend of decreased frequencies at lower specificity gRNAs (Fig. 7B). Evaluation of on-target indel editing in HSPCs had a range of 4.4–88.4% indel editing for the DSB-based HiFi Cas9 editor, 0.0–2.3% indel editing for ABE editors, and 0.0–3.8% indel editing for CBE editors, and 0.0–0.36% indel editing for Prime Editors (Fig. 7C). Intended on-target cumulative base editing ranged from 16.5 to 75.9% for ABEs, and from 0.78 to 32.9% for CBEs (Fig. 7D). No significant base editing was observed for either the DSB-based *S.p.* Cas9 editors or SSB prime editor, as would be expected (Source Data). Only very low total editing was observed for any prime editing conditions at any sites (Fig. 7C; Source Data). For this reason, we excluded further evaluation of the prime editor. Since we show evidence of on-target indel editing with the PE construct, we hypothesize the lack of intended activity is due to the need for substantial pegRNA optimization to achieve successful prime editing, as has been previously reported³⁸. For all editors in HSPCs, similar trends were observed to HEK293-Cas9 that the lower specificity gRNAs *PDCDI* site 8 and *CYP2C18* had trends with decreased editing, suggesting that these sites may be overall less effective at the on-target, potentially due to competing off-targets (Fig. 7). We conclude from this that on-target editing was successful in all conditions except prime editing, with highly variable frequencies as may be expected without substantial optimization.

Comparative analysis of DSB editors in HEK293-Cas9 and HSPCs (off-target)

A range of nominated off-targets for the six gRNAs were selected from two orthogonal methods for downstream confirmation: UNCOVERseq nominations and in silico nominations. A range of 26–201 putative editing sites were interrogated per multiplexed amplicon rhAmpSeq panel with an UNCOVERseq: in silico nomination origin split ranging from 53.8 to 100% for the interrogated target lists (Fig. 7E). Panels were sequenced with a goal of at least reaching 1000× coverage per off-target to ensure adequate sensitivity for calling significantly edited off-targets. Sequencing the six panels demonstrated we were able to achieve a median read depth per assay ranging between 34,000× and 76,000×, with consistency in coverage between edited and corresponding control samples (Supplementary Fig. 12A). The number of off-targets reaching sufficient coverage (>1000×) ranged from 92 to 100% per panel, with a range of 0–15 targets not being at sufficient coverage per panel (Supplementary Fig. 12B).

To determine the frequency of UNCOVERseq HEK293-Cas9 nominations that convert to empirically edited sites in variable DSB editing contexts, we compared this frequency for *S.p.* Cas9 both in HEK293-Cas9 and HSPCs (HiFi Cas9). For HEK293-Cas9, a range of 54.5% of nominated off-targets all the way to 100% of off-targets had confirmed editing ranging from 0.02 to 95% indel editing, demonstrating the true positive rate for UNCOVERseq-nominated sites remains high even with only a single replicate in the appropriately paired confirmation context (Fig. 7F; Supplementary Fig. 13A–F). For HSPCs, a range of 2.4–34.5% of nominated targets were successfully confirmed per gRNA, with confirmed indel editing ranging from 0.06 to 88% (Fig. 7F; Supplementary Fig. 14A–F). Nomination: confirmation frequencies trended to increase as gRNA specificity decreased, suggesting that the method is still successfully nominating relevant off-targets, but that these sites likely no longer exceed detectable frequencies or are no longer edited in the higher genome editing specificity context of HSPCs delivered a HiFi Cas9 mRNA editor (Fig. 7F). Furthermore, at higher gRNA specificities the only nominated target

being confirmed is the on-target site for HiFi Cas9 in HSPCs (Supplementary Fig. 14A–F).

Off-targets that were confirmed were compared to the list of those that would have been dropped given a different previously evaluated alignment method (Regex method; Fig. 1). A range of 3–23 bona fide off-targets per gRNA in HEK293-Cas9 were successfully nominated using the UNCOVERseq alignment method that were missed using GUIDE-seq analysis Regex method, with a range of observed indel editing from 0.02 to 68% (Supplementary Fig. 15A). In HSPCs, 1 bona fide HiFi Cas9 off-target was identified with a frequency of 0.01% with our alignment method that was missed using the Regex alignment method (Supplementary Fig. 15B). This demonstrates that DSB off-target sites that were nominated due to differences in alignment criteria can result in bona fide off-target indel editing in both HEK293-Cas9 and HSPCs, further emphasizing the importance of the improvements implemented in UNCOVERseq.

Comparative analysis of non-DSB editors in HSPCs (off-target)

Off-targets were simultaneously confirmed for both indel and base editing in the non-DSB treatments for HSPCs (ABE and CBE). Similarly, we interrogated the frequency that UNCOVERseq HEK293-Cas9 nominations convert to empirically edited sites in HSPCs being delivered a base editor. For ABE treatments, a range of 2.4–29% of nominated targets had confirmed editing ranging from 0.53 to 75.9% cumulative ABE editing (Fig. 7G–H; Supplementary Fig. 16A–F). For CBE treatments, a range of 2.4% of nominated targets to 14.5% of targets had confirmed editing ranging from 0.51 to 32.9% CBE editing (Fig. 7G; Fig. 7I; Supplementary Fig. 17A–F). Significant indel editing was observed for all ABE and CBE editing treatments, with largely only the on-target gRNA containing indels at higher specificity gRNAs (Fig. 7C; Supplementary Fig. 18A–F; Supplementary Fig. 19A–F). Off-target indel frequencies for ABE treatments ranged from 0.02 to 0.88% indels across different gRNAs (Supplementary Fig. 18A–F). Interestingly, three off-target sites were found to generate indel events at the higher specificity *TRAC7* gRNA under ABE treatment conditions, which lacked any significant off-targets in paired DSB Cas9 treatment (Supplementary Fig. 18D). Off-target indel frequencies for CBE treatments ranged from 0.08 to 0.66% indels across different gRNAs (Supplementary Fig. 19A–F). Generally, it was observed that indel and base editing frequencies were lower in CBE-treated samples in comparison to ABE-treated samples (Fig. 7I, J; Supplementary Fig. 16A–F; Supplementary Fig. 17A–F), although this could be a result of lower overall activity of the base editor instead of off-target propensity.

When comparing the list of confirmed ABE/CBE off-targets to those that would have been excluded given a different alignment method during nomination we found 1 bona fide off-target of the *PDCD1* gRNA that was identified for both ABE and CBE treatments with a frequency range of 0.5–3.1% base editing that was missed using the Regex alignment method (Supplementary Fig. 15B). This demonstrates that off-target sites that were nominated due to differences in alignment criteria can also result in bona fide off-target base editing activity for both ABE and CBE editors in HSPCs.

To investigate relationships between DSB indels, SSB indels, and base editing, we binned confirmed base editing off-targets based on their presence of indels in either DSB or SSB systems. Base editing with the highest frequencies (median 20.6% and 2.3% for ABE and CBE, respectively), were found to coincide with indel editing for both DSB and SSB systems (Fig. 7H). Interestingly, only ABE treatments were found to have an increased frequency of base editing at SSB only sites, with eight detected SSB-only off-targets with a median 13.1% cumulative base editing compared to zero sites for CBE (Fig. 7). This may coincide to activity differences, as both ABE editing/indel activity was generally higher than CBE editing/indel activity across the different sites (Supplementary Figs. 16–19). DSB-only and sites with no evidence of significant indel editing were present in confirmed sites for both ABE

and CBE treatments, albeit with lower median cumulative base editing frequencies (Fig. 7). The on-target indel and base editing activity of the different gRNAs were rank-order correlated ($r = 0.66–0.89$), suggesting that indel editing frequencies may be predictive of base editing frequencies (Supplementary Fig. 20). Similarly, off-target DSB indel editing frequencies from HiFi Cas9 demonstrated rank-order correlation with off-target base editing frequencies ($r = 0.77–0.78$) at sites that had significant DSB indel editing frequencies in HSPCs (Fig. 7J). This provides evidence that DSB editing may be indicative of base editing activity, meaning that DSB-nominated sites are meaningful for interrogation in the context of both indel and base editing off-target assessment for both ABE and CBE modalities.

Comparative translocation analysis and overall editing burden across editing modalities

To investigate differential frequencies of editor modalities to generate large structural variants (>0.1% frequencies) in HSPCs, we investigated the previously described six sites for on-target:off-target and off-target:off-target translocations using amplicon sequencing. Of the six interrogated gRNAs, only the *PDCD1* gRNA had detectable translocations, with two out of three of the translocations being shared between the *S.p.* Cas9 and ABE conditions (Supplementary Fig. 21). Shared translocations included a fusion of OTE132 to OTE94 and OTE160 to OTE158, with comparable average frequencies ranging between 1.0–1.7% and 0.3–0.5%, respectively (Supplementary Fig. 21). The on-target site was not a translocation partner in any significant events, which is likely due to the low frequency of on-target indels for the *PDCD1* gRNA relative to off-target indel activity (Fig. 7B, C; Supplementary Fig. 14A; Supplementary Fig. 18A). The overall estimated translocation burden for this gRNA was estimated to be an average of 1.4% translocations for *S.p.* Cas9 and an average of 2.2% translocations for ABE conditions (Supplementary Fig. 21). This suggests that translocations are either below 0.1% or not occurring in healthy HSPC donors across higher gRNA specificities. However, it is noteworthy that they are still occurring for both SSB and DSB modalities, as has been previously observed³⁹.

When calculating the normalized risk of cumulative off-target frequencies (indels, base edits, and translocations) across editor modalities throughout our six gRNAs displaying a range of specificities, off-target ratios were observed for the *PDCD1* gRNA (specificity ratio = 0.001) over a range 9.4–89.0 off-target events per 1 on-target event (Supplementary Fig. 21). Off-target ratios for the *CYP2C18* gRNA (specificity ratio = 0.290) ranged from 0.07 to 0.76 off-target events per 1 on-target event (Supplementary Fig. 21). Trends consistently showed that overall off-target burden of high-fidelity DSB editors was actually decreased in comparison to SSB base editors for the cumulative frequency of event types monitored using this strategy (Supplementary Fig. 21). Even though the *B2M* gRNA was considered higher-specificity (specificity ratio = 0.785), a single significant ABE off-target was observed for this treatment contributing to a higher ratio (Supplementary Fig. 21). This highlights that even higher-specificity gRNAs may generate observable off-targets in clinically relevant cell types, and that the off-target burden of high-fidelity DSB enzymes can actually be lower than that of current generation SSB base editing modalities in translational contexts.

Discussion

Our study presents an improved, end-to-end characterized in cellulo method for the nomination of off-target sites in CRISPR experiments that we collectively refer to as UNCOVERseq (v1.0). This method leverages several technological improvements to collectively streamline the in cellulo nomination process, improve NGS data quality, and increase the number of high-confidence nominated sites through computational analysis improvements compared to other previously published methods. By demonstrating recommended operational

conditions that can allow the experiments to be performed independent of cell context with controls grounded in empirical data, we provide a framework to ensure translation to different treatment modalities with quantifiable levels of performance from experiment to experiment. Furthermore, we demonstrate the workflow is capable of nominating relevant unique and shared off-targets for both DSB-based and SSB-based CRISPR editing systems and demonstrate correlations between DSB formation and the frequency of a site to be edited by ABE/CBE editors. Finally, we perform a thorough assessment of off-targets nominated by different methods to ground both UNCOVERseq and other methods with measured analytical sensitivity and precision to demonstrate that UNCOVERseq outperforms other available nomination methods.

To ensure all relevant off-targets are assessed, high analytical sensitivity is a critical off-target nomination metric. However, accurate calculations of false-negative rates from nomination methods have been challenged by technical difficulties in obtaining an empirically defined gold-standard of all true-positive off-targets. In this work, we create a starting point for a gold-standard dataset for measuring sensitivity by selecting common gRNAs to multiple nomination methods and performing multiplexed ultra-deep NGS sequencing to confirm off-targets nominated across diverse methods in a promiscuous editing system. This leads us to find that UNCOVERseq is the most sensitive gene editing off-target nomination method currently available (>95% sensitivity) and demonstrate that many other existing methods have sensitivity gaps leading to missed nomination of ~20% to ~80% of off-targets that can exhibit significant editing in cells. Previous work has led to a perception that *in vitro* biochemical methods are inherently more sensitive than *in cellulo* methods as evidenced by: (1) true positive sites captured by *in vitro* methods like CHANGE-seq that are missed with GUIDE-seq and (2) multiple accounts of *in cellulo* methods being largely a subset of *in vitro* results^{6,7}. While we demonstrate that previously published accounts of *in cellulo* methods were not as sensitive as UNCOVERseq, it is not clear whether this is due to insufficient operational conditions (read depth, library complexity, etc.) to maximize capabilities of many assays as opposed to the technical improvements that confer enhanced nomination capabilities to UNCOVERseq. To reduce risk of false negatives, we demonstrate a generalizable strategy for *in cellulo* off-target nomination where high gDNA input and promiscuous editing conditions are used to greatly amplify nomination signal to reproducibly detect sub-0.05% editing events in UNCOVERseq data while still retaining sites derived from higher fidelity modalities and primary cell lines. This finding is in agreement with previous comparative work of *in cellulo* tag-based methods in HSPCs²⁰, meaning it likely extends to additional cell types as well. We also find that *in vitro* biochemical assays do not currently sensitively cover the full range of true positive off-target sites as evidenced by (1) missed nomination of ~20–60% of significantly edited EMX1 and FANCF off-targets and (2) missed sites that were generated from highly reproduced UNCOVERseq nomination. A likely explanation is that *in vitro* systems may lack or modify key features of biological systems that have been shown or hypothesized to have impacts on genome editing, including genomic accessibility¹⁵, local DNA topology^{40,41}, macromolecular crowding⁴², and non-physiological buffer systems^{43,44}. Together, these factors in addition to our findings indicate that *in vitro* assays may currently underestimate certain biologically relevant off-target events due to sub-optimal operating conditions, reinforcing that assay sensitivity is not an intrinsic property of assay class. Overall, our findings ground the sensitivity of multiple nomination methods in empirical data and highlight the importance of establishing empirically defined, modality-appropriate conditions for each platform. Future work should focus on expanding the size of a gold-standard dataset with paired nomination and confirmation, in addition to further quantifying and decreasing the limit of detection

for confirming off-target editing from targeted sequencing to limit potential false negatives.

Precision is another important metric for off-target nomination methods to appropriately select sites for downstream confirmation. Some off-target nomination methods produce extensive, low-precision site lists that are costly to confirm, and provide limited actionable correlation with cellular editing outcomes, which complicates both regulatory evaluation and the development of effective risk mitigation strategies. Using UNCOVERseq with sufficient replication, we show that nominated sites can achieve very high precision, enabling highly sensitive nomination while producing approximately an order of magnitude less nominated sites than low precision *in vitro* methods. We also find that sites nominated by UNCOVERseq have nomination frequencies that strongly correlate with observed editing frequencies—an association that *in vitro* biochemical methods generally fail to match when compared with *in cellulo* and *in situ* approaches. This type of actionable correlation enables acceleration of translational programs with integrated, risk-based decision making that incorporates both the likelihood of editing and its potential biological impact. As a result, UNCOVERseq substantially improves method-specific site prioritization, helping to filter or rank large off-target lists based on biologically relevant metrics. The importance of this step of the process has already been realized and exemplified in the de-risking of translational *in vivo* correctional gene therapies, further demonstrating clinical impact⁴⁵. Thus, by providing a more accurate and biologically relevant map of off-target activity, UNCOVERseq enables therapeutic developers to better quantify and characterize the risk profiles of candidate guides and editing systems. Additional work should further explore whether the output of UNCOVERseq can be used to better refine the sensitivity and precision of faster turn-around time *in silico* nomination methods to further bolster the translational process in an orthogonal manner.

Using a simple prioritization method based on frequency, replication, and high-level indicators of risk (exonic regions vs. intergenic, etc.), we demonstrate UNCOVERseq nominations with recommended experimental structures can result in manageable sequencing panel sizes for downstream confirmation (e.g., <250 sites across a range of specificities). We demonstrate that replicates are important in the off-target nomination process to ensure that sites meeting basic prioritization criteria are appropriately identified, with >25% of putative targets being missed without use of biological replicates. However, to better understand risk after off-target nomination and confirmation across methods, a more standardized scoring system to prioritize off-targets is needed in the future. The fields of oncology and heritable diseases have encountered similar issues and derived guidelines including tiered scoring systems from the American College of Medical Genetics (ACMG) and Association of Molecular Pathology (AMP) and modifications leveraging these criteria^{21,46,47}. Gene editing may be able to leverage some of these learnings but will face unique challenges in categorization of off-target risk since even off-targets in intergenic space during the nomination phase can be at risk for known structural variations derived from DSBs and SSBs. This includes events such as translocations⁴⁸, loss of heterozygosity⁴⁹ (LoH), aneuploidy⁵⁰, and other large variants like multi-kilobase deletions⁵¹. Given this, it seems likely that probability, frequency, and even proximity to other coding regions will be important criteria for triaging off-targets for assessment. While our work shows that UNCOVERseq can provide predictive values for probability and frequency of off-target events, future work should target determining generalized annotation prioritization and risk-tiering scoring schemas that can be broadly applied across technologies.

Benchmarked nomination and confirmation methodologies are needed for both DSB and SSB-based editing modalities. By selecting a variable range of gRNA specificities, we demonstrate that even in

popular ex vivo models like HSPCs with mRNA delivery, high specificity gRNAs are still sensitive to both SSB indels and base editing off-target effects at frequencies >0.01% and 0.5%, respectively. Furthermore, we demonstrate that indel editing and base editing are rank-order correlated across 34 base editing on/off-targets, supporting the idea that DSB-based nomination methods are effective tools for nominating both indel and base editing activity. Base editing-specific nomination methods, such as CHANGE-seq-BE, have been developed to target base editing events, while demonstrating unique off-target confirmation findings in comparison to published methods⁵². It is our belief that orthogonal methods for nominating both base editing events and indel events are likely needed for the foreseeable future until more complete cross-modality benchmarking can be done, especially given some of our findings that some UNCOVERseq-nominated sites generate confirmable indels only in conditions using the SSB base editing modalities in translational cellular contexts.

We envision UNCOVERseq coupled with promiscuous conditions being a powerful tool to help sensitively identify CRISPR-Cas off-targets for interrogation during pre-clinical development phases. By enhancing the quality of off-target nomination and grounding NGS operational conditions in empirical data, we believe that UNCOVERseq improves informed risk assessment of gene editing in translational systems. In a clinical development pipeline, this improved off-target resolution facilitates more rational guide RNA selection, more targeted mitigation approaches (e.g., guide redesign, high-fidelity nucleases, delivery modulation), and more efficient toxicology study design. Ultimately, the ability to characterize off-target effects with high precision strengthens the regulatory framework for genome editing therapeutics, accelerates candidate advancement, and reduces uncertainty in long-term safety projections. As gene therapy continues to shift toward in vivo editing and systemic delivery modalities, the need for such robust, context-relevant off-target analysis will only grow—positioning UNCOVERseq as a powerful enabling technology in the path toward safe, effective, and durable genomic medicines.

Methods

Human cell culture and transfection (K562 and HEK293-Cas9)

K562 (ATCC; CCL-243) and HEK293-Cas9 (ATCC; CRL-1573) cells were cultured in Iscove's Modified Dulbecco's Medium (IMDM; ATCC) and Eagle's Minimum Essential Medium (EMEM; ATCC) supplemented with 10% FBS at 37 °C with 5% CO₂. Ribonucleoprotein (RNP) complexes were formed by mixing Alt-R™ *S.p.* Cas9 Nuclease V3 (IDT) or SpyFi™ Cas9 Nuclease (Aldevron) and Alt-R CRISPR-Cas9 sgRNA (gRNA, IDT) and incubating for 20 min at room temperature (Molar Ratio: 1:1.2, Cas9:gRNA). For each transfection, 8.0×10^5 cells were washed with 1X phosphate-buffered saline, resuspended in 20 µL of solution SF (Lonza). For K562 cells, RNP complexes at 4 µM were combined with 4 µM of the dsODN (Supplementary Data 1) into the SF solution, while for the HEK293-Cas9 cells, 5 µM gRNA and 0.5 µM dsODN were added to the SF solution. This mixture was transferred into 1 well of a 96-well Nucleocuvette plate (Lonza) and electroporated using program FF-120 (K562) or DS-150 (HEK293-Cas9). Two nucleofections per replicate were performed and each treatment done in triplicate. Following electroporation, cells were transferred to a 6-well plate preheated with either IMDM or EMEM and were incubated at 37 °C with 5% CO₂ for 72 h. After incubation, genomic DNA (gDNA) was extracted using either the Purelink™ Pro 96 Genomic DNA Purification kit or the Monarch™ Spin gDNA Extraction Kit (New England Biolabs) according to the manufacturer's instructions, eluted in low-EDTA TE buffer (IDT, 11-05-01-05), and quantified using a NanoDrop 8000 UV-Vis Spectrophotometer (ND-8000-GL).

Primary T-cell culture and transfection

Frozen human primary pan-T cells (STEMCELL Technologies; 70024) from 2 unique human donors were thawed in ImmunoCult-XF T Cell

Expansion Medium including 300IU IL-2 (Cytiva) and activated with 10 µL/mL TransAct, human, T cell activator (Miltenyi Biotec) for 48 h. To prepare for transfection using Lonza 96-well plate 4-D Nucleofector system, cells were counted, pelleted using centrifugation (300 × *g*, 10 min at room temperature), and washed gently with 10 mL 1X phosphate-buffered saline. Cells were again pelleted and resuspended in Lonza Nucleofection Solution P3 at 2.5×10^6 cells/mL. For each electroporation, 5 µL of RNP complex and 3 µL dsODN was added to 20 µL of cells in P3 (5×10^5 cells/nucleofection) for a final concentration of 4 µM RNP (1:1.2 ratio of Cas9 to gRNA) and 1–4 µM dsODN. Where tag was not included, 3 µL of IDT Alt-R Cas9 Electroporation Enhancer was added for 3 µM final concentration to achieve a fixed final nucleofection reaction volume of 28 µL. Each reaction was mixed by pipetting and 25 µL was transferred to an electroporation cuvette plate. The cells were electroporated according to the manufacturer's protocol using the Amaxa 96-well Shuttle and nucleofection protocol 96-EH-140. After electroporation, the cells were resuspended in 75 µL pre-warmed IL-2 culture media in the electroporation cuvette. Triplicate aliquots of 25 µL of recovered cells were further cultured in 175 µL pre-warmed IL-2 media with TransAct. Cells were incubated for 72 h, after which gDNA was isolated and quantified.

iPSC culture and transfection

iPSCs from fibroblasts (Coriell Institute, GM23338) were cultured in mTeSR™ Plus media (Stemcell Technologies) at 37 °C with 5% CO₂. RNPs were formed as described above. For transfection using Lonza 96-well plate 4-D Nucleofector system, cells were detached using ReLeSR™ (Stemcell Technologies) and washed with 1X phosphate-buffered saline. Cells were resuspended in P3 buffer at 2×10^5 cells/nucleofection. CRISPR reagents at required final concentrations (4 µM RNP; 0.5 µM dsODN) were added to the mix to make a final volume of 25 µL, and of which 20 µL was transferred to the nucleocuvette for electroporation. The nucleocuvette plate was electroporated using code CA-137. After the nucleofection, cells were recovered and plated in complete mTeSR Plus medium with 1X CloneR™ 2 supplement (Stemcell Technologies). Recovery media was added to the electroporated cells to achieve a final volume of 100 µL, and 25 µL of this was added to 175 µL media per replicate well for final plating in a vitronectin-coated 96-well plate. During recovery and growth at 37 °C with 5% CO₂ for up to 96–120 h, media changes were performed as desired and/or following manufacturer's protocols for media and CloneR 2 supplement. gDNA extraction and quantification was performed as described above.

Off-target nomination with UNCOVERseq

500 ng of purified gDNA was enzymatically fragmented and adapter-ligated using the xGen™ DNA Library Prep EZ UNI kit along with the xGen Deceleration Module (IDT, xGen DNA Library Prep EZ UNI 96 rxn, 10009822; xGen Deceleration Module 96 rxn, 10009823) according to the manufacturer's instructions and cleaned with AMPure XP beads (Beckman). Following fragmentation and adapter ligation, rhPCR was performed using rhAmpSeq™ Library Mix 1 (IDT) to amplify the DNA in a single tube using a forward primer specific to the P5 adapter, a reverse primer specific for top and bottom strand of the integrated dsODN tag, and an adapter-blocking oligo corresponding to each strand of the dsODN³¹. Following PCR, samples were diluted 1:40 with nuclease-free water and used in a second PCR with rhAmpSeq™ Library Mix 2 (IDT) that added a unique P7 adapter to each library. Libraries were then cleaned with AMPure XP beads and run on an Agilent Fragment Analyzer for library quality assessment. All libraries were quantified with the Qubit 1X dsDNA HS Assay kit (Invitrogen) and pooled in equimolar amounts. All libraries were sequenced using an Illumina MiSeq or NextSeq2000 instrument with 150-bp paired-end reads. All oligonucleotides used in preparation of UNCOVERseq libraries are listed in Supplementary Data 1. All oligonucleotides used

in this study (Supplementary Data 1) were synthesized by Integrated DNA Technologies.

Vector construct and in vitro transcription of modified mRNA

Genes encoding HiFi Cas9⁵³, ABE8e³⁷, AncBE4max³⁶, and PE2²⁸ were each cloned into a plasmid vector with a dT7 promoter followed by a 5'UTR, Kozak sequence, ORF, and 3'UTR for subsequent in vitro transcription. Chemically modified Cas9, ABE8e, AncBE4max and PE2 mRNA was transcribed in vitro from the PCR templates with full substitution of uridine by N¹-methylpseudouridine. Co-transcriptional capping was achieved using CleanCap AG analog (TriLink), yielding a 5' Cap 1 structure. Transcription reactions were carried out with the HiScribe T7 High Yield RNA Synthesis Kit (New England Biolabs) in 0.5× transcription buffer supplemented with 4 mM CleanCap AG. Following synthesis, mRNAs were purified using the Monarch Spin RNA Cleanup Kit (New England Biolabs). PCR templates incorporated mammalian-optimized UTRs (TriLink) and a 120-nucleotide poly(A) tail.

HSPC culture and transfection

CD34⁺ HSPCs from a single human donor (Fred Hutch Cancer Center, RO04089) were cultured at 1×10⁵ cells/mL in StemSpan SFEM II medium supplemented with 100 ng/mL SCF, 100 ng/mL TPO, 100 ng/mL FLT3L, 100 ng/mL IL-6, 20 µg/mL streptomycin, and 20 U/mL penicillin. Cultures were maintained at 37 °C in a humidified incubator with 5% CO₂. For electroporation, Cas9 mRNA and gRNA were mixed at a 1:1 weight ratio (3 µg each per reaction). ABE, CBE, and PE2 mRNAs were used at equimolar amounts to Cas9 (3.4 µg ABE, 4 µg CBE, and 6 µg PE2), and pegRNA was added at the same molar amount as gRNA (4.2 µg). HSPCs were resuspended in 20 µL Lonza P3 buffer and electroporated using a Lonza 4D-Nucleofector (program DZ-100). After electroporation, cells were plated at 1×10⁵ cells/mL in HSPC medium.

Computational analysis—nomination

Following NGS, Illumina adapters and UMIs were identified and annotated using Picard MarkIlluminaAdapters. Tag sequences were identified and trimmed using Cutadapt v4.2⁵⁴. Sequencing reads were aligned to hg38 (GRCh38.p12) reference genome using BWA mem⁵⁵ v0.7.15 and UMI consensus reads were generated based on consensus from a single-strand (minimum UMI consensus size = 1) using fgbio v0.7.0 (<https://github.com/fulcrumgenomics/fgbio>). Nomination of candidate off-target sites began by using mapped UMI consensus reads to create a flanked search space (±40 bp) to perform alignment between the guide and empirical target region using a global implementation (+2/-1/-10/-1 match/mismatch/gap open/gap extension; penalize_end_gaps = False) of the Needleman-Wunsch alignment^{56,57}. After a candidate match to the gRNA spacer region was identified in the sequencing data, nominated off-target sites were identified using a hypergeometric test with multiple testing correction (Benjamini and Hochberg; FDR < 0.05) by comparing individual treatment samples and pooled control samples for significant differences in representation between the two. We used the following criteria to nominate off-target sites from this analysis for verification: (1) at least one sample nominated a given site with NGS evidence on both sides of the cut site, (2) Levenshtein distance < 7 as determined post-alignment, and (3) significant adjusted *p*-value when comparing the frequency of the event to the pooled control(s). Nominated on/off-target sites had additional meta-data added based on alignment/genomic context and were placed into described Tiers based on this meta-data (Supplementary Data 2).

Library preparation—confirmation

Genomic DNA was extracted from control and genome-edited cells as described above. Libraries for amplicon NGS were prepared using a previously described rhAmpSeq amplification-based method (IDT) using 100 ng of gDNA input⁵⁸. Briefly, the first round of PCR was

performed using target-specific primers. A second round of PCR was used to incorporate P5 and P7 Illumina adapters to the ends of the amplicons for universal amplification. Libraries were purified using Agencourt AMPure XP system (Beckman Coulter) and quantified by Qubit 1X dsDNA HS Assay kit (Invitrogen) before sequencing on the Illumina NextSeq platform (150-bp paired end reads). Read demultiplexing was performed on the resulting BCL files using Picard v2.18.9 (<https://github.com/broadinstitute/picard>) IlluminaBasecallsToFastq.

Computational analysis—confirmation

Analysis of the sequencing data to identify confirmed off-target editing at the nominated sites was performed using CRISPAItRations v1.2.1⁵⁹. This analysis comprised two parallel workflows: identification of indels at the position of the DSB/SSB, and identification of base editor-induced A>G (ABE) or C>T (CBE) transversions in the relevant base-editing window.

For identifying indels, the window for event quantification was centered on the canonical cut site and events quantified utilizing the default window size for Cas9 (8 bp). To determine whether indels found in the sequencing data could result from bona fide off-target cleavage, indels were grouped by location relative to the cut site (prioritizing minimum distance to cut site) followed by fitting counts of events to a negative binomial model with a Wald test for significance in each location bin per off-target using the DESeq2 package³⁵ within IDT's OTEasy tool (Schmaljohn et al., Manuscript in Preparation). For classification of indel off-target editing, the tool requires: (1) sufficient read coverage for the site (>1000×) in all replicates, (2) significant edits to occur at or adjacent to the cut site after optimal alignment, (3) the classified cumulative significant edits to exceed 0.01%, (4) the comparison of treatment/control samples at the site to have a significant adjusted *p*-value (*p* < 0.05), and (5) an average coverage frequency of at least 5× the ascribed cumulative frequency observed (e.g., for 0.1% editing, at least 5000× coverage). Blunt dsDNA tag integration quantification for quality control analysis were classified by aligning quantified insertions >10 bp in size from the CRISPAItRations output to the expected dsDNA tag using the biopython implementation of the Needleman-Wunsch aligner⁵⁶. Alignments with an alignment score greater than 50 were quantified as a tag integration event, and events were considered imperfect if any base was mutated or missing from the expected dsDNA sequence alignment.

For identifying base editing-generated off-target effects, the window for event quantification was centered in the middle of canonical base-editing window between position +5/+6 of the spacer (5' to 3') with a 5 bp window for quantification. To determine significant base-editing transitions resulting in off-target editing, all individual events that contained an ABE (A>G or T>C) or CBE (C>T or G>A) transition were grouped according to unique base editing events in the window and fitting counts of events to a negative binomial model with a Wald test for significance in each location bin per off-target using the DESeq2 package⁶⁰ within IDT's OTEasy tool (Schmaljohn et al., Manuscript in Preparation). For classification of adenine base editing at off-targets, the tool requires: (1) sufficient read coverage for the site (>1000×) in all replicates, (2) the classified cumulative significant edits to exceed 0.5%, (3) the comparison of treatment/control samples at the site to have a significant adjusted *p*-value (*p* < 0.05), and (4) an average coverage frequency of at least 5× the ascribed cumulative frequency observed.

Computational analysis—translocations

To quantify translocations from editing, Primer Anchored Statistical Translocation Analysis (PASTA) was used (Kurgan et al., Manuscript in Preparation). This analysis is only performed on the amplicon sequencing pools containing the on-target edit (because multiplexed amplification is a requirement for event detection using the method), and reactions not containing the on-target edit are unlikely to have any

significant translocation events. To quantify translocations, expected primers were identified in reads using fg-idprimer (<https://github.com/fulcrumgenomics/fg-idprimer>; $-k=6$, $-K=8$, $-S=5$, $--max-mismatch-rate=0.07$). Following this, treatment/control pairs had their counts paired and primer count frequencies subjected to a one-tailed hypergeometric test with Benjamini-Hochberg correction (statsmodel v0.15.0; default settings) to calculate an adjusted p -value (padj). Unexpected primer pairs with padj < 0.01 with no flags were classified as a translocation and had the translocation frequency (P) calculated using the following equation:

$$P_t = \frac{n_t}{f_{total}} + \frac{n_t}{r_{total}}$$

Where ‘ n ’ is equal to the count of the unexpected primer pair of interest, ‘ t ’ is the significant translocation being interpreted, ‘ f ’ is the total count of the shared forward primer events excluding the count participating in the ‘ n ’ translocation event, and ‘ r ’ is the total count of shared reverse primer events excluding the count participating in the ‘ n ’ translocation event. The translocation frequency is then adjusted by the background level frequency in the control by subtracting any translocation frequency observed in the control sample from the treatment frequency. Total translocation burden (B) was calculated using the following equation:

$$B = 1 - \prod_t^{t_n} (1 - P_t)$$

Where ‘ t ’ is equal to a significant translocation, and ‘ tn ’ is equal to the last significant translocation of all translocations. All translocations for the purposes of this equation are assumed to be occurring independently. Using the method, translocations are quantified if: (1) the estimated frequency exceeds 0.1% of editing, (2) the translocation has a significant p -value ($p < 0.01$), and (3) the translocation is found to meet these criteria in all replicates.

Statistics and reproducibility

No statistical method was used to predetermine sample size. No data were excluded from the analyses. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment. Statistical analysis was done using specified tests on a per analysis basis using the GraphPad PRISM data analysis suite. Most experiments were done at least in biological triplicate, but experiment-specific replication and the use of biological or technical replicates is specified.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Sequencing data generated in this study have been deposited in the SRA database under accession code [PRJNA1337486](https://www.ncbi.nlm.nih.gov/sra/PRJNA1337486). Processed data using the UNCOVERseq method are available for download through provided Supplementary Data. These datasets are provided exclusively for non-commercial research purposes. Use of the data for any commercial applications—including but not limited to training or developing predictive models, designing tools for gene editing applications, or integrating into commercial software or services—is strictly prohibited without prior written permission from the authors. Source data are provided with this paper.

References

- Jinek, M. et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012).
- Wang, J. Y. & Doudna, J. A. CRISPR technology: a decade of genome editing is only the beginning. *Science* **379**, eadd8643 (2023).
- Hsu, P. D. et al. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* **31**, 827–832 (2013).
- Fu, Y. et al. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol.* **31**, 822–826 (2013).
- 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).
- Tsai, S. Q. et al. CIRCLE-seq: a highly sensitive in vitro screen for genome-wide CRISPR-Cas9 nuclease off-targets. *Nat. Methods* **14**, 607–614 (2017).
- Lazzarotto, C. R. et al. CHANGE-seq reveals genetic and epigenetic effects on CRISPR-Cas9 genome-wide activity. *Nat. Biotechnol.* **38**, 1317–1327 (2020).
- Cameron, P. et al. Mapping the genomic landscape of CRISPR-Cas9 cleavage. *Nat. Methods* **14**, 600–606 (2017).
- Kim, D. et al. Digenome-seq: genome-wide profiling of CRISPR-Cas9 off-target effects in human cells. *Nat. Methods* **12**, 237–243 (2015).
- Dobbs, F. M. et al. Precision digital mapping of endogenous and induced genomic DNA breaks by INDUCE-seq. *Nat. Commun.* **13**, 3989 (2022).
- Cradick, T. J., Qiu, P., Lee, C. M., Fine, E. J. & Bao, G. COSMID: a web-based tool for identifying and validating CRISPR/Cas off-target sites. *Mol. Ther. Nucleic Acids* **3**, e214 (2014).
- Bae, S., Park, J. & Kim, J.-S. Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics* **30**, 1473–1475 (2014).
- Chaudhari, H. G. et al. Evaluation of homology-independent CRISPR-Cas9 off-target assessment methods. *CRISPR J.* **3**, 440–453 (2020).
- Cromer, M. K. et al. Comparative analysis of CRISPR off-target discovery tools following ex vivo editing of CD34+ hematopoietic stem and progenitor cells. *Mol. Ther.* **31**, 1074–1087 (2023).
- Horlbeck, M. A. et al. Nucleosomes impede Cas9 access to DNA in vivo and in vitro. *Elife* **5**, e12677 (2016).
- Cancellieri, S. et al. Human genetic diversity alters off-target outcomes of therapeutic gene editing. *Nat. Genet.* **55**, 34–43 (2023).
- Jennings, L. J. et al. Guidelines for validation of next-generation sequencing-based oncology panels. *J. Mol. Diagn.* **19**, 341–365 (2017).
- Yang, Z.-X. et al. OliTag-seq enhances in cellulo detection of CRISPR-Cas9 off-targets. *Commun. Biol.* **7**, 1–15 (2024).
- Nobles, C. L. et al. iGUIDE: an improved pipeline for analyzing CRISPR cleavage specificity. *Genome Biol.* **20**, 14 (2019).
- Shapiro, J. et al. Increasing CRISPR efficiency and measuring its specificity in HSPCs using a clinically relevant system. *Mol. Ther. Methods Clin. Dev.* **17**, 1097–1107 (2020).
- Li, M. M. et al. Standards and guidelines for the interpretation and reporting of sequence variants in cancer. *J. Mol. Diagn.* **19**, 4–23 (2017).
- U.S. Food and Drug Administration. Human gene therapy products incorporating human genome editing: Guidance for industry. <https://www.fda.gov/media/156894/download> (2024).
- Zhang, L. et al. AsCas12a ultra nuclease facilitates the rapid generation of therapeutic cell medicines. *Nat. Commun.* **12**, 3908 (2021).
- Liu, G., Lin, Q., Jin, S. & Gao, C. The CRISPR-Cas toolbox and gene editing technologies. *Mol. Cell* **82**, 333–347 (2022).
- Zetsche, B. et al. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* **163**, 759–771 (2015).

26. Gaudelli, N. M. et al. Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature* **551**, 464–471 (2017).
27. Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A. & Liu, D. R. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* **533**, 420–424 (2016).
28. Anzalone, A. V. et al. Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* **576**, 149–157 (2019).
29. Newby, G. A. et al. Base editing of haematopoietic stem cells rescues sickle cell disease in mice. *Nature* **595**, 295–302 (2021).
30. Arbab, M. et al. Base editing rescue of spinal muscular atrophy in cells and in mice. *Science* **380**, eadg6518 (2023).
31. Dobosy, J. R. et al. RNase H-dependent PCR (rhPCR): improved specificity and single nucleotide polymorphism detection using blocked cleavable primers. *BMC Biotechnol.* **11**, 80 (2011).
32. Tian, R. et al. Massively parallel CRISPR off-target detection enables rapid off-target prediction model building. *Med* **4**, 478–492.e6 (2023).
33. Yan, W. X. et al. BLISS is a versatile and quantitative method for genome-wide profiling of DNA double-strand breaks. *Nat. Commun.* **8**, 15058 (2017).
34. Musunuru, K. et al. In vivo CRISPR base editing of PCSK9 durably lowers cholesterol in primates. *Nature* **593**, 429–434 (2021).
35. Kim, D. & Kim, J.-S. DIG-seq: a genome-wide CRISPR off-target profiling method using chromatin DNA. *Genome Res* **28**, 1894–1900 (2018).
36. Koblan, L. W. et al. Improving cytidine and adenine base editors by expression optimization and ancestral reconstruction. *Nat. Biotechnol.* **36**, 843–846 (2018).
37. Richter, M. F. et al. Phage-assisted evolution of an adenine base editor with improved Cas domain compatibility and activity. *Nat. Biotechnol.* **38**, 883–891 (2020).
38. Mathis, N. et al. Predicting prime editing efficiency and product purity by deep learning. *Nat. Biotechnol.* <https://doi.org/10.1038/s41587-022-01613-7> (2023).
39. Fiumara, M. et al. Genotoxic effects of base and prime editing in human hematopoietic stem cells. *Nat. Biotechnol.* **42**, 877–891 (2024).
40. Newton, M. D. et al. Negative DNA supercoiling induces genome-wide Cas9 off-target activity. *Mol. Cell* **83**, 3533–3545.e5 (2023).
41. Eggers, A. R. et al. Rapid DNA unwinding accelerates genome editing by engineered CRISPR-Cas9. *Cell* **187**, 3249–3261.e14 (2024).
42. Collette, D., Dunlap, D. & Finzi, L. Macromolecular crowding and DNA: bridging the gap between in vitro and in vivo. *Int. J. Mol. Sci.* **24**, 17502 (2023).
43. Nguyen, G. T. et al. CRISPR-Cas12a exhibits metal-dependent specificity switching. *Nucleic Acids Res* **52**, 9343–9359 (2024).
44. Nguyen, G. T., Raju, A., Schelling, M. A. & Sashital, D. G. Rapid CRISPR-Cas9 target-strand nicking can provide phage resistance by reducing DNA abundance. *Nucleic Acids Res.* **53**, gkaf900 (2025).
45. Musunuru, K. et al. Patient-specific in vivo gene editing to treat a rare genetic disease. *N. Engl. J. Med.* **392**, 2235–2243 (2025).
46. Richards, S. et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet. Med.* **17**, 405–424 (2015).
47. Tavtigian, S. V. et al. Modeling the ACMG/AMP variant classification guidelines as a bayesian classification framework. *Genet. Med.* **20**, 1054–1060 (2018).
48. Ghezraoui, H. et al. Chromosomal translocations in human cells are generated by canonical nonhomologous end-joining. *Mol. Cell* **55**, 829–842 (2014).
49. Alanis-Lobato, G. et al. Frequent loss of heterozygosity in CRISPR-Cas9-edited early human embryos. *Proc. Natl. Acad. Sci.* **118**, e2004832117 (2021).
50. Cullot, G. et al. CRISPR-Cas9 genome editing induces megabase-scale chromosomal truncations. *Nat. Commun.* **10**, 1136 (2019).
51. Kosicki, M., Tomberg, K. & Bradley, A. Repair of double-strand breaks induced by CRISPR-Cas9 leads to large deletions and complex rearrangements. *Nat. Biotechnol.* **36**, 765–771 (2018).
52. Lazzarotto, C. R. et al. Sensitive and Unbiased Genome-Wide Profiling of Base-Editor-Induced Off-Target Activity Using CHANGE-Seq-BE. *Nat. Biotechnol.* <https://doi.org/10.1038/s41587-025-02948-7> (2026).
53. Vakulskas, C. A. et al. A high-fidelity Cas9 mutant delivered as a ribonucleoprotein complex enables efficient gene editing in human hematopoietic stem and progenitor cells. *Nat. Med* **24**, 1216–1224 (2018).
54. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.* **17**, 10–12 (2011).
55. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
56. Pj, C. et al. Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* **25**, 1422 (2009).
57. Needleman, S. B. & Wunsch, C. D. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol.* **48**, 443–453 (1970).
58. Jacobi, A. M. et al. Simplified CRISPR tools for efficient genome editing and streamlined protocols for their delivery into mammalian cells and mouse zygotes. *Methods* **121–122**, 16–28 (2017).
59. Kurgan, G. et al. CRISPRatRations: a validated cloud-based approach for interrogation of double-strand break repair mediated by CRISPR genome editing. *Mol. Ther. Methods Clin. Dev.* **21**, 478–491 (2021).
60. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

Author contributions

K.J.K. performed UNCOVERseq library preparation optimization and nomination using both GUIDE-seq and UNCOVERseq in HEK293-Cas9 cells. K.J. performed all HSPC editing experiments with different editing modalities. H.Z., E.S., T.O., G.K. performed UNCOVERseq computational data analysis and method optimization. S.W., S.C., E.S., T.O. and R.C. optimized UNCOVERseq analysis infrastructure. A.S.P. performed positive control wet lab confirmation experiments. B.T. and K.M. performed UNCOVERseq in iPSCs and T-cells. K.J.K., M.S., R.T., M.B., A.J., G.R., M.K.C., and G.K. provided wet lab guidance and experimental design guidance. G.K. and M.M. provided computational analysis guidance and support. K.J.K. and G.K. drafted the manuscript. All authors reviewed the drafts and performed editorial work.

Competing interests

Products and tools supplied in this manuscript by IDT are for research use only and not intended for diagnostic or therapeutic purposes. Purchaser and/or user are solely responsible for all decisions regarding the use of these products and any associated regulatory or legal obligations. For informational use only. The data provided are for informational use only and should not be used as the sole basis for any critical decision making. The data generated are based on assay procedures that have not undergone full validation: formal design and development activities are on-going. K.J.K., B.T., M.L.S., G.L.K., S.C., R.C., E.S., T.O., K.M., A.S.P., H.Z., M.B., M.S.M., S.W., R.T., G.R., and A.M.J. are employees of IDT, which offers reagents and services for sale similar to some described in the manuscript. M.K.C. has equity in Kamau Therapeutics. The remaining authors claim no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-026-74623-7>.

Correspondence and requests for materials should be addressed to Gavin L. Kurgan.

Peer review information *Nature Communications* thanks Ciaran Lee and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2026