Prime Editing-Based Resistance Profiling of ABL1 Variants against Kinase Inhibitors

Jung, Yusang

Department of Medicine

Graduate School

Yonsei University

Prime Editing-Based Resistance Profiling of ABL1 Variants against Kinase Inhibitors

Advisor Hyongbum Henry Kim

A Dissertation Submitted
to the Department of Medicine
and the Committee on Graduate School
of Yonsei University in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy in Medical Science

Jung, Yusang

June 2025

Prime Editing-Based Resistance Profiling of ABL1 Variants against Kinase Inhibitors

This Certifies that the Dissertation of Jung, Yusang is Approved

Committee Chair	**Hae-Jeong Park**
Committee Member	**Hyongbum Henry Kim**
Committee Member	**Sung-Rae Cho**
Committee Member	**Tae-Min Kim**
Committee Member	**Dongwoo Chae**

Department of Medicine
Graduate School
Yonsei University
June 2025

ACKNOWLEDGEMENTS

I’m very grateful to Younggwang Kim and Seungho Lee for helpful discussions and technical advice. I also would like to thank Younghye Kim, Seonmi Park, and Gahyun Baek for assisting with the experiments. This work was supported, in part, by a National Research Foundation of Korea grant funded by the Korean government (MSIT) (2022R1A3B1078084 (H.H.K.), 2018R1A5A2025079 (H.H.K.)), the Bio and Medical Technology Development Program of the National Research Foundation (NRF) funded by the Korean government (MSIT) (2022M3A9E4017127 (H.H.K.), 2022M3A9F3017506 (H.H.K.) and (RS-2023-00260968) (H.H.K.)), the Bio & Medical Technology Development Program of the National Research Foundation (NRF) funded by the Korean government (MSIT) (NRF-2021R1A2C3011992), the Yonsei Signature Research Cluster Program of 2023-22- 0012 (H.H.K.), the Brain Korea 21 FOUR Project for Medical Science (Yonsei University College of Medicine), the SNUH Kun-hee Lee Child Cancer and Rare Disease Project, Republic of Korea (22B-000-0101 (H.H.K.)), the Yonsei Fellow Program, funded by Lee Youn Jae, and the Korea Health Technology R&D Project funded by the Ministry of Health and Welfare, Republic of Korea (HI21C1314 (H.H.K.)); a grant of the MD-PhD/Medical Scientist Training Program (Y.J.) through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea.

TABLE OF CONTENTS

LIST OF FIGURES

ABSTRACT

Prime editing-based resistance profiling of ABL1 variants against kinase inhibitors

In chronic myeloid leukemia (CML), therapeutic strategies frequently involve tyrosine kinase inhibitors (TKIs) aimed at suppressing the constitutively active ABL1 kinase resulting from the BCR-ABL1 fusion, but some patients become resistant, often due to ABL1 mutations. The effects of many of the ABL1 kinase mutations on resistance against each TKI are unclear, which makes treatment selection challenging. Here, I used prime editing to generate 97% (= 2,802/2,892) of all possible single nucleotide variants in the sequence encoding the ABL1 kinase domain, which encode 98% (= 1,954/1,998) of all possible corresponding single amino acid variants, and evaluated their effects on resistance to the five TKIs in clinical use in CML-relevant K562 cells. This comprehensive resistance map will help in drug selection for CML patients based on ABL1 mutations, facilitating precision medicine.

Key words : chronic myeloid leukemia; prime editing; variants of uncertain significance; ABL1; single nucleotide variants; tyrosine kinase inhibitors

1. INTRODUCTION

The majority of chronic myeloid leukemia (CML) cases involve the BCR-ABL1 fusion, leading to aberrant activation of ABL1 kinase. Tyrosine kinase inhibitors (TKIs), including imatinib, exert their therapeutic effects in CML by suppressing the hyperactive ABL1 kinase. Most TKIs, with the notable exception of the recently introduced asciminib, function through ATP-competitive inhibition, occupying the ATP-binding site of ABL1 and thereby blocking its activation (1). Despite the significant success of TKIs in CML treatment (2-4), a subset of patients fails to attain optimal responses, necessitating treatment optimization (5). The primary mechanism of resistance to TKIs involves point mutations in the ABL1 kinase domain (for brevity, hereafter ABL1 kinase)(6). Notably, 30 to 40 percent of patients with suboptimal responses harbor low-level resistance mutations, which may become dominant without alterations in the treatment strategy(7). As a result, both the National Comprehensive Cancer Network(8) and European LeukemiaNet(9) recommend mutation screening in cases of treatment failure. Despite ongoing research, a substantial proportion of ABL1 kinase mutations have yet to be functionally characterized and are currently designated as variants of uncertain significance (VUS) with respect to TKI resistance. Consequently, establishing an extensive resistance landscape covering the full spectrum of potential ABL1 kinase mutations against the five clinically utilized TKIs would be highly informative. This includes more precise TKI selection for individual patients and the development of enhanced strategies to combat drug resistance.

While clinical studies involving a substantial number of patients can identify resistant variants, this approach is time-consuming and costly (7). The expenses further escalate as the evaluation encompasses an increasing number of variant and TKI combinations. Therefore, using clinical studies to assess the resistance of all potential ABL1 kinase variants against the therapeutically used TKIs would be prohibitively expensive. A more cost-effective alternative for determining the resistance of ABL1 variants involves utilizing cultured CML-derived cells encoding these variants at their endogenous genomic sites.

In this investigation, prime editing(10) was employed to generate and assess the functional impact of 97% (= 2,802/2,892) of all potential single nucleotide variants (SNVs) in the ABL1 gene. These

variants encompassed 98% (= 1,954/1,998) of all conceivable single amino acid variants (SAAVs). The evaluation focused on their effects concerning resistance to the five clinically utilized TKIs, namely imatinib, nilotinib, bosutinib, ponatinib, and asciminib. The study was conducted in K562 cells, initially derived from a CML patient. By addressing the existing gap in TKI resistance profiles for ABL1 variants, this research may substantially contribute to more informed and tailored TKI selection in clinical settings, thereby optimizing treatment efficacy.

2. MATERIALS AND METHODS

2.1. Cell lines and culture

K562 (CRL-3343) and all cell lines derived from K562 were cultured in RPI 1640 medium (Welgene) supplemented with 2.05mM L-glutamine, 25mM HEPES, 1mM sodium pyruvate, 10% fetal bovine serum (RDT), and 1% penicillin-streptomycin (Gibco). HEK293T cells (ATCC) were maintained separately in DMEM (Gibco) supplemented with 10% fetal bovine serum (RDT). Maintenance of all cell lines involved keeping them by maintaining cultures at sub-confluent levels (<80%), incubating at 37°C in a humidified atmosphere with 5% CO_2, and routinely passaging every 48 hours. The concentrations of polybrene, puromycin, and blasticidin were as follows, unless otherwise specified: 4 μg/ml; 0.8 μg/ml; 30 μg/ml, respectively.

2.2. Lentivirus production

Lentiviral vectors were produced using a previously established protocol. HEK293T cells were seeded into 150-mm culture dishes at a density of 1×10^7 cells per dish approximately 18 hours before transfection. Transfection was carried out using Lipofectamine 3000 (Invitrogen) according to the manufacturer's instructions. Six hours after transfection, the medium was replaced with fresh culture medium. Viral supernatants were harvested 48 hours post-transfection, clarified by centrifugation at 1,000 × g for 3 minutes, passed through a 0.45 μm low protein-binding Millex-HV membrane filter (Millipore), aliquoted, and stored at −80 °C until use.

2.3. Transduction of lentivirus

K562 or KCL22 cells were transduced in 12-well plates at a density of 1.5×10^6 cells per well, using 1.5 mL of culture medium supplemented with polybrene and 0.5 mL of thawed lentiviral particles. The plates were centrifuged at 640 × g for 2 hours at 32 °C to enhance viral infection, followed by incubation at 37 °C for an additional 6 hours. Subsequently, cells from three wells were combined and transferred into a 150-mm dish containing fresh medium. For high-throughput experiments, transduction efficiency was evaluated 48 hours later by measuring the proportion of RFP-positive cells via fluorescence microscopy.

2.4. Preparation of empty vector for the engineered-prime editing guide RNA (epegRNA) delivery system

To construct the pLenti-gRNA-Puro-RFP plasmid, the lentiviral vector pLenti_gRNA-Puro (Addgene #84752) was first linearized using XhoI and MluI restriction enzymes. An mRFP cassette, sourced from the Cas9-2A-mRFP-2A-Puro vector (Toolgen), was then inserted via ligation. Following this, epegRNA sequences were cloned into the resulting construct to enable epegRNA expression.

2.5. Construction of MMR-deficient cell lines expressing prime editor

To establish cell lines stably expressing PEmax, K562 and KCL22 cells were transduced with a lentiviral construct encoding PE2max-P2A-BSD (Addgene #191102), which had been concentrated fivefold using VivaSpin columns (Sigma-Aldrich). Following transduction, cells were subjected to blasticidin selection for seven days. Successful integration and expression of PEmax were verified via PCR and Sanger sequencing. To further generate PE4max-expressing cells, a second lentiviral vector harboring hMLH1dn–P2A–eGFP (Addgene #191104) was introduced into the PEmax-stable cells, and GFP-positive populations were isolated by fluorescence-activated cell sorting one week post-transduction. For MSH6 gene knockout, both K562-PE4max and KCL22-PE4max cells were co-transfected with the base editor plasmid pCMV-BE4max-3×HA (Addgene #112096) and a pLenti-gRNA-Puro-RFP construct carrying an sgRNA targeting MSH6 (5’-gAATCCcAAGCCCACGTTAG-3’), using Lipofectamine 2000. The transfected plates were centrifuged at 640 × g for 2 hours at 32 °C,

then incubated for an additional 6 hours at 37 °C. Two days later, RFP-expressing cells were sorted via FACS and clonally expanded. Clones harboring homozygous premature stop codons in exon 4 of the MSH6 gene, confirmed by deep sequencing, were designated K562-PE4K and KCL22-PE4K.

2.6. Design of library for epegRNA abundance-based screening (Library A)

Initially, I performed targeted deep sequencing on the protein-coding sequence of ABL1 exons 4-9 in K562 cells to validate its alignment with the reference genome sequence registered at the National Library of Medicine (NCBI). Subsequently, an oligo pool was devised, encompassing sequences for epegRNAs designed to induce all possible SNVs in exons 4-9 of ABL1. Additionally, five sets of control guide RNAs were included in the pool: epegRNAs designed to generate SNVs previously demonstrated to confer TKI resistance (positive controls), sgRNAs, sham-editing epegRNAs, non-targeting epegRNAs, and epegRNAs for knock-out of essential genes. Furthermore, a library of prime editing guide RNA (pegRNA) -encoding and target sequence pairs from a prior study was incorporated(11).

Two BsmBI cut sites were integrated between the spacer and the RTT-PBS region of the epegRNA to facilitate subsequent insertion of a scaffold sequence during the cloning process. Except for the pegRNAs in the paired library, all other epegRNAs feature a tevopreQ1 motif and an optimal linker chosen using pegLIT(12) attached to their 3' ends. To mitigate potential PCR length bias resulting from variations in the length of the RTT-PBS region in different epegRNA sequences during the oligo pool amplification step, a random sequence was added to each oligo, adjusting the length to 235 nucleotides.

2.7. Design of libraries for endogenous region sequencing-based screening (Libraries E4-E9)

The design of libraries for endogenous region sequencing followed a similar approach to that employed for library A in the epegRNA abundance screening. In these libraries, three epegRNAs were chosen for inducing each SNV based on scores from the DeepPrime-FT: DLD1-PE2max model. To implement endogenous region sequencing, the 'SynonymousPE'

module of the Python package 'GenET' was utilized to introduce additional synonymous mutations in the epegRNA RTTs. Priority was given to introducing synonymous mutations in the protospacer adjacent motif (PAM) (GG) sequence of the epegRNA. If introducing synonymous mutations at the PAM position was not feasible, they were introduced in the left homology arm (LHA) or, if necessary, in the right homology arm (RHA), further from the intended edit. If a newly introduced synonymous mutation was in the RHA, the RTT's length was extended by the distance between the synonymous mutation and the intended edit, irrespective of its effect on the PAM site. When multiple synonymous mutations meeting the criteria were available, the one closest to the edit position was selected, minimizing changes to the RTT's GC content. If the synonymous mutation occurred within the intron where it could potentially affect splicing (within 5 nt of the exon), a different position for the synonymous mutation was chosen. epegRNAs for which introducing additional synonymous mutations within the RTT was not possible were excluded, and those with the next highest DeepPrime scores were selected.

In the positive control group of epegRNAs used in library A and those designed to induce additional SNVs showing log-fold changes exceeding 0.5 and P values less than 0.05 in epegRNA abundance-based analysis, the sequences were made five times more abundant in the oligo pool. Two BsmBI cut sites were integrated between the spacer and the RTT-PBS region of each epegRNA to allow for subsequent insertion of the scaffold sequence during cloning. A tevopreQ1 motif and an optimal linker selected using pegLIT(12) were attached to the 3’ end of each epegRNA. To address potential PCR length bias arising from variations in the lengths of the RTT-PBS region in different epegRNA sequences during the oligo pool amplification step, a random sequence was added to each oligo, adjusting the length to 199 nucleotides, mirroring the method used for library A.

2.8. Plasmid library preparation

Plasmid libraries were constructed using a modified two-step cloning approach based on a previously described protocol. In the first step, oligonucleotide pools were amplified by PCR and integrated into BsmBI-digested pLenti-gRNA-Puro-RFP vectors through HiFi DNA Assembly. The second step involved the insertion of an enhanced sgRNA scaffold using an additional round of BsmBI-mediated cloning. To eliminate any remaining vector backbone

impurities, the resulting plasmid library was subjected to a final BsmBI digestion, followed by purification through isopropanol precipitation.

2.9. Assessing the resistance of cells with a single SNV

Top-ranked epegRNAs targeting Y253H, E255V, V299L, and T315I—selected using the DeepPrime algorithm—were each cloned into BsmBI-digested pLenti-gRNA-Puro-RFP backbones. The assembly strategy involved four distinct fragments: spacer oligonucleotides, an optimized SpCas9 sgRNA scaffold, a 3′ RTT-PBS extension specific to each epegRNA, and a tevopreQ1-modified poly-T tail, all designed with compatible overhangs for seamless ligation. Lentiviral particles were subsequently generated and transduced into K562-PE4K cells at a multiplicity of infection of 1. After puromycin selection for 7 days, transduced cells were plated in 96-well formats at a density of 3,000 cells per well and exposed to either DMSO or TKIs for 72 hours. Cell viability was then measured using the CCK-8 assay (Dojindo), with absorbance readings obtained at 450 nm. Viability data were normalized against DMSO-treated ABL1 wild-type controls.

2.10. High-throughput evaluation of integrated epegRNA sequence enrichment

Experiments were independently repeated twice, with initial cell numbers adjusted to provide ~10,000 cells per epegRNA. To control integration frequency, viral transduction was conducted at a MOI of 0.5, yielding approximately 5,000 transduced cells per epegRNA. Following a 72-hour infection period, cells were subsequently cultured for 10 more days in medium containing 0.3 μg/ml puromycin following media replacement. After this 10-day puromycin selection, to ensure sufficient prime editing efficiency, the culture was continued for 7 more days in newly replenished medium. Following expansion, the cell population was divided into a drug-treated group and an untreated control, with each arm maintained at a scale preserving ~5,000x library coverage. The applied drug concentrations were as follows: imatinib (SelleckChem, S2475) at 15 nM; nilotinib (SelleckChem, S1033) at 2.5 nM; bosutinib (SelleckChem, S1014) at 2.5 nM; ponatinib (SelleckChem, S1490) at 0.2 nM. Cells were continuously passaged for an additional 10 days before being harvested for genomic

DNA isolation.

2.11. High-throughput evaluation of endogenous genomic regions

Cells were seeded and lentivirus transduction were conducted following the procedures outlined in the integrated epegRNA enrichment assay described earlier. Five days post-infection, the culture medium was replaced with puromycin-containing medium, and cells were maintained under selection for an additional 7 days. After withdrawal of puromycin, cells were cultured in fresh medium for a further 8 days. The population was then split into drug-treated and untreated control arms. Drugs were administered at the following concentrations: imatinib (SelleckChem, S2475) at 100 nM; nilotinib (SelleckChem, S1033) at 2.5 nM; bosutinib (SelleckChem, S1014) at 2.5 nM; ponatinib (SelleckChem, S1490) at 0.2 nM; asciminib (SelleckChem, S8555) at 5 nM. Cells underwent continued passaging for 10 days prior to being harvested for genomic DNA purification.

2.12. Genomic DNA preparation and deep sequencing

Genomic DNA was isolated using the Wizard Genomic DNA Purification Kit (Promega). To evaluate the integration of epegRNAs, approximately 1,152 μg of genomic DNA was processed per sample, corresponding to roughly 10 μg per million cells, to achieve a sequencing depth exceeding 10,000×. A total of 192 individual PCR reactions (each 40 μL in volume) were conducted using 6 μg of template DNA per reaction, 100 nM of target-specific primers, and 2× Taq PCR Smart mix (SolGent). Thermal cycling conditions included an initial denaturation at 95 °C for 10 minutes, followed by 26 amplification cycles of 95 °C for 20 seconds, 61 °C for 20 seconds, and 72 °C for 20 seconds, with a final elongation step at 72 °C for 3 minutes. Resulting amplicons were pooled, purified using a commercial kit (iNtRON Biotechnology), and size-selected by agarose gel electrophoresis. For quantification of editing outcomes at endogenous loci, >5,000× library coverage was maintained. Two-step PCR amplification was performed using PrimeSTAR® GXL polymerase (Takara). In the first round, genomic DNA was distributed into multiple 40 μL reactions containing 10 μg of input DNA and 10 pmol of gene-specific primers. The thermal profile consisted of an initial denaturation at 98 °C for 5 minutes, followed by 26 cycles of 98 °C for 20 seconds, 60 °C for

20 seconds, and 68 °C for 1 minute, concluding with a 3-minute extension at 68 °C. In the second PCR, 125 ng of purified product from the first step was used as template across five 20 μL reactions, each containing 20 pmol of Illumina indexing primers and amplified for 8 additional cycles. Final amplicons were cleaned using the MEGAquick-Spin purification kit and submitted for next-generation sequencing on the NovaSeq platform (Illumina).

2.13. epegRNA abundance-based analysis

An 8-nt UMI was inserted upstream of the next-generation sequencing primer site to facilitate lineage tracing of transduced cells. UMI quantification and assignment were carried out using in-house Python pipelines based on sorting barcodes. To correct for sequencing-induced UMI errors, I applied a directional adjacency method for UMI deduplication and consolidation(14). UMIs associated with each epegRNA were stratified into four discrete groups. Considering that a given SNV could be introduced by multiple (typically two or three) epegRNAs, MAGeCK analysis was performed across 8–12 UMI–epegRNA pairs per variant. Read counts were normalized to RPM, and adjusted expression levels were obtained by computing the mean RPM across the respective UMI groups.

For analytic purposes, epegRNAs were regarded as functional equivalents of sgRNAs, and each SNV or SAAV was assigned as a surrogate gene label for resistance profiling. Differential representation was assessed using MAGeCK v0.5.9.3(15), comparing variant frequencies between treated and untreated samples, and between day 0 and day 10, generating positive/negative P values and the log2-fold change (LFC) of a SNV or SAAV. To visualize the results, a volcano plot was generated with the x-axis representing the LFC and the y-axis indicating the $-\log_{10}$ transformed RRA P value. For each SNV or SAAV, the lower of the positive or negative P values was selected for display. Variants were then stratified into three classification groups:

i. Resistant: Variants (SNVs or SAAVs) exhibiting a LFC exceeding the 99.7th percentile and an RRA P value below the 0.3rd percentile of the negative control distribution were classified as resistant.

ii. Intermediate: Variants with an LFC above the 95th percentile and an RRA P value below the 5th percentile of the negative controls, but not meeting the threshold for resistance, were

designated as intermediate.

iii. Sensitive: SNVs or SAAVs that did not fall into either the 'Resistant' or 'Intermediate' categories.

Negative controls included epegRNAs designed to induce synonymous mutations, sham-editing epegRNAs, non-targeting epegRNAs, and sgRNAs lacking the PBS and RTT.

2.14. Data analysis and variant filtering

To detect SNVs generated by libraries E4 through E9, a customized SNV reference library was constructed based on the coding sequence of *ABL1* (NC_000009.12, positions 130713043 to 130887675), incorporating only the designed SNVs along with a synonymous marker mutation. NGS reads were aligned against this reference, and only those exhibiting exact matches were considered valid. Reads that matched the wild-type sequence without any variation were categorized as unedited. To accurately differentiate genuine prime editing outcomes from artifacts arising due to incomplete reverse transcription template integration, sequencing errors, or biases introduced during library preparation, statistical significance was evaluated using Fisher's exact test. Odds ratios and Bonferroni-adjusted *P* values were computed by comparing read distributions between DMSO-treated cells harvested on day 10 and unedited negative control samples.

$$Odds\ ratio = \frac{\frac{(Reads\ of\ SNV\ of\ DMSO - treated\ cells\ at\ day\ 10 + 1)}{(Reads\ of\ wild\ type\ sequence\ of\ DMSO - treated\ cells\ at\ day\ 10 + 1)}}{\frac{(Reads\ of\ SNV\ in\ unedited\ cells + 1)}{(Reads\ of\ wild\ type\ sequence\ in\ unedited\ cells + 1)}}$$

SNVs were considered confidently edited if they exhibited an odds ratio greater than 2 and a P value below 0.05, thereby meeting thresholds for statistical significance and ensuring robust detection within the sequencing dataset.

2.15. Calculation and normalization of adjusted LFCs and resistance scores

To determine the SNV-based adjusted LFC, I initially aggregated the counts of variants inducing the same SNV from the subset that met the criteria for true edited SNVs. Subsequently, I converted this sum to RPM. Concurrently, to assess the impact of drug treatment, LFC values were calculated for each SNV by comparing allele frequencies observed in day 10 drug-exposed cells with those in matched untreated samples. The obtained LFC values underwent normalization for further analysis.

To minimize confounding effects from local sequence context, I adjusted the LFC of each SNV by referencing the LFCs of synonymous variants, which are assumed to reflect neutral editing outcomes. A regressed baseline for synonymous SNVs was computed at each codon position within exons, based on the lowest 15th percentile of synonymous mutation LFCs. Normalized LFC values were then derived by subtracting the position-specific regressed LFC of synonymous SNVs and subsequently dividing by the standard deviation of synonymous SNV LFCs within each exon, enabling comparative analysis across exons. The average of normalized LFC for each replicate was calculated to obtain the adjusted LFC of the SNVs.

$$Normalized\ LFC = \frac{LFC\ of\ each\ SNV - Synonymous\ LFC\ (LOWESS\ regressed)}{Standard\ deviation\ of\ LFC\ of\ synonymous\ SNV}$$

To evaluate the impact of SAAVs, I computed the average of normalized LFC values for SNVs inducing the same SAAV, resulting in the normalized LFCs for SAAVs. Subsequently, the averaged normalized LFCs of SAAVs for each replicate were determined to derive the ultimate resistance scores of SAAVs.

2.16. SNV and SAAV drug resistance classifications

Resistance classification for variants in the E4–E9 libraries was performed by referencing exon-wise distributions of adjusted LFCs for synonymous SNVs, which served as neutral baselines. Functional annotation was subsequently assigned according to the following criteria:

i. Resistant: SNVs were classified as resistant if their adjusted LFC scores exceeded the 99.7th percentile of synonymous SNV distributions in both replicates.

ii. Sensitive: SNVs were designated as sensitive if their adjusted LFC scores were below the 95th percentile of the synonymous SNV distributions across both replicates.

iii. Intermediate: SNVs that did not align with either the 'Resistant' or 'Sensitive' categories.

For the classification based on SAAVs, resistance scores of SAAVs were utilized instead of the adjusted LFC scores. The criteria for classification in each dataset remained consistent with those employed in the SNV-based categorization.

2.17. Generation of transformed Ba/F3 cells

Mutant variants of *ABL1* were generated via PCR-based mutagenesis and subsequently subcloned into the BlpI-linearized MCSV-(pBabe mcs)-human p210BCR-ABL-IRES-GFP vector (Addgene #79248). Retroviral particles were produced by transfecting 2×10^6 Platinum-A packaging cells, cultured in 60-mm dishes, with the MCSV plasmids (Addgene #24828) using Lipojet (SignaGene) according to the manufacturer's instructions. Viral supernatants were harvested 48 hours post-transfection, filtered through 0.45 μm membranes to remove cellular debris, and used immediately for downstream transduction. For retroviral delivery, 2×10^6 Ba/F3 murine pro-B cells were seeded into 6-well plates in high-glucose RPMI supplemented with 8 μg/mL polybrene and 10 ng/mL IL-3. The plates were centrifuged at 2,000 × g for 90 minutes at 25°C to facilitate transduction, followed by incubation for 48 hours. Selection of successfully transduced cells was achieved by culturing in IL-3–depleted medium, allowing enrichment of clones harboring *ABL1* mutations conferring cytokine-independent growth.

2.18. Cell viability assays in Ba/F3-derived cells

Drug response was assessed by treating Ba/F3-derived cells with serially diluted TKIs in 96-well plates over a 72-hour period. Cell viability was assessed using the CellTiter-Glo luminescent assay (Promega), and luminescence signals were subsequently measured to quantify cell survival. GI_{50} values were obtained through nonlinear curve fitting in GraphPad Prism.

2.19. Quantification and statistical analysis

All statistical analyses detailed in this study were conducted using Python software packages.

2.20. Data visualization

Data visualization was carried out using Python and Prism software. Variant mapping onto protein structures was conducted in PyMOL (version 2.5.5), referencing the crystal structure available under PDB accession code 8SSN (ABL1 complex with nilotinib and asciminib).

3. RESULTS

3.1. Generation of K562-PE4K cells

To generate and assess ABL1 kinase variants, K562 cells (mainly triploid), originally derived from a CML patient, were utilized. K562 cells harbor a Philadelphia chromosome with a BCR/ABL fusion gene (**Fig. 1a**), a prevalent mutation causing CML, and are responsive to TKIs. Initial steps involved the creation of PE4max-expressing K562 cells through lentiviral delivery of sequences encoding PEmax and dominant negative MLH1 (MLH1dn)(16) (**Fig. 1b**). Considering that MSH6 plays a role in substitution mismatch repair(17) and that MSH6 knockdown enhances prime editing efficiencies(16), MSH6 was knocked out in PE4max-expressing K562 cells using cytosine base editing (**Fig. 1b-1c**). The resulting single-cell-derived clone was designated K562-PE4K (K562 cells expressing PE4max with MSH6 knockout). To identify the optimal timeframe for efficient prime editing, editing efficiencies were monitored for up to 20 days using a previously reported pairwise library of pegRNA and corresponding target sequences named Library-Small(11) in K562-PE4K cells. The results were compared with efficiencies in K562-PEmax and K562-PE4max cells, serving as controls. The findings indicated that average prime editing efficiency increased over time in all tested cell lines, with the highest efficiency observed in K562-PE4K cells. In this context, the efficiency in K562-PE4K cells was 3.4-fold higher than that in K562-PEmax cells and 1.4-fold higher than that in K562-PE4max cells after 20 days (**Fig. 1d**). Subsequent studies induced prime editing in K562-PE4K cells for a duration of 20 days based on these results.

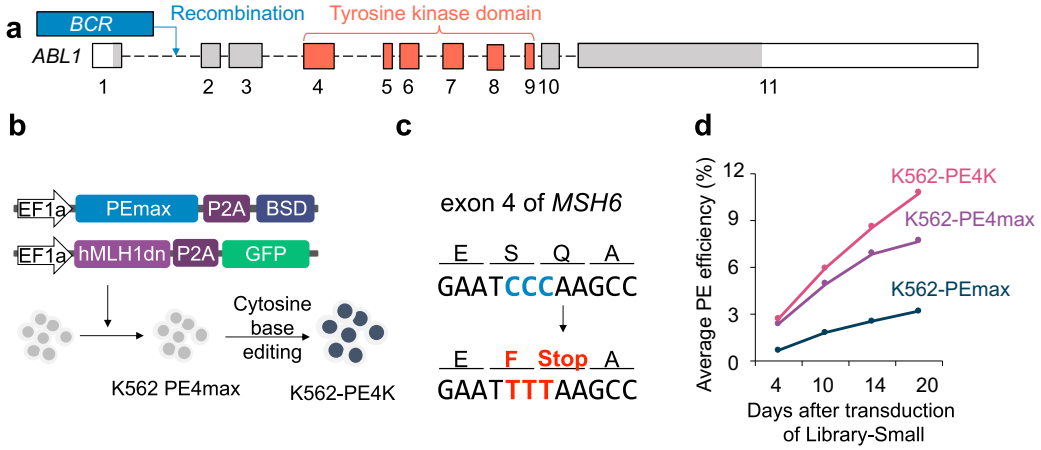


Fig. 1. Generation of K562-PE4K cells. (a) The structure of ABL1 after the BCR-ABL1 fusion. The coding sequence for the tyrosine kinase domain spans exons 4–9. Gray and orange-colored regions represent protein coding sequences. (b) Generation of K562-PE4K cells. MSH6 was knocked out using cytosine base editing. (c) Introduction of a premature termination codon into the MSH6 gene in K562-PE4K cells via cytosine base editing. (d) Mean prime editing efficiencies assessed using the Library-Small, which comprises 160 matched pairs of pegRNA constructs and corresponding target sequences, over time in three cell lines expressing different prime editing systems.

3.2. Evaluation of ABL1 kinase variants using epegRNA abundance-based analysis

Five TKIs targeting the BCR/ABL1 fusion protein bind to the ABL1 kinase, encoded by exons 4 to 9 of ABL1. The majority of reported BCR/ABL1 mutations leading to TKI resistance are situated within this kinase domain(6). Therefore, I introduced SNVs in exons 4 to 9 of ABL1 (**Fig. 1a**), encompassing a coding sequence of 964 base pairs (= 273 (exon 4) + 85 (exon 5) + 178 (exon 6) + 185 (exon 7) + 153 (exon 8) + 90 (exon 9)). Utilizing DeepPrime-FT(11), a highly accurate deep learning model for predicting pegRNA efficiencies, I designed and selected 8,673 epegRNAs(12) (= 964 bp × 3 SNVs/bp × 2-3 epegRNAs/SNV). The predicted median scores for these epegRNAs varied from 43 to 54 (average: 50) depending on the analyzed exon (**Fig. 2a**), suggesting overall high prime editing efficiencies.

I constructed a lentiviral library named 'library A', encompassing sequences encoding various guide RNAs. This library included 8,673 epegRNAs designed for introducing SGE in the ABL1 kinase-encoding sequences. Additionally, it comprised 190 epegRNAs targeting essential genes, inducing 1-bp insertions (resulting in frame-shift mutations) to serve as positive controls for depletion. Negative controls consisted of 81 sgRNAs lacking the primer binding site and reverse transcription template found in epegRNAs, 270 sham-editing epegRNAs that would not induce any changes, and 200 non-targeting epegRNAs. The library also included 44 pegRNA and corresponding target sequence pairs serving as sensors for prime editing. This library was transduced into K562-PE4K cells at a multiplicity of infection of 0.5, followed by the removal of untransduced cells using puromycin. The expression of GFP (a marker of MLH1dn expression) and RFP (a marker of epegRNA expression) was observed under a fluorescent microscope (**Fig. 2b**). Over a 20-day period, the median and average observed prime editing efficiencies in K562-PE4K cells expressing the sensor pegRNA and target pairs were 68% and 58%, respectively, indicating robust prime editing (**Fig. 2c**).

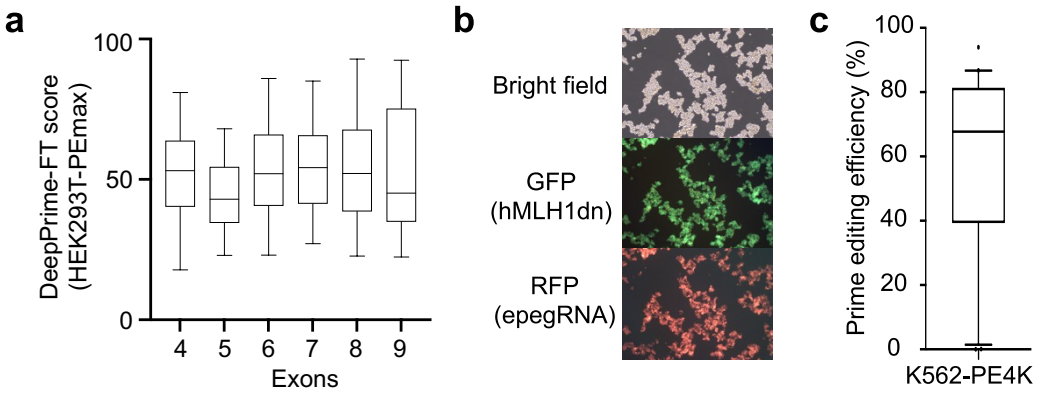


Fig. 2. Production of library A. (a) Distribution of DeepPrime-FT scores for the epegRNAs designed to generate SNVs in exons encoding the ABL1 kinase. (b) Fluorescence microscopy of K562-PE4K cells to visualize green fluorescent protein (GFP) and red fluorescent protein (RFP), markers of hMLH1dn and epegRNA expression, respectively. (c) Experimentally measured prime editing efficiencies at the 44 sensor pairs of pegRNA-encoding and target sequences in K562-PE4K cells 20 days after the lentiviral transduction of the pairs.

Twenty days after introducing library A into K562-PE4K cells, I subjected the resulting prime-edited cells to treatment with one of the five TKIs, namely imatinib, nilotinib, bosutinib, or ponatinib, or DMSO as a vehicle control for a duration of 10 days (**Fig. 3a**). To assess the functional impact of variants introduced through prime editing, I conducted deep sequencing to quantify the abundance of various guide RNAs, representing the frequency of cells expressing the edited RNA, at day 0 (20 days post-delivery of library A) and day 10 (30 days post-delivery of library A). Subsequently, I calculated the LFC in the number of cells containing each epegRNA over the 10-day period. As expected, the LFCs of negative control guide RNAs, including those inducing synonymous mutations, sham-editing epegRNAs, non-targeting epegRNAs, and sgRNAs, were distributed near zero (**Fig. 3b**). In contrast, epegRNAs inducing frame-shifting mutations in 19 essential genes and nonsense mutations in ABL1 exhibited significant depletion, with median LFCs of -0.58 (for those targeting essential genes; $P = 1.2\times10^{-88}$ vs. those inducing synonymous mutations) and -0.38 (for those inducing nonsense mutations in ABL1; $P = 2.9\times10^{-56}$ vs. those inducing synonymous mutations), respectively.

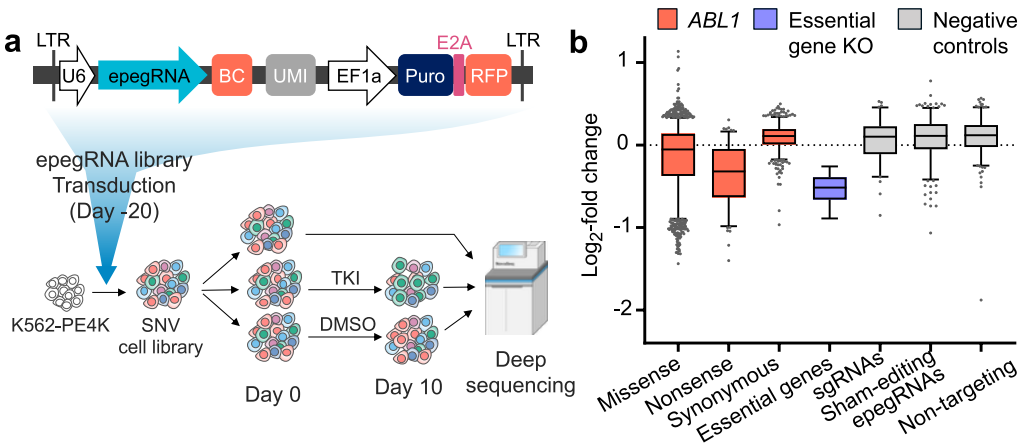


Fig. 3. Schematic overview and quality check of epegRNA abundance-based analysis. (a) Overview of high-throughput generation and evaluation of ABL1 kinase variants. A SNV cell library was generated using prime editing in K562-PE4K cells. Fold changes in the abundance of epegRNAs

targeting ABL1 kinase (red), epegRNAs targeting essential genes (blue), and negative control epegRNAs (gray) from day 0 to day 10. Box plots depict the 25th, 50th, and 75th percentiles, with whiskers spanning from the 5th to the 95th percentile. Outliers are indicated separately.

To assess the resistance induced by SNVs, I calculated the LFCs of the two to three epegRNAs designed to induce specific SNVs, along with the corresponding P values using MAGeCK(15) (**Fig. 4a-4e**). Subsequently, I classified the epegRNAs and the corresponding SNVs and SAAVs into three categories: "resistant" (LFCs > 99.7th percentile of negative control guide RNAs, P value < 0.3rd percentile of negative controls), "intermediate" (LFCs between the 99.7th and 95th percentiles of negative control LFCs, P values between the 0.3rd and 5th percentiles of negative controls), and "sensitive" (remaining epegRNAs), following established criteria. I identified a total of 58, 3, 1, and 32 SAAVs causing resistance to imatinib, nilotinib, bosutinib, and ponatinib, respectively. Among these, 6 (10%), 1 (33%), 1 (100%), and 18 (56%) of the resistant SAAVs were previously unreported, revealing a total of 26 (= 6 + 1 + 1 + 18) novel resistant SAAVs (**Fig. 4f**).

However, the T315I mutation, known as a 'gatekeeper mutation' with demonstrated resistance against five TKIs(1), was categorized as a sensitive SAAV in my evaluation with the four TKIs, indicating potential false negative outcomes. To identify the root cause of such discrepancies, I individually assessed the prime editing efficiencies of the epegRNA anticipated to be the most effective among the three epegRNAs designed to induce T315I (epegRNA-T315I) and other epegRNAs designed to induce SNVs classified as positive (resistant) in K562-PE4K cells. Notably, I observed that the efficiency of epegRNA-T315I was only 0.60%, whereas the epegRNAs targeting positive SNVs exhibited efficiencies of 49% (Y253H), 42% (E255V), and 61% (V299L) (**Fig. 5a**). Additionally, K562-PE4K cells transduced with epegRNA-T315I did not demonstrate resistance to imatinib, nilotinib, and bosutinib, in contrast to cells transduced with epegRNA-Y253, -E255V, and -V299L, which exhibited resistance to the respective TKIs (i.e., Y253H and E255V induced imatinib and nilotinib resistance, and V299L induced bosutinib resistance) (**Fig. 5b**).

Collectively, these findings indicate that epegRNA abundance-based analysis may yield false negatives, primarily attributed to insufficient prime editing efficiency. The accuracy of

this approach needs enhancement before functional assessments can be deemed clinically relevant.

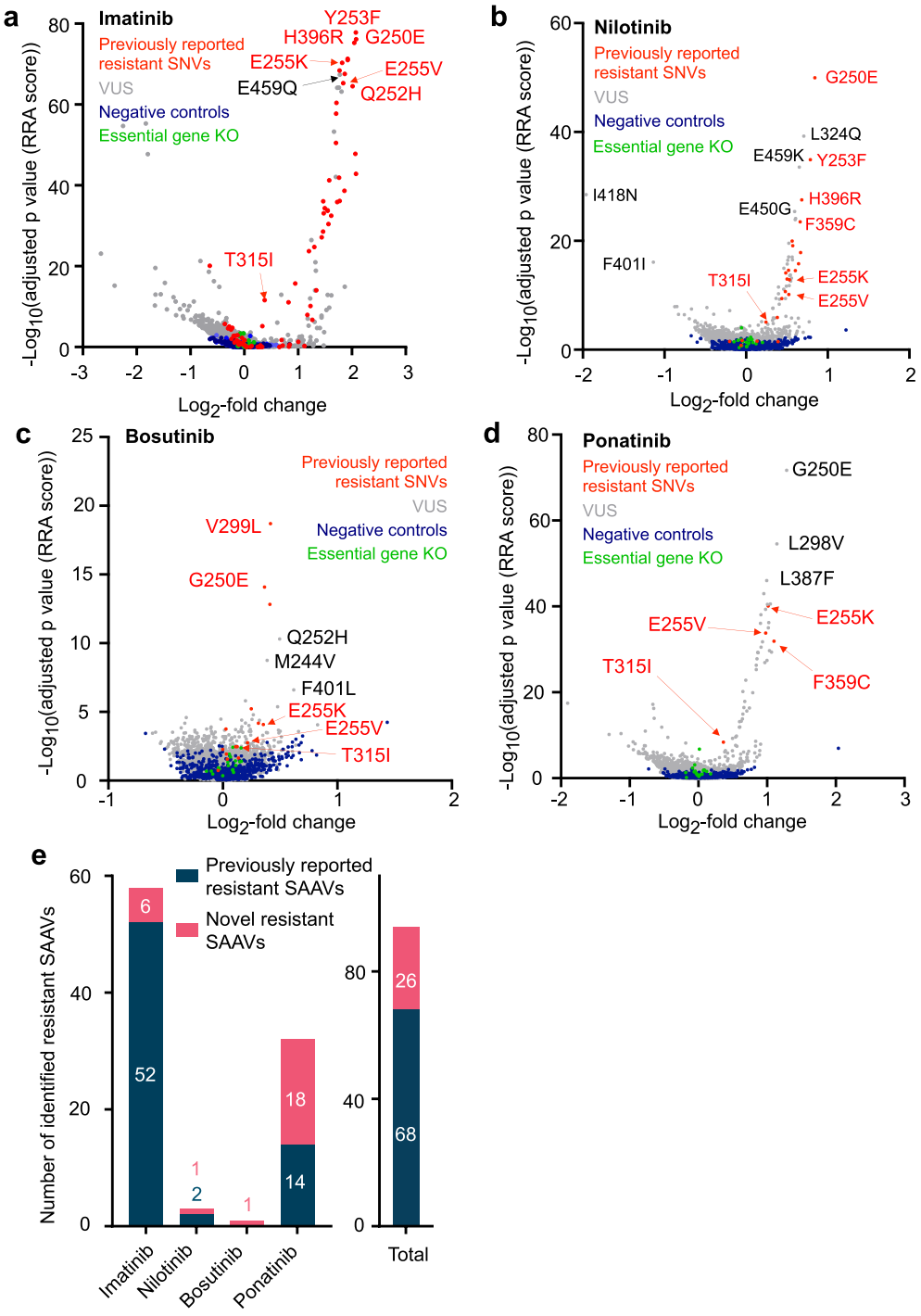


Fig. 4. Evaluation of ABL1 kinase variants using epegRNA abundance-based analysis. (a-e) Volcano plots showing log2-fold changes and P values of epegRNAs generating SNVs in ABL1 after treatments with imatinib (a), nilotinib (b), bosutinib (c), and ponatinib (d). (e) The number of previously reported or novel resistant SAAVs identified by the epegRNA abundance-based analyses.

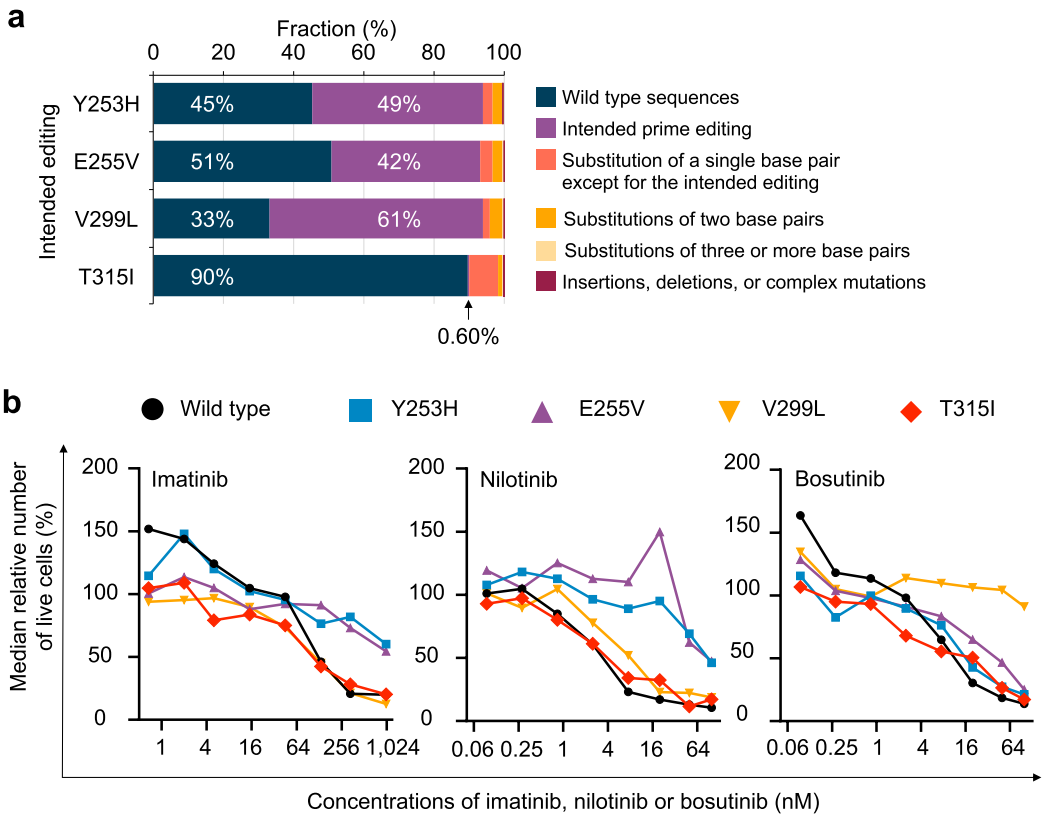


Fig. 5. False negative findings from epegRNA abundance-based analysis. (a) Prime editing efficiencies determined by deep sequencing in K562-PE4K cells after transduction of individual epegRNAs intended to generate codons for the indicated SAAVs. (b) Median relative numbers of live K562-PE4K cells containing ABL1WT (wild-type), Y253H, E255V, and T315I SAAVs after treatment with imatinib (left), nilotinib (middle), or bosutinib (right).

3.3. Generation and functional evaluations of the ABL1 variants by direct

sequencing of prime-edited endogenous regions

To improve the precision of SGE-based functional assessments, I opted for direct sequencing of the endogenous coding regions rather than relying on epegRNA abundance-based analyses. In implementing this approach, I introduced a synonymous mutation alongside the intended mutation to facilitate accurate identification of the intended SNV from sequencing reads(18). Moreover, the inclusion of an additional base pair modification was anticipated to enhance the efficacy of the intended editing by mitigating the activity of the mismatch repair system(16). A total of 8,676 epegRNAs were designed to induce all possible SNVs in the ABL1 kinase region (exons 4 to 9). Separate epegRNA libraries were created for each exon, denoted as E4, E5, through E9, corresponding to the targeted exon.

SGE was induced by separately transducing each library (E4, E5, through to E9) into K562-PE4K cells, followed by a 20-day culture to ensure adequate prime editing. When examining the prime editing efficiency over a 20-day period in exon 8 as a representative region in K562-PEmax, K562-PE4max, and K562-PE4K cells, I observed an increase over time in the percentage of deep sequencing reads containing both the intended edit and the intended synonymous edit in all tested cell lines. At 20 days after the transduction of lentiviral library E8, the efficiency was 6.1 (= 29.8%/4.87%)- and 1.9 (= 29.8%/15.8%)-fold higher in K562-PE4K cells compared to K562-PEmax and K562-PE4max cells, respectively (**Fig. 6a**). Subsequently, these cell libraries, harboring the introduced SNVs, were cultured for an additional 10 days under seven experimental conditions: a control condition without any treatment and under treatment with five different TKIs, namely imatinib, nilotinib, bosutinib, ponatinib, and asciminib (**Fig. 3a**). Deep sequencing analysis of genomic DNA from these cells at day 10 (corresponding to 30 days post-transduction) revealed that the average frequency of double-hit reads—defined as reads containing both the intended nucleotide substitution and the accompanying synonymous mutation—was 0.027%, ranging from 0.00% to 1.6% (**Fig. 6b**). This represented a substantial 179-fold increase relative to the baseline frequency of 1.5×10^{-4} % observed in unedited control cells. Importantly, double-hit SNV reads accounted for 17% to 33% (mean 25%) of total exon-aligned reads in the prime-edited cell, which is on average 133 (= 25%/0.19%)-fold higher than the range from 0.0089% to 0.66% (mean 0.19%) in the unedited control cells (**Fig. 6c**).

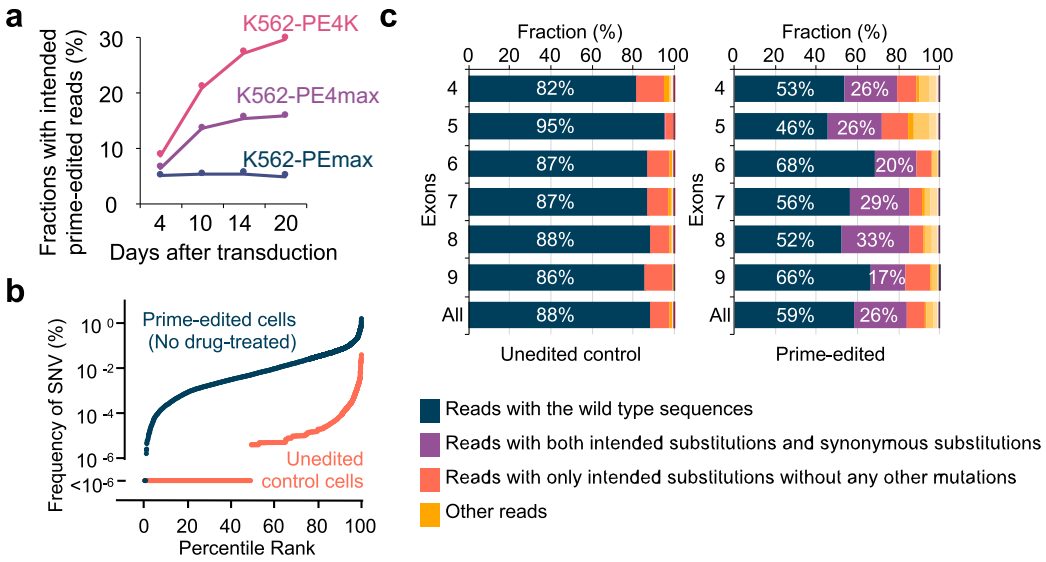


Fig. 6. Construction of library E4 – E9. (a) Fractions of reads with the intended prime-edit in exon 8 of ABL1 in cells expressing different prime editing systems after the transduction of library E8. (b) Comparison of SNV frequencies in exonic regions between untreated control cells and prime-edited cells, assessed 30 days after introducing libraries E4 and E5 through E9 without TKI exposure. (c) Relative abundance of reads containing the designed SNV and/or associated synonymous mutation in K562-PE4K cells prior to editing (left, control) and at 30 days after transduction with exon-specific ABL1 libraries.

To evaluate the effectiveness of prime editing in inducing specific substitutions, I quantified the frequency of double-hit SNVs by comparing read counts in prime-edited cells to those in control cells. The odds ratio and statistical significance of the differences were calculated using Fisher's exact test. An intended SNV was considered to be reliably introduced if it satisfied the following conditions: an odds ratio greater than 2, a P value less than 0.05, and the absence of any additional mutations at other positions across all replicates. After excluding 90 non-significant SNVs, I identified 2,802 significant SNVs, constituting 97% of all 2,892 possible SNVs. Among these, 1,954 were SAAVs, representing 98% of all 1,998 possible SAAVs (**Fig. 7a-7d**).

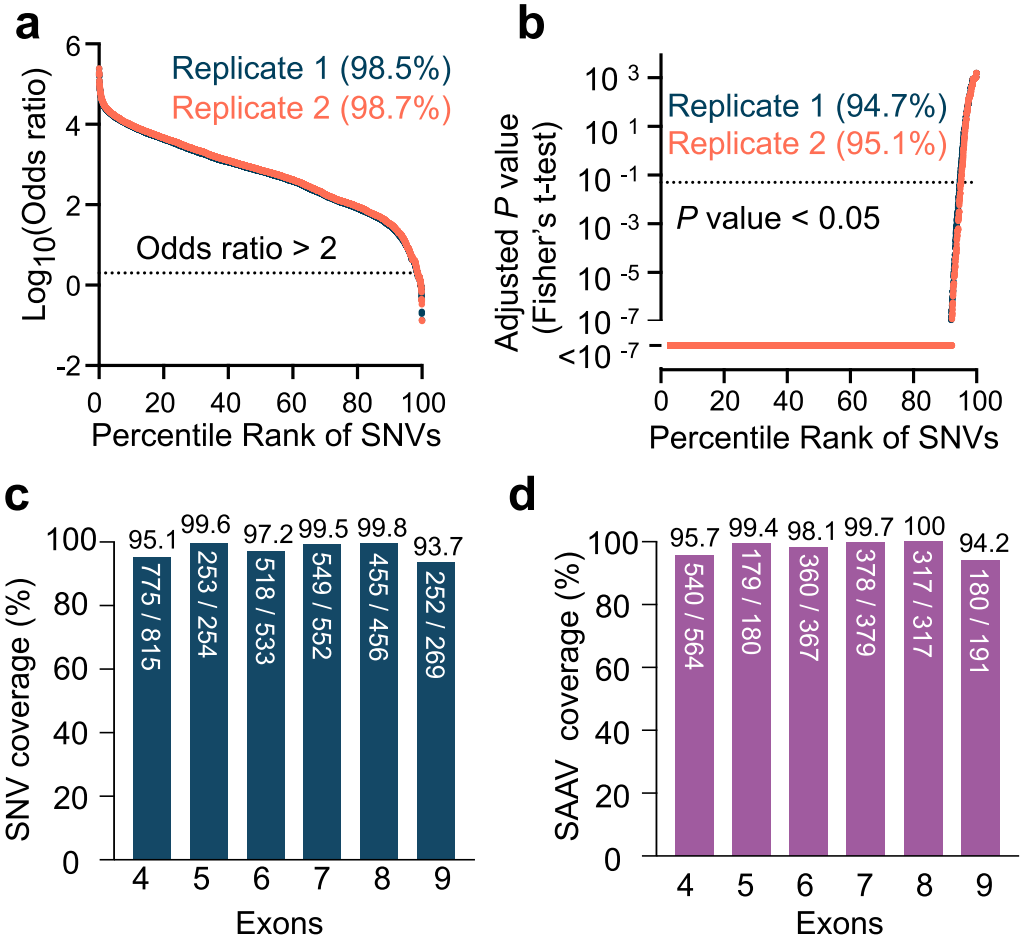


Fig. 7. Composition of library E4 – E9. (a–b) Odds ratios (a) and P values (b) for SNVs located within exons 4 through 9 of ABL1 in K562-PE4K cells following transduction with libraries E4 and E5–E9. Dashed lines represent the significance thresholds: an odds ratio of 2 (a) and a P value of 0.05 (b). Data from two independent replicates are shown in blue and red. (c–d) Summary of the number and representation of SNVs (c) and SAAVs (d) classified as significantly introduced and successfully identified, in comparison to the total number of possible SNVs and SAAVs across each exon of ABL1.

3.4. Comprehensive resistance profiles of 2,802 SNVs in ABL1 against

five TKIs

Utilizing normalized LFCs obtained by 15% LOWESS modeling, SNVs were categorized as "resistant" (normalized LFC surpassing the 99.7th percentile of synonymous SNVs in both replicates), "sensitive" (normalized LFC lower than the 95th percentile of synonymous SNVs in both replicates), and "intermediate" (remaining SNVs). Resistance scores for each SAAV were determined by averaging the adjusted LFC values of SNVs inducing the same SAAV. My analysis revealed minimal influence of the estimated relative prime editing efficiencies on both the adjusted LFCs and resistance scores of the variants, indicating a negligible bias attributable to editing efficiency (**Fig. 8a-8b**). The counts for sensitive, intermediate, and resistant SNVs and SAAVs were as follows: for imatinib, 2,379, 276, and 147 SNVs (1,550, 267, and 137 SAAVs); for nilotinib, 2,456, 275, and 71 SNVs (1,619, 254, and 81 SAAVs); for bosutinib, 2,489, 301, and 12 SNVs (1,642, 297, and 15 SAAVs); for ponatinib, 2,492, 303, and 7 SNVs (1,655, 284, and 15 SAAVs); for asciminib, 2,221, 318, and 263 SNVs (1,437, 273, and 244 SAAVs), respectively. Among the SAAVs inducing resistance to imatinib, nilotinib, bosutinib, and asciminib, 66 (48%), 59 (73%), 10 (67%), 6 (40%), and 220 (90%), respectively, were not reported previously. These findings present a comprehensive overview of resistance profiles for 98% (= 9,770/9,990) of all possible SAAV and TKI combinations (**Fig. 9a-9f**).

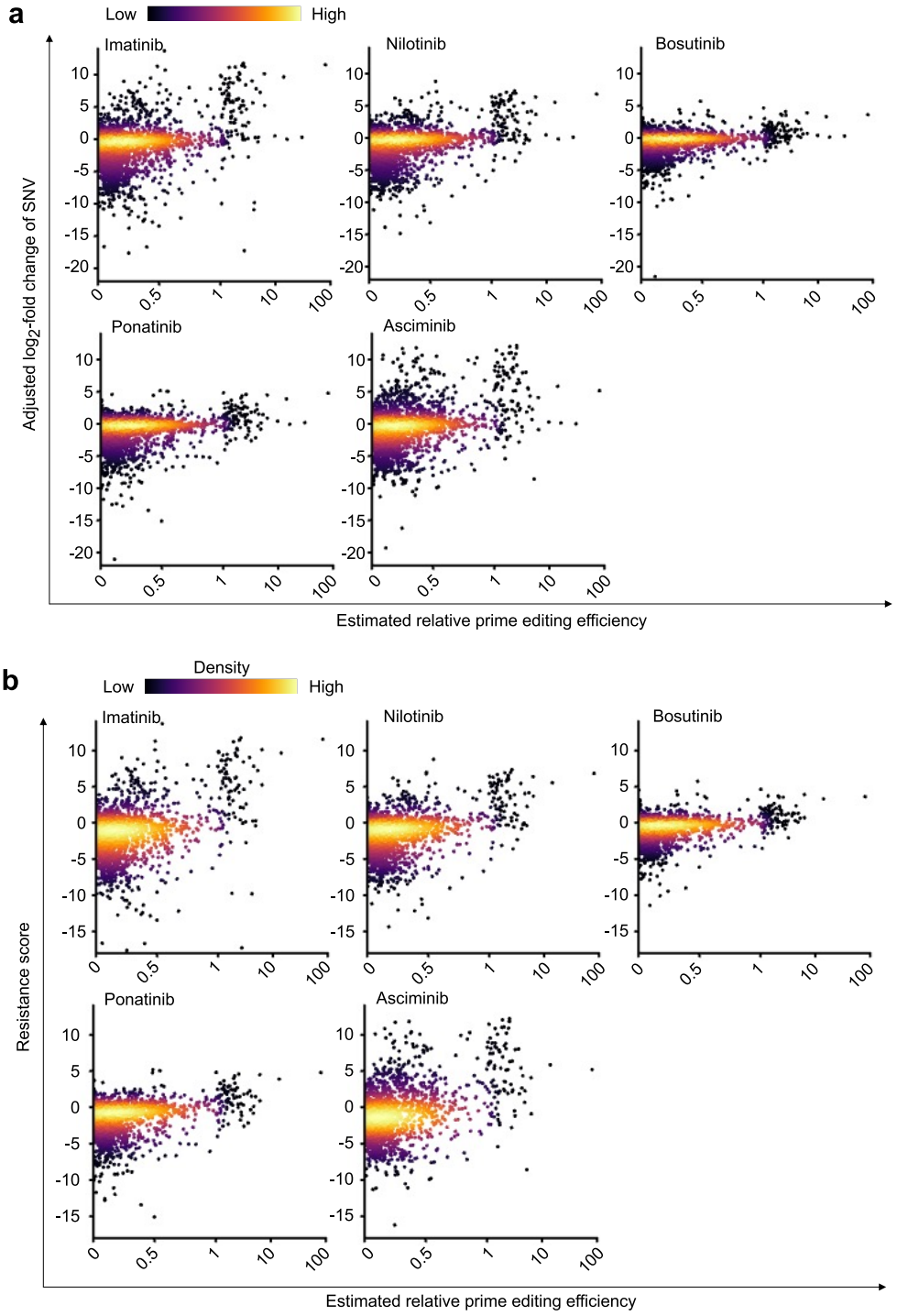
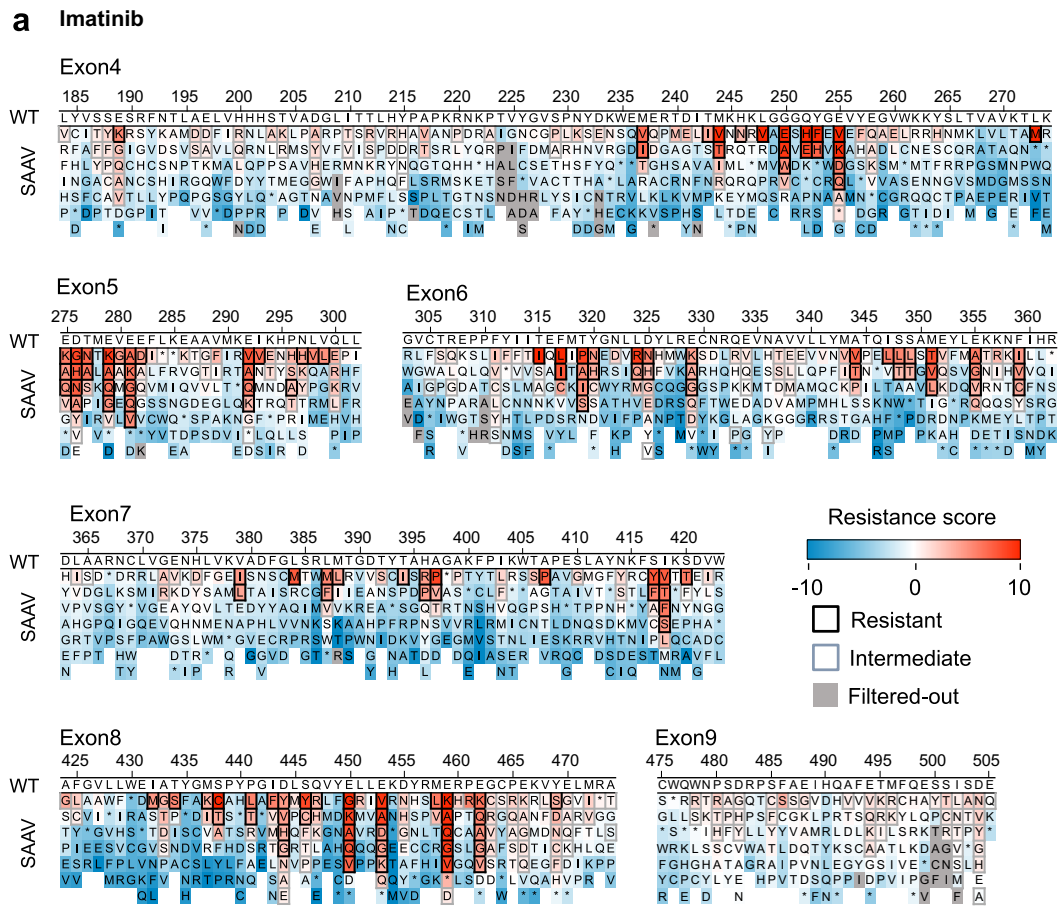
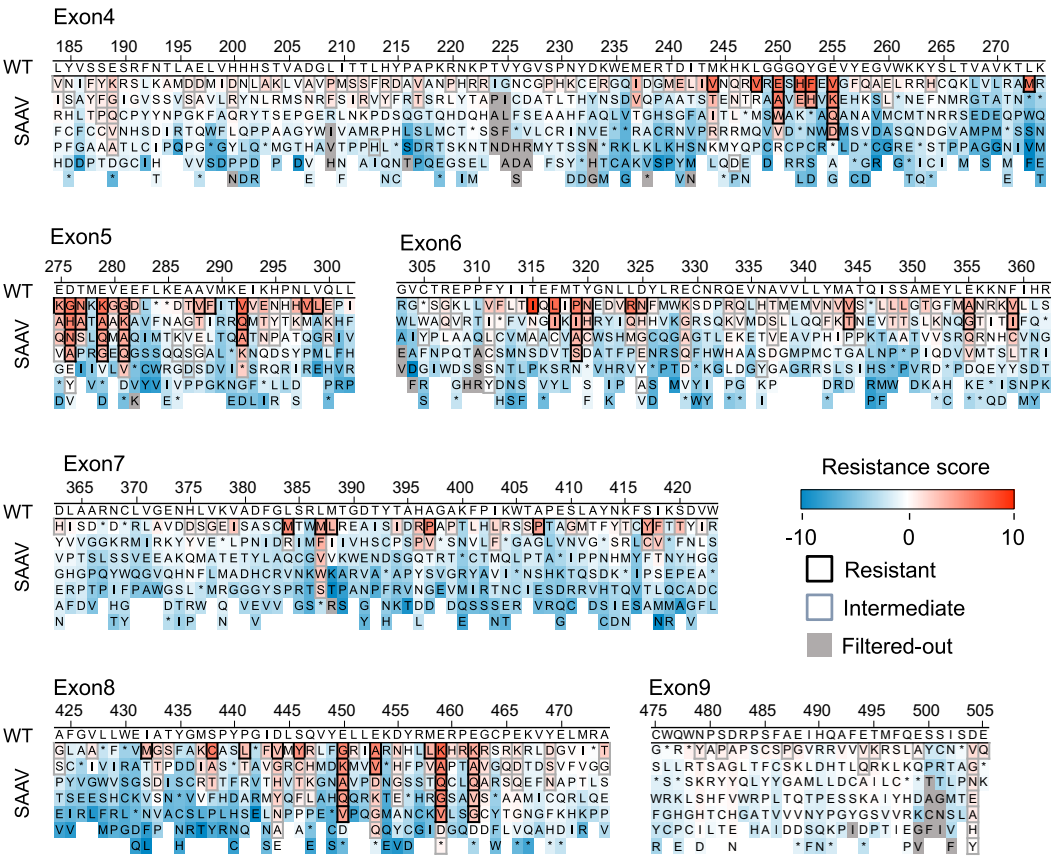


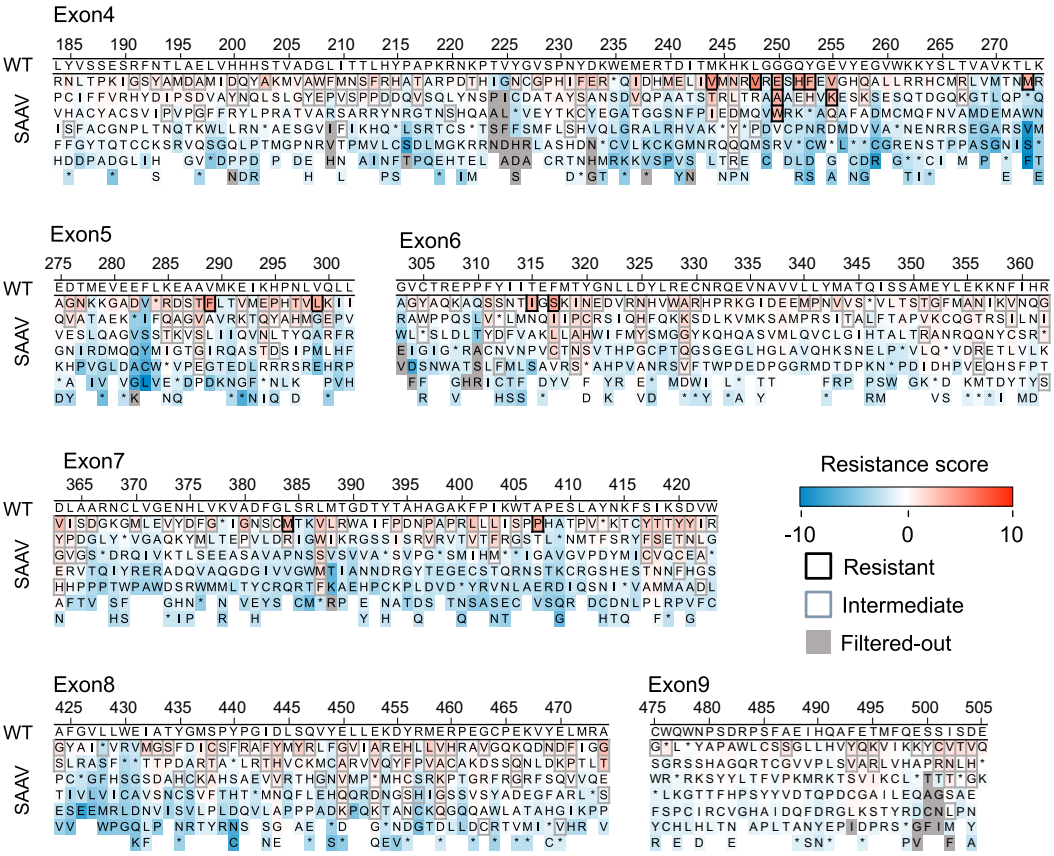
Fig. 8. Impact of prime editing efficiency on TKI resistance assessments. To examine the potential effect of prime editing efficiency on resistance evaluation, I analyzed the relationship between estimated relative editing efficiency and either adjusted LFC (for SNVs, panel a) or resistance score (for SAAVs, panel b) across five TKIs. Relative prime editing efficiency for each variant was derived by normalizing the proportion of sequencing reads corresponding to each SNV or SAAV within the endogenous region against their corresponding epegRNA read abundance, adjusted by total read counts in each dataset. The x-axis represents editing efficiency on a linear scale from 0 to 1 and a logarithmic scale from 1 to 100. Each data point is color-coded based on local point density.



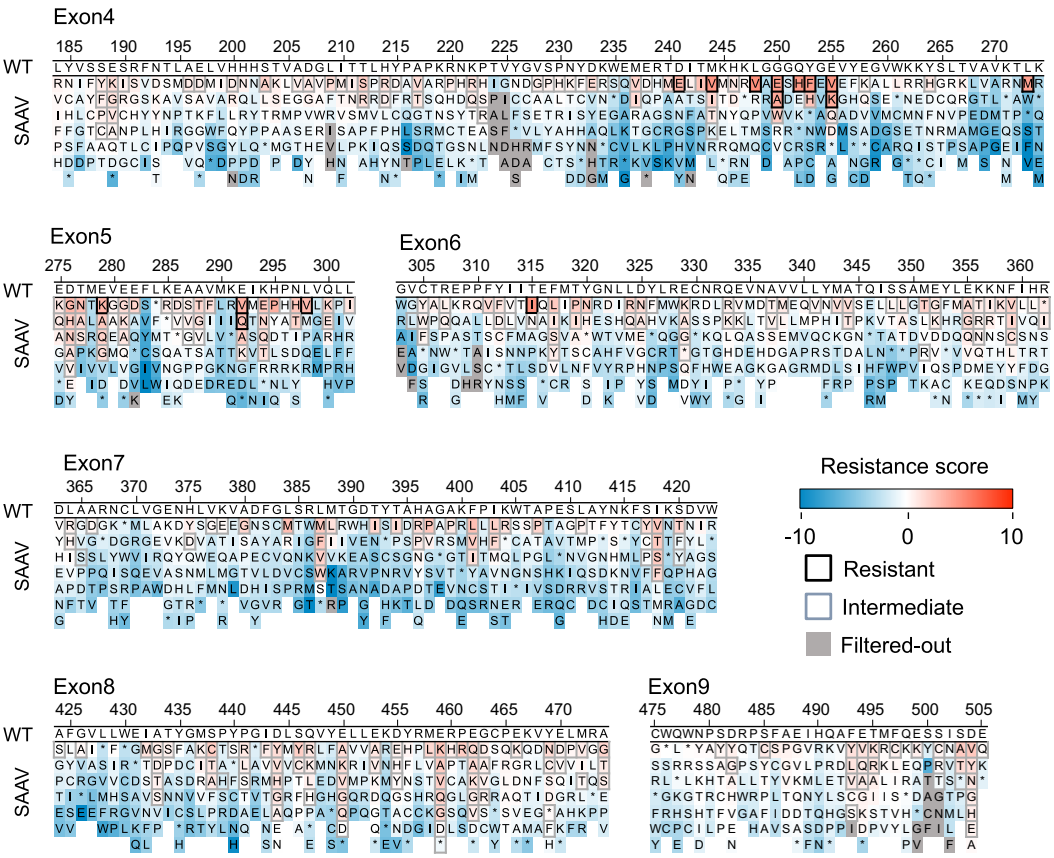
b **Nilotinib**



c **Bosutinib**



d **Ponatinib**



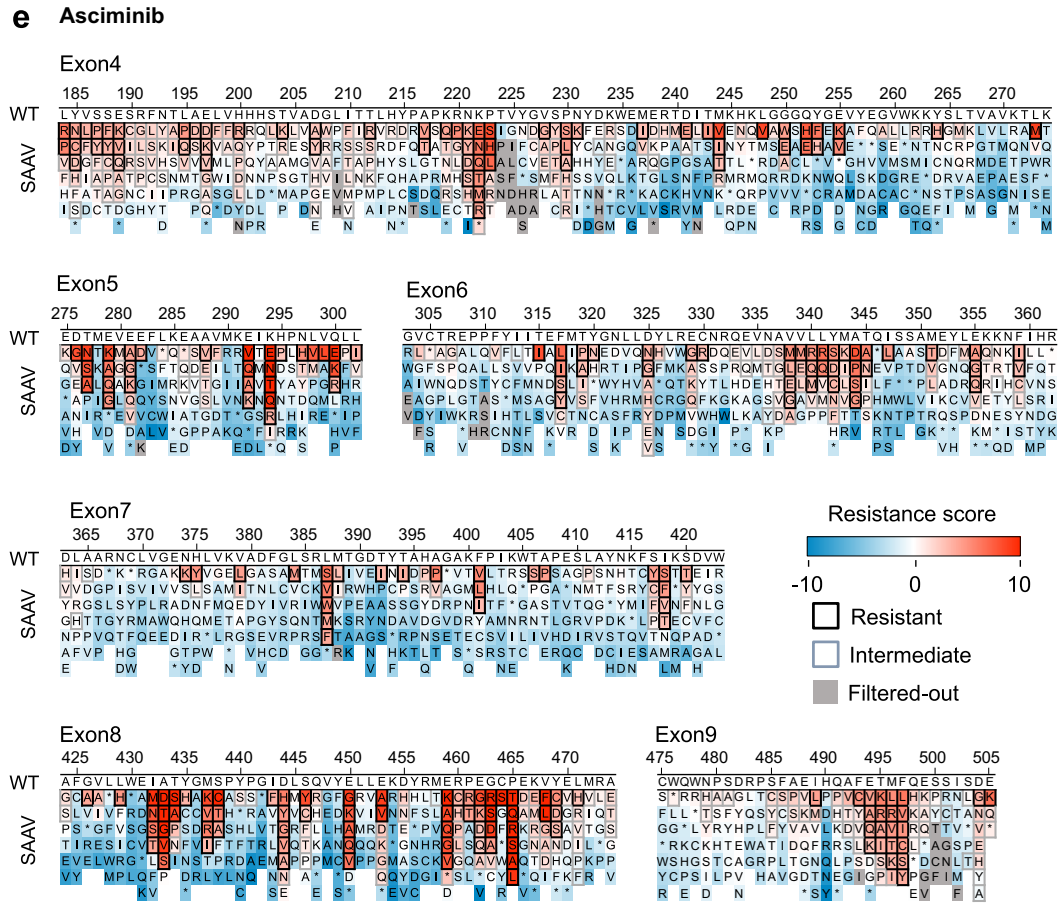


Fig. 9. Resistance scores of 1,954 SAAVs against each of the five TKIs. Heatmaps show resistance scores of SAAVs in the ABL1 kinase (encoded by exons 4 – 9) against imatinib (a), nilotinib (b), bosutinib (c), ponatinib (d), and asciminib (e) determined using the endogenous region sequencing approach in K562-PE4K cells. SAAVs conferring resistant and intermediate phenotypes are highlighted with black and gray outlines, respectively

Key observations include the following. When evaluating adjusted LFCs for SNVs and resistance scores of SAAVs across different TKIs, I noted a tendency for relatively strong correlations in resistance scores between imatinib and nilotinib, while correlations were lower between asciminib and the other TKIs (**Fig. 10a-10c**). These variations in resistance score correlations can be attributed, at least in part, to the distinct TKI binding sites within

the ABL1 kinase and the diverse types of TKIs (**Fig. 10d**). Imatinib and nilotinib, both classified as type II TKIs, share a common chemical scaffold and bind to the ATP pocket of the ABL1 kinase in its inactive conformation(1, 19, 20). In contrast, asciminib, the sole type IV TKI examined in this study, binds to the myristoylation pocket(21), while all other TKIs used here target the ATP pocket in ABL1. For instance, residue P465 is situated within the myristoylation pocket, and resistance scores for the P465T/Q/R variants were elevated specifically for asciminib, not for the other TKIs (**Fig. 10c and Fig. 11**).

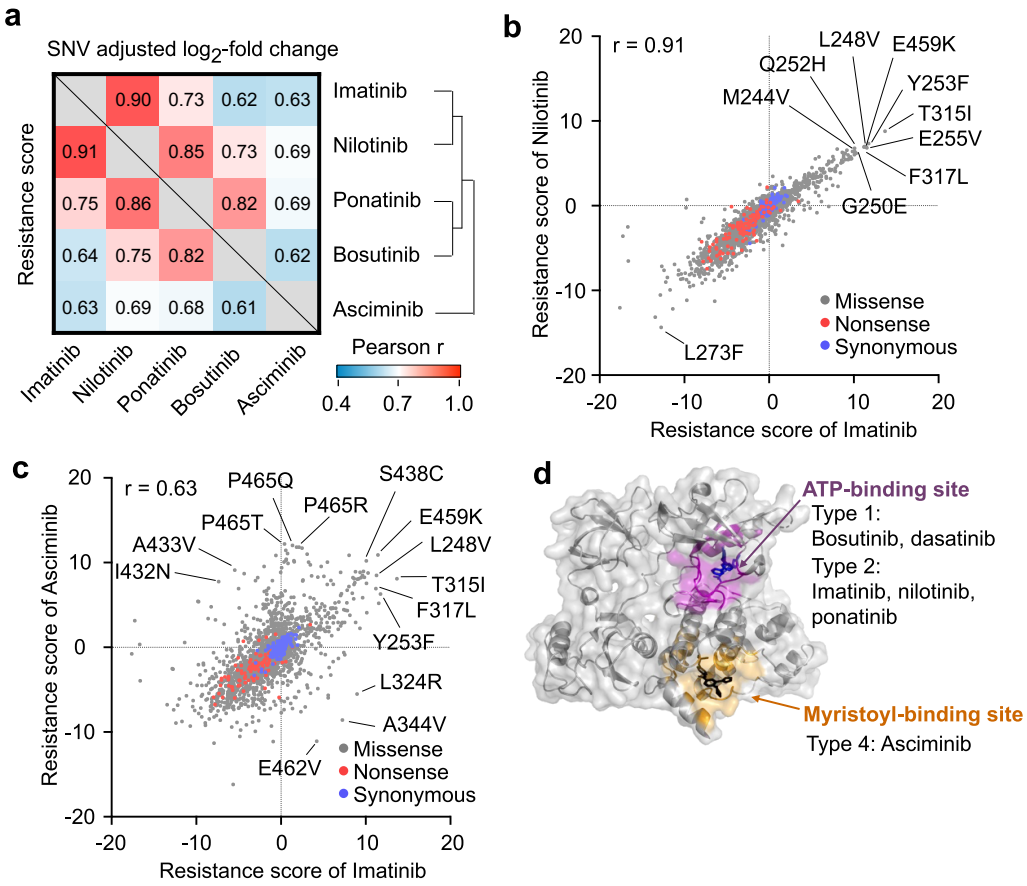


Fig. 10. Comprehensive resistance profiles of 2,802 SNVs in ABL1 against five TKIs. (a) Correlations between the adjusted LFCs of SNVs (upper right) and resistance scores of SAAVs (lower left) against five TKIs. (b-c) Resistance scores of SAAVs in ABL1 against imatinib and nilotinib (b) and against imatinib and asciminib (c). (d) Three-dimensional structure of the ABL1

core displaying the binding sites of six TKIs (PDB: 8SSN).

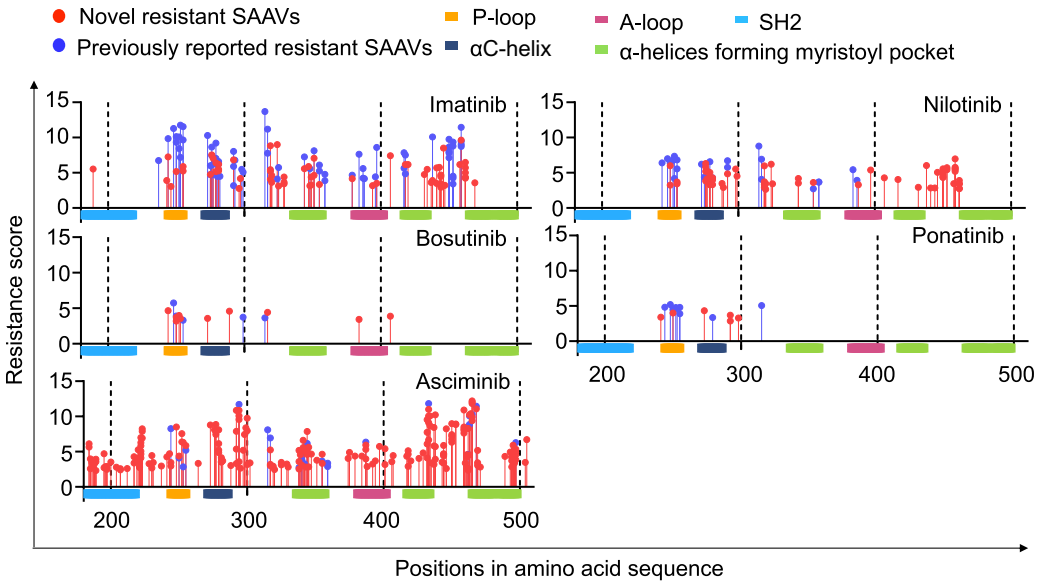


Fig. 11. SAAVs conferring resistance against TKIs. Lollipop plots of SAAVs encoded in ABL1 exons 4 to 9 and their resistance scores against imatinib, nilotinib, bosutinib, ponatinib, and asciminib determined in K562-PE4K cells. For brevity, only SAAVs classified as resistant are plotted.

The ABL1 kinase's phosphate binding loop (p-loop), spanning residues M244 to E255, has been associated with the survival of CML patients(22); however, the precise underlying mechanism remains unclear. The M244V variant, known to confer resistance to imatinib, nilotinib, bosutinib, ponatinib, and asciminib(5, 23), was validated in my study to induce resistance across all tested TKIs (**Fig. 12)**. Furthermore, M244T/I exhibited resistance or intermediate effects for all TKIs. L248V, reported to cause resistance to imatinib, nilotinib, ponatinib, and bosutinib(24, 25), demonstrated resistance against all TKIs, including asciminib. G250E, associated with resistance to imatinib, nilotinib, bosutinib, and ponatinib(1, 9, 25-28), showed resistance for all TKIs. Q252K/L/P/R/Stop and Y253C/D/N/S/Stop were classified as sensitive for all TKIs, while Q252H and Y253F were resistant for all TKIs, consistent with previous reports on Q252H-induced resistance to imatinib, nilotinib, and ponatinib(5, 25, 26). E255K/V variants were resistant for most TKIs,

confirming previous findings except for E255K, which was resistant to asciminib(26, 29, 30). Novel resistant variants, E255D/Q, were identified as conferring resistance specifically to imatinib and nilotinib, distinguishing them from other TKIs. In summary, my findings highlight several resistant mutations within the p-loop of the ABL1 kinase, suggesting a potential association with the adverse prognosis observed in patients harboring mutations in this region.

L273M was determined to be resistant to all tested TKIs, consistent with previous reports highlighting its role in conferring resistance to imatinib and nilotinib(5, 31). Intriguingly, other variants at this site, including L273F/S/V/W/Stop, were classified as sensitive to all TKIs. T315I, a recognized 'gatekeeper mutation' associated with resistance to all TKIs(5, 26, 29, 30), was also confirmed as resistant in my results. (**Fig. 12**).

As previously mentioned, asciminib stands apart from conventional TKIs by binding to the myristoylation pocket, a deep cavity situated in the C-lobe of the kinase (**Fig. 10d**). Notably, asciminib exhibits the highest count of resistant SAAVs, surpassing the number associated with imatinib, the TKI with the second-highest count of resistant SAAVs, by a factor of 1.8 (= 244/137). Moreover, these resistant SAAVs were dispersed extensively across the ABL1 kinase.

Among the 1,954 identified SAAVs, eight exhibited resistance to all five TKIs. Notably, six of these (M244V, L248V, G250E, Q252H, Y253F, E255K) were situated in the p-loop. The remaining two, L273M and the gatekeeper mutation T315I, were located outside the p-loop. A majority of variants with premature stop codons (94% = 499/530) were categorized as sensitive, while the remaining 5.8% (= 31/530) were classified as intermediate (**Fig. 12**).

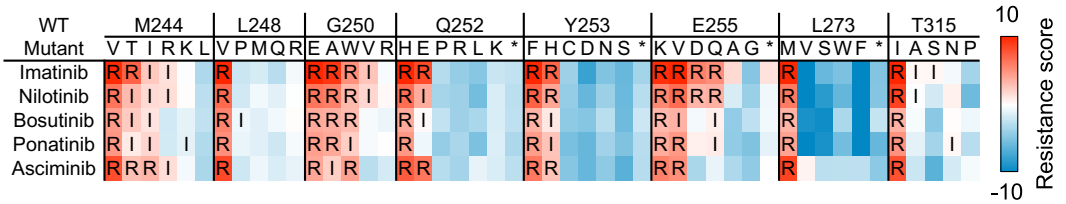


Fig. 12. Resistance scores of SAAVs induced at eight notable positions in ABL1 against five TKIs. A heatmap of resistance scores and functional classifications of SAAVs at M244, L248, G250, Q252, Y253, E255, L273, and T315 against imatinib, nilotinib, bosutinib, ponatinib, and asciminib.

3.5. Validation of TKI resistance using Ba/F3 cell-based assays

To conventionally assess TKI resistance conferred by a subset of newly characterized SNVs, I introduced 12 selected variants—comprising both previously uncharacterized and newly identified resistant mutations—into murine Ba/F3 cells via retroviral transduction of ABL1 constructs carrying the respective SNVs (**Fig. 13a**). Following transformation, the cells were exposed to serial dilutions of various TKIs to determine the half-maximal growth inhibitory concentrations (GI_{50}). A strong concordance was observed between the GI_{50} values obtained from these assays and the resistance scores derived from my high-throughput screening (**Fig. 13b**), with Spearman correlation coefficients (R) ranging from 0.63 to 0.86 (mean R = 0.80), thereby supporting the robustness and reliability of my primary screening approach.

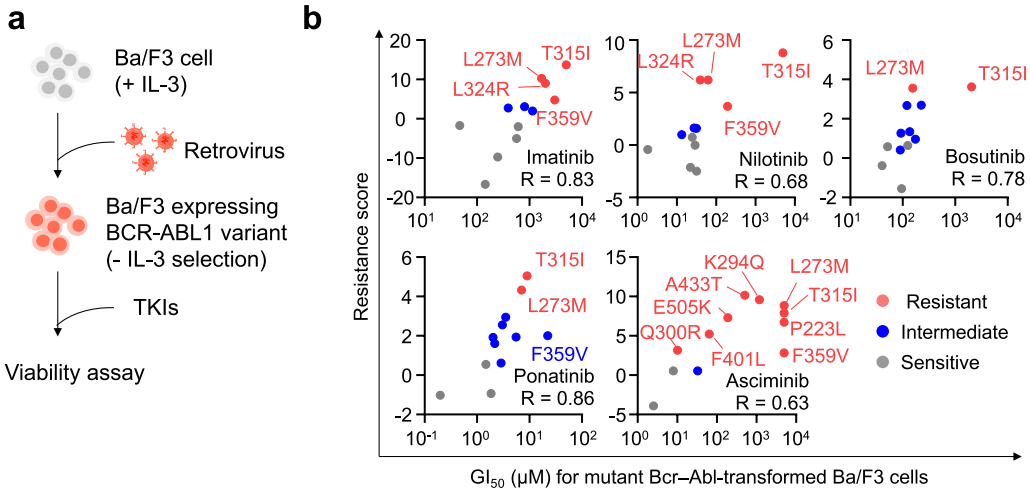


Fig. 13. Functional assessment of TKI resistance using murine Ba/F3 cells. (a) Diagram outlining the experimental workflow for validating variant function in Ba/F3 cells. Each cell population was individually transduced with retroviral constructs encoding one of 12 selected BCR-ABL1 SNVs, followed by treatment with five distinct TKIs. Cellular viability was subsequently assessed to evaluate drug response.

(b) Comparison of resistance scores obtained from high-throughput prime-edited K562-PE4K assays with the GI_{50} values measured in Ba/F3 cells expressing mutant ABL1. Spearman correlation coefficients (R) indicate the degree of concordance between the two assay systems. Each variant is color-coded according to its resistance classification in the K562-PE4K screen (gray: sensitive, green: intermediate, red: resistant).

3.6. Accuracy of direct sequencing of prime-edited endogenous regions vs. epegRNA abundance based analyses in ABL1

The correlations between LFCs of replicates in the direct sequencing of endogenous regions were found to be higher compared to epegRNA abundance-based analyses (**Fig. 14a-14b**). Among the 1,954 SAAVs, 194 and 28 were represented by two and three SNVs, respectively. When assessing the correlations between LFCs of paired members (each trio consisting of three pairs), the endogenous region sequencing approach exhibited stronger correlations (Pearson's correlation coefficient r, mean = 0.79, ranging from 0.63 to 0.87) in contrast to epegRNA abundance-based analyses (r, mean = 0.53, ranging from 0.18 to 0.87) (**Fig. 14c-14d**). Additionally, the separations of LFCs for the endogenous region sequencing approach were more distinct than those for epegRNA abundance-based analyses (**Fig. 15a**). These findings suggest that the signal-to-noise ratios for the endogenous region sequencing approach are generally higher than those for epegRNA abundance-based analysis.

Consistent with these observations, the SNV encoding T315I, a well-known SAAV associated with resistance to all TKIs used in the treatment of CML patients, exhibited the largest adjusted LFC and was consequently ranked first to twelfth in terms of its impact on resistance in the endogenous region sequencing approach. In contrast, it was ranked between 157th and 542nd for epegRNA abundance-based analyses (**Fig. 15a**). The relatively low signal-to-noise ratios observed in epegRNA abundance-based analyses may be attributed to the fact that the intended prime edit was not induced in a substantial fraction of cells transduced with specific epegRNAs. This led to a relatively low correlation between log-fold changes obtained from endogenous target site sequencing and those determined by sequencing to assess epegRNA abundance (r = from 0.19 to 0.67, mean: 0.45; **Fig. 15b**).

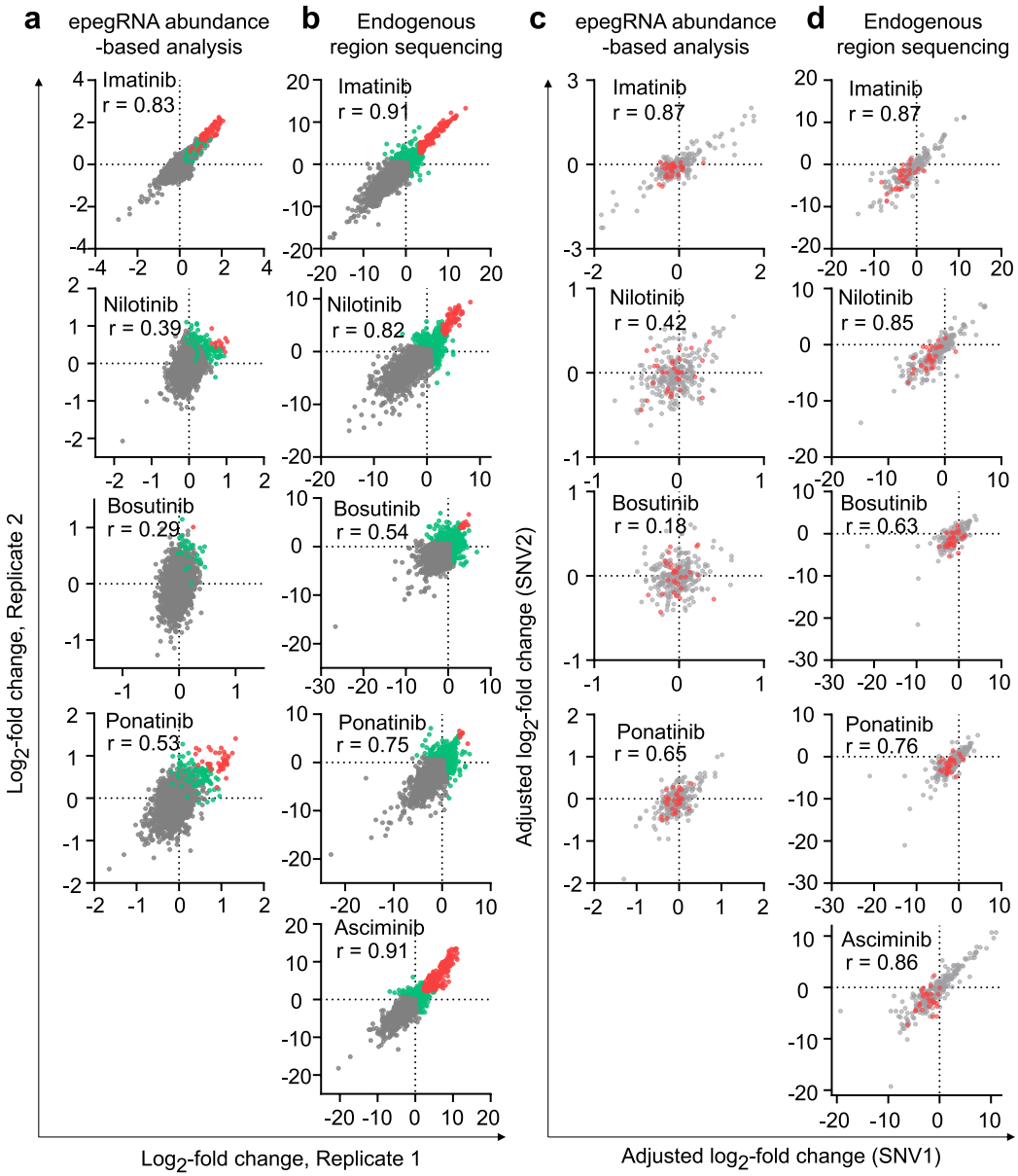


Fig. 14. Correlations between replicates or SNVs encoding the same SAAVs in the epegRNA abundance-based (a, c) and endogenous regions sequencing (b, d) analyses. (a-b) Correlations between normalized LFCs of SNVs from two replicates independently treated with the indicated TKI in the epegRNA abundance-based (a) and endogenous region sequencing (b) analyses. The classification of each SNV is indicated by the dot color (gray: sensitive, green: intermediate, red:

resistant). (c-d) Correlations between adjusted LFCs following treatment with the indicated TKI in pairs of SNVs encoding the same SAAVs in the epegRNA abundance-based (c) and endogenous region sequencing (d) analyses. The colors of the dots represent their SNV categories (gray: missense, red: nonsense).

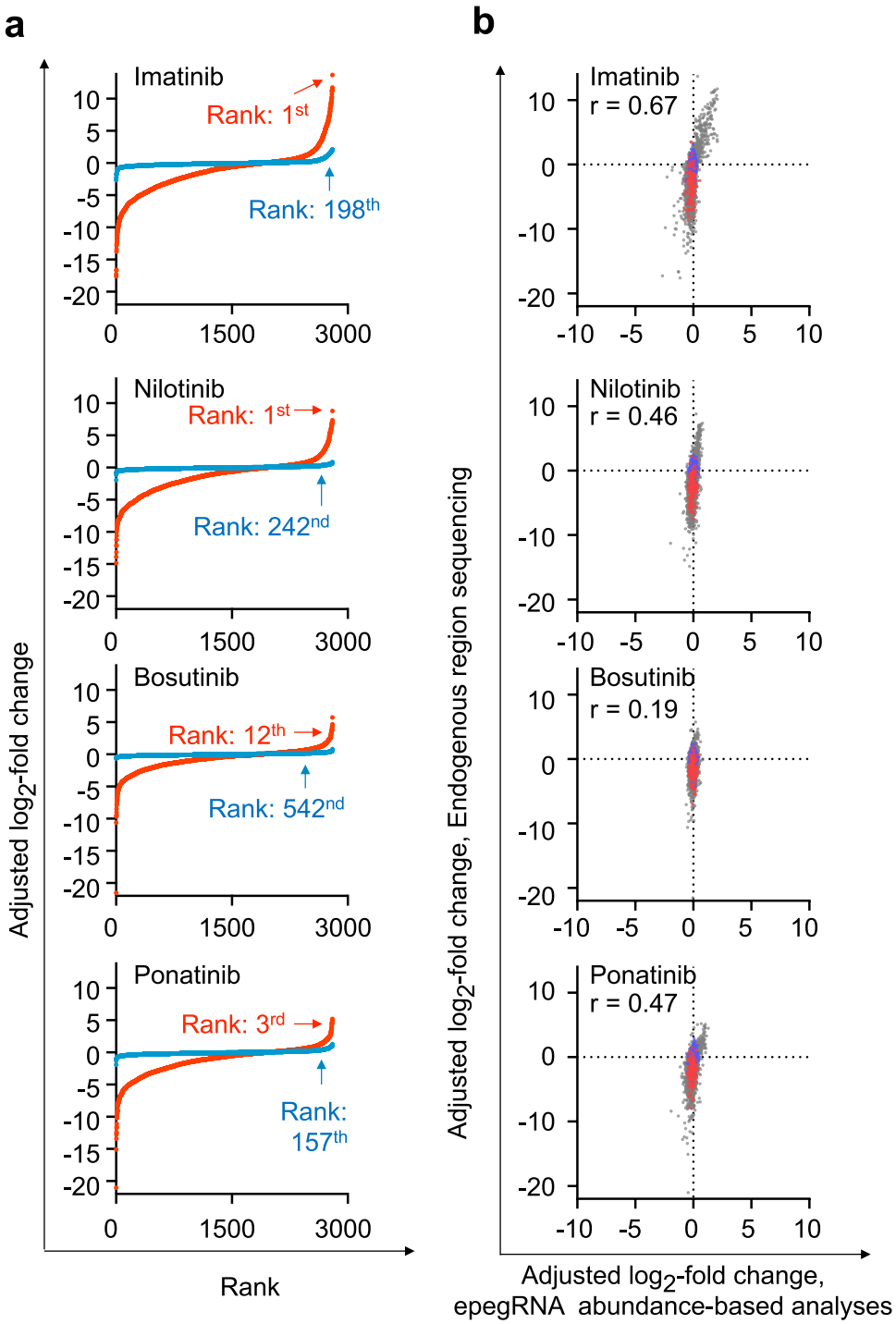


Fig. 15. The range and correlation of adjusted LFCs from endogenous region sequencing and epegRNA abundance-based analyses. (a) The range of adjusted LFCs of SNV-containing cells treated with the indicated TKI determined by the endogenous region sequencing and epegRNA abundance-based approaches. In the rank plots, the results of endogenous region sequencing and epegRNA abundance-based analyses are represented in red and blue, respectively. The arrows indicate the ranks of the SNV corresponding to T315I. (b) Correlations between adjusted LFCs of SNV-containing cells treated with the indicated TKI determined by the epegRNA abundance-based (x-axis) and endogenous region sequencing (y-axis) approaches. In the scatter plot, the colors of the dots represent the mutational consequence (gray: missense, red: nonsense, blue: synonymous).

I conducted a comparative analysis of functional classifications obtained from endogenous region sequencing and epegRNA abundance-based analyses. Out of the 248 pairs involving TKIs and SAAVs identified as resistant by endogenous region sequencing, only 70 SAAVs (28%) were characterized as resistant to the corresponding TKI in epegRNA abundance-based analyses, while 117 (47%) were erroneously labeled as sensitive (**Fig. 16a**). These findings suggest a notable occurrence of false negatives in epegRNA abundance-based analyses. Conversely, among the 6,466 pairs of TKIs and SAAVs categorized as sensitive by endogenous region sequencing, 6,429 (99%) were consistently identified as sensitive by epegRNA abundance-based analyses (**Fig. 16a**).

Collectively, my findings indicate that the endogenous region sequencing approach offers enhanced accuracy and higher signal-to-noise ratios compared to epegRNA abundance-based analyses. This suggests that utilizing direct sequencing of endogenous sites for the identification of prime-edited cells leads

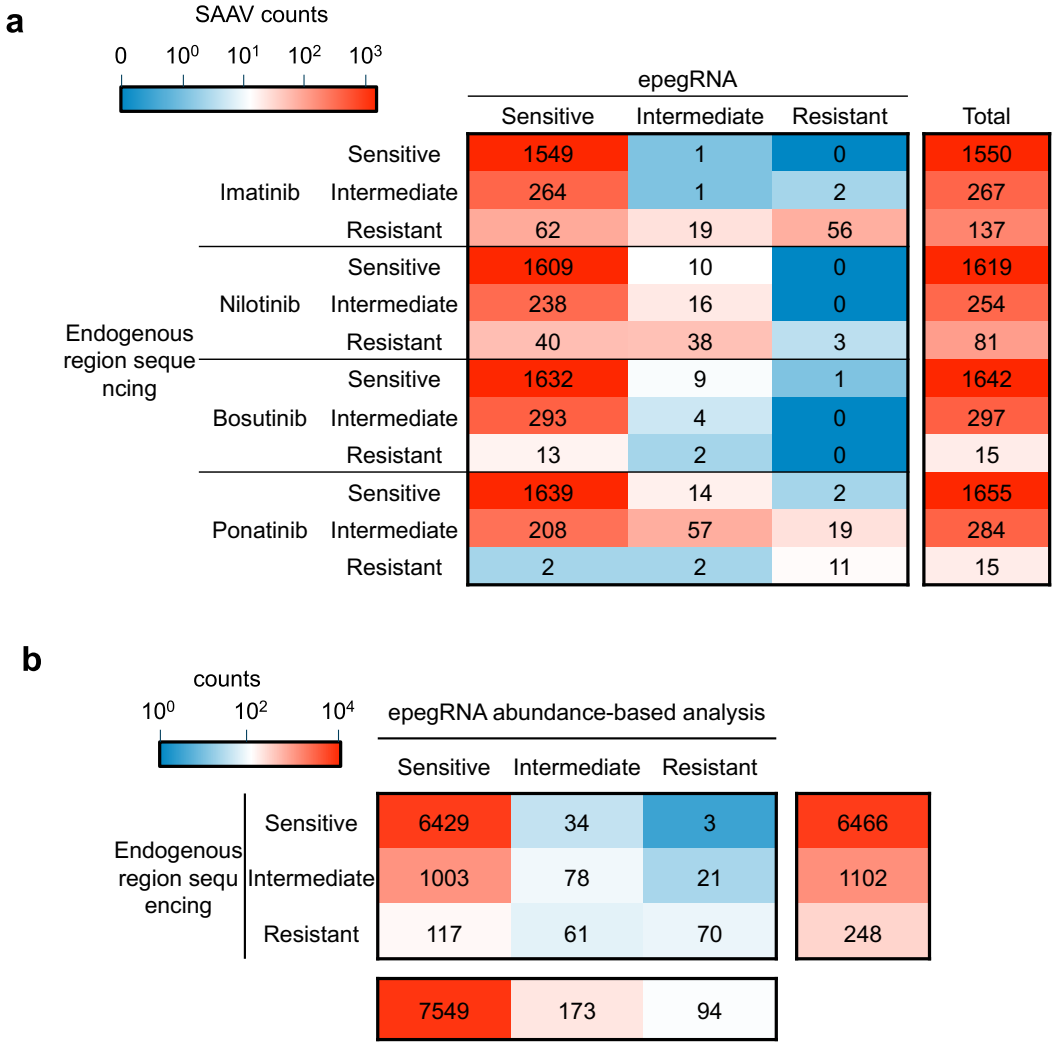


Fig. 16. Comparison of classification results determined by the epegRNA abundance-based and endogenous region sequencing analyses. (a) A heatmap that compares the functional classification results of SAAVs with regards to their effect on resistance against the indicated TKIs determined using epegRNA abundance- and endogenous region sequencing-based analyses. (b) A heatmap that compares the functional classification results for SAAVs and TKIs determined by the two methods. This heatmap is the summary of results shown in (a).

4. DISCUSSION

In the event of a CML patient experiencing a relapse during TKI therapy and subsequently being identified with a novel ABL1 mutation, it is plausible that the mutation has conferred resistance to the TKI. However, substantiating this association necessitates evaluating the TKI response in either a sufficient number of patients harboring the mutation or in cells containing the mutation. Notably, given the limited clinical guidelines for TKI selection, determining the most suitable TKI for use in such patients has often been challenging. In such scenarios, clinicians now have the option to make informed decisions on the appropriate TKI based on this study.

To enhance the efficiency of prime editing in this high-throughput assessment of the ABL1 kinase, I employed DeepPrime(11) to design epegRNAs(12). Additionally, I knocked out MSH6 and expressed MLH1dn to suppress mismatch-repair systems in the analyzed cells(16). Furthermore, I introduced a synonymous substitution, along with the mutation of interest, using prime editing to facilitate the identification of cells containing the desired edit. The identification of cells with the intended edits was accomplished through the sequencing of endogenous regions, enabling the exclusion of cells where the intended edit was not successfully introduced from the analysis.

For the direct translation of the functional evaluation results of VUS into clinical applications for human patients, precision is of utmost importance. In this study, I provide strong evidence that direct sequencing of endogenous genomic regions offers superior performance compared to recently adopted approaches relying on guide RNA abundance as a proxy for editing outcomes (32-36). Notably, even with the incorporation of cutting-edge techniques such as DeepPrime(11) to optimize prime editing guide RNA efficiencies, my findings underscore the superiority of the endogenous region sequencing approach in terms of accuracy.

Asciminib, designed to target the myristoylation pocket, was developed to address TKI resistance observed with other TKIs(1, 21). However, my study reveals that asciminib is linked to the highest number of resistant SAAVs, albeit demonstrating relatively limited overlap with those imparting resistance to the other four TKIs that bind to the ATP binding pocket. This extensive array of resistant SAAVs aligns with recent literature highlighting an escalating number of SAAVs conferring resistance against asciminib(9, 30, 37-40).

My study has certain limitations that should be noted. The investigation was carried out exclusively in a singular cell line, K562-PE4K, aligning with the approach taken in previous

studies that focused on functional evaluations of SNVs(35, 41). Additionally, for the high-throughput assessment, I administered a single dose of each TKI.

The lack of detailed insights into how various mutations confer resistance against TKIs has posed a challenge in determining the most effective therapy CML patients harboring these mutations. In my study, I systematically evaluated 2,802 (97%) of the possible SNVs within the ABL1 kinase sequence, encompassing 1,954 (98%) of the possible SAAVs, to assess their sensitivity to five TKIs commonly used in clinical practice for treating CML patients. This comprehensive information offers valuable guidance for selecting TKIs tailored to patients with specific point mutations in the ABL1 kinase, thereby promoting precision medicine in CML. Furthermore, by extending this approach to other genes, I may address the challenges associated with VUSs, potentially ushering in a new era of precise and individualized medical interventions.

5. CONCLUSION

This study employed prime editing to comprehensively evaluate the functional impact of 2,802 SNVs in the ABL1 kinase, encompassing 98% of possible SAAVs. This extensive resistance profile against five clinically used TKIs in CML provides valuable insights for precision medicine, aiding clinicians in selecting optimal therapeutic strategies tailored to individual patients based on ABL1 mutations. The study highlights the efficacy of direct sequencing of prime-edited endogenous regions for accurate functional evaluations and presents a comprehensive resource for understanding drug resistance mechanisms in CML.

REFERENCES

1. Braun TP, Eide CA, Druker BJ. Response and Resistance to BCR-ABL1-Targeted Therapies. Cancer Cell. 2020 Apr 13;37(4):530-42.
2. Druker BJ, Talpaz M, Resta DJ, Peng B, Buchdunger E, Ford JM, et al. Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia. N Engl J Med. 2001 Apr 5;344(14):1031-7.
3. Saglio G, Kim DW, Issaragrisil S, le Coutre P, Etienne G, Lobo C, et al. Nilotinib versus imatinib for newly diagnosed chronic myeloid leukemia. N Engl J Med. 2010 Jun 17;362(24):2251-9.
4. Kantarjian H, Shah NP, Hochhaus A, Cortes J, Shah S, Ayala M, et al. Dasatinib versus imatinib in newly diagnosed chronic-phase chronic myeloid leukemia. N Engl J Med. 2010 Jun 17;362(24):2260-70.
5. Soverini S, Mancini M, Bavaro L, Cavo M, Martinelli G. Chronic myeloid leukemia: the paradigm of targeting oncogenic tyrosine kinase signaling and counteracting resistance for successful cancer therapy. Mol Cancer. 2018 Feb 19;17(1):49.
6. Sierra JR, Cepero V, Giordano S. Molecular mechanisms of acquired resistance to tyrosine kinase targeted therapy. Mol Cancer. 2010 Apr 12;9:75.
7. Soverini S, Bavaro L, De Benedittis C, Martelli M, Iurlo A, Orofino N, et al. Prospective assessment of NGS-detectable mutations in CML patients with nonoptimal response: the NEXT-in-CML study. Blood. 2020 Feb 20;135(8):534-41.
8. Deininger MW, Shah NP, Altman JK, Berman E, Bhatia R, Bhatnagar B, et al. Chronic Myeloid Leukemia, Version 2.2021, NCCN Clinical Practice Guidelines in Oncology. J Natl Compr Canc Netw. 2020 Oct 1;18(10):1385-415.
9. Cross NCP, Ernst T, Branford S, Cayuela JM, Deininger M, Fabarius A, et al. European LeukemiaNet laboratory recommendations for the diagnosis and management of chronic myeloid leukemia. Leukemia. 2023 Oct 4.
10. Anzalone AV, Randolph PB, Davis JR, Sousa AA, Koblan LW, Levy JM, et al. Search-and-replace genome editing without double-strand breaks or donor DNA. Nature. 2019 Dec;576(7785):149-57.
11. Yu G, Kim HK, Park J, Kwak H, Cheong Y, Kim D, et al. Prediction of efficiencies for diverse prime editing systems in multiple cell types. Cell. 2023 May 11;186(10):2256-72.e23.
12. Nelson JW, Randolph PB, Shen SP, Everette KA, Chen PJ, Anzalone AV, et al. Engineered pegRNAs improve prime editing efficiency. Nat Biotechnol. 2022 Mar;40(3):402-10.
13. Dang Y, Jia G, Choi J, Ma H, Anaya E, Ye C, et al. Optimizing sgRNA structure to improve CRISPR-Cas9 knockout efficiency. Genome Biol. 2015 Dec 15;16:280.
14. Smith T, Heger A, Sudbery I. UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. Genome Res. 2017 Mar;27(3):491-9.
15. Li W, Xu H, Xiao T, Cong L, Love MI, Zhang F, et al. MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. Genome Biol. 2014;15(12):554.
16. Chen PJ, Hussmann JA, Yan J, Knipping F, Ravisankar P, Chen PF, et al. Enhanced prime editing systems by manipulating cellular determinants of editing outcomes. Cell. 2021 Oct 28;184(22):5635-52.e29.
17. Tosti E, Katakowski JA, Schaetzlein S, Kim HS, Ryan CJ, Shales M, et al. Evolutionarily conserved genetic interactions with budding and fission yeast MutS identify orthologous relationships in mismatch repair-deficient cancer cells. Genome Med. 2014;6(9):68.

18. Kim Y, Oh H-C, Lee S, Kim HH. Saturation resistance profiling of EGFR variants against tyrosine kinase inhibitors using prime editing. bioRxiv. 2023:2023.12.03.569825.
19. Weisberg E, Manley PW, Breitenstein W, Brüggen J, Cowan-Jacob SW, Ray A, et al. Characterization of AMN107, a selective inhibitor of native and mutant Bcr-Abl. Cancer Cell. 2005;7(2):129-41.
20. Benn CL, Dawson LA. Clinically Precedented Protein Kinases: Rationale for Their Use in Neurodegenerative Disease. Front Aging Neurosci. 2020 Sep 2;12.
21. Wylie AA, Schoepfer J, Jahnke W, Cowan-Jacob SW, Loo A, Furet P, et al. The allosteric inhibitor ABL001 enables dual targeting of BCR-ABL1. Nature. 2017 Mar 30;543(7647):733-7.
22. Nicolini FE, Corm S, Le QH, Roche-Lestienne C, Preudhomme C. The prognosis impact of BCR-ABL P-loop mutations: worse or not worse? Leukemia. 2007 Feb;21(2):193-4.
23. O'Hare T, Walters DK, Stoffregen EP, Jia T, Manley PW, Mestan J, et al. In vitro activity of Bcr-Abl inhibitors AMN107 and BMS-354825 against clinically relevant imatinib-resistant Abl kinase domain mutants. Cancer Res. 2005 Jun 1;65(11):4500-5.
24. Redaelli S, Piazza R, Rostagno R, Sassone M, Magistroni V, Perini P, et al. Determination of the activity profile of bosutinib, dasatinib and nilotinib against 18 imatinib resistant Bcr/Abl mutants. Blood. 2008;112(11):3220.
25. Patel AB, O'Hare T, Deininger MW. Mechanisms of Resistance to ABL Kinase Inhibition in Chronic Myeloid Leukemia and the Development of Next Generation ABL Kinase Inhibitors. Hematol Oncol Clin North Am. 2017 Aug;31(4):589-612.
26. Redaelli S, Piazza R, Rostagno R, Magistroni V, Perini P, Marega M, et al. Activity of bosutinib, dasatinib, and nilotinib against 18 imatinib-resistant BCR/ABL mutants. J Clin Oncol. 2009 Jan 20;27(3):469-71.
27. Lindström HJG, Friedman R. Rotating between ponatinib and imatinib temporarily increases the efficacy of imatinib as shown in a chronic myeloid leukaemia model. Sci Rep. 2022;12(1):5164.
28. Benjamin C, Murugan S, Hoosen S, Rapiti N. Chronic myeloid leukemia kinase domain mutations: A retrospective descriptive study on the therapeutic and prognostic significance in patients at King Edward VIII Hospital, KwaZulu-Natal, South Africa. Health Sci Rep. 2023 Jul;6(7):e1376.
29. Redaelli S, Mologni L, Rostagno R, Piazza R, Magistroni V, Ceccon M, et al. Three novel patient-derived BCR/ABL mutants show different sensitivity to second and third generation tyrosine kinase inhibitors. Am J Hematol. 2012;87(11):E125.
30. Eide CA, Zabriskie MS, Savage Stevens SL, Antelope O, Vellore NA, Than H, et al. Combining the Allosteric Inhibitor Asciminib with Ponatinib Suppresses Emergence of and Restores Efficacy against Highly Resistant BCR-ABL1 Mutants. Cancer Cell. 2019 Oct 14;36(4):431-43 e5.
31. Soverini S, Gnani A, Colarossi S, Castagnetti F, Abruzzese E, Paolini S, et al. Philadelphia-positive patients who already harbor imatinib-resistant Bcr-Abl kinase domain mutations have a higher likelihood of developing additional mutations associated with resistance to second- or third-line tyrosine kinase inhibitors. Blood. 2009 Sep 3;114(10):2168-71.
32. Chardon FM, Suiter CC, Daza RM, Smith NT, Parrish P, McDiarmid T, et al. A multiplex, prime editing framework for identifying drug resistance variants at scale. bioRxiv. 2023:2023.07.27.550902.
33. Ren X, Yang H, Nierenberg JL, Sun Y, Chen J, Beaman C, et al. High-throughput PRIME-editing screens identify functional DNA variants in the human genome. 2023;83:4633-45.
34. Gould SI, Wuest AN, Dong K, Johnson GA, Hsu A, Narendra VK, et al. High throughput evaluation of genetic variants with prime editing sensor libraries. bioRxiv. 2023:2022.10.26.513842.

35. Kim Y, Lee S, Cho S, Park J, Chae D, Park T, et al. High-throughput functional evaluation of human cancer-associated mutations using base editors. Nat Biotechnol. 2022 Jun;40(6):874-84.
36. Sanchez-Rivera FJ, Diaz BJ, Kastenhuber ER, Schmidt H, Katti A, Kennedy M, et al. Base editing sensor libraries for high-throughput engineering and functional analysis of cancer-associated single nucleotide variants. Nat Biotechnol. 2022 Jun;40(6):862-73.
37. Lee BJ, Shah NP. Identification and characterization of activating ABL1 1b kinase mutations: impact on sensitivity to ATP-competitive and allosteric ABL1 inhibitors. Leukemia. 2017 May;31(5):1096-107.
38. Hughes TP, Mauro MJ, Cortes JE, Minami H, Rea D, DeAngelo DJ, et al. Asciminib in Chronic Myeloid Leukemia after ABL Kinase Inhibitor Failure. N Engl J Med. 2019 Dec 12;381(24):2315-26.
39. Mauro MJ, Hughes TP, Kim DW, Rea D, Cortes JE, Hochhaus A, et al. Asciminib monotherapy in patients with CML-CP without BCR::ABL1 T315I mutations treated with at least two prior TKIs: 4-year phase 1 safety and efficacy results. Leukemia. 2023 Mar 22.
40. Hochhaus A, Réa D, Boquimpani C, Minami Y, Cortes JE, Hughes TP, et al. Asciminib vs bosutinib in chronic-phase chronic myeloid leukemia previously treated with at least two tyrosine kinase inhibitors: longer-term follow-up of ASCEMBL. Leukemia. 2023 Mar;37(3):617-26.
41. Findlay GM, Daza RM, Martin B, Zhang MD, Leith AP, Gasperini M, et al. Accurate classification of BRCA1 variants with saturation genome editing. Nature. 2018 Oct;562(7726):217-22.

Abstract in Korean

프라임 편집을 이용한 ABL1 변이에 대한 키나제 억제제 저항성 프로파일링

만성 골수성 백혈병 치료에서는 BCR-ABL1 융합 유전자에 의해 비정상적으로 활성화된 ABL1 키나제를 표적으로 하는 티로신 키나제 억제제가 이용된다. 하지만, 일부 환자들은 ABL1 변이로 인해 내성을 보이게 되는데, 각 티로신 키나제 억제제에 대한 많은 ABL1 키나제 변이의 영향이 명확하지 않아 약물 선택이 어렵다. 이 논문에서는 프라임 편집을 이용하여 ABL1 키나제 도메인을 코딩하는 염기에 가능한 모든 단일 염기 변이 중 97% (= 2,802/2,892), 단일 아미노산 변이 중 98% (= 1,954/1,998) 를 유도하였고, 인간 유래 만성 골수성 백혈병 세포주인 K562 세포에서 임상에서 이용중인 다섯 종류의 티로신 키나제 억제제에 대한 저항성의 영향을 평가하였다. 이 포괄적인 저항성 프로파일은 ABL1 변이에 따른 만성 골수성 백혈병 환자들의 약물 선택을 돕고, 정밀 의학을 촉진시킬 것으로 예상된다.

핵심되는 말 : 만성 골수성 백혈병; 프라임 편집; 임상적 의미가 불확실한 변이; ABL1; 단일 염기 변이; 티로신 키나제 억제제

PUBLICATION LIST

Jung, Y.*, Yu, G.*, Oh, H.-C., Lee, J.-H., Jeon, E., Bae, J., Sim, T. & Kim, H. H.#. Comprehensive resistance profiling of chronic myeloid leukemia associated ABL1 variants against five tyrosine kinase inhibitors using prime editing. *Nature Biomedical Engineering* (Accepted)