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

AUTHORS

Ivan Sokirniy[†,1] Haider Inam[†,1,2], David McCandlish[3], and Justin Pritchard[1,2]

† The first authors contributed equally to this work

1. Huck Institute for the Life Sciences, University Park, PA 16802
2. Department of Biomedical Engineering, University Park, PA 16802
3. Simons Center for Quantitative Biology, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724

ABSTRACT

Variant annotation is a crucial objective in mammalian functional genomics. Deep Mutational Scanning (DMS) 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 base editing measurements can annotate variant function and the extent of downstream experimental validation required. This study presents the first direct comparison of DMS and BE in the same lab and cell line. Results indicate that focusing on the most likely edits and highest efficiency sgRNAs enhances 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 the variants created in the pool, rather than sgRNA abundance, can recover high-quality variant annotation measurements in multiplexed pools. Taken together, our data show an impressive degree of correlation between base editor data and gold standard deep mutational scanning.

Correspondence to:

Justin R. Pritchard PhD

Department of Biomedical Engineering

Pennsylvania State University

211 Wartik Lab

University Park, PA 16802

Jrp94@psu.edu

INTRODUCTION

A major goal of mammalian functional genomics is to understand the genotype-phenotype relationship (1). While loss-of-function screens in mammalian cells rose to prominence with the invention of RNAi based screening tools (2–4) in the early 2000's and later CRISPR/CRISPRi (5–8), the high throughput annotation of coding variants in mammalian cell lines is a more recent innovation (9, 10).

A leading technology in the field of high throughput variant annotation is a technique known as deep mutational scanning (DMS) (10). Deep mutational scanning libraries tend to involve 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 (11–13), or into a safe harbor "landing pad"(14–16). While DMS is capable of providing comprehensive measurements of variant effects and has been used beyond variant annotations in efforts to understand protein structure-function (10) and in the computational prediction of variant function (17, 18), 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.

As a functional genomic technology, base editing is in its early stages. Base editing (BE) screens use nCas9 to target a deaminase to a specific site in the genome and generate transition mutations (C>T or A>G) (19, 20). BE screens use a surrogate measure of genotype by sequencing the guide RNA (gRNA) sequence, which allows BE screens to measure phenotypes across the genome (21–23). 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 (21, 24–27). However, base editor screens also present certain challenges. The primary challenges are: 1. BE efficiency (only a portion of individual cells

harboring an sgRNA are likely to be edited and some cell lines do not edit well (28)), 2. 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 (29, 30)), 3. 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 (31, 32)), and 4. protospacer adjacent motif (PAM) requirements (PAMs limit where sgRNAs targets (33) and PAM-less variants appear to have decreased efficiency (34–36)). These issues have led the field to view BE screens as a method for initially identify 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 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 (37–39). 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 (40, 41).

Here we perform the first direct comparison between base editing and deep mutational scanning 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. 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 (42).

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 (SM) 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 >1000X, on to 15-cm LB agar plates with ampicillin. After 48 hours at 30°C, colonies were scrapped off the agar and plasmid DNA was extracted using Omega Bio-Tek's E.Z.N.A. Plasmid DNA Midi kit. HEK293Ts were transfected with 35 ug of SM ABL library and 10 ug of helper plasmids (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 (Cytivia SH30027.02). After 36 hours viral RPMI media was used to infect BaF3s, in the presence of 1 ug/mL mouse IL-3 (peprotech 213-13) and 6 ug/mL polybrene, at a low multiplicity of infection. After another 36 hours, BaF3s were maintained in RMPI with 1 ug/mL IL-3. Infected cells were enriched by fluorescence-activated cell sorting 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 less than 90%, then viable cells were enriched by Ficol-Paque (Cytiva).

Single Strand Consensus

High quality genomic DNA was extracted by Monarch Genomic DNA Purification Kit. Then an adapted duplex sequencing workflow was used to determine accurate variant distributions (42). For each time point 20 ug of genomic DNA specifically was digested just outside of the mutagenized region of ABL kinase cDNA using Cas9. After end-repair and A-tailing, UMI ligation was used to label single molecules of DNA. Then NGS indices were added by PCR to help deconvolute samples after pooling. Oligo enrichment and PCR were used to further increase the abundance of the mutagenized region in the pooled samples.

One-hundred-fifty-nucleotide paired-end sequencing of the UMIs and mutagenized region was done on an Illumina NovaSeq 6000. For each sample, dunovo 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 5 mismatches would undergo variant calling and annotation using a custom R script. Briefly, for each alignment, variants were converted from the MDZ read tag and normalized to read depth at that position. Mutant growth rates were calculated using exponential growth equation and the mutant allele frequency (MAF):

$$growth\ rate = \ln\left(\frac{MAF_1 \times Count_1}{MAF_0 \times Count_0}\right) \div (Time_1 - Time_0),$$

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 (43). Z-score cutoff for

DMS data was determined by fit mean and standard deviation of the wild type like component of the mixture distribution.

Base Editor Library Preparation, Screen, and Sequencing

BCR-ABL tiling gRNA sequences were generated by CHOP-CHOP (44) 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 was made by deleting the U6 sgRNA cassette from pRDA_479 (Addgene #179099) (45) using NEB KLD (NEB #M0554S).

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 BaF3s were EGFP+. Then they were infected with either ABE8e SpG or CBEd SpG and selected with 1 ug/mL puromycin for 5 days. In a 15-cm dish HEK293Ts were transfected with 60 ug of the BCR-ABL and control sgRNA libraries and 40 ug helper plasmids (1:1:1:1) using calcium phosphate. The next day DMEM media was changed to RPMI. Independent infections of pUltra BCR-ABL WT ABE8e BaF3s were done at low multiplicity of infection and >500X coverage. Transfected HEK23Ts 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. Genomic DNA was extracted using phenol-chloroform (46) and quantified by Qubit. Staggered PCR was used to extract sgRNA sequences from 10 ug of genomic DNA, as described previously (47) with two modifications: custom forward staggered primers were used (Supplemental Table 1) and PCR amplicons were gel extracted using Omega Bio-Tek Gel Extraction kit.

sgRNA Analysis

Guide RNA amplicons were trimmed with CutAdapt (48) to remove the U6 promoter and gRNA scaffold. The remaining sequences were aligned to a reference list of guides using Bowtie (49). Guide counts were established by Counter from the Collections python package. pyDESeq2 (50) was used to

determine the fold change in guides between time points. The fold change was converted to growth rates using the formula above assuming wild-type growth rate. Z-score cutoff for ABE data was based on the human targeting (AAVS1, CCR5, and ROSA26) control sgRNA.

Verification Screen, Sequencing, and Analysis

A pool of 20 guides were synthesized by IDT and cloned into lenti-sgRNA hygro vector using golden gate sites (47, 51). After making BCR-ABL WT ABE8e BaF3 cells as above, IL-3 was returned to the RPMI media. The lentivirus library of sgRNAs was made by calcium phosphate transfection of HEK293T. Two independent infections of 1.5 million BCR-ABL WT ABE8e BaF3 cells grew for 3 days before the start of 1mg/mL hygromycin selection. After 6 days, hygromycin was removed, a pellet was saved for sgRNA sequencing, and the cells recovered for 2 days. After the 2-day recovery, IL-3 was removed, and a pellet was saved for mutation sequencing. Six and eight days after IL-3 withdrawal pellets were saved for mutation sequencing. Nineteen days after IL-3 removal pellets were saved for sgRNA sequencing. Mutation and sgRNA sequencing were performed as described above.

A custom Python script was used to determine the frequency of in phase or *cis* mutations made by ABE. Briefly, UMI_tools (52) was used to move 24 nucleotide UMI in to read header. At least 5 million paired end reads were aligned to the ABL kinase after 12 nucleotides were hard clipped from the 5' end using Bowtie2 (53). To deduplicate barcodes, each sample's barcodes was grouped based on alignment with an allowance of one mismatch using Bowtie (49). Only the most frequent UMI in each group continued to the rest of the analysis. Mismatches with a quality score greater than 30 were grouped by UMI. A mutation was called and counted if the mismatch is observed in more than 80% of reads in the UMI, and the UMI contains more than 2 reads. Fold change was determined by pyDESeq2 (50) and converted to growth rate using WT growth rate. Background mutation frequency was set based the 95th percentile of non-(A > G) and non-(T > C) mutant frequency outside of the 2 to 12 nucleotide sgRNA

editing window. All A > G mutations within the targeting strand and 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 were predicted by BE-HIVE (54). Only edits with a frequency greater than 0.05 and sgRNAs with completely matched DMS measurement were allowed. The growth rate of an sgRNA is sum of the relative growth rates of edited and un-edited cells:

$$sgRNA\ growth\ rate = (edit \times edit\ growth\ rate) + (undedited\ proportion \times 0.055)$$

In the case of edits across multiple amino acids, a null model of mutant interactions predicts the growth rate of the multi amino acid mutant as the product of individual mutants divided by the *WT* growth rate multiplied by the *WT* growth rate (0.055 $hour^{-1}$) (55). In other words, two deleterious *cis* mutations are expected to be more deleterious than the individual mutants. If there was an epistatic interaction, the growth rates of the double mutant would differ from the null model.

RESULTS AND DISCUSSION

DMS Data is 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: 1. It is an important oncogene that is of interest to applied and basic research groups. 2. Its structure has been solved many times in many conformations (56, 57). 3. There are years of experiments that can be drawn from to gain confidence in the resultant data (57–59).

To perform the deep mutational scan (DMS) in BCR-ABL, we designed a saturation mutagenesis library spanning amino acid residues 242 to 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 contained 1441 AA variants and was transduced into Ba/F3s in the presence of IL-3. After infection, fluorescence activated cell sorting (FACS) was used to enrich for infected cells. Following recovery from sorting, the Ba/F3 cells were screened for six days with and 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 (42). 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 (Figure 1A). Initially, the library contained ninety seven percent of all possible single residue changes from AA 242-320.

When the mutant in BCR-ABL impairs kinase function, then that cell will exhibit a reduced growth rate. After sequencing, we found that mutant growth rates form a bimodal distribution of counts, consistent with measured distributions of fitness effects in other DMS screen (12, 15, 16). One of the peaks in this bimodal distribution of growth rates centered at

approximately 0.02 $hour^{-1}$ and constituted the set of deleterious mutations, while the other distribution mode was centered at the known *WT* growth rate of 0.055 $hour^{-1}$ and harbored AAs that are "wild type like fitness" (Figure 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 "wild type like" distribution. This specific cutoff is used to be consistent with prior literature in the sgRNA community and to draw a line that can be approximated in both studies using approximately analogous statistical and experimental criteria (Figure 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 $hour^{-1}$) (Figure 1C). Moreover, the systematic insertion of prolines during DMS studies causes a discernible "proline band" in high quality studies (12, 60).We clearly observe a proline band in our data(Figure 1C). Beyond these critical depletion signals, a known flexible and non-conserved region in ABL1 (i.e. residues 262-267 between the β2 and β3 strands) did not deplete.

More systematically, functionally important residues are typically evolutionarily conserved and their mutagenesis tends to be deleterious in DMS studies (10, 16). Consistent with this, there is a -0.74 correlation ($p < 0.001$) between the mean mutant growth rate at a residue and the evolutionary conservation scores of that residue from ConSurf (56, 61–63)(Figure 1C, bottom row of heatmap, Supplemental Figure 1). Moreover, from a biophysical

perspective, solvent accessible residues tend to be more tolerant of mutagenesis and are less conserved (64). We observe this trend in Figure 1C, 1D and Supplementary Figure 1.

The ABL N-lobe is composed of conserved structural elements that include the alpha-helix called the αC-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 3 conserved Glycines show strong depletion phenotypes. Additionally, the alternating banding pattern in anti-parallel β -strands 1 and 2 (residues 242-262, Figure 1D) is explained by essential residues facing towards the substrate pocket (blue residues, Figure 1D). Moreover, the αC-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, Figure 1E), while the outward facing residues show tolerance to mutations (red, Figure 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 Base Editing 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 3115 sgRNAs across the entire BCR-ABL cDNA, including regions beyond the original DMS library. The BE phenotype was determined by the change in sgRNA abundance

following exponential growth in cell culture (Figure 2A). Looking across the entire length of the protein (Figure 2B), we performed a sliding window estimate of the proportion of strongly negative Z-scores across the protein (see Methods). 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, Dbl homology (DH) domain, pleckstrin 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 (65, 66) , SH2 (66–68), Kinase (69), and FAB domains (70, 71) show significant depletion. While the role of the DH domain is not well understood, our base editor data suggests that it is important for growth factor independence (Figure 2B) (72, 73). Finally to call depletion "hits" we used a –2 Z-score based on the distribution of negative-control guides to be consistent with Hannah et al. (22) (Figure 2C). Applying this criteria yields 387 hits, or approximately 12.4% of all BCR-ABL sgRNAs (consistent with others (22, 45)). Most notably, there is a strong over-representation of depletion phenotypes for guides that target the ABL 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 buried region of the αC-helix produce deleterious phenotypes. (Figure 2D).

Side by side comparison of ABE predicted edits and DMS Screens

One naïve approach to determine the degree of correlation between our ABE screen and our deep mutational scan is to compare the growth rates of the sgRNAs (and their predicted variants) to the growth rates of the mutations that they target as determined by the

deep mutational scan. Of the 3115 original sgRNAs screened (spanning over 2,000 AAs), 118 sgRNAs spanned the DMS region in the N-lobe of the ABL kinase. Of these 118 sgRNAs, 80sgRNAs had complete variant information in the DMS screen (Figure 2D). The predicted edits of sgRNAs predominantly occur within an editing window (31, 74). While the editing window is in the non-targeting strand, we (and others) refer to the window relative to the guide sequence. Initially we focused our analysis on a broad window of predicted variants between positions 2 and 12 in these 80 sgRNAs (74, 75). Plotting the sgRNA data from Figure 2D against the DMS data from Figure 1 yields the plot in Figure 3A. In Figure 3A, each dot is an individual variant, so that a single sgRNA appears as a row of dots corresponding to each of its possible edits. It is also important to note in Figure 3A that the dynamic range of the Y-axis is reduced (there is an ~ 0.045-0.06 $hour^{-1}$ Y-axis measurement range for base editing versus ~0.01 to 0.06 $hour^{-1}$ for DMS on the X-axis). This is likely because an sgRNA read can count edited and unedited cells. This naïve plotting/analysis of the combined data yields a modest, but significant, Pearson correlation of 0.35 (Figure 3A). While interesting, this level of correlation means that direct annotation of individual variants from a raw sgRNA experiment could be unwise without follow up experiments with single sgRNAs that are tested individually.

Given this modest correlation when considering all possible putative variants from the sgRNAs, we examined the hypothesis that the predicted editing window from 2-12bp is too wide and includes too many low efficiency edits. As described by other groups (22, 75, 76) most of the base edits (76%) occur between a narrower range of 4 and 8 nucleotides. Applying this reduced window size as a “likely” edits filter improves the Pearson correlation from 0.35 to 0.52 and improves odds ratio of hits versus non-hits to 7.1 (Figure 3B). The cost of this improved

agreement is the removal of 26 variants from our analysis that might have been correctly annotated as deleterious (Figure 3A,B and Supplemental Figure 2).

After applying the “likely” edits filter to generate Figure 3B we still identify 13 false negative and 22 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 (47). Thus, we examined the consequences of adding an additional filter requiring an sgRNA efficiency score greater than 50 (Figure 3C, Supplemental Figure 2S) (47). This simple additional cutoff further improves the Pearson correlation between sgRNA predicted edit growth rates and DMS data to 0.59 and eliminates 9 of the 13 remaining false negative variants (Supplemental Figure 2S).

Having dramatically reduced the issues with false negative sgRNAs, we turned our attention to the false positive variants in Figure 3B (lower right quadrant, pink points). Twenty-two false positive variants were present among the likely edits of Figure 3B. These 22 variants are predicted to be made by 15 individual sgRNAs. Interestingly, 10 of these 15 sgRNAs are multi-edit sgRNAs that are predicted to also make true positives (blue variant dots). Compared to the wild-type like sgRNAs in the upper right quadrant (orange dots) portion of Figure 3B, false negative sgRNAs are significantly confounded by the presence of the true positive (blue dot) deleterious mutations potentially created by the same sgRNA (Fisher’s exact test, OR = 5, p = 0.018). 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 to 8 nucleotide editing region. The addition of this “single likely edit” filter enhances the correlation to 0.64 (p-value < 0.001) and the OR for hits to 44 (p-value < 0.001) (Figure 3D). For single 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, the use

of single edits eliminates 14 out of a total of 39 original sgRNA hits (Supplementary Figure 3). For a more detailed accounting of the filters and the numbers of both sgRNAs and variants that are filtered see Supplementary Figures 2 and 3.

Since filtering out sgRNAs based on sgRNA efficiency, edit probability, and the number of edits results in more confident single variant annotation at the expense of the total number of sgRNA hits we wanted to figure out how to get correlated data from multi-edit sgRNAs (Figure 3D). At a first pass, it seems like one might create variant-level interpretations for multi-edit sgRNA hits by predicting the ensemble of sgRNA edits and their abundance in the population using machine learning algorithms like BE-HIVE (54). However, while BE-HIVE can predict the ensemble of mutations (Figure 3E), it can't predict the proportion of the measured phenotype that is attributable to those mutations. We reasoned that as a first step, we needed to know if multi-edit sgRNA dynamics might be predicted from the combined dynamics of the polyclonal edits they produce. To answer this, we obtained the set of BE-HIVE variant predictions for each single sgRNA (a vector of the predicted proportions of all the variants in the polyclonal population). Then we calculated the inner product of this vector of allele frequencies with the corresponding vector of gold standard DMS growth rates (see Methods). This produced a simple estimate of the expected multi-edit sgRNA growth rate. While this estimate does not lead to better variant annotation directly from sgRNA data, it does ask: "how much of the phenotype of an sgRNA is from the editing window, and how much is due to off target editing?". In other words, reproducing the sgRNA growth rates from the BE-HIVE weighted DMS growth rates can tell us whether the sgRNA dynamics are due to "in window" editing effects or not. We observed a robust correlation of 0.57 (p-value < 0.001) between the

predicted and actual sgRNA growth rates (Figure 3F) and an RMSE of a linear model of 0.0016. Because this result was a significant correspondence, we reasoned that enough of the sgRNA signal was due to on-target editing that it might be possible to perform variant validation experiments with medium-throughput pools of sgRNAs by directly sequencing the edited variants from the genomic DNA of the pool.

Deep Sequencing of ABE mutant pools annotates sgRNA induced variants in a multiplexed manner

Because our BE-HIVE regression results suggested in-window editing drives multi-edit sgRNA growth rates (Figure 3F), we reasoned that validating the variants produced by sgRNA hits might be sped up via the direct sequencing of the variants in genomic DNA during pooled validation experiments. In this experiment, instead of validating sgRNA-variant relationships in a one-by-one manner, we propose to make medium-throughput validation pools of sgRNAs. The variants made in validation pools would then be directly measured by UMI corrected sequencing of the editing target instead of indirectly measuring the sgRNA.

To test this idea, we selected 20 sgRNAs in the N-lobe with a range of predicted fitness measurements. We examined our ability to correctly determine whether a variant is deleterious, and we aimed to explain sgRNA measurements using the direct sequencing of the ensemble of individual variants that are created by an sgRNA (Figure 4A).

The measured sgRNA growth rates from this pool showed a strong correlation ($r=0.79$, $p < 0.001$) with the high-throughput screen from Figure 2, confirming robust sgRNA phenotypes. This suggested that the validation pool measurements were similar to the original

high throughput measurements and added confidence to this new dataset. To further examine the relationship between the growth rates of the edited variants and the growth rates of the sgRNAs that generate them, we plotted the correlation between sgRNA growth rates and the measured mutant growth rates (Figure 4B). This yielded a high correlation of r=0.75 (Figure 4B). Thus, medium scale measurements of variants correlate with measurements of sgRNAs. This validates that the fitness effects of the sgRNAs are largely due to on-target editing.

We also examined the correlation between the variant growth rate measurements in the medium-throughput pool and the original DMS measurements in Figure 1. Here we found that there was a high correlation (Pearson's r=0.77) between the direct variant measurements and the gold standard DMS measurements from Figure 1 (Figure 4C). This suggests that the direct sequencing of edited variants in resequencing pools can approach the quality of gold standard DMS data. Among the 19 variants we detected, only F283L showed a spurious depletion phenotype that deviated from observed DMS measurements. This single false positive variant could be due to off-target editing confounding the variant interpretation or be misannotated in the DMS screen. Interestingly, Figure 4C only examined the *single* mutant edits that we detected by deep sequencing, but we also identified double mutants (*in cis*) in multi edit sgRNA windows. These double mutants exhibited growth rates that did not suggest epistasis (i.e. the fitness defects of two deleterious edits combined independently and predictably under a product model of mutant interaction (55) (see methods)) (Supplemental Table 2). Strong correlations between direct variant measurements and the presence of only a single false positive edit suggests that when variants are sequenced directly from a base editor pool, the resultant fitness measurements are highly concordant with gold standard data.

CONCLUSION

In conclusion, we find that while DMS data has remarkable depth and richness, sgRNA libraries can be cloned and screened rapidly. When comparing the data from the two methods in the same cell line and controlling for the effects of differences due to genomic context we find that the agreement between the two methods can be impressively high considering the large methodological differences and the potential off-target effects of base editors. This high agreement suggests that in the subset of sgRNAs with single predicted edits, variant annotation might proceed with reasonably high confidence directly from base editor screens. In complex cases of sgRNAs with multiple-edits and double mutations, targeted resequencing of mutant pools is a promising approach to directly validate sgRNA predicted variant phenotypes and to improve the throughput of variant annotation across the genome.

ACKNOWLEDGEMENTS

We would like to acknowledge members of the Pritchard lab for constructive criticisms of the manuscript, and the Penn State Flow Cytometry Core for help with sorting cells. Protocol diagrams were made with BioRender. 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.

CONFLICT OF INTEREST

JRP is a co-founder of RedAce Bio. JRP was a co-founder and consultant for Theseus Pharmaceuticals. JRP held equity in Theseus Pharmaceuticals. JRP holds equity in MOMA therapeutics and RedAce Bio. JRP has consulted/consults for MOMA therapeutics, Curie.Bio, Third Rock Ventures, Takeda Pharmaceuticals, Galapagos Pharmaceuticals, and Roche/Genentech. JRP has received honoraria and travel expenses from Roche/Genentech, Third Rock Ventures, and Theseus Pharmaceuticals.

Figure 1. Functional landscape of ABL N-Lobe (A) Schematic of deep mutational scanning. After lentiviral integration of EGFP-P2A-BCR-ABL, BaF3s 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 ABL N-lobe is bimodal. Two skewed gaussians are fit to determine the variation in deleterious (blue) and "wild-type like" (orange) mutations. The dotted line represents a -2 Z-score threshold with respect to the "wild type-like" distribution. (C) Heatmap of the growth rate of mutations at each position in ABL1 N-lobe. Black dot represents WT positions. Missing data is 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"). Tolerance/sensitivity to mutagenesis is projected onto two key structural features of the ABL N-lobe (PDB 6XR6): the (D) anti-parallel beta-sheet, and the (E) αC-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.

Figure 2. Adenosine Base Editor screen of full length of BCR-ABL. (A) Schematic of adenosine base editor screen. Three days after infection with sgRNA library, EGFP-P2A-BCR-ABL1 ABE8e BaF3 were selected with 1mg/mL hygromycin for 6 days and pelleted. Guides were PCR amplified and sequenced. (B) Sliding window of 40 sgRNAs estimate of the proportion of BCR-ABL sgRNAs that drop out more than a Z-score of -4 of the non-targeting control sgRNA growth rate. (C) Kernel density estimate of growth rate distributions of non-targeting control, and BCR-ABL sgRNA libraries. Dashed grey line represents a –2 Z-score of the targeting control. (D) Lollipop plot displays dropout of each sgRNA across the ABL kinase domain. Dashed grey line represents a –2 Z-score of the targeting control.

Figure 3. Comparison of Adenosine Base Editor sgRNA growth rate and their respective mutation growth rates from Deep Mutational Scan. Each dot represents a mutation an sgRNA is predicted to make. Dashed lines represent –2 Z-score of the non-deleterious distribution and non-targeting 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 does not deplete in their respective screen, in other words both are non-deleterious, then they are colored orange. If they both are deleterious, or true positive, then they are colored blue. If an sgRNA deletes, but the predicted edit does not deplete, a false positive, then the dot is colored in the pink. 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 edit 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 small screen of 20 sgRNAs targeting ABL kinase, where the edits and sgRNA are sequenced after IL-3 withdrawal. (B) Comparison of measured sgRNA and single amino acid edit growth rates. If the mutation is deleterious or non-deleterious in the DMS data, then it is marked by a X or closed circle, respectively. If the different mutations are made by the same sgRNA, then they are connected by a grey line. (C) Represents a direct comparison between the measured growth rate of single amino acid mutants during the DMS and 20 sgRNA ABE screens.

Figure 1

Sokirniy & Inam

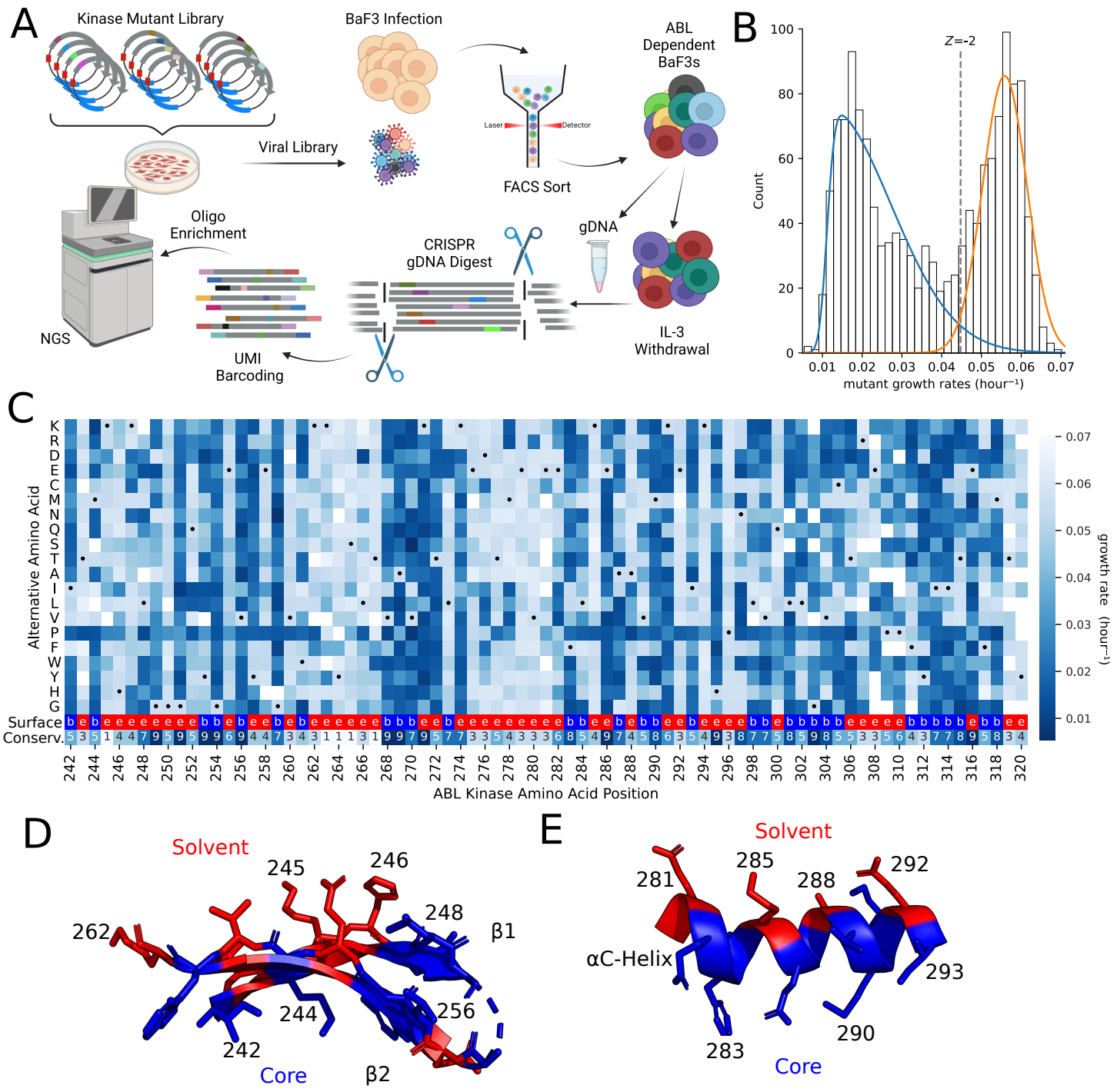


Figure 2 Sokirniy & Inam

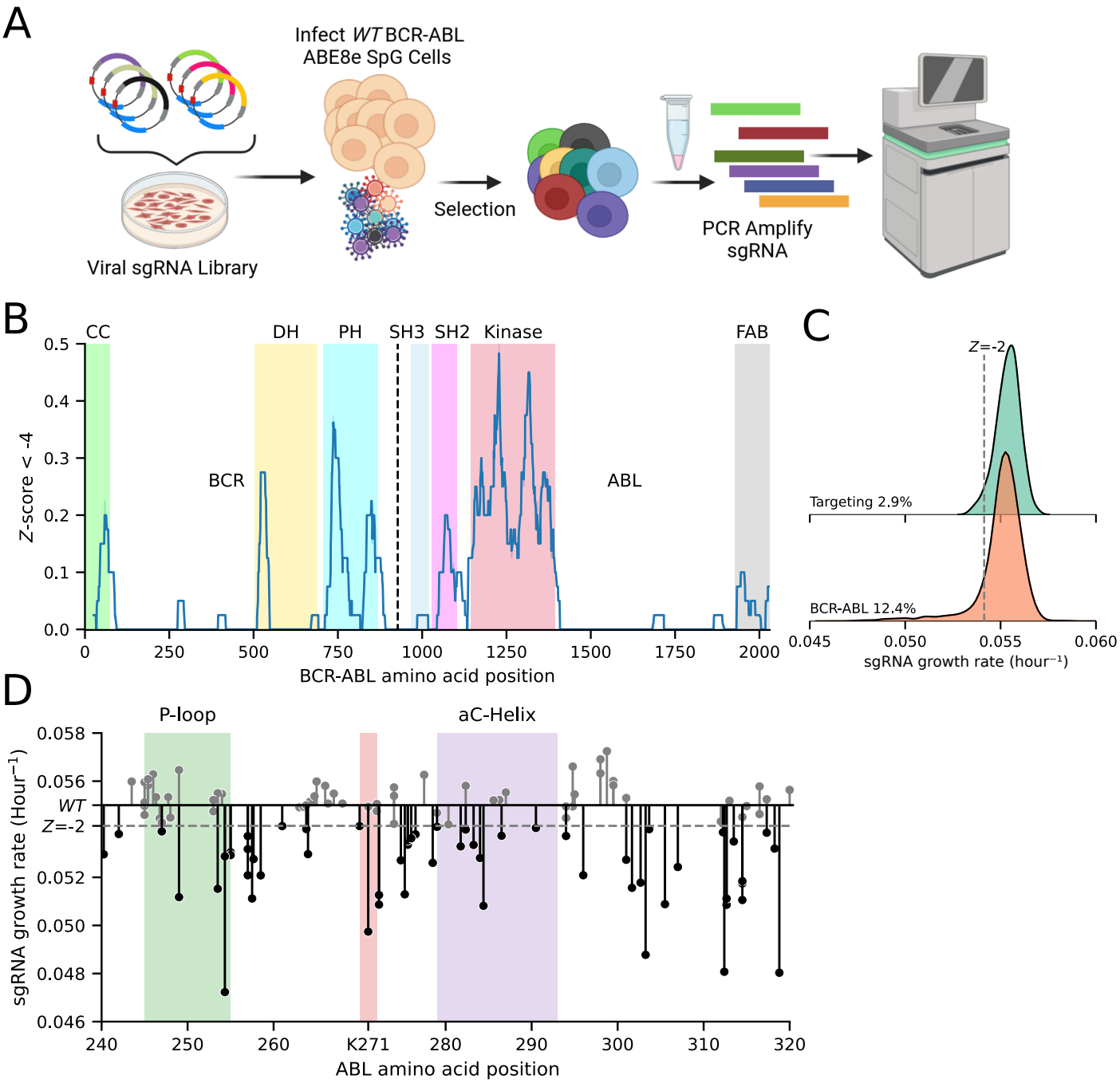


Figure 3

Sokirniy & Inam

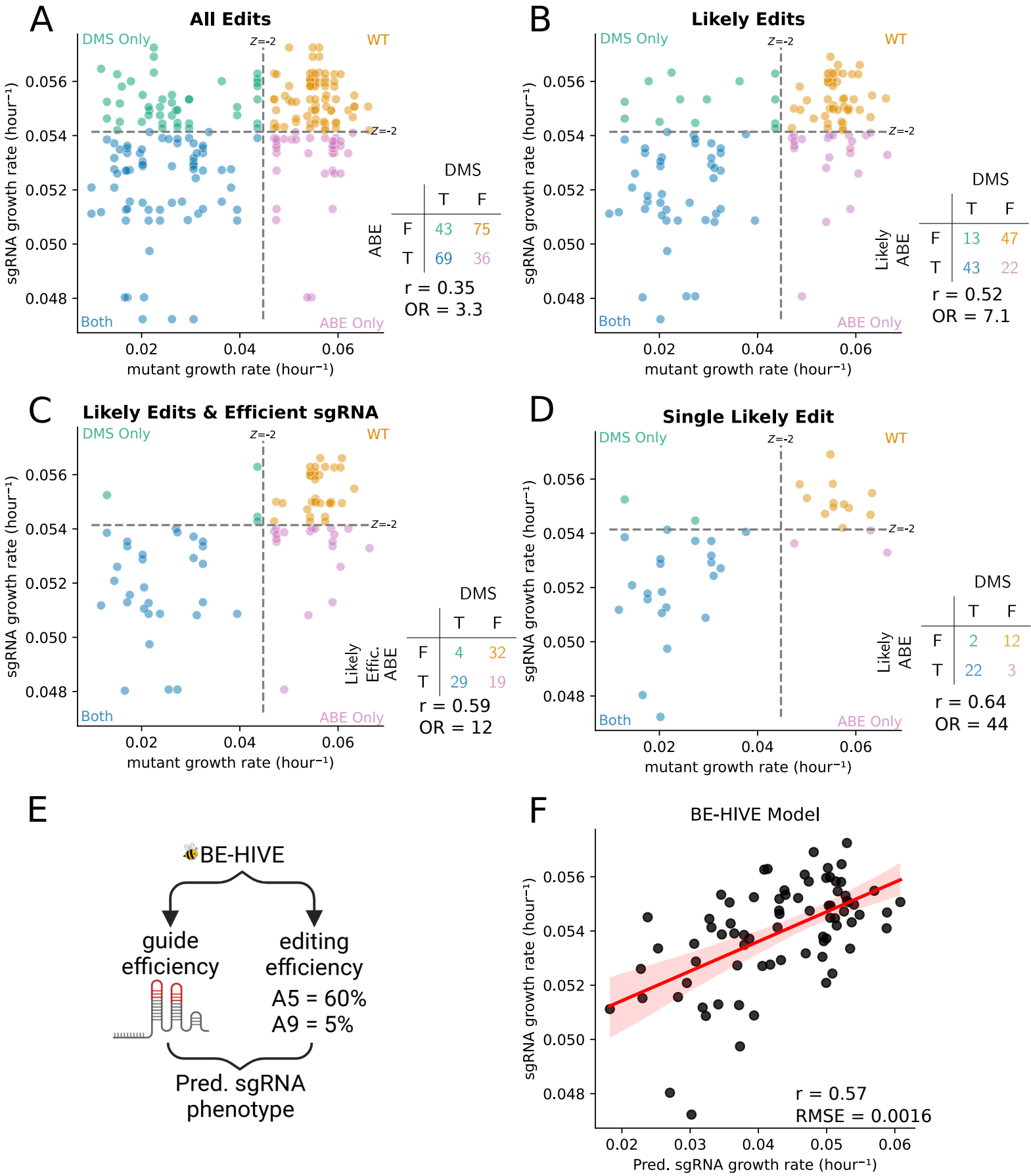
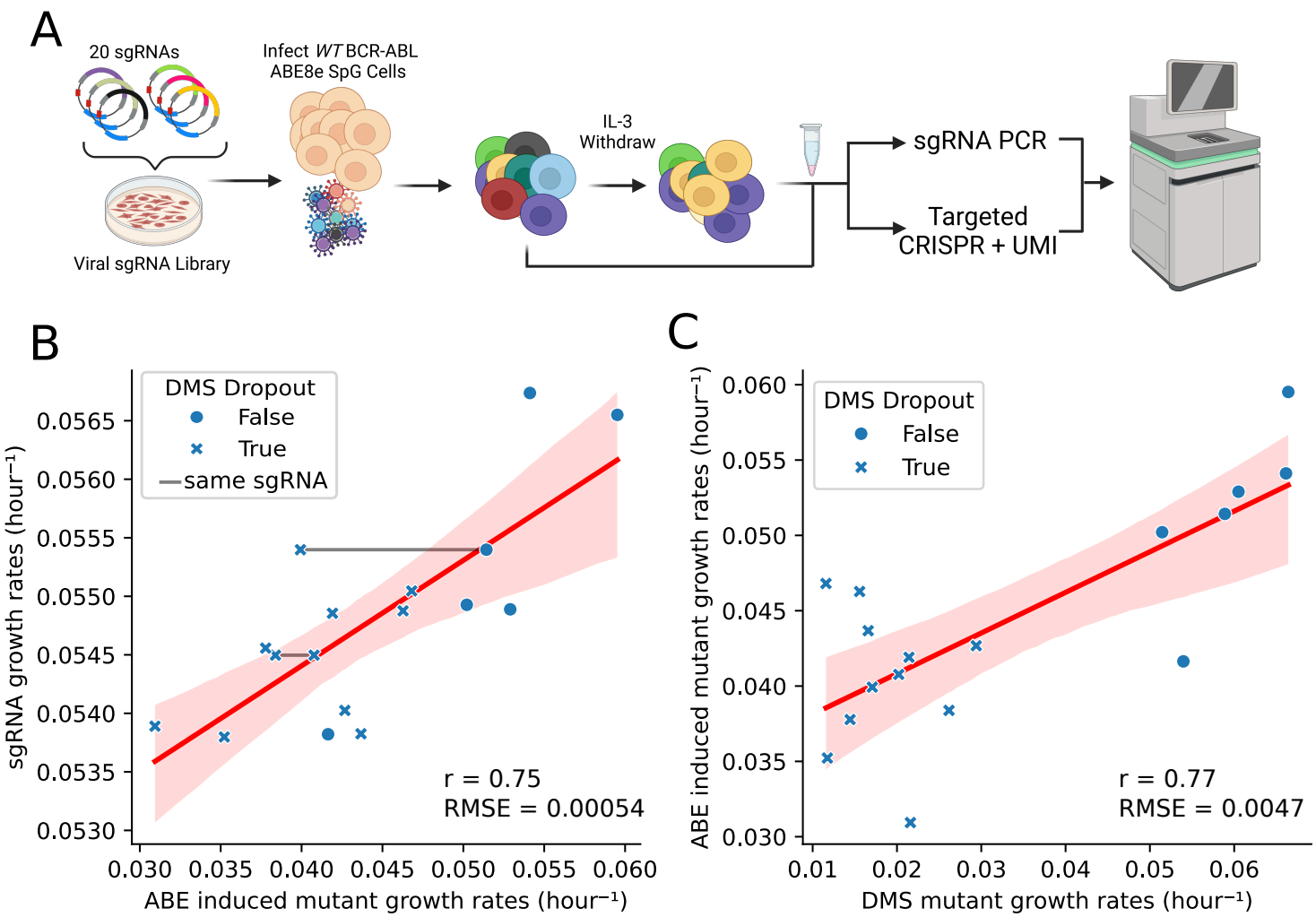


Figure 4

Sokirniy & Inam



REFERENCES

1. Kinney,J.B. and McCandlish,D.M. (2019) Massively Parallel Assays and Quantitative Sequence-Function Relationships. *Annu Rev Genomics Hum Genet*, **20**, 99–127.

2. Elbashir,S.M., Harborth,J., Lendeckel,W., Yalcin,A., Weber,K. and Tuschl,T. (2001) Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature*, **411**, 494–498.

3. Root,D.E., Hacohen,N., Hahn,W.C., Lander,E.S. and Sabatini,D.M. (2006) Genome-scale loss-of-function screening with a lentiviral RNAi library. *Nat Methods*, **3**, 715–719.

4. Fire,A., Xu,S., Montgomery,M.K., Kostas,S.A., Driver,S.E. and Mello,C.C. (1998) Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature*, **391**, 806–811.

5. Ran,F.A., Hsu,P.D., Wright,J., Agarwala,V., Scott,D.A. and Zhang,F. (2013) Genome engineering using the CRISPR-Cas9 system. *Nat Protoc*, **8**, 2281–2308.

6. Shalem,O., Sanjana,N.E., Hartenian,E., Shi,X., Scott,D.A., Mikkelson,T., Heckl,D., Ebert,B.L., Root,D.E., Doench,J.G., *et al.* (2014) Genome-Scale CRISPR-Cas9 Knockout Screening in Human Cells. *Science*, **343**, 84–87.

7. Jinek,M., Chylinski,K., Fonfara,I., Hauer,M., Doudna,J.A. and Charpentier,E. (2012) A programmable dual RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*, **337**, 816–821.

8. Gilbert,L.A., Horlbeck,M.A., Adamson,B., Villalta,J.E., Chen,Y., Whitehead,E.H., Guimaraes,C., Panning,B., Ploegh,H.L., Bassik,M.C., *et al.* (2014) Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell*, **159**, 647–661.

9. Fowler,D.M., Araya,C.L., Fleishman,S.J., Kellogg,E.H., Stephany,J.J., Baker,D. and Fields,S. (2010) High Resolution Mapping of Protein Sequence–Function Relationships. *Nat Methods*, **7**, 741–746.

10. Fowler,D.M. and Fields,S. (2014) Deep mutational scanning: a new style of protein science. *Nat Methods*, **11**, 801–807.

11. Jia,X., Burugula,B.B., Chen,V., Lemons,R.M., Jayakody,S., Maksutova,M. and Kitzman,J.O. (2021) Massively parallel functional testing of MSH2 missense variants conferring Lynch syndrome risk. *Am J Hum Genet*, **108**, 163–175.

12. Estevam,G.O., Linossi,E.M., Macdonald,C.B., Espinoza,C.A., Michaud,J.M., Coyote-Maestas,W., Collisson,E.A., Jura,N. and Fraser,J.S. (2023) Conserved regulatory motifs in the juxtamembrane domain and kinase N-lobe revealed through deep mutational scanning of the MET receptor tyrosine kinase domain. *eLife*, **12**.

13. Hayes,T.K., Aquilanti,E., Persky,N.S., Yang,X., Kim,E.E., Brenan,L., Goodale,A.B., Alan,D., Sharpe,T., Shue,R.E., *et al.* (2024) Comprehensive mutational scanning of EGFR reveals TKI sensitivities of extracellular domain mutants. *Nat Commun*, **15**, 2742.

14. Matreyek,K.A., Stephany,J.J. and Fowler,D.M. (2017) A platform for functional assessment of large variant libraries in mammalian cells. *Nucleic Acids Research*, **45**, e102.

15. Clausen,L., Voutsinos,V., Cagiada,M., Johansson,K.E., Grønbæk-Thygesen,M., Nariya,S., Powell,R.L., Have,M.K.N., Oestergaard,V.H., Stein,A., *et al.* (2024) A mutational atlas for Parkin proteostasis. *Nat Commun*, **15**, 1541.

16. Grønbæk-Thygesen,M., Voutsinos,V., Johansson,K.E., Schulze,T.K., Cagiada,M., Pedersen,L., Clausen,L., Nariya,S., Powell,R.L., Stein,A., *et al.* (2024) Deep mutational scanning reveals a correlation between degradation and toxicity of thousands of aspartoacylase variants. *Nat Commun*, **15**, 4026.

17. Gelman,S., Fahlberg,S.A., Heinzelman,P., Romero,P.A. and Gitter,A. (2021) Neural networks to learn protein sequence–function relationships from deep mutational scanning data. *Proceedings of the National Academy of Sciences*, **118**, e2104878118.

18. Livesey,B.J. and Marsh,J.A. (2020) Using deep mutational scanning to benchmark variant effect predictors and identify disease mutations. *Mol Syst Biol*, **16**, e9380.

19. Komor,A.C., Kim,Y.B., Packer,M.S., Zuris,J.A. and Liu,D.R. (2016) Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature*, **533**, 420–424.

20. Gaudelli,N.M., Komor,A.C., Rees,H.A., Packer,M.S., Badran,A.H., Bryson,D.I. and Liu,D.R. (2017) Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature*, **551**, 464–471.

21. Ma,Y., Zhang,J., Yin,W., Zhang,Z., Song,Y. and Chang,X. (2016) Targeted AID-mediated mutagenesis (TAM) enables efficient genomic diversification in mammalian cells. *Nat Methods*, **13**, 1029–1035.

22. Hanna,R.E., Hegde,M., Fagre,C.R., DeWeirdt,P.C., Sangree,A.K., Szegletes,Z., Griffith,A., Feeley,M.N., Sanson,K.R., Baidi,Y., *et al.* (2021) Massively parallel assessment of human variants with base editor screens. *Cell*, **184**, 1064-1080.e20.

23. Cuella-Martin,R., Hayward,S.B., Fan,X., Chen,X., Huang,J.-W., Taglialatela,A., Leuzzi,G., Zhao,J., Rabadan,R., Lu,C., *et al.* (2021) Functional interrogation of DNA damage response variants with base editing screens. *Cell*, **184**, 1081-1097.e19.

24. Hess,G.T., Tycko,J., Yao,D. and Bassik,M.C. (2017) Methods and applications of CRISPR-mediated base editing in eukaryotic genomes. *Mol Cell*, **68**, 26–43.

25. Gapinske,M., Luu,A., Winter,J., Woods,W.S., Kostan,K.A., Shiva,N., Song,J.S. and Perez-Pinera,P. (2018) CRISPR-SKIP: programmable gene splicing with single base editors. *Genome Biol*, **19**, 107.

26. Kluesner,M.G., Lahr,W.S., Lonetree,C., Smeester,B.A., Qiu,X., Slipek,N.J., Claudio Vázquez,P.N., Pitzen,S.P., Pomeroy,E.J., Vignes,M.J., *et al.* (2021) CRISPR-Cas9 cytidine and adenosine base editing of splice-sites mediates highly-efficient disruption of proteins in primary and immortalized cells. *Nat Commun*, **12**, 2437.

27. Nambiar,T.S., Baudrier,L., Billon,P. and Ciccia,A. (2022) CRISPR-Based Genome Editing Through the Lens of DNA Repair. *Mol Cell*, **82**, 348–388.

28. Sánchez-Rivera,F.J., Diaz,B.J., Kastenhuber,E.R., Schmidt,H., Katti,A., Kennedy,M., Tem,V., Ho,Y.-J., Leibold,J., Paffenholz,S.V., *et al.* (2022) Base editing sensor libraries for high-throughput engineering and functional analysis of cancer-associated single nucleotide variants. *Nat Biotechnol*, **40**, 862–873.

29. Fiumara,M., Ferrari,S., Omer-Javed,A., Beretta,S., Albano,L., Canarutto,D., Varesi,A., Gaddoni,C., Brombin,C., Cugnata,F., *et al.* (2024) Genotoxic effects of base and prime editing in human hematopoietic stem cells. *Nat Biotechnol*, **42**, 877–891.

30. Grünewald,J., Zhou,R., Garcia,S.P., Iyer,S., Lareau,C.A., Aryee,M.J. and Joung,J.K. (2019) Transcriptome-wide off-target RNA editing induced by CRISPR-guided DNA base editors. *Nature*, **569**, 433–437.

31. Rees,H.A. and Liu,D.R. (2018) Base editing: precision chemistry on the genome and transcriptome of living cells. *Nat Rev Genet*, **19**, 770–788.

32. Wang,Q., Yang,J., Zhong,Z., Vanegas,J.A., Gao,X. and Kolomeisky,A.B. (2021) A general theoretical framework to design base editors with reduced bystander effects. *Nat Commun*, **12**, 6529.

33. Collias,D. and Beisel,C.L. (2021) CRISPR technologies and the search for the PAM-free nuclease. *Nat Commun*, **12**, 555.

34. Walton,R.T., Christie,K.A., Whittaker,M.N. and Kleinstiver,B.P. (2020) Unconstrained Genome Targeting with near-PAMless Engineered CRISPR-Cas9 Variants. *Science*, **368**, 290–296.

35. Christie,K.A., Guo,J.A., Silverstein,R.A., Doll,R.M., Mabuchi,M., Stutzman,H.E., Lin,J., Ma,L., Walton,R.T., Pinello,L., *et al.* (2023) Precise DNA Cleavage using CRISPR-SpRYgests. *Nat Biotechnol*, **41**, 409–416.

36. Vicencio,J., Sánchez-Bolaños,C., Moreno-Sánchez,I., Brena,D., Vejnar,C.E., Kukhtar,D., Ruiz-López,M., Cots-Ponjoan,M., Rubio,A., Melero,N.R., *et al.* (2022) Genome editing in animals with minimal PAM CRISPR-Cas9 enzymes. *Nat Commun*, **13**, 2601.

37. Krill-Burger,J.M., Dempster,J.M., Borah,A.A., Paolella,B.R., Root,D.E., Golub,T.R., Boehm,J.S., Hahn,W.C., McFarland,J.M., Vazquez,F., *et al.* (2023) Partial gene suppression improves identification of cancer vulnerabilities when CRISPR-Cas9 knockout is pan-lethal. *Genome Biology*, **24**, 192.

38. Evers,B., Jastrzebski,K., Heijmans,J.P.M., Grernrum,W., Beijersbergen,R.L. and Bernards,R. (2016) CRISPR knockout screening outperforms shRNA and CRISPRi in identifying essential genes. *Nat Biotechnol*, **34**, 631–633.

39. Morgens,D.W., Deans,R.M., Li,A. and Bassik,M.C. (2016) Systematic comparison of CRISPR-Cas9 and RNAi screens for essential genes. *Nat Biotechnol*, **34**, 634–636.

40. Stransky,N., Ghandi,M., Kryukov,G.V., Garraway,L.A., Amzallag,A., Pruteanu-Malinici,I., Haber,D.A., Ramaswamy,S., Benes,C.H., Lehár,J., *et al.* (2015) Consistency of drug profiles and predictors in large-scale cancer cell line data. *Nature*, **528**, 84–87.

41. Haibe-Kains,B., El-Hachem,N., Birkbak,N.J., Jin,A.C., Beck,A.H., Aerts,H.J.W.L. and Quackenbush,J. (2013) Inconsistency in large pharmacogenomic studies. *Nature*, **504**, 389–393.

42. Nachmanson,D., Lian,S., Schmidt,E.K., Hipp,M.J., Baker,K.T., Zhang,Y., Tretiakova,M., Loubet-Senear,K., Kohrn,B.F., Salk,J.J., *et al.* (2018) 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.*, **28**, 1589–1599.

43. Virtanen,P., Gommers,R., Oliphant,T.E., Haberland,M., Reddy,T., Cournapeau,D., Burovski,E., Peterson,P., Weckesser,W., Bright,J., *et al.* (2020) SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods*, **17**, 261–272.

44. Labun,K., Montague,T.G., Gagnon,J.A., Thyme,S.B. and Valen,E. (2016) CHOPCHOP v2: a web tool for the next generation of CRISPR genome engineering. *Nucleic Acids Res*, **44**, W272–W276.

45. Sangree,A.K., Griffith,A.L., Szegletes,Z.M., Roy,P., DeWeirdt,P.C., Hegde,M., McGee,A.V., Hanna,R.E. and Doench,J.G. (2022) Benchmarking of SpCas9 variants enables deeper base editor screens of BRCA1 and BCL2. *Nat Commun*, **13**, 1318.

46. Koh,C.M. (2013) Isolation of genomic DNA from mammalian cells. *Methods Enzymol*, **529**, 161–169.

47. Doench,J.G., Fusi,N., Sullender,M., Hegde,M., Vaimberg,E.W., Donovan,K.F., Smith,I., Tothova,Z., Wilen,C., Orchard,R., *et al.* (2016) Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat Biotechnol*, **34**, 184–191.

48. Martin,M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, **17**, 10–12.

49. Langmead,B., Trapnell,C., Pop,M. and Salzberg,S.L. (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*, **10**, R25.

50. Muzellec,B., Teleńczuk,M., Cabeli,V. and Andreux,M. (2023) PyDESeq2: a python package for bulk RNA-seq differential expression analysis. *Bioinformatics*, **39**, btad547.

51. Read,A., Gao,S., Batchelor,E. and Luo,J. (2017) Flexible CRISPR library construction using parallel oligonucleotide retrieval. *Nucleic Acids Res*, **45**, e101.

52. Smith,T., Heger,A. and Sudbery,I. (2017) UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. *Genome Res*, **27**, 491–499.

53. Langmead,B. and Salzberg,S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods*, **9**, 357–359.

54. Arbab,M., Shen,M.W., Mok,B., Wilson,C., Matuszek,Ż., Cassa,C.A. and Liu,D.R. (2020) Determinants of Base Editing Outcomes from Target Library Analysis and Machine Learning. *Cell*, **182**, 463-480.e30.

55. Mani,R., St.Onge,R.P., Hartman,J.L., Giaever,G. and Roth,F.P. (2008) Defining genetic interaction. *Proceedings of the National Academy of Sciences*, **105**, 3461–3466.

56. Xie,T., Saleh,T., Rossi,P. and Kalodimos,C.G. (2020) Conformational states dynamically populated by a kinase determine its function. *Science*, **370**, eabc2754.

57. Nagar,B., Hantschel,O., Seeliger,M., Davies,J.M., Weis,W.I., Superti-Furga,G. and Kuriyan,J. (2006) Organization of the SH3-SH2 unit in active and inactive forms of the c-Abl tyrosine kinase. *Mol Cell*, **21**, 787–798.

58. Rudkin,G.T., Hungerford,D.A. and Nowell,P.C. (1964) DNA Contents of Chromosome Ph1 and Chromosome 21 in Human Chronic Granulocytic Leukemia. *Science*, **144**, 1229–1232.

59. Nagar,B., Hantschel,O., Young,M.A., Scheffzek,K., Veach,D., Bornmann,W., Clarkson,B., Superti-Furga,G. and Kuriyan,J. (2003) Structural Basis for the Autoinhibition of c-Abl Tyrosine Kinase. *Cell*, **112**, 859–871.

60. Newberry,R.W., Leong,J.T., Chow,E.D., Kampmann,M. and DeGrado,W.F. (2020) Deep Mutational Scanning Reveals the Structural Basis for α-Synuclein Activity. *Nat Chem Biol*, **16**, 653–659.

61. Ashkenazy,H., Abadi,S., Martz,E., Chay,O., Mayrose,I., Pupko,T. and Ben-Tal,N. (2016) ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res*, **44**, W344-350.

62. Yariv,B., Yariv,E., Kessel,A., Masrati,G., Chorin,A.B., Martz,E., Mayrose,I., Pupko,T. and Ben-Tal,N. (2023) Using evolutionary data to make sense of macromolecules with a 'face-lifted' ConSurf. *Protein Sci*, **32**, e4582.

63. Ashkenazy,H., Erez,E., Martz,E., Pupko,T. and Ben-Tal,N. (2010) ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res*, **38**, W529-533.

64. Kalman,M. and Ben-Tal,N. (2010) Quality assessment of protein model-structures using evolutionary conservation. *Bioinformatics*, **26**, 1299–1307.

65. McWhirter,J.R., Galasso,D.L. and Wang,J.Y. (1993) A coiled-coil oligomerization domain of Bcr is essential for the transforming function of Bcr-Abl oncoproteins. *Mol Cell Biol*, **13**, 7587–7595.

66. Smith,K.M., Yacobi,R. and Van Etten,R.A. (2003) Autoinhibition of Bcr-Abl through Its SH3 Domain. *Mol Cell*, **12**, 27–37.

67. Mayer,B.J., Jackson,P.K., Van Etten,R.A. and Baltimore,D. (1992) Point mutations in the abl SH2 domain coordinately impair phosphotyrosine binding in vitro and transforming activity in vivo. *Mol Cell Biol*, **12**, 609–618.

68. Lamontanara,A.J., Georgeon,S., Tria,G., Svergun,D.I. and Hantschel,O. (2014) The SH2 domain of Abl kinases regulates kinase autophosphorylation by controlling activation loop accessibility. *Nat Commun*, **5**, 5470.

69. Schindler,T., Bornmann,W., Pellicena,P., Miller,W.T., Clarkson,B. and Kuriyan,J. (2000) Structural Mechanism for STI-571 Inhibition of Abelson Tyrosine Kinase. *Science*, **289**, 1938–1942.

70. Woodring,P.J., Litwack,E.D., O'Leary,D.D.M., Lucero,G.R., Wang,J.Y.J. and Hunter,T. (2002) Modulation of the F-actin cytoskeleton by c-Abl tyrosine kinase in cell spreading and neurite extension. *J Cell Biol*, **156**, 879–892.

71. Zheng,R., Wei,W., Liu,S., Zeng,D., Yang,Z., Tang,J., Tan,J., Huang,Z. and Gao,M. (2024) The FABD domain is critical for the oncogenicity of BCR/ABL in chronic myeloid leukaemia. *Cell Communication and Signaling*, **22**, 314.

72. Kin,Y., Li,G., Shibuya,M. and Maru,Y. (2001) The Dbl Homology Domain of BCR Is Not a Simple Spacer in P210BCR-ABL of the Philadelphia Chromosome *. *Journal of Biological Chemistry*, **276**, 39462–39468.

73. Reckel,S., Gehin,C., Tardivon,D., Georgeon,S., Kükenshöner,T., Löhr,F., Koide,A., Buchner,L., Panjkovich,A., Reynaud,A., *et al.* (2017) Structural and functional dissection of the DH and PH domains of oncogenic Bcr-Abl tyrosine kinase. *Nat Commun*, **8**, 2101.

74. Richter,M.F., Zhao,K.T., Eton,E., Lapinaite,A., Newby,G.A., Thuronyi,B.W., Wilson,C., Koblan,L.W., Zeng,J., Bauer,D.E., *et al.* (2020) Phage-assisted evolution of an adenine base editor with improved Cas domain compatibility and activity. *Nat Biotechnol*, **38**, 883–891.

75. Marquart,K.F., Allam,A., Janjuha,S., Sintsova,A., Villiger,L., Frey,N., Krauthammer,M. and Schwank,G. (2021) Predicting base editing outcomes with an attention-based deep learning algorithm trained on high-throughput target library screens. *Nat Commun*, **12**, 5114.

76. Pallaseni,A., Peets,E.M., Koeppel,J., Weller,J., Vanderstichele,T., Ho,U.L., Crepaldi,L., van Leeuwen,J., Allen,F. and Parts,L. (2022) Predicting base editing outcomes using position-specific sequence determinants. *Nucleic Acids Res*, **50**, 3551–3564.