

A side-by-side comparison of variant function measurements using deep mutational scanning and base editing

Ivan Sokirniy^{1,†}, Haider Inam^{1,2,†}, Marta Tomaszekiewicz^{1,2}, Joshua Reynolds^{1,2}, David McCandlish³, Justin Pritchard^{1,2,*}

¹Huck Institute for the Life Sciences, University Park, PA 16802, United States

²Department of Biomedical Engineering, University Park, PA 16802, United States

³Simons Center for Quantitative Biology, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, United States

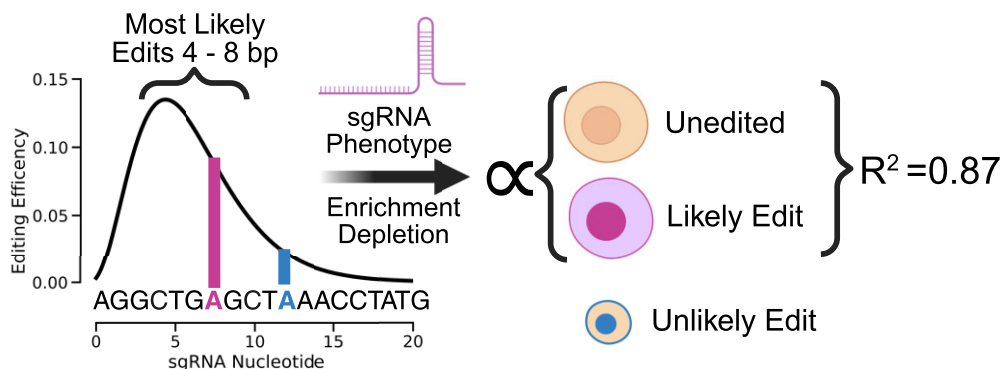
*To whom correspondence should be addressed. Email: Jrp94@psu.edu

†The first two authors should be regarded as Joint First Authors.

Abstract

Variant annotation is a crucial objective in mammalian functional genomics. Deep mutational scanning (DMS) using saturation libraries of complementary DNAs (cDNAs) is a well-established method for annotating human gene variants, but CRISPR base editing (BE) is emerging as an alternative. However, questions remain about how well high-throughput BE measurements can annotate variant function and the extent of downstream experimental validation required. This study is the first direct comparison of cDNA DMS and BE in the same lab and cell line. We focus on how well short guide RNA (sgRNA) depletion or enrichment is explained by the predicted edits within the editing “window” defined by the sgRNA. The most likely predicted edits enhance the agreement between a “gold standard” DMS dataset and a BE screen. A simple filter for sgRNAs making single edits in their window could sufficiently annotate a large proportion of variants directly from sgRNA sequencing of large pools. When multi-edit guides are unavoidable, directly measuring edits in medium-sized validation pools can recover high-quality variant annotation data. Our data show a surprisingly high degree of correlation between base editor data and gold standard DMS. We suggest that the main variable measured in base editor screens is the desired base edits.

Graphical abstract



Introduction

A major goal of mammalian functional genomics is to understand the genotype-phenotype relationship [1]. However, achieving this lofty goal requires large-scale experiments that can, in parallel, examine both genotypes and phenotypes in pooled assays. While gene-level loss-of-function screens in mammalian cells rose to prominence with the invention of RNAi-based screening tools [2–4] in the early 2000s, and eventually CRISPR/CRISPRi [5–8] in 2012–2013, the high throughput annotation of individual coding variants in mammalian cell lines is a more recent innovation [9, 10].

High-throughput variant annotation studies in mammalian cell systems include CRISPR-mediated HDR, cDNA library introduction, base editing (BE) by cytosine or adenosine base editors, and prime editing. For the purposes of this introduction, we have included [Supplementary Table S1](#) highlighting the strengths and weaknesses of these approaches alongside key references from the present literature. Importantly, HDR is a powerful tool but it has not been widely adopted outside of haploid HAP1 cells [11–15]. cDNA libraries can be introduced by viral vectors into virtually any mammalian cell line or incorporated into landing pads [16–18] in some cell

Received: February 25, 2025. Revised: July 1, 2025. Editorial Decision: July 3, 2025. Accepted: July 21, 2025

© The Author(s) 2025. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

lines and are easy to synthesize externally, but the artificial expression context is a source of concern. BE is becoming a remarkably efficient way to edit variants in the endogenous genomic context [19, 20], libraries of sgRNAs are easy to clone [21], and moving between cell lines is well demonstrated [22], but it suffers from a limited mutational repertoire and the experimental complexity of potentially making multiple mutations in a single editing window. Prime editing is becoming much more efficient [23], but it still requires MLH1 deficiency [24, 25] and requires pre-identification of strong epegRNA sequences for highest efficiency [23, 26]. Thus, at present, the two methods that are demonstrated to be the most useful across many wild type (WT) mammalian cell lines are cDNA libraries and base editor libraries. As mammalian geneticists and cell biologists work in over 1000 different cell lines, we focused this paper on the most portable technologies that have been shown by others to allow for the measurement of variant function in many different cell lines with little cell-type-specific modifications. Because of this, we have chosen to compare BE to cDNA-based deep mutational scanning (DMS).

DMS is a leading technology in the field of high-throughput variant annotation [9, 10]. DMS libraries can involve heterologous expression of large cDNA libraries of single amino acid mutations that encompass all 20 possible amino acids at every position. When performed in mammalian cell lines, these libraries are introduced via a transduction [27–29] or into a safe harbor “landing pad” [16–18]. While DMS is capable of providing comprehensive measurements of variant effects [10, 30, 31], DMS has nonetheless been difficult to scale to large genes or to multi-gene families, and the measurements may not reflect the effects of the same mutations at the endogenous genomic locus. Moreover, the technical challenges involved in DMS can lead to variable dataset quality [32].

As a functional genomic technology, BE is at development stage. BE screens use nCas9 to target a deaminase to a specific site in the genome and generate transition mutations (C > T for cytosine base editors—CBEs) or A > G for adenine base editors—ABEs) [19, 20]. BE screens use a surrogate measure of genotype by sequencing the short guide RNA (sgRNA) sequence, which allows BE screens to measure phenotypes across the genome [33–35]. Moreover, base editor screens have other major advantages that can include the ability to edit at the endogenous genomic locus and the ability to identify splicing defects [33, 36–39]. However, base editor screens also present certain challenges. The primary challenges are (i) BE efficiency (only a portion of individual cells harboring an sgRNA are likely to be edited, and some cell lines do not edit well [40]); (ii) off-target editing (while editing is largely constrained to a small window within the sgRNA non-target sequence, some off-target editing is likely occurring [41, 42]); (iii) bystander editing (when more than one possible edit occurs in an on-target editing window, the amino acid variant(s) made are more challenging to infer [43, 44]); and (iv) protospacer adjacent motif (PAM) requirements (PAMs limit where sgRNAs target [45], and current PAM-less Cas9 variants appear to have decreased efficiency [46–48]). These issues have led the field to view BE screens as a method for initial identification of interesting variants and regions but with limited capability to directly annotate loss-of-function phenotypes.

When competing high-throughput measurement methodologies can generate similar data, it can be extremely useful to directly benchmark these methods against each other. For

instance, the direct comparison of CRISPR Cas9 LOF screens and RNAi screens suggested that CRISPR is a more sensitive and specific technique for identifying essential LOF phenotypes, but that RNAi screens can help understand the dosage sensitivities of essential genes and can sometimes rescue false negatives in CRISPR screens for a subset of biological functions [49–51]. Additionally, the direct comparisons of high throughput drug sensitivity measurements found that the precise metrics and methods that are used in comparing datasets can create different conclusions on dataset reliability and usability. Together, these high-profile efforts highlight the importance of a careful comparison of high throughput datasets using multiple metrics and the public dissemination of the resultant data [52, 53].

Here we perform the first direct comparison between BE and DMS in the same cell line in the same lab. To accomplish this, we use the Ba/F3 cell system. This allows for a direct comparison and eliminates differences in genetic context as a confounding variable driving the differences in measurements between the approaches. Using this system, we identify specific data filters that generate largely matching conclusions about the phenotypes of loss-of-function variants. Furthermore, we demonstrate that applying these data filters enhances the correlation within the endogenous genomic context, indicating their robustness across different cell models. We also identify a two-step high throughput workflow for base editor screens that can streamline the validation of variant interpretation in pools by directly sequencing the edited variant fraction with error corrected sequencing [54, 55].

Materials and methods

DMS library preparation and screen

BCR-ABL cDNA was cloned downstream of EGFP in the pUltra (Addgene #24129) lentiviral vector by GenScript to make pUltra BCR-ABL WT (Addgene #210432). Twist Bioscience generated a saturating mutagenesis library of single amino acid changes in the N-lobe of the ABL kinase domain. NEB Stable chemically competent (NEB #C3040I) cells were transformed with the SM library, with a coverage of >1000×, onto 15-cm LB agar plates with ampicillin. After 48 h at 30°C, colonies were scrapped off the agar and plasmid DNA was extracted using Omega Bio-Tek E.Z.N.A. Plasmid DNA Midi kit (#D6904). HEK293Ts were transfected with 35 µg of saturating mutagenesis ABL library and 10 µg of helper plasmids (1:1:1:1) in 10-cm dishes using Thermo Fisher Lipofectamine 3000 (5 Lipo : 1 DNA). The next day, the media was changed to fresh RPMI (Cytiva SH30027.02). After 36 h, viral RPMI media was used to infect Ba/F3s, in the presence of 1 µg/ml mouse IL-3 (peprotech 213-13) and 6 µg/ml polybrene, at a low multiplicity of infection. After another 36 h, Ba/F3s were maintained in RPMI with 1 ng/ml IL-3. Infected cells were enriched by fluorescence-activated cell sorting (FACS) on EGFP at the Penn State Flow Cytometry Core facility.

At the start of the DMS screen, IL-3 was removed, and 30 million cells were saved to establish a baseline mutation frequency. Approximately 5 million cells were treated with DMSO for 6 days. Media was refreshed on day 3. Cell count was tracked by a BD Accuri C6 Plus flow cytometer. Cells were maintained in exponential phase. If cell viability was <90%, then viable cells were enriched by Ficol-Paque (Cytiva).

DMS library preparation and single-strand consensus sequencing

High-quality genomic DNA was extracted by Monarch Genomic DNA Purification Kit (NEB #T3010S). Then, a modified and scaled-up CRISPR-DS workflow was used to determine accurate variant distributions [56]. Equimolar crRNAs (IDT, 5'-caagtgggagatgaacgca-3', 5'-catgacctacgggaacctcc-3') were pooled and combined with tracrRNA (IDT) (final concentration: 10 μ M each). Guide RNA (gRNA) duplexes were formed by heating to 95°C for 5 min, followed by cooling to room temperature for 5 min. RNPs were assembled by incubating 10 μ l of gRNA duplexes with 1.6 μ l HiFi Cas9 (IDT), 3 μ l 1 \times CutSmart buffer (NEB), and 15.4 μ l nuclease-free water for 20 min at room temperature. Genomic DNA (20 μ g, resuspended in 1 \times CutSmart buffer) was digested by adding 20 μ l of pre-assembled RNPs and incubating at 37°C for 1 h. Proteinase K (10 μ l, 20 mg/ml, NEB) was then added, and the sample incubated at 56°C for 10 min. Undigested high-molecular-weight DNA was removed using 0.5 \times AMPure XP beads. A second purification with 1.8 \times AMPure XP beads was performed to remove short fragments, and DNA was eluted in 30 μ l TE. DNA was quantified using a Qubit fluorometer. A-tailing, unimolecular identifier (UMI) ligation, hybridization capture with custom probes, and PCR amplification were performed as described previously [57, 58].

One-hundred-fifty-nucleotide paired-end sequencing of the UMIs and mutagenized region was done on an Illumina NovaSeq 6000. For each sample, Du Novo [59] was used to generate error-corrected single-strand consensus from the UMI barcodes. Then bwa-mem2 was used to align the census to human ABL cDNA. After filtering out for mouse ABL reads, aligned reads with less than five mismatches would undergo variant calling and annotation using a custom R script. Briefly, for each alignment, variants were converted from the MDZ read tag. Mutant growth rates were calculated using exponential growth equation and the mutant allele frequency (MAF):

$$\text{growth rate} = \ln \left(\frac{\text{MAF}_1 \times \text{Count}_1}{\text{MAF}_0 \times \text{Count}_0} \right) \div (\text{Time}_1 - \text{Time}_0) \quad (1)$$

where the subscripts 0 and 1 denote the initial and final time point, respectively. Cell counts and the splitting ratio were used to account for dilution due to cell splitting during the DMS screen. Time is measured in units of hours. Skewed Gaussian mixture models were fit over the bimodal distribution of mutant growth rates using the Curve_fit function from the Scipy Python package [60]. Z-score cutoff for DMS data was determined by fit mean and standard deviation of the WT like component of the mixture distribution.

Base editor library preparation, screen, and sequencing

BCR-ABL tiling gRNA sequences were generated by CHOP-CHOP [61] with “NGN” PAM setting. Guides were cloned into lenti-sgRNA hygro vector (Addgene #104991) by GenScript to make the BCR-ABL sgRNA library. ABE8e SpG plasmid was made by deleting the U6 sgRNA cassette from pRDA_479 (Addgene #179099) [62] using NEB KLD (NEB #M0554S). For the CBEd SpG plasmid, a highly efficient and specific CBE from TadCBEd (Addgene #193835) [63] was subcloned to replace the ABE. The nickase (nSpG) plasmid was created by deleting the editor by NEB KLD.

Ba/F3s infected with pUltra BCR-ABL WT (Addgene #210432) were allowed to grow in the absence of IL-3. After 5 days >95% of the Ba/F3s were EGFP+. Then they were infected with either ABE8e SpG (Addgene #235044), CBEd SpG (Addgene #235045), or nSpG (Addgene #235046) and selected with 1 μ g/ml puromycin for 5 days. K562 ABE8e SpG were created similarly. In a 15-cm dish HEK293Ts were transfected with 60 μ g of the BCR-ABL and control sgRNA libraries and 40 μ g of helper plasmids (1:1:1:1) using calcium phosphate. The next day DMEM media was changed to RPMI. Three independent infections of Ba/F3 pUltra BCR-ABL WT ABE8e cells were performed at low multiplicity of infection and >500 \times coverage. CBEd SpG and nSpG carrying cells were infected similarly. Transfected HEK293Ts were saved to establish a baseline sgRNA frequency. Three days after sgRNA library infection, the cells were selected with 1 mg/ml hygromycin for 6 days and pelleted.

The K562 screen largely mirrored the Ba/F3 BCR-ABL screen, with both screens lasting 9 days. K562 cells were initially infected with ABE8e SpG and selected with 1 μ g/ml puromycin for 5 days, consistent with the Ba/F3 protocol. Subsequently, these K562 ABE8e SpG cells were infected with the BCR-ABL library, also as described previously, but selected with 200 μ g/ml hygromycin for 6 days instead. For the comparison between the K562 and Ba/F3 screens, we used a frequency cutoff of 1:10 000 to identify high-confidence sgRNAs.

Genomic DNA was extracted using phenol-chloroform [64] and quantified by Qubit. Staggered PCR was used to extract sgRNA sequences from 10 μ g of genomic DNA, as described previously [21], with two modifications: custom forward staggered primers were used, and PCR amplicons were gel extracted using Omega Bio-Tek Gel Extraction kit (#D6294).

sgRNA analysis

gRNA amplicons were trimmed with CutAdapt [65] to remove the U6 promoter and gRNA scaffold. The remaining sequences were aligned to a reference list of guides using Bowtie [66]. Guide counts were established by Counter from the collections Python package. pyDESeq2 [67] was used to determine the fold change in guides between time points. The fold change was converted to growth rates using Equation 1 above assuming WT growth rate. Z-score cutoff for ABE data was based on the human targeting (AAVS1, CCR5, and ROSA26) negative control sgRNA.

Verification screen, sequencing, and analysis

After making three independent Ba/F3 BCR-ABL WT ABE8e cell lines as above, IL-3 was returned to the RPMI media. We then cloned 14 pools of sgRNAs into a lenti-sgRNA hygro vector using Golden Gate sites or Gibson assembly [21, 68]. These sgRNAs targeted the ABL kinase domain (amino acids 242–320), aligning with the DMS region. To prevent overlapping edits and ensure individual sgRNA assignment within each pool, we staggered each sgRNA. In total, 71 ABL sgRNAs were evaluated, with five negative control sgRNAs (targeting mROSA26) spiked into each pool at ~12%. The lentivirus library of sgRNAs was made by lipofection of HEK293T in 10-cm dishes as above. We performed three independent infections using 1.5 million Ba/F3 BCR-ABL WT ABE8e cells. Cells were cultured for 3 days before initiating

Table 1. Polymerase chain reaction protocol for edit amplification

Temperature (°C)	Time	Cycles
95	2 min	1
95	10 s	
58	10 s	25
70	5 s	
65	5 min	1

1 mg/ml hygromycin selection. After 6 days, hygromycin and IL-3 were withdrawn, and a cell pellet was collected. This initial pellet was used for both sgRNA and edit sequencing. Nine days post-IL-3 withdrawal, further cell pellets were harvested for both sgRNA and edit sequencing. Guide sequencing was performed as described above [21]. Paired end 300 × 300 edit sequencing was carried out using combinatorial dual-index next-generation sequencing primers (Supplementary data S1) on NextSeq 2000 using P1 XLEAP-SBS reagents. To mitigate the risk of index hopping, we incorporated a randomized stagger, ranging from 0 to 7 nucleotides, upstream of the primer binding site. This design ensures that each unique index combination corresponds to a distinct read start and end location. One microgram of genomic DNA was used for both sgRNA and edit sequencing. To prepare the sample for edit sequencing, the target region was amplified with high-fidelity UltraII Q5 polymerase (NEB M0544L) according to the following PCR protocol (Table 1).

A custom Python script was used to determine the frequency of in-phase or *cis* mutations made by ABE. Briefly, reads were trimmed and aligned to ABL cDNA using Bowtie 2 [69]. A mismatch was classified as a bona fide edit only if it possessed a quality score exceeding 20 in both the forward and reverse sequencing reads. Background mutation frequency was set based on the triplicate mean of the 95th percentile of mutant frequency in lentivirus-integrated BCR-ABL from a negative control library containing only mROSA26-targeting sgRNAs. We used an sgRNA frequency cutoff of 1% to identify high-confidence sgRNAs. Fold change of sgRNAs and edits was determined by pyDESeq2 [67]. Guide and edit growth rates were calculated by normalizing them against the known WT BCR-ABL Ba/F3 growth rate of 0.055 h⁻¹. Specifically, the WT growth rate for guide analysis was derived from mROSA26 spike-in data, while for edit analysis, it was based on reads containing no mismatches. All A > G mutations within the editing window of an sgRNA were linked to that sgRNA.

BE-HIVE weighted model

In order to determine how well the on-target editing sgRNA phenotype can be recapitulated, sgRNA and editing efficiency was predicted by BE-HIVE [70]. Only edits with a frequency >0.05 and sgRNAs with completely matched DMS measurement were allowed. We model the growth rate of an sgRNA as the weighted sum of the growth rates of edited and unedited cells:

$$\text{sgRNA growth rate} = (\text{edit proportion} \times \text{edit growth rate}) + (\text{unedited proportion} \times 0.055) \quad (2)$$

In the case of simultaneous edits in *cis* across (*n*) multiple amino acids, a null model of mutant interactions predicts the growth rate of the multi-amino acid mutant as the product of individual mutant phenotypes [phenotype is growth rates (*r_n*) divided by the WT growth rate (0.055 h⁻¹)] multiplied by the

WT growth rate [71]:

$$\text{WT} \prod_{i=1}^n \frac{r_n}{\text{WT}} \quad (3)$$

In other words, under our null model if Mutant A confers a growth rate that is 0.7× the WT growth rate and Mutant B confers the same disadvantage of 0.7× the WT growth rate, then the compound AB mutant is expected to reduce the 0.7-fold growth rate by another 0.7-fold for an expected compound mutant growth rate of 0.49× the WT growth rate. If there were an epistatic interaction, the growth rates of the double mutant would differ from this expectation.

Results and discussion

DMS data are high quality and correlates with evolutionary conservation, secondary structure, and function

For our comparison experiments, we chose to use a fit-for-purpose cellular system, where the growth of the cell line, Ba/F3, depends upon the presence of IL-3 in the media until an activated tyrosine kinase is added. This system allows us to introduce the DMS library in the same genomic context as the target of the base editor screen. Heterologous expression of a tyrosine kinase cDNA integrated into the genome is advantageous with respect to our study design because all variants (DMS or BE) are evaluated using the same promoter. We use the BCR-ABL oncogene as the activated tyrosine kinase for several reasons: (i) It is an important oncogene that is of interest to applied and basic research groups. (ii) Its structure has been solved many times in many conformations [72–74]. (iii) There are years of experiments that can be drawn from to gain confidence in the resultant data [73, 75, 76].

To perform the DMS in BCR-ABL, we designed a saturation mutagenesis library spanning amino acid residues 242–320 in the N-lobe of the ABL1 kinase domain. This region was chosen for its reasonable size and the presence of known structural features that include the p-loop, the gatekeeper, and the alpha-C helix. This library was transduced into Ba/F3s in the presence of IL-3. After infection, FACS was used to enrich for infected cells. Following recovery from sorting, the Ba/F3 cells were screened for 6 days without IL3. In this negative selection screen, variants expressing non-functional copies of the BCR-ABL fusion protein deplete from the population. After the screen, we perform a sensitive barcoded sequencing protocol where the heterologous copy is specifically excised from the genome using Cas9-sgRNA complexes that are specific to the cDNA [56]. These genomic fragments are then ligated with unique molecular identifiers to allow for the deconvolution of library variants and the elimination of PCR and sequencing noise (Fig. 1A) [57, 58]. Genomic sequencing confirmed that the cell population contained a library representing 97% of all possible single amino acid variants spanning ABL1 residues 242–320.

When a mutant in BCR-ABL impairs kinase function, then the host cell will exhibit a reduced growth rate. After sequencing, we found that mutant growth rates form a bimodal distribution of counts (Fig. 1B), consistent with measured distributions of fitness effects in other DMS screens [17, 18, 28]. One of the peaks in this bimodal distribution of growth rates centered at ~0.02 h⁻¹ and constituted the set of deleterious mutations, while the other distribution mode was centered at

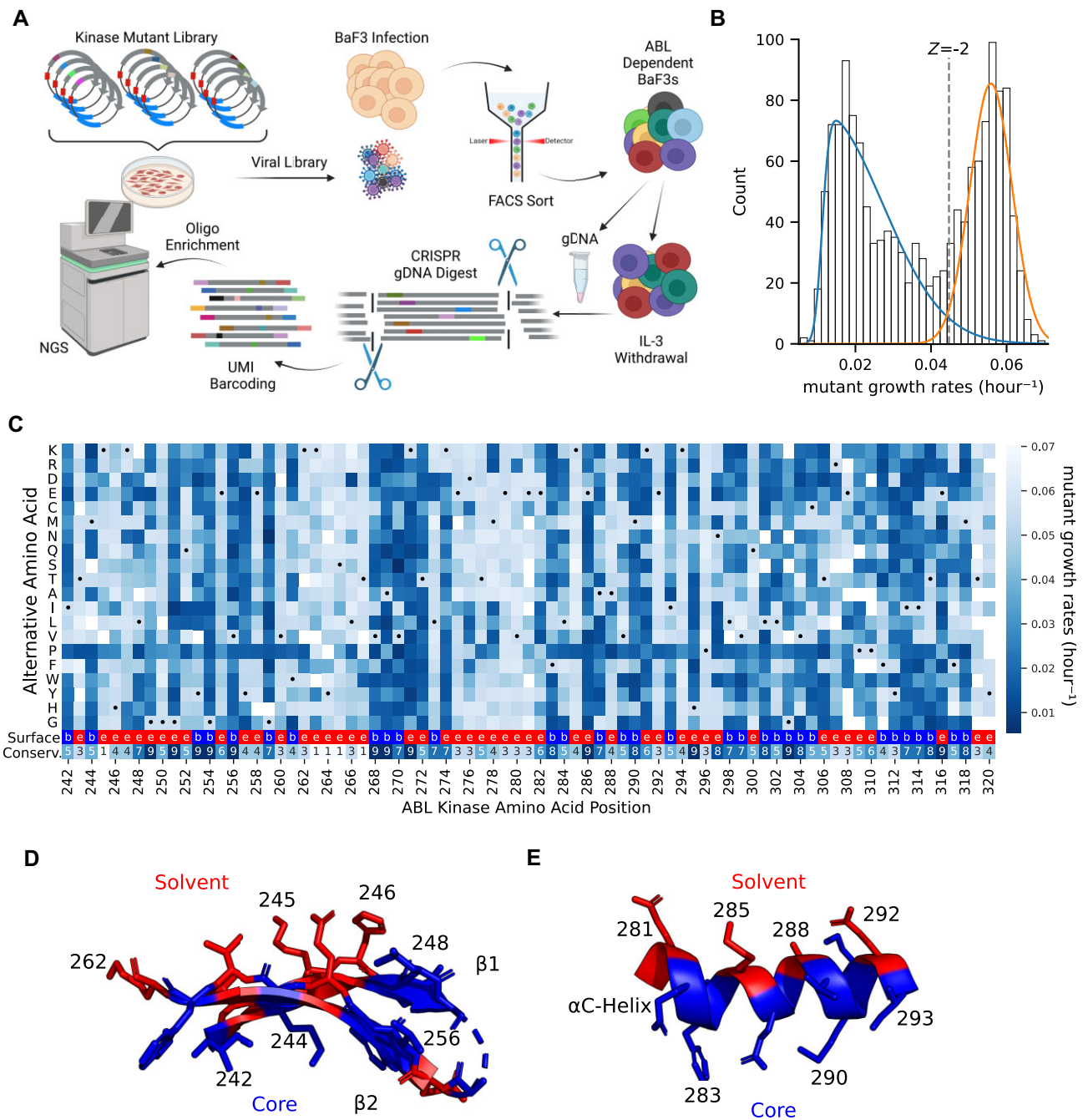


Figure 1. Functional landscape of ABL N-Lobe. **(A)** Schematic of DMS. After lentiviral integration of EGFP-P2A-BCR-ABL, Ba/F3s were sorted to enrich EGFP + infected cells. Cells were pelleted before and 6 days after IL-3 withdrawal. After targeted genomic DNA digest by Cas9, single molecules of DNA were labeled by UMI ligation. Then biotinylated oligo baits were used to enrich the mutagenized region. **(B)** The distribution of mutant growth rates in the ABL1 N-lobe is bimodal. Two skewed Gaussians are fit to determine the variation in deleterious (blue) and "WT-like" (orange) mutations. The dotted line represents a -2 Z-score threshold with respect to the "WT-like" distribution. **(C)** Heatmap of the growth rate of mutations at each position in ABL1 N-lobe. Black dot represents WT positions. Missing data are in white. Second to the last row of the heatmap provides surface exposure information because solvent exposed residues (red, "e") tend to be more tolerant of mutations than buried residues (blue, "b"). The bottom row of the heatmap indicates the evolutionary conservation for each residue on a scale of 1 (low conservation) to 9 (high conservation). Tolerance/sensitivity to mutagenesis is projected onto two key structural features of the ABL1 N-lobe (PDB 6XR6): the **(D)** anti-parallel beta-sheet, and the **(E)** α C-Helix. If mean growth rate of alternative alleles at a residue is less than the -2 Z-score cutoff, then the residue is colored blue. In contrast, if the mean growth rate of alternative alleles is greater than the -2 Z-score cutoff, then it is colored in red. ($N = 2$).

the known WT growth rate of 0.055 h^{-1} and harbored AAs that are “WT-like fitness” (Fig. 1B). To call the “hits” of specific amino acid variants that are required for growth, we use a Z-score cutoff of -2 for the “WT-like” distribution. This specific cutoff was selected to be consistent with prior literature in the sgRNA community [34, 35, 62, 77] and to draw a line that can be approximated in both studies using approximately analogous statistical and experimental criteria (Fig. 1B). Using this cutoff, we estimate that 56% of measured variants in the N-lobe of the ABL kinase impair kinase function.

Initial observations of the deleterious residues in the DMS dataset are consistent with prior knowledge. For instance, the catalytic lysine K271 is required for kinase activity and mutations at that position strongly deplete (measured growth rate averaged 0.019 h^{-1}) (Fig. 1C) [74]. Moreover, the systematic insertion of prolines during DMS studies causes a discernible “proline band” in high quality studies [28, 78]. We clearly observe a proline band in our data (Fig. 1C). Beyond these critical depletion signals, a known flexible and non-conserved region in ABL1 (i.e. residues 262–267 between the $\beta 2$ and $\beta 3$ strands) did not deplete.

More systematically, functionally important residues are typically evolutionarily conserved and their mutagenesis tends to be deleterious in DMS studies [10, 18]. Consistent with this, there is a -0.74 correlation ($P < .001$) between the mean mutant growth rate at a residue and the evolutionary conservation scores of that residue from ConSurf [72, 79–81] (Fig. 1C, bottom row of heatmap, Supplementary Fig. S1). Moreover, from a biophysical perspective, solvent accessible residues tend to be more tolerant of mutagenesis and are less conserved [82]. We observe this trend in Fig. 1C and D, and Supplementary Fig. S1.

The ABL1 N-lobe is composed of conserved structural elements that include the alpha-helix called the α C-helix, the five-stranded antiparallel β -sheet, the GxGxxG motif of the P-loop, as well as several catalytically important residues. Investigation of these key structural features reveals patterns that are consistent with known structure-function relationships. The GxGxxG of the P-loop spans residues 249–254 and all three conserved glycines show strong depletion phenotypes. Additionally, the alternating banding pattern in anti-parallel β -strands 1 and 2 (residues 242–262, Fig. 1D) is explained by essential residues facing toward the substrate pocket (blue residues, Fig. 1D). Moreover, the α C-helix, dynamically transitions from its “in” to its “out” state during kinase activation. This highly conserved regulatory dynamic requires support from the underlying hydrophobic core. Thus, the residues that interact with the core of the protein show the expected pattern of intolerance (blue, Fig. 1E), while the outward-facing residues show tolerance to mutations (red, Fig. 1E). It is the connection to known structure, function, and conservation data that gives us high confidence in the validity of the DMS data as a gold standard for comparison with sgRNA data.

An adenosine BE screen isolates functionally important domains and residues

One of the major benefits of base editor (BE) screens is the ability to rapidly screen across a protein’s entire length. To demonstrate this breadth and speed, we rapidly cloned a tiled library of 3535 sgRNAs across the entire BCR-ABL cDNA, including regions beyond the original DMS library. The BE phenotype was determined by the change in sgRNA abun-

Table 2. Hit rate among across different editors in Ba/F3 BCR-ABL screen

Editor	Library	Depleting guides	
		False	True
ABE	BCR-ABL	3488	47
	Control	1489	0
CBE	BCR-ABL	3531	4
	Control	1489	0
nSpG	BCR-ABL	3535	0
	Control	1489	0

Depleting guides are defined as guides with a \log_2 fold change of less than zero and an adjusted P -value of $< .05$.

dance following exponential growth in cell culture (Fig. 2A). We initially screened three different editors: ABE, CBE, and an nickase (nSpG) version of Cas9. While ABE had a high hit rate, CBE and nSpG had substantially lower hit rates (Table 2). Importantly, the CBE hits occur with sgRNAs that are expected to make edits in key residues, but both CBE and nSpG appeared limited in the breadth of phenotypes discovered. Based on this result, we focused our variant comparisons on our ABE data. Looking across the entire length of the protein (Fig. 2B), we performed a sliding window estimate of the proportion of strongly negative Z-scores across the protein with the ABE data (see “Materials and methods” section). This sliding window estimate highlights the regions of the protein that are unusually enriched for the presence of highly essential residues. Colored regions corresponding to the coiled-coil (CC) domain, double homology (DH) domain, plekstrin homology (PH) domain, Src-homology 3 (SH3), Src-homology 2 (SH2), the kinase domain, and F-actin binding domain (FAB) are highlighted. Highly significant domains that have been previously implicated in transformation, such as the CC [83, 84], SH2 [84–86], Kinase [87], and FAB domains [88, 89] show significant depletion. While the role of the DH domain is not understood, our base editor data suggests that it is important for growth factor independence (Fig. 2B) [90, 91].

While the matching genomic context of Ba/F3 (i.e. heterologous expression of BCR-ABL) for BE is useful for the comparison of BE and DMS, it is possible that editing variants in a Ba/F3 cell will yield a different result than editing variants in a K562 cell, which harbors true endogenous expression of BCR-ABL. To test for endogenous versus exogenous phenotypic differences, we also performed the same ABE screen in K562 cells, a human leukemia cell line, and we observed strongly correlated results with our Ba/F3 ABE screen (Pearson $r = 0.76$, Fig. 2C). Most of the spread in the points was due to sgRNAs that did not score in either screen, suggesting that noise in variants with no measured effects drove the correlation value.

Finally, to call depletion “hits” in our Ba/F3 screen we used a -2 Z-score based on the distribution of negative-control guides to be consistent with Hannah *et al.* [34] (Fig. 2C). Applying this criteria yields 387 hits, or $\sim 12.4\%$ of all BCR-ABL sgRNAs (Fig. 2D), consistent with others [34, 62]. Most notably, there is a strong over-representation of depletion phenotypes for guides that target the ABL1 kinase domain, with 38% of kinase domain guides depleting below a -2 Z-score. In concordance with the DMS screen above, the guides that can target the conserved regions of the P-loop, Lys 271, and the

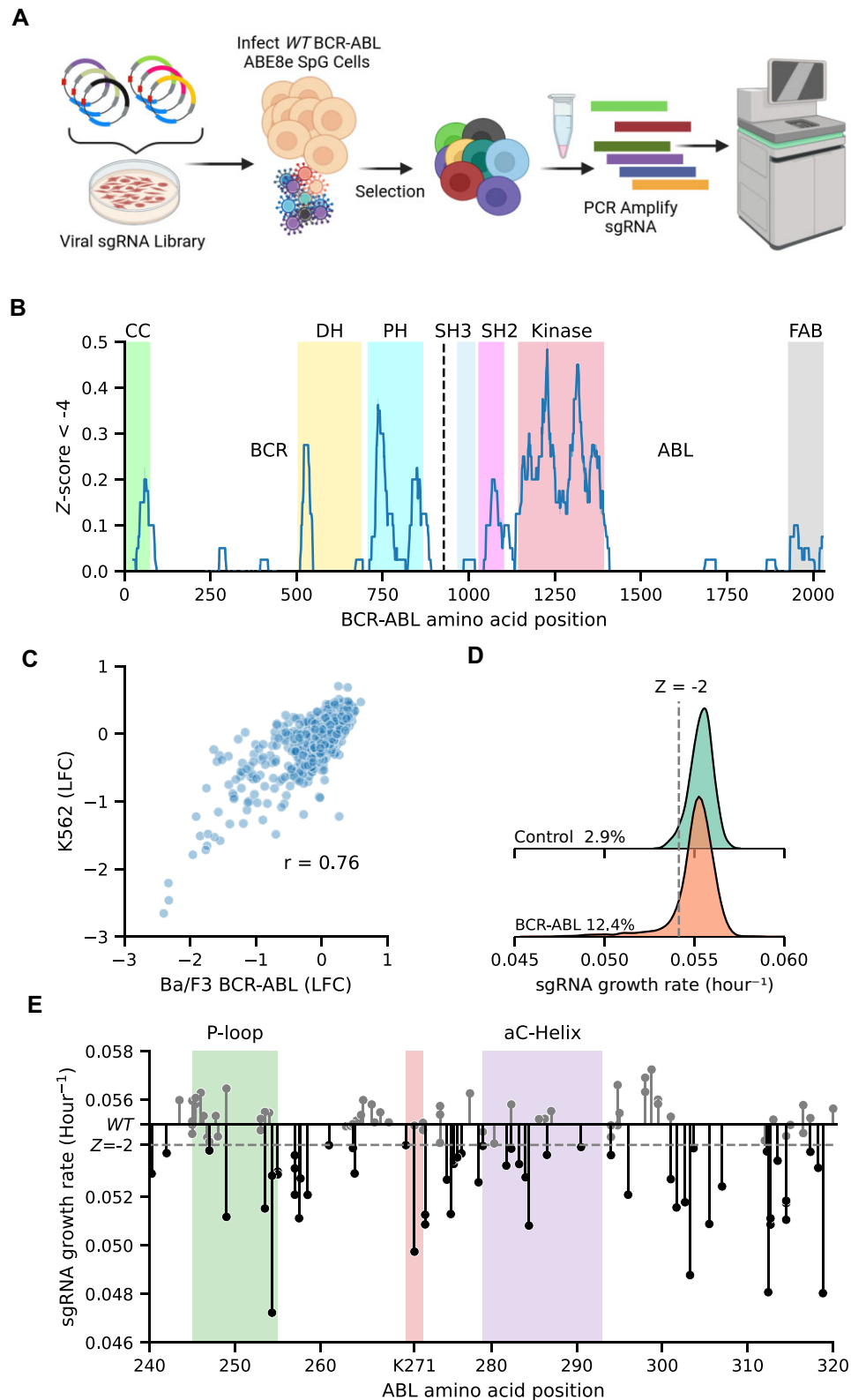


Figure 2. Adenosine base editor screen of full length of BCR-ABL. **(A)** Schematic of adenine base editor screen. Three days after infection with BCR-ABL sgRNA library, Ba/F3 EGFP-P2A-BCR-ABL ABE8e cells were selected with 1 mg/ml hygromycin for 6 days and pelleted. Guides were PCR-amplified and sequenced. **(B)** A sliding window analysis using a window size of 40 sgRNAs. In a window we quantified the proportion of BCR-ABL sgRNAs that drop out more extreme than a Z-score of -4 of the negative control sgRNA growth rate ($N = 3$). **(C)** Correlation of the same ABE BCR-ABL screens performed in K562s ($N = 2$) and Ba/F3 expressing BCR-ABL ($N = 3$). **(D)** Kernel density estimate of growth rate distributions of non-targeting control, and BCR-ABL sgRNA libraries. Dashed gray line represents a -2 Z-score of the targeting control. **(E)** Lollipop plot displays dropout of each sgRNA across the ABL1 kinase domain. Dashed gray line represents a -2 Z-score of the targeting control ($N = 3$).

buried region of the α C-helix produce deleterious phenotypes (Fig. 2E).

Side-by-side comparison of ABE predicted edits and DMS screens

To assess the correlation between our ABE screen and our DMS, we compared growth rates of each individual sgRNA (from the ABE screen) to the growth rates of the variants that sgRNA was predicted to create (from the DMS). We focused on 80/118 sgRNAs for which all predicted edits have variant information in the DMS screen. (Figs 1C and 2E).

The predicted edits of the sgRNAs predominantly occur within an editing window [43, 92]. While the editing window is in the non-targeting strand, we (and others) refer to the window relative to the guide sequence position. Most edits occur between positions 2 and 12 [92, 93]. Plotting sgRNA data against DMS data (Fig. 3A), each dot represents an individual variant, with each sgRNA appearing as a row of dots for all of its possible edits. An explanation of each quadrant is available in [Supplementary Fig. S2](#). Notably, the dynamic range of the Y-axis is reduced for BE compared to DMS (there is an $\sim 0.045\text{--}0.06\text{ h}^{-1}$ Y-axis measurement range for BE versus $\sim 0.01\text{--}0.06\text{ h}^{-1}$ for DMS on the X-axis). This is likely because counting an sgRNA read can count either edited or unedited cells.

This initial analysis yields a modest but significant Pearson correlation of 0.35 (Fig. 3A). This indicates a relationship but suggests caution in directly annotating individual variants from a raw sgRNA experiment without further validation of the variants with individual sgRNAs. This view is in line with the current practice in the field.

Considering the modest correlation with all putative variants, we hypothesized that the 2–12 bp editing window was too broad. By focusing on a narrower, more efficient 4–8 nucleotide editing window [34, 93, 94], the Pearson correlation improved from 0.35 to 0.51, and the odds ratio (OR) of hits versus non-hits increased to 6.9 (Fig. 3B). This improved agreement came at the cost of removing 25 potentially correctly annotated variants (Fig. 3A and B, and [Supplementary Fig. S3](#)).

After applying the “likely” edits filter to generate Fig. 3B, we still identify 13 false negative and 24 false positive variants in our plots. In terms of these remaining false negatives, an sgRNA could fail to deplete when sgRNA sequences are of low efficiency [21]. Thus, we examined the consequences of adding an additional filter requiring an sgRNA efficiency score >50 (Fig. 3C and [Supplementary Fig. S4](#)) [21]. This simple additional cutoff further improves the Pearson correlation between sgRNA predicted edit growth rates and DMS data to 0.56, and eliminates 9 of the 13 remaining false negative variants ([Supplementary Fig. S3](#)).

Next, focusing on the false positive variants in Fig. 3B (blue dots), we identified 15 sgRNAs that are predicted to make 24 false positive variants. Interestingly, 10 of the 15 sgRNAs are multi-edit sgRNAs that are predicted to also make true positives edits (dark orange variant dots). One simple way to reduce the effect of multi-edit ambiguity is to only examine the sgRNAs that are predicted to make a single edit in the 4–8 nucleotide editing region. The addition of this “single likely edit” filter enhances the Pearson correlation to 0.64 (P -value $< .001$) and the OR for hits to 44 (P -value $< .001$) (Fig. 3D). For single-likely-edit sgRNAs there is a true positive rate of

0.88 and an accuracy of 0.87 with respect to gold standard DMS data. However, this filter removed 15 out of 40 original sgRNA hits, i.e. the vastly increased specificity comes at a 38% loss in sensitivity. For detailed information on filters and filtered sgRNAs/variants, see [Supplementary Figs S3 and S4](#).

In order to expand the comparison to another cell line that expresses BCR-ABL from an endogenous genomic context, we performed the same comparison as in Fig. 3B and D, but we compared our DMS data in Ba/F3 to our ABE BE data in K562 cells. Interestingly, applying the same editing window filters from Fig. 3 to the K562 data yielded similar increases in Pearson correlation ([Supplementary Fig. S5](#)). This finding suggests that changing the genomic context and cell line does not dramatically change the results of this loss-of-function screen. It also suggests that the difference in the technique is a larger source of variation than the difference in genomic context.

Given that filtering sgRNAs by efficiency, edit probability, and the number of edits improves annotation confidence at the expense of total hits, we sought to utilize data more effectively from multi-edit sgRNAs (Fig. 3D). At a first pass, it seems like one might create variant-level interpretations for multi-edit sgRNA hits just by predicting the ensemble of sgRNA edits and their abundance in the population using machine learning algorithms like BE-HIVE [70]. However, while BE-HIVE can predict the ensemble of mutations (Fig. 3E), it can not predict the proportion of the measured phenotype that is attributable to those mutations. Therefore, we investigated whether multi-edit sgRNA dynamics could be predicted from the combined dynamics of their polyclonal edits. We used the BE-HIVE predicted allele frequencies vector (to approximate the structure of the polyclonal population), and simply assumed that all of the predicted variants grow according to the gold standard DMS growth rates. This approach estimated the expected multi-edit sgRNA growth rate as a function of the individual edits (Equation 3). This also provides insights into the contribution of in-window editing versus off-target effects to sgRNA dynamics. We observed a moderate Pearson correlation of 0.57 (P -value $< .001$) between the predicted and actual sgRNA growth rates (Fig. 3F). This suggests on-target editing is a major contributor to sgRNA growth rates. It also supports the feasibility of variant validation experiments with pooled sgRNAs by directly sequencing edited variants from genomic DNA. We will test this hypothesis in the next section.

Pools of multi-edit sgRNAs create a pool of edited variants that can be measured by error-corrected deep sequencing of genomic DNA

Our BE-HIVE analysis suggested that in-window editing drives multi-edit sgRNA growth rates (Fig. 3F). Therefore, we hypothesized that variant validation could be accelerated by directly sequencing variants in genomic DNA during pooled experiments. While doing this in a larger pool with thousands of sgRNAs is challenging because of the error rates involved in variant detection, sequencing the genomic DNA in pools with ~ 100 sgRNAs is approachable with simple error correction schemes like TileSeq [54, 55]. This can be compared to the standard validation approach where sgRNA-variant relationships are tested in a one-by-one manner. The proposed approach involves medium-throughput validation pools of sgRNAs, and measuring the resultant variants directly in gDNA via error-corrected sequencing of the editing target. To test this, we selected 71 sgRNAs with varying predicted fitness

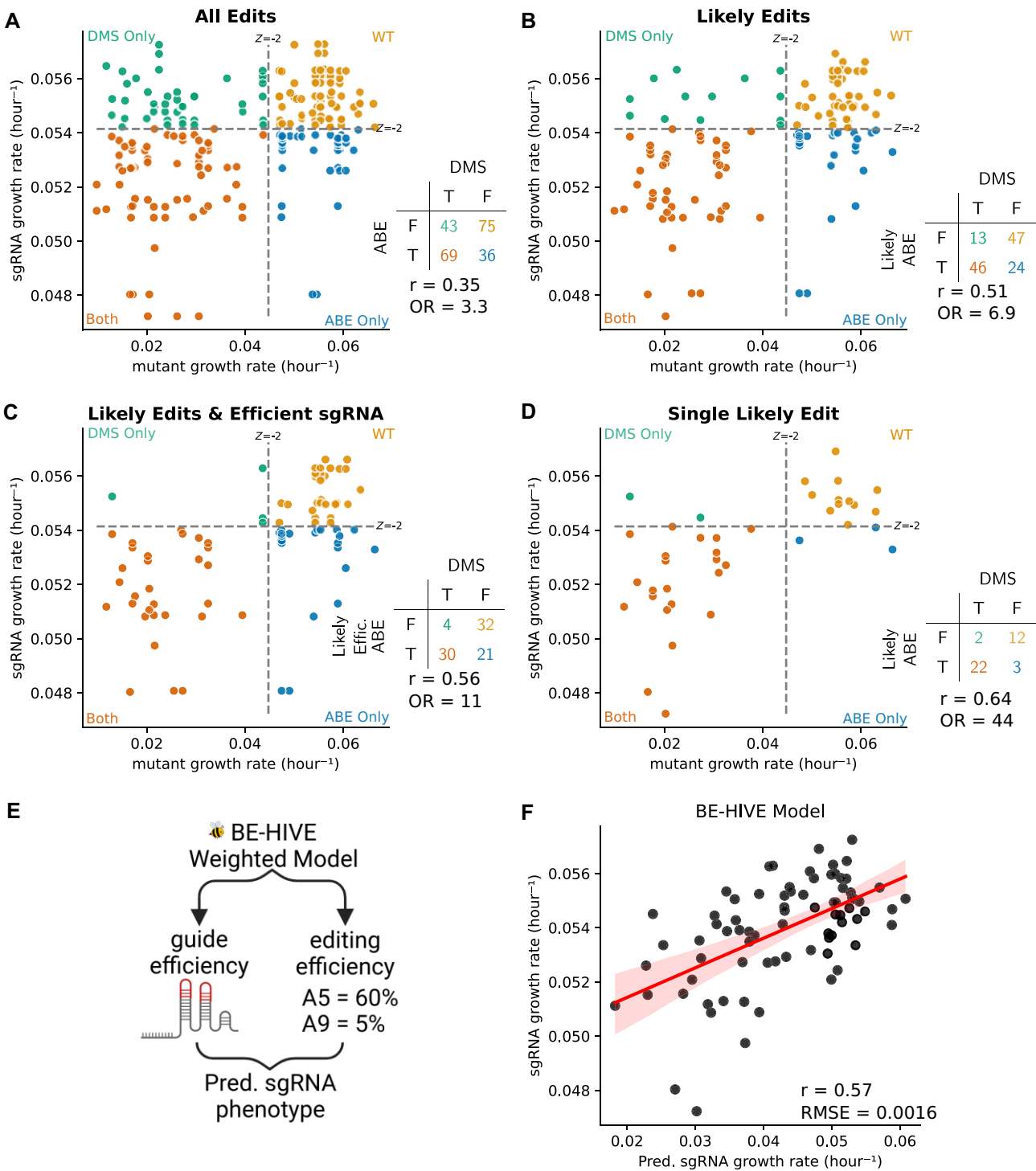


Figure 3. Comparison of adenosine base editor sgRNA growth rate and their respective mutation growth rates from DMS. Each dot represents a mutation an sgRNA is predicted to make. Dashed lines represent -2 Z-score of the non-deleterious distribution, and negative control sgRNA for the DMS and ABE screens, respectively. These cutoffs are used to define if an sgRNA or mutation is deleterious. If an sgRNA and its mutation do not deplete in their respective screen, in other words, both are non-deleterious, then they are colored yellow. If they both are deleterious, or true positive, then they are colored orange. If an sgRNA depletes, but the predicted edit does not deplete, a false positive, then the dot is colored in the blue. If a sgRNA fails to deplete, and the predicted mutation(s) are deleterious, a false negative, then that point is colored in green. **(A)** Shows all possible edits between nucleotides 2 and 12. **(B)** Shows only the most likely edits, those between nucleotides 4 and 8. **(C)** Shows only sgRNA predicted to be efficient, and edits between nucleotides 4 and 8. **(D)** Shows sgRNAs that are predicted to make only a single edit between nucleotides 4 and 8. **(E)** The distribution of edits can be estimated by machine learning model called BE-HIVE. **(F)** Correlation between predicted sgRNA growth rate and observed sgRNA growth rate. The x-axis shows the predicted growth rate of each sgRNA based on a weighted sum of the probability edit(s), and the effect of that edit(s) from DMS data. The y-axis shows the measured growth rate of the efficiently editing sgRNAs from the ABE screen.

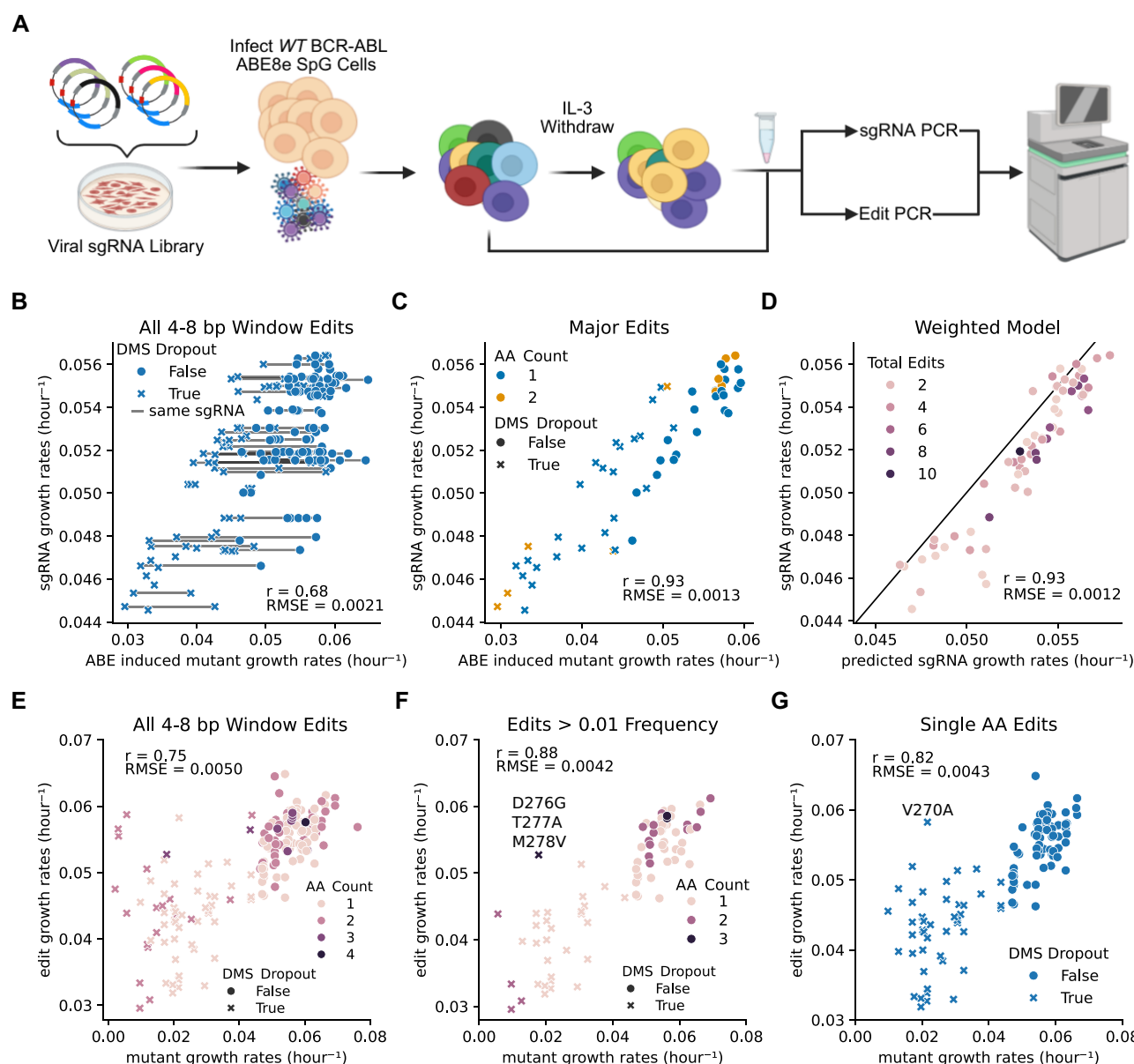


Figure 4. Medium-throughput pooled adenosine base editor screen. **(A)** Schematic of medium-scale validation screen of 71 sgRNAs targeting ABL1 kinase, where the edits and sgRNA are sequenced after IL-3 withdrawal. **(B, C)** Growth rates of sgRNA-induced edits. Each dot represents a specific edit and its measured growth rate, while each "X" or filled circle indicates whether the corresponding mutation was deleterious or nondeleterious in a prior DMS experiment. **(B)** Growth rates for all detected edits within the 4–8 nucleotide editing window of their respective sgRNAs. Gray lines connect edits generated by the same sgRNA. **(C)** Highlights the most prevalent edits, defined as those occurring at a frequency of over 50% of all edits within the sgRNA's editing window. **(D)** Weighted model for sgRNA growth rate. The sgRNA growth rate was predicted by a weighted sum of the growth rates of all edits within its 4–8 bp editing window, with weights corresponding to the frequency of each edit. The black line indicates a perfect correlation between the predicted and experimentally measured sgRNA growth rates. **(E–G)** Comparison of growth rates from the pooled ABE screen and DMS data. These panels directly compare the growth rates of specific amino acid mutants measured in the pooled ABE screen and the prior DMS experiment. The amino acid (AA) count represents the number of amino acids simultaneously edited. **(E)** All edits detected within the 4–8 bp editing window. **(F)** Highlights high-confidence growth rate measurements by applying a stringent edit frequency cutoff of 0.01, and **(G)** focuses exclusively on single amino acid edits to enable a direct comparison between the ABE screen and DMS data ($N = 3$).

measurements, aiming to determine if a variant is deleterious, and to explain sgRNA measurements through direct sequencing of the polyclonal pool (Fig. 4A). Sixty-three of the 71 selected sgRNAs made edits that were observable above the background mutation rate.

The measured sgRNA growth rates strongly correlated (Pearson $r = 0.87$, $P < .001$) with the sgRNA growth rates measured in the high-throughput screen from Fig. 2, confirming that the medium-throughput validation pool behaves

similarly to the larger pool (Supplementary Fig. S6A). Sequencing the edits for each sgRNA revealed a wide range in the number of edits associated with individual sgRNAs (Supplementary Fig. S6B and C). Further analysis revealed a high Pearson correlation ($r = 0.68$) between sgRNA growth rates and the directly measured mutant growth rates (Fig. 4B), directly confirming that the fitness effects of sgRNAs are due to on-target editing. This correlation rose to 0.93 when considering only the major edit (defined as an edit whose fre-

quency accounts for over 50% of an sgRNA's total edits) (Fig. 4C). Even when incorporating all edits occurring within the likely editing window (nucleotides 4–8) using the weighted model from BE-HIVE (Fig. 3F; Equation 3), the correlation remained at 0.93 (Fig. 4D). This strong correlation suggests that while an sgRNA can generate multiple edits, a large majority of its phenotype is explained by these major edits.

If we directly compare edits made by ABE at a permissive variant frequency cutoff of 0.00015, based on the 95th percentile of mutant frequency detected in the absence of ABL1 targeting sgRNAs, and the mutations by DMS, we observe a Pearson correlation of 0.75 (Fig. 4E). This indicated that the direct sequencing of edited variants in resequencing pools can approach the quality of gold standard DMS data. However, some false negative edits appear to persist. This means that we directly observe edits that fail to drop out despite DMS data indicating they should. One explanation is that these edits are sequencing errors. Applying a more stringent background mutation cutoff of 1% of the relevant reads can eliminate these low-confidence measurements, which in turn improves the Pearson correlation to 0.88. The sole remaining strong outlier is a triple mutant (D276G T277A M278A) (Fig. 4F). While our compound mutation model (Equation 3) [71] predicts this triple mutant should deplete because M278V is strongly deleterious per DMS data, it's possible that the addition of D276G and T277A rescue M278V's deleterious effects on kinase function. To avoid potential confounding epistatic effects of compound mutations we also directly compare single amino acids mutants only, which still maintains a strong Pearson correlation of 0.82 (Fig. 4G).

Conclusion

Thousands of human cell lines are routinely used to understand gene function in diverse tissue types, genetic backgrounds, and disease states [95, 96]. To plan how to perform variant annotation in these diverse contexts, the mammalian cell biologist/molecular geneticist requires a comparison of the variant editing tools that, at present, appear most suited to tackling the experimental challenges presented by human cell line diversity. Briefly, while HDR is a tremendous method that has made a huge impact, its use has largely been confined to haploid cell lines. Prime editing, while historically inefficient, has seen recent breakthroughs [23, 25]. However, these breakthroughs have required MLH1 KO [24] and the careful preselection of epegRNA sequences, resulting in an experimental method with uncertain transferability in WT cell lines [26]. [Supplementary Table S1](#) has a summary of current methods alongside their strengths and weaknesses, while acknowledging that the field is changing rapidly. To enable the use of variant annotation tools by a broader community of mammalian cell biologists and geneticists, we sought to focus our benchmarking on two sets of tools (cDNA–DMS and BE) that have been already used in many different mammalian cell lines for functional genomic screens [22, 28, 31, 32, 34, 35, 40, 41, 62], and therefore represent the tools that are likely to be usable by the broadest set of audiences and biological questions.

In conclusion, our study provides a comprehensive comparison of DMS and CRISPR-based BE for variant annotation. We demonstrate that while DMS offers unparalleled depth and structural resolution, BE screening provides a rapid, broad, and efficient alternative at the cost of mutation diver-

sity. By analyzing both methods in the same cellular and genomic context, we achieved a surprisingly high degree of correlation between the two, despite their inherent methodological differences and the potential for off-target effects and elevated mutation rates in BE. This robustness can clearly be seen in Fig. 4 which shows that sgRNA enrichment or depletion is largely interpretable as the composite sum of on-target edits within an editing window ($r = 0.93$).

We show that incorporating filters for sgRNA editing efficiency and reducing multi-edit ambiguity enhances the correlation between BE and DMS data. Our findings further reveal that variant annotation can be achieved directly from sequenced sgRNAs in BE screens when focusing on sgRNAs with single predicted edits within a narrow and efficient editing window. This streamlines the annotation process, particularly for variants exhibiting strong phenotypes. It also means that these single-edit sgRNAs sequenced directly from a pool can immediately classify a specific variant as deleterious and would be expected to create a ~13% false positive rate directly from high-throughput screens. We also present a rapid way to deconvolute medium-throughput pools with error-corrected deep sequencing. Because sgRNAs are challenging to interpret when they are predicted to make multiple edits in their core editing window, we propose a two-step validation workflow using error-corrected deep sequencing. Error correction is necessary in medium-throughput pools because even a 100 sgRNA pool with editing efficiencies between 10% and 50% will create variant frequencies below the limit of detection (0.1%–1%) of conventional pooled NGS [55]. Figure 4 suggests that sequencing the variants with error correction in medium-sized validation pools of sgRNAs improves the quantitative correlation between base-edited variant annotation and cDNA-based DMS. We suggest the following workflow for base editor dropout screens: (i) Do a large unbiased screen (thousands of sgRNAs) where sequencing the sgRNA is used to deconvolute the pool. (ii) For sgRNAs that make single edits in the core editing window, annotate the predicted variant function directly from the screen in step i. (iii) Do a second validation screen of only the hits from step i using pool sizes of ~100 sgRNAs. Directly sequence all the edits in these pools using error corrected TileSeq [54]. (iv) Use the direct sequencing data to validate all multi-edit variants.

One of the closest comparison results to the present study is that of Hanna *et al.* [34], who compared saturating genome editing [12] (akin to HDR-based DMS, performed at the University of Washington) with BE in BRCA1 in HAP1 cells performed at the Broad Institute. They reported a modest Pearson correlation ($r = 0.44$) [34]. A second similar study performed a comparison between cDNA-based DMS and prime editing in TP53 in A549 cells treated with a small molecule that selects for TP53 loss-of-function mutations [26]. However, the Pearson correlation they observed in this TP53 study was extraordinarily poor, with the authors arguing that supraphysiological overexpression of the TP53 mutant in the cDNA-based DMS screen was the cause.

Our comparison and our findings are different from both of these prior results in different systems. In our tightly controlled dropout screens, we observe a higher correlation than Hannah *et al.* [34] when limiting our comparisons to the core editing window. When comparing our K562 and Ba/F3 BE results (in the absence of splicing altering mutations), we suggest

that the high correlation suggests that endogenous genomic context plays a minor role in variant annotation in our system. Thus, not all overexpression systems may have the same downsides that were observed in the case of TP53 [26]. We speculate that the dominant negative effects of TP53 are behind the importance of genome context in the A549 system. Our findings also extended to gain-of-function experiments, where imatinib treatment of Ba/F3 BCR-ABL cells showed comparable improvements (Supplementary Fig. S7). This was contingent on incorporating filters for likely edits and restricting analysis to single nucleotide edits within the 4–8 bp core editing window.

However, for complex cases involving multi-edit sgRNAs and double mutations, direct sequencing of mutant pools offers a robust validation strategy. By directly measuring the variants generated by pooled sgRNAs, we confirmed that the fitness effects observed in BE screens are primarily due to on-target editing in the sgRNA editing window. This approach allows for accurate variant annotation even in challenging scenarios, while maintaining a higher throughput.

Overall, our study also suggests that because these tools are measuring the same underlying phenomena, the complementary strengths versus weaknesses of cDNA DMS and BE screening for variant annotation might be combined for maximum impact in human cell biology and genetics across large regions of proteins, or entire protein networks. BE libraries can be rapidly deployed in a range of cell lines for editing at the endogenous locus. DMS in mammalian cells can then give structural resolution and detail in proteins. By strategically combining these two powerful tools, researchers can achieve efficient variant characterization across the genome with BE and structural detail with DMS. This will accelerate our understanding of gene function, structure-function, and disease mechanisms in diverse human cell lines.

Acknowledgements

We would like to acknowledge members of the Pritchard lab for constructive criticisms of the manuscript, The Pennsylvania State University Flow Cytometry Core for help with sorting cells, and The Pennsylvania State University Genomics Core and Genomics Research Incubator for next generation sequencing.

Author contributions: Ivan Sokirniy (Conceptualization [equal], Data curation [equal], Formal analysis [equal], Investigation [equal], Methodology [equal], Software [equal], Validation [lead], Visualization [lead], Writing—original draft [equal], Writing—review & editing [equal]), Haider Inam (Conceptualization [equal], Data curation [equal], Formal analysis [equal], Investigation [equal], Methodology [equal], Software [equal], Writing—original draft [equal], Writing—review & editing [supporting]), Marta Tomaszewicz (Data curation [supporting], Investigation [supporting], Methodology [supporting], Writing—review & editing [supporting]), Joshua Reynolds (Data curation [supporting], Investigation [supporting], Methodology [supporting], Writing—review & editing [supporting]), David McCandlish (Formal analysis [supporting], Funding acquisition [supporting], Writing—review & editing [supporting]), Justin Robert Pritchard (Conceptualization [equal], Funding acquisition [lead], Project administration [equal], Writing—original draft [equal], Writing—review & editing [equal]).

Supplementary data

Supplementary data is available at NAR online.

Conflict of interest

J.R.P. is a co-founder of RedAce Bio. J.R.P. is a co-founder and consultant for Theseus Pharmaceuticals. J.R.P. received equity from Theseus Pharmaceuticals, MOMA therapeutics and RedAce Bio. J.R.P. has consulted/consults for MOMA therapeutics, CurieBio, Third Rock Ventures, Takeda Pharmaceuticals, Galapagos Pharmaceuticals, and Roche/Genentech. J.R.P. has received honoraria and travel expenses from Roche/Genentech, Third Rock Ventures, and Theseus Pharmaceuticals. H.I., M.T., and J.R. are co-founders of Atlas Biotech.

Funding

This work was in part funded by an NSF Modulus Grant MCB 2141650 (J.R.P., D.W.). NIH Grant T32GM108563 (I.S.), NCI U01 Synthetic Biology Grant U01CA265709 (J.R.P.), NSF RECODE Grant CBET 2033673 (J.R.P.), NIH Grant R35 GM133613 (D.M.M.) and additional funding from the Simons Center for Quantitative Biology at Cold Spring Harbor Laboratory. Funding to pay the Open Access publication charges for this article was provided by NSF Modulus Grant MCB 2141650.

Data availability

All sequencing data used in this study can be examined at Short Read Archives BioProject PRJNA1218278. All code used to analyze base editor and DMS data is available on [GitHub](#) and [Figshare](#) with the DOIs 10.6084/m9.figshare.29497922, 10.6084/m9.figshare.29497925, 10.6084/m9.figshare.29525993 and 10.6084/m9.figshare.29565386.

References

1. Kinney JB, McCandlish DM. Massively parallel assays and quantitative sequence-function relationships. *Annu Rev Genom Hum Genet* 2019;20:99–127. <https://doi.org/10.1146/annurev-genom-083118-014845>
2. Elbashir SM, Harborth J, Lendeckel W *et al.* Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* 2001;411:494–8. <https://doi.org/10.1038/35078107>
3. Root DE, Hacohen N, Hahn WC *et al.* Genome-scale loss-of-function screening with a lentiviral RNAi library. *Nat Methods* 2006;3:715–9. <https://doi.org/10.1038/nmeth924>
4. Fire A, Xu S, Montgomery MK *et al.* Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 1998;391:806–11. <https://doi.org/10.1038/35888>
5. Ran FA, Hsu PD, Wright J *et al.* Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* 2013;8:2281–308. <https://doi.org/10.1038/nprot.2013.143>
6. Shalem O, Sanjana NE, Hartenstein E *et al.* Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* 2014;343:84–7. <https://doi.org/10.1126/science.1247005>
7. Jinek M, Chylinski K, Fonfara I *et al.* A programmable dual RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 2012;337:816–21. <https://doi.org/10.1126/science.1225829>

8. Gilbert LA, Horlbeck MA, Adamson B *et al.* Genome-scale CRISPR-mediated control of gene repression and activation. *Cell* 2014;159:647–61. <https://doi.org/10.1016/j.cell.2014.09.029>
9. Fowler DM, Araya CL, Fleishman SJ *et al.* High resolution mapping of protein sequence–function relationships. *Nat Methods* 2010;7:741–6. <https://doi.org/10.1038/nmeth.1492>
10. Fowler DM, Fields S. Deep mutational scanning: a new style of protein science. *Nat Methods* 2014;11:801–7. <https://doi.org/10.1038/nmeth.3027>
11. Maes S, Deploey N, Peelman F *et al.* Deep mutational scanning of proteins in mammalian cells. *Cell Reports Methods* 2023;3:100641. <https://doi.org/10.1016/j.crmeth.2023.100641>
12. Findlay GM, Daza RM, Martin B *et al.* Accurate classification of BRCA1 variants with saturation genome editing. *Nature* 2018;562:217–22. <https://doi.org/10.1038/s41586-018-0461-z>
13. Buckley M, Terwagne C, Ganner A *et al.* Saturation genome editing maps the functional spectrum of pathogenic VHL alleles. *Nat Genet* 2024;56:1446–55. <https://doi.org/10.1038/s41588-024-01800-z>
14. Radford EJ, Tan H-K, Andersson MHL *et al.* Saturation genome editing of DDX3X clarifies pathogenicity of germline and somatic variation. *Nat Commun* 2023;14:7702. <https://doi.org/10.1038/s41467-023-43041-4>
15. Findlay GM, Boyle EA, Hause RJ *et al.* Saturation editing of genomic regions by multiplex homology-directed repair. *Nature* 2014;513:120–3. <https://doi.org/10.1038/nature13695>
16. Matreyek KA, Stephany JJ, Fowler DM. A platform for functional assessment of large variant libraries in mammalian cells. *Nucleic Acids Res* 2017;45:e102. <https://doi.org/10.1093/nar/gkx183>
17. Clausen L, Voutsinos V, Cagiada M *et al.* A mutational atlas for Parkin proteostasis. *Nat Commun* 2024;15:1541. <https://doi.org/10.1038/s41467-024-45829-4>
18. Grønbaek-Thygesen M, Voutsinos V, Johansson KE *et al.* Deep mutational scanning reveals a correlation between degradation and toxicity of thousands of aspartoacylase variants. *Nat Commun* 2024;15:4026. <https://doi.org/10.1038/s41467-024-48481-0>
19. Komor AC, Kim YB, Packer MS *et al.* Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 2016;533:420–4. <https://doi.org/10.1038/nature17946>
20. Gaudelli NM, Komor AC, Rees HA *et al.* Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature* 2017;551:464–71. <https://doi.org/10.1038/nature24644>
21. Doench JG, Fusi N, Sullender M *et al.* Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat Biotechnol* 2016;34:184–91. <https://doi.org/10.1038/nbt.3437>
22. Coelho MA, Strauss ME, Watterson A *et al.* Base editing screens define the genetic landscape of cancer drug resistance mechanisms. *Nat Genet* 2024;56:2479–92. <https://doi.org/10.1038/s41588-024-01948-8>
23. Cirincione A, Simpson D, Yan W *et al.* A benchmarked, high-efficiency prime editing platform for multiplexed dropout screening. *Nat Methods* 2025;22:92–101. <https://doi.org/10.1038/s41592-024-02502-4>
24. Chen PJ, Hussmann JA, Yan J *et al.* Enhanced prime editing systems by manipulating cellular determinants of editing outcomes. *Cell* 2021;184:5635–52. <https://doi.org/10.1016/j.cell.2021.09.018>
25. Yan J, Oyler-Castrillo P, Ravisankar P *et al.* Improving prime editing with an endogenous small RNA-binding protein. *Nature* 2024;628:639–47. <https://doi.org/10.1038/s41586-024-07259-6>
26. Gould SI, Wuest AN, Dong K *et al.* High-throughput evaluation of genetic variants with prime editing sensor libraries. *Nat Biotechnol* 2024. <https://doi.org/10.1038/s41587-024-02172-9>
27. Jia X, Burugula BB, Chen V *et al.* Massively parallel functional testing of MSH2 missense variants conferring Lynch syndrome risk. *Am J Hum Genet* 2021;108:163–75. <https://doi.org/10.1016/j.ajhg.2020.12.003>
28. Estevam GO, Linossi EM, Macdonald CB *et al.* Conserved regulatory motifs in the juxtamembrane domain and kinase N-lobe revealed through deep mutational scanning of the MET receptor tyrosine kinase domain. *eLife* 2023;12:RP91619.
29. Hayes TK, Aquilanti E, Persky NS *et al.* Comprehensive mutational scanning of EGFR reveals TKI sensitivities of extracellular domain mutants. *Nat Commun* 2024;15:2742. <https://doi.org/10.1038/s41467-024-45594-4>
30. Gelman S, Fahlberg SA, Heinzelman P *et al.* Neural networks to learn protein sequence–function relationships from deep mutational scanning data. *Proc Natl Acad Sci USA* 2021;118:e2104878118. <https://doi.org/10.1073/pnas.2104878118>
31. Livesey BJ, Marsh JA. Using deep mutational scanning to benchmark variant effect predictors and identify disease mutations. *Mol Syst Biol* 2020;16:e9380. <https://doi.org/10.15252/msb.20199380>
32. Rubin AF, Gelman H, Lucas N *et al.* A statistical framework for analyzing deep mutational scanning data. *Genome Biol* 2017;18:150. <https://doi.org/10.1186/s13059-017-1272-5>
33. Ma Y, Zhang J, Yin W *et al.* Targeted AID-mediated mutagenesis (TAM) enables efficient genomic diversification in mammalian cells. *Nat Methods* 2016;13:1029–35. <https://doi.org/10.1038/nmeth.4027>
34. Hanna RE, Hegde M, Fagre CR *et al.* Massively parallel assessment of human variants with base editor screens. *Cell* 2021;184:1064–80.e20. <https://doi.org/10.1016/j.cell.2021.01.012>
35. Cuella-Martin R, Hayward SB, Fan X *et al.* Functional interrogation of DNA damage response variants with base editing screens. *Cell* 2021;184:1081–97. <https://doi.org/10.1016/j.cell.2021.01.041>
36. Hess GT, Tycko J, Yao D *et al.* Methods and applications of CRISPR-mediated base editing in eukaryotic genomes. *Mol Cell* 2017;68:26–43. <https://doi.org/10.1016/j.molcel.2017.09.029>
37. Gapinske M, Luu A, Winter J *et al.* CRISPR-SKIP: programmable gene splicing with single base editors. *Genome Biol* 2018;19:107. <https://doi.org/10.1186/s13059-018-1482-5>
38. Kluesner MG, Lahr WS, Lonetree C *et al.* CRISPR-Cas9 cytidine and adenosine base editing of splice-sites mediates highly-efficient disruption of proteins in primary and immortalized cells. *Nat Commun* 2021;12:2437. <https://doi.org/10.1038/s41467-021-22009-2>
39. Nambiar TS, Baudrier L, Billon P *et al.* CRISPR-based genome editing through the lens of DNA repair. *Mol Cell* 2022;82:348–88. <https://doi.org/10.1016/j.molcel.2021.12.026>
40. Sánchez-Rivera FJ, Diaz BJ, Kasthuber ER *et al.* Base editing sensor libraries for high-throughput engineering and functional analysis of cancer-associated single nucleotide variants. *Nat Biotechnol* 2022;40:862–73. <https://doi.org/10.1038/s41587-021-01172-3>
41. Fiumara M, Ferrari S, Omer-Javed A *et al.* Genotoxic effects of base and prime editing in human hematopoietic stem cells. *Nat Biotechnol* 2024;42:877–91. <https://doi.org/10.1038/s41587-023-01915-4>
42. Grünwald J, Zhou R, Garcia SP *et al.* Transcriptome-wide off-target RNA editing induced by CRISPR-guided DNA base editors. *Nature* 2019;569:433–7. <https://doi.org/10.1038/s41586-019-1161-z>
43. Rees HA, Liu DR. Base editing: precision chemistry on the genome and transcriptome of living cells. *Nat Rev Genet* 2018;19:770–88. <https://doi.org/10.1038/s41576-018-0059-1>
44. Wang Q, Yang J, Zhong Z *et al.* A general theoretical framework to design base editors with reduced bystander effects. *Nat Commun* 2021;12:6529. <https://doi.org/10.1038/s41467-021-26789-5>
45. Collias D, Beisel CL. CRISPR technologies and the search for the PAM-free nuclease. *Nat Commun* 2021;12:555. <https://doi.org/10.1038/s41467-020-20633-y>

46. Walton RT, Christie KA, Whittaker MN *et al.* Unconstrained genome targeting with near-PAMless engineered CRISPR-Cas9 variants. *Science* 2020;368:290–6. <https://doi.org/10.1126/science.aba8853>
47. Christie KA, Guo JA, Silverstein RA *et al.* Precise DNA cleavage using CRISPR-SpRYgests. *Nat Biotechnol* 2023;41:409–16. <https://doi.org/10.1038/s41587-022-01492-y>
48. Vicencio J, Sánchez-Bolaños C, Moreno-Sánchez I *et al.* Genome editing in animals with minimal PAM CRISPR-Cas9 enzymes. *Nat Commun* 2022;13:2601. <https://doi.org/10.1038/s41467-022-30228-4>
49. Krill-Burger JM, Dempster JM, Borah AA *et al.* Partial gene suppression improves identification of cancer vulnerabilities when CRISPR-Cas9 knockout is pan-lethal. *Genome Biol* 2023;24:192. <https://doi.org/10.1186/s13059-023-03020-w>
50. Evers B, Jastrzebski K, Heijmans JPM *et al.* CRISPR knockout screening outperforms shRNA and CRISPRi in identifying essential genes. *Nat Biotechnol* 2016;34:631–3. <https://doi.org/10.1038/nbt.3536>
51. Morgens DW, Deans RM, Li A *et al.* Systematic comparison of CRISPR-Cas9 and RNAi screens for essential genes. *Nat Biotechnol* 2016;34:634–6. <https://doi.org/10.1038/nbt.3567>
52. Stransky N, Ghandi M, Kryukov GV *et al.* Consistency of drug profiles and predictors in large-scale cancer cell line data. *Nature* 2015;528:84–7.
53. Haibe-Kains B, El-Hachem N, Birkbak NJ *et al.* Inconsistency in large pharmacogenomic studies. *Nature* 2013;504:389–93. <https://doi.org/10.1038/nature12831>
54. Weile J, Sun S, Cote AG *et al.* A framework for exhaustively mapping functional missense variants. *Mol Syst Biol* 2017;13:957. <https://doi.org/10.15252/msb.20177908>
55. Zhang T-H, Wu NC, Sun R. A benchmark study on error-correction by read-pairing and tag-clustering in amplicon-based deep sequencing. *BMC Genomics* 2016;17:108. <https://doi.org/10.1186/s12864-016-2388-9>
56. Nachmanson D, Lian S, Schmidt EK *et al.* Targeted genome fragmentation with CRISPR/Cas9 enables fast and efficient enrichment of small genomic regions and ultra-accurate sequencing with low DNA input (CRISPR-DS). *Genome Res* 2018;28:1589–99. <https://doi.org/10.1101/gr.235291.118>
57. Schmitt MW, Kennedy SR, Salk JJ *et al.* Detection of ultra-rare mutations by next-generation sequencing. *Proc Natl Acad Sci USA* 2012;109:14508–13. <https://doi.org/10.1073/pnas.1208715109>
58. Kennedy SR, Schmitt MW, Fox EJ *et al.* Detecting ultralow-frequency mutations by Duplex sequencing. *Nat Protoc* 2014;9:2586–606. <https://doi.org/10.1038/nprot.2014.170>
59. Stoler N, Arbeithuber B, Guiblet W *et al.* Streamlined analysis of duplex sequencing data with Du Novo. *Genome Biol* 2016;17:180. <https://doi.org/10.1186/s13059-016-1039-4>
60. Virtanen P, Gommers R, Oliphant TE *et al.* SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods* 2020;17:261–72. <https://doi.org/10.1038/s41592-019-0686-2>
61. Labun K, Montague TG, Gagnon JA *et al.* CHOPCHOP v2: a web tool for the next generation of CRISPR genome engineering. *Nucleic Acids Res* 2016;44:W272–6. <https://doi.org/10.1093/nar/gkw398>
62. Sangree AK, Griffith AL, Szegetes ZM *et al.* Benchmarking of SpCas9 variants enables deeper base editor screens of BRCA1 and BCL2. *Nat Commun* 2022;13:1318. <https://doi.org/10.1038/s41467-022-28884-7>
63. Neugebauer ME, Hsu A, Arbab M *et al.* Evolution of an adenine base editor into a small, efficient cytosine base editor with low off-target activity. *Nat Biotechnol* 2023;41:673–85. <https://doi.org/10.1038/s41587-022-01533-6>
64. Koh CM. Isolation of genomic DNA from mammalian cells. *Methods Enzymol* 2013;529:161–9.
65. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet j* 2011;17:10–2. <https://doi.org/10.14806/ej.17.1.200>
66. Langmead B, Trapnell C, Pop M *et al.* Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 2009;10:R25. <https://doi.org/10.1186/gb-2009-10-3-r25>
67. Muzellec B, Teleńczuk M, Cabeli V *et al.* PyDESeq2: a python package for bulk RNA-seq differential expression analysis. *Bioinformatics* 2023;39:btad547. <https://doi.org/10.1093/bioinformatics/btad547>
68. Read A, Gao S, Batchelor E *et al.* Flexible CRISPR library construction using parallel oligonucleotide retrieval. *Nucleic Acids Res* 2017;45:e101. <https://doi.org/10.1093/nar/gkx181>
69. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 2012;9:357–9. <https://doi.org/10.1038/nmeth.1923>
70. Arbab M, Shen MW, Mok B *et al.* Determinants of base editing outcomes from target library analysis and machine learning. *Cell* 2020;182:463–80. <https://doi.org/10.1016/j.cell.2020.05.037>
71. Mani R, RP St. Onge, Hartman JL *et al.* Defining genetic interaction. *Proc Natl Acad Sci USA* 2008;105:3461–6. <https://doi.org/10.1073/pnas.0712255105>
72. Xie T, Saleh T, Rossi P *et al.* Conformational states dynamically populated by a kinase determine its function. *Science* 2020;370:eabc2754. <https://doi.org/10.1126/science.abc2754>
73. Nagar B, Hantschel O, Seeliger M *et al.* Organization of the SH3-SH2 unit in active and inactive forms of the c-abl tyrosine kinase. *Mol Cell* 2006;21:787–98. <https://doi.org/10.1016/j.molcel.2006.01.035>
74. Kornev AP, Haste NM, Taylor SS *et al.* Surface comparison of active and inactive protein kinases identifies a conserved activation mechanism. *Proc Natl Acad Sci USA* 2006;103:17783–8. <https://doi.org/10.1073/pnas.0607656103>
75. Rudkin GT, Hungerford DA, Nowell PC. DNA contents of chromosome Ph1 and chromosome 21 in human chronic granulocytic leukemia. *Science* 1964;144:1229–32. <https://doi.org/10.1126/science.144.3623.1229>
76. Nagar B, Hantschel O, Young MA *et al.* Structural basis for the autoinhibition of c-Abl Tyrosine Kinase. *Cell* 2003;112:859–71. [https://doi.org/10.1016/S0092-8674\(03\)00194-6](https://doi.org/10.1016/S0092-8674(03)00194-6)
77. Munoz DM, Cassiani PJ, Li L *et al.* CRISPR screens provide a comprehensive assessment of cancer vulnerabilities but generate false-positive hits for highly amplified genomic regions. *Cancer Discov* 2016;6:900–13. <https://doi.org/10.1158/2159-8290.CD-16-0178>
78. Newberry RW, Leong JT, Chow ED *et al.* Deep mutational scanning reveals the structural basis for α -synuclein activity. *Nat Chem Biol* 2020;16:653–9. <https://doi.org/10.1038/s41589-020-0480-6>
79. Ashkenazy H, Abadi S, Martz E *et al.* ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res* 2016;44:W344–50. <https://doi.org/10.1093/nar/gkw408>
80. Yariv B, Yariv E, Kessel A *et al.* Using evolutionary data to make sense of macromolecules with a ‘face-lifted’ ConSurf. *Protein Sci* 2023;32:e4582. <https://doi.org/10.1002/pro.4582>
81. Ashkenazy H, Erez E, Martz E *et al.* ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res* 2010;38:W529–33. <https://doi.org/10.1093/nar/gkq399>
82. Kalman M, Ben-Tal N. Quality assessment of protein model-structures using evolutionary conservation. *Bioinformatics* 2010;26:1299–307. <https://doi.org/10.1093/bioinformatics/btq114>
83. McWhirter JR, Galasso DL, Wang JY. A coiled-coil oligomerization domain of Bcr is essential for the transforming function of Bcr-Abl oncoproteins. *Mol Cell Biol* 1993;13:7587–95.
84. Smith KM, Jacobi R, Van Etten RA. Autoinhibition of Bcr-Abl through its SH3 domain. *Mol Cell* 2003;12:27–37. [https://doi.org/10.1016/S1097-2765\(03\)00274-0](https://doi.org/10.1016/S1097-2765(03)00274-0)

85. Mayer BJ, Jackson PK, Van Etten RA *et al.* Point mutations in the abl SH2 domain coordinately impair phosphotyrosine binding in vitro and transforming activity *in vivo*. *Mol Cell Biol* 1992;12:609–18.
86. Lamontanara AJ, Georgeon S, Tria G *et al.* The SH2 domain of Abl kinases regulates kinase autophosphorylation by controlling activation loop accessibility. *Nat Commun* 2014;5:5470. <https://doi.org/10.1038/ncomms6470>
87. Schindler T, Bornmann W, Pellicena P *et al.* Structural mechanism for STI-571 inhibition of abelson tyrosine kinase. *Science* 2000;289:1938–42. <https://doi.org/10.1126/science.289.5486.1938>
88. Woodring PJ, Litwack ED, O'Leary DDM *et al.* Modulation of the F-actin cytoskeleton by c-abl tyrosine kinase in cell spreading and neurite extension. *J Cell Biol* 2002;156:879–92. <https://doi.org/10.1083/jcb.200110014>
89. Zheng R, Wei W, Liu S *et al.* The FABD domain is critical for the oncogenicity of BCR/ABL in chronic myeloid leukaemia. *Cell Commun Signal* 2024;22:314. <https://doi.org/10.1186/s12964-024-01694-8>
90. Kin Y, Li G, Shibuya M *et al.* The Dbl homology domain of BCR is not a simple spacer in P210BCR-ABL of the Philadelphia chromosome. *J Biol Chem* 2001;276:39462–8. <https://doi.org/10.1074/jbc.M105484200>
91. Reckel S, Gehin C, Tardivon D *et al.* Structural and functional dissection of the DH and PH domains of oncogenic bcr-abl tyrosine kinase. *Nat Commun* 2017;8:2101. <https://doi.org/10.1038/s41467-017-02313-6>
92. Richter MF, Zhao KT, Eton E *et al.* Phage-assisted evolution of an adenine base editor with improved Cas domain compatibility and activity. *Nat Biotechnol* 2020;38:883–91. <https://doi.org/10.1038/s41587-020-0453-z>
93. Marquart KF, Allam A, Janjuha S *et al.* Predicting base editing outcomes with an attention-based deep learning algorithm trained on high-throughput target library screens. *Nat Commun* 2021;12:5114. <https://doi.org/10.1038/s41467-021-25375-z>
94. Pallaseni A., Peets EM, Koeppl J *et al.* Predicting base editing outcomes using position-specific sequence determinants. *Nucleic Acids Res* 2022;50:3551–64. <https://doi.org/10.1093/nar/gkac161>
95. Jin H, Zhang C, Zwahlen M *et al.* Systematic transcriptional analysis of human cell lines for gene expression landscape and tumor representation. *Nat Commun* 2023;14:5417. <https://doi.org/10.1038/s41467-023-41132-w>
96. Geraghty RJ, Capes-Davis A, Davis JM *et al.* Guidelines for the use of cell lines in biomedical research. *Br J Cancer* 2014;111:1021–46. <https://doi.org/10.1038/bjc.2014.166>