Simultaneous single-cell CRISPR, RNA, and ATAC-seq enables multiomic CRISPR screens to identify gene regulatory relationships

Kaivalya Shevade[1,2*], Yeqing Angela Yang[1,3*#], Kevin Feng[1,2], Karl Mader[1,2], Volkan Sevim[1,3], Jacob Parsons[4], Gunisha Arora[1,3], Hasnaa Elfawy[1,2], Rachel Mace[1,2], Scot Federman[1,2], Rustam Esanov[1,3], Shawn Shafer[1,3], Eric D Chow[2,5#], Laralynne Przybyla[1,2]

Affiliations:

[1]*Laboratory for Genomics Research, San Francisco, CA, 94158, USA*

[2]*Department of Biochemistry and Biophysics, University of California, San Francisco, CA 94158, USA*

[3]*GSK, San Francisco, CA, USA*

[4]*GSK, Collegeville, PA, USA*

[5]*Center for Advanced Technology, University of California, San Francisco, CA 94158*

**These authors contributed equally to the work*

#Corresponding authors. Correspondence and requests for materials should be addressed to Eric Chow (email: Eric.Chow@ucsf.edu) and Yeqing Angela Yang (email: Angela.x.Yang@gsk.com)

Abstract

The ability to identify gene functions and interactions in specific cellular contexts has been greatly enabled by functional genomics technologies. CRISPR-based genetic screens have proven invaluable in elucidating gene function in mammalian cells. Single-cell functional genomics methods, such as Perturb-seq and Spear-ATAC, have made it possible to achieve high-throughput mapping of the functional effects of gene perturbations by profiling transcriptomes and DNA accessibility, respectively. Combining single-cell chromatin accessibility and transcriptomic data via multiomic approaches has facilitated the discovery of novel cis and gene regulatory interactions. However, pseudobulk readouts from cell populations can often cloud the interpretation of results due to a heterogeneous response from cells receiving the same genetic perturbation, which could be mitigated by using transcriptional profiles of single cells to subset the ATAC-seq data. Existing methods to capture CRISPR guide RNAs to simultaneously assess

the impact of genetic perturbations on RNA and ATAC profiles require either cloning of gRNA libraries in specialized vectors or implementing complex protocols with multiple rounds of barcoding. Here, we introduce CAT-ATAC, a technique that adds CRISPR gRNA capture to the existing 10X Genomics Multiome assay, generating paired transcriptome, chromatin accessibility and perturbation identity data from the same individual cells. We demonstrate up to 77% guide capture efficiency for both arrayed and pooled delivery of lentiviral gRNAs in induced pluripotent stem cells (iPSCs) and cancer cell lines. This capability allows us to construct gene regulatory networks (GRNs) in cells under drug and genetic perturbations. By applying CAT-ATAC, we were able to identify a GRN associated with dasatinib resistance, indirectly activated by the HIC2 gene. Using loss of function experiments, we further validated that the gene, ZFPM2, a component of the predicted GRN, also contributes to dasatinib resistance. CAT-ATAC can thus be used to generate high-content multidimensional genotype-phenotype maps to reveal novel gene and cellular interactions and functions.

Introduction

Since the discovery of CRISPR-Cas9 genome editing over a decade ago, researchers have been adopting the system in a wide range of applications due to its high efficiency and ability to target specific genomic loci. CRISPR-mediated perturbations (knockout, interference, activation) provide a powerful tool to interrogate gene functions in a systematic manner. Given its ease of programmability, CRISPR-Cas9 can be combined with genome-wide guide RNA (gRNA) libraries for conducting unbiased and high-throughput genetic screens[1]. Different phenotypic assays can be utilized to evaluate the outcome of a pooled CRISPR screen, including gRNA representation analysis and single-cell evaluation[2,3,4]. By using single-cell RNA-seq (scRNA-seq) as a readout for pooled CRISPR screens, methods such as Perturb-seq[5,6], CRISP-seq[7], and CROP-seq[8] allow us to dive deep into the relationship between every genetic perturbation and its corresponding transcriptomic outcomes. These functional genomic studies provide information-rich phenotypic profiling that can significantly accelerate our understanding of biological and pathological processes.

In addition to transcriptomic changes, epigenomic effects are often as informative and crucial to understanding transcriptional regulation in a complex system. The dynamics of chromatin accessibility, dictated by transcription factor occupancy and epigenetic mechanisms such as histone modifications and DNA methylation, underlie many fundamental biological processes in

63 development and disease[9,10]. Thus, single-cell ATAC-seq (Assay for Transposase-Accessible
64 Chromatin with sequencing)[11], which measures chromatin accessibility in individual cells, has
65 been implemented as a readout for pooled CRISPR screens, and is represented by methods such
66 as Perturb-ATAC[12], CRISPR-sciATAC[13], and Spear-ATAC[14]. These methods employ various
67 single-cell platforms and have guide capture efficiencies ranging from 40% to 90%. As an example
68 of their utility, recently published atlases detailing single-cell chromatin accessibility in various
69 tissues offer a abundant insights into cell type-specific cis-regulatory elements and how they can
70 be used to predict links between non-coding disease risk variants to downstream affected genes
71 and functions[15,16].
72 Single-cell profiling techniques such as these can provide novel insights into gene regulatory
73 networks that drive several complex biological processes, with applications to developmental
74 trajectories and disease mechanisms. Drug resistance to chemotherapeutics and/or targeted
75 therapies is a known problem in cancer and can account for almost 90% of cancer-associated
76 deaths[17]. Chronic myeloid leukemia (CML) develops from a recombination between
77 chromosomes 9 and 22 creating a short chromosome 22 (Philadelphia chromosome) in myeloid
78 cells that results in a BCR-ABL gene fusion, causing constitutive activation of the Abl tyrosine
79 kinase. Dasatinib is an FDA-approved drug for the treatment of CML. Dasatinib, like most other
80 treatments for CML, is a broad-spectrum tyrosine kinase inhibitor (TKI)[18]. The development of
81 TKIs has increased 5-year survival rate from 22% in the mid 1970s to 69%[19]. However, there is
82 no complete cure for the disease. Although dasatinib and other TKIs effectively block the Abl
83 kinase, drug resistance ultimately leads to treatment failure. In as many as 40% of cases of clinical
84 TKI failure, cancer cells exhibit sustained BCR-ABL inhibition[20] but instead activate alternative
85 pathways which help them survive in presence of the drug. Identification of BCR-ABL-
86 independent mechanisms of drug resistance could lead to new therapeutic targets for CML[21,22].
87 Therefore, understanding the dynamics of transcriptome and DNA accessibility in response to
88 genetic and drug perturbation might help us piece together the novel gene regulatory relationships
89 that underlie drug resistance by providing a deeper understanding of the direct and indirect effects
90 of alternative pathway activation.
91 To interrogate both the transcriptional profile as well as the epigenetic landscape in CRISPR
92 screens, we have developed CAT-ATAC (CRISPR And Transcriptomics-Assay for Transposase-
93 Accessible Chromatin). This method simultaneously profiles the transcriptome, chromatin

94 accessibility, and CRISPR gRNA expression within single cells by leveraging the Multiome assay
95 kit from 10X Genomics and is compatible with the widely used Perturb-seq dual-guide vector
96 backbone as well as other gRNA vectors. The pairing of the two layers of data, transcriptional and
97 regulatory states, at single-cell resolution provides a multidimensional phenotypic readout that will
98 enable CRISPR-based screens in a range of applications intended to uncover gene-regulatory
99 networks. In this study, we developed CAT-ATAC in iPSCs and used it to elucidate a novel gene
100 regulatory network that underlies dasatinib drug resistance in CML cells.

101 **Results**

102 **CAT-ATAC methodology and proof-of-concept**

103 CAT-ATAC is built on the 10X Genomics Multiome assay, which uses a droplet-based method to
104 capture both RNA and chromatin fragments. In the workflow (Figure 1A), dual-guide direct
105 capture constructs[23] are introduced into cells that express Cas9-CRISPR machinery (eg, CRISPRn,
106 CRISPRi, or CRISPRa). CAT-ATAC reagents are combined with tagmented nuclei and passed
107 through the 10X Genomics single-cell controller to compartmentalize each nucleus into an oil
108 droplet containing a mix of reagents, enzymes and oligo primers. Through downstream
109 thermocycling reactions and enrichment steps, the mRNA transcripts, ATAC fragments, and the
110 gRNA transcripts within each cell are barcoded with a unique sequence and constructed into three
111 libraries, one each for assaying RNA, ATAC and gRNA. Figure 1B details the steps involved in
112 capturing gRNAs, which involves the spike-in of a pre-annealed reverse transcription (RT) primer
113 with a splint oligo to reverse transcribe the gRNA. The splint oligo is complementary to the spacer
114 sequence on the ATAC-capture oligo of the multiome gel beads and serves as a bridge between
115 the downstream gRNA RT products and the gel bead. The RT primer contains a sequence
116 complementary to the capture sequence (CS1) found in gRNAs expressed from the dual-guide
117 Perturb-seq gRNA construct[24], a UMI, and the Nextera R1 sequence and forms a duplex with the
118 splint oligo through a partial Nextera R1 sequence. During RT and template switching, the spiked-
119 in primer initiates first strand cDNA synthesis of the gRNA followed by template switching.
120 Subsequent rounds of PCR amplification add appropriate adaptors and sample indexes to make
121 the final single-cell gRNA library.

122 For proof-of-concept, we performed a pilot experiment (Figure 1C). We used a human induced
123 pluripotent stem cell (iPSC) line engineered with dox-inducible CRISPRa-VPR and delivered 4
124 dual gRNA constructs into the cells by lentiviral transduction in an arrayed format. Each dual

125 gRNA construct expresses 2 different activating guide sequences for the same target, in this case,
126 3 transcription factors (NEUROG2, GATA5, MEOX1) and a non-targeting control (NT); guide
127 sequences were selected from the top two performing guides in the Weissman V2 CRISPRa library
128 as listed in Supplemental Table 1. We activated transcription factors in the pilot experiment to
129 induce directed differentiation through overexpression of lineage-specific transcription factors[25],
130 based in part on prior work that established sets of key regulators leading to directed differentiation
131 in iPSCs[26,27]. NEUROG2 is a well-established driver of neuronal cell fate[28], whereas GATA5 and
132 MEOX1 overexpression caused loss of pluripotency markers in iPSCs[26] but led to an undefined
133 cell fate. By turning on the expression of these TFs in iPSCs through CRISPR-mediated activation,
134 we expect the cells to exhibit changes in RNA and ATAC profiles that are distinct from those
135 observed in the parental iPSCs, making this an ideal system to evaluate our CAT-ATAC method.
136 After dox induction of CRISPR activation targeting these transcription factors in iPSCs for 7 days,
137 the cells were pooled and analyzed by CAT-ATAC. It is worth noting that excess cells were
138 cryopreserved and used in subsequent iterations for optimization and no loss in performance was
139 observed, demonstrating the flexibility and utility of this assay.
140 We obtained a total of 6,656 cells with both RNA and ATAC profiles after data preprocessing and
141 filtering, of which 23% of cells could not be assigned to a gRNA identity, 30% of cells were
142 assigned to a single gRNA, 35% of cells were assigned to a pair of protospacers (PS) with the same
143 target, and 11% were classified as doublets which contained more than one gRNA with
144 incongruous targets that could be from droplets containing multiple cells (Figure 1D). The heatmap
145 in Figure 1E depicts guide expression in all 6,656 cells with their gRNA identity labeled in each
146 row. Upon filtering for only cells expressing paired gRNAs (n = 2,361 cells), we plotted the guide
147 classification in a UMAP embedding (Figure 1F) and observed 4 distinct clusters corresponding
148 to the 4 groups of cells receiving a different guide construct, indicating faithful gRNA
149 identification. Figure 1G shows the RNA expression of the target genes on the guide UMAP with
150 successful overexpression due to CRISPR-mediated gene activation. POU5F1 (gene name for
151 OCT4) is a canonical iPSC marker, and its expression can be seen in the negative control gRNA
152 cluster as well as the MEOX1 cluster but is reduced in NEUROG2 and GATA5. Violin plots in
153 Supplemental Figure 1A further demonstrate that TF upregulation is only observed in the cluster
154 receiving the corresponding gRNAs, and POU5F1 is downregulated in both the GATA5 and
155 NEUROG2 group, but not the MEOX1 group, suggesting that unlike cells in the former two

156 groups, MEOX1 cells were not fully differentiated and may not yet have exited their pluripotent
157 state, or may be on a lineage trajectory that requires sustained expression of POU5F1. Guide count
158 rank plot in Figure 1H shows a median of 108 guide UMIs per cell. This is lower than the reported
159 guide UMI of ~1000 in Perturb-seq[24] due to differences in the assay chemistries and cell lines used
160 but is sufficient to assign guide identity to a majority of the cells with high confidence, supported
161 by the target gene expression data for each gRNA cluster.
162 Next, we proceeded with further evaluations of the pilot CAT-ATAC data by integrating both the
163 RNA and ATAC profiles for each cell. Shown in Supplemental Figure 1B, we performed
164 clustering and constructed UMAP embeddings with just the RNA, ATAC, or weighted nearest
165 neighbor (WNN) analysis[29]. We classified the cells into 5 clusters, including one for each of the
166 overexpressed TF, one for iPSCs, and another termed as intermediate due to the lack of expression
167 of iPSC markers or targeted TFs. Motif enrichment was carried out for each cluster using Signac
168 [30,31], and we observed a high level of correlation between RNA expression of the TF (in purple)
169 and enrichment of its motif (in brown) (Figure 1I). Similarly, transcription factor footprinting
170 analysis indicates increased Tn5 accessibility flanking the motif locations (Figure 1J). We
171 compared CAT-ATAC to other recently published single-cell RNA-seq, ATAC-seq and joint
172 multiome technologies and found better performance of RNA and ATAC capture, including
173 increased RNA UMIs and genes per cell (Supplemental Figure 1C and D), as well as increased
174 ATAC fragments and peaks per cell (Supplemental Figure 1E and F). This indicated that CAT-
175 ATAC can successfully be used to analyze perturbation-specific effects on RNA and ATAC
176 profiles of cells.

177 **Genome-wide CRISPRi screen identifies gene knockdowns that affect sensitivity to**
178 **dasatinib.**

179 Next, we sought to integrate and apply the RNA and ATAC data to identify how genes affected
180 by specific perturbations drive the mechanisms of cancer drug resistance and identify targets that
181 can disrupt the transcriptional programs responsible. For this application, we targeted CML
182 resistance mechanisms to TKIs that are driven by activation of pathways independent of BCR-
183 ABL mutations. We performed a genome-wide CRISPRi screen to identify genetic modifiers of
184 sensitivity to the TKI dasatinib and hypothesized that some of these genetic modifiers must control
185 transcriptional programs driving resistance.

186 The screen was performed in K562 cells, a CML cell line that harbors a BCR-ABL mutation, and
187 the experimental design is outlined in Supplemental Figure 2A. After observing high concordance
188 between the replicates (Figure 2A) we identified 164 dasatinib sensitizing gene knockdowns and
189 800 dasatinib de-sensitizing gene knockdowns that were differentially enriched using an adjusted
190 p-value cutoff of 0.05 and a phenotype score cutoff of +/- 0.1 (Figure 2B). These include
191 previously reported known TKI sensitizing genes BCL2L1, MCL1 and PI3K-AKT pathway genes
192 as well as desensitizing genes BAX and LZTR1 (Figure 2B and Supplemental Figure 2B)[32,21,33].
193 We decided to focus on sensitizing gene hits as we hypothesized that some of these genes might
194 constitute the cellular response to dasatinib and potentially be responsible for resistance due to
195 aberrant pathway activation. We filtered down the list of 164 sensitizing gene hits to 60 hits by
196 first removing 61 genes which had significant growth phenotype scores in our vehicle control
197 sample or which were identified as a core essential gene by Hart et. al. [34] and by keeping only the
198 genes which have nuclear function to focus on potential core components of gene regulatory
199 networks (Supplemental Figure 2C). We performed arrayed validation for these 60 gene hits by
200 transducing K562 cells expressing CRISPRi machinery with an arrayed dual-guide library for the
201 60 genes and 2 non-targeting controls such that each vector targeting one gene had the 2 top
202 performing guides from the pooled survival screen (Figure 2C). Following FACS analysis we
203 calculated percent increase in cell death and confirmed that all the gene knockdowns sensitized
204 cells to dasatinib when compared to NTCs (Figure 2D). GO analysis indicated that these 60 genes
205 broadly included functional classification of transcriptional regulation, protein ubiquitination and
206 RNA binding (Figure 2E). Of the 20 genes with transcriptional control functions there were 7
207 genes with known transcriptional repression function, 3 genes with known transcriptional activator
208 function and 10 genes with unknown transcriptional activity. The transcriptional networks
209 controlled by these 20 genes could regulate cellular resistance to dasatinib and reveal new potential
210 therapeutic targets. We hypothesized the global changes to the transcriptomic and DNA
211 accessibility profiles after knockdown would shed light on the underlying regulatory networks and
212 effectors that confer resistance to dasatinib. Hence, we used CAT-ATAC to evaluate the effect of
213 gene perturbations simultaneously on transcriptome and DNA accessibility from the same cells.
214 **Identification of perturbation-specific effects on dasatinib-induced alterations to DNA**
215 **accessibility and transcription.**

K562 cells expressing CRISPRi machinery were transduced with a pooled lentiviral library containing 26 pairs of gRNAs. The 26-element library comprised 20 dual-guide vectors targeting the 20 transcriptional regulatory genes identified above and 6 non-targeting gRNA pairs. The gRNA sequences used are listed in Supplemental Table 1. We harvested the cells 48hr after dasatinib treatment to minimize loss due to cell death and investigate the earliest transcriptional and DNA accessibility changes that happen upstream of cell death (Figure 3A). The experimental design consisted of 2 biological replicates resulting in 4 pooled CAT-ATAC samples, two treated with dasatinib and two treated with DMSO as a vehicle control. We processed the 4 samples individually using the Seurat[30] analysis package and achieved 72-77% efficiency of guide assignment across the 4 samples (Supplemental Figure 3A). The UMAP shows high concordance between the replicates (Figure 3B). After guide assignment and filtering, we observed significant knockdowns (median KD 85.12% +/- 10% SD) for all the targeted genes except for those where the percentage of cells expressing the target gene was very low (Figure 3C). Concomitantly we saw a decrease in promoter accessibility for many of the targeted genes based on the scATAC data (Figure 3D and Supplemental Figure 3B).

Erythroid differentiation is one of the hallmarks of TKI treatment[35]. To assess whether dasatinib treatment in the experiment produced the expected effect on cells, we used the genes identified by Velten et al. 2017[36] to assess erythroid differentiation in response to dasatinib treatment. We calculated a module score for these genes and saw an enrichment of the module in the dasatinib treated cells (Figure 3E) confirming the erythroid differentiation response and our ability to successfully assess transcriptional responses to dasatinib treatment in the context of genetic perturbations. We performed differential gene expression and chromatin accessibility analysis and observed perturbation-specific changes in gene expression and ATAC peaks (Supplemental Table 2, Supplemental Table 3, Supplemental Figure 4A and 4B). We observed that cells harboring knockdowns for the genes HIC2, KMT2B, YEATS4, PIAS1, PQBP1 and GPBP1L1 formed clusters when compared to the non-targeting guide expressing cells (Supplemental Figure 3C). Amongst these genes, HIC2 knockdown had the strongest effect on both the transcriptome and ATAC-seq profiles especially in the dasatinib-treated group (Figure 3F), leading us to focus on identifying genes downstream of HIC2 that might confer resistance to dasatinib. ChromVar[37] can take sparse chromatin accessibility data from single cells as input and predict enrichment of transcription factor motifs. Using ChromVar[37] we identified perturbation-specific enrichment of

247 transcription factor motifs (Figure 3G). It has been previously reported that HIC2 has repressor
248 activity[38]. Transcriptional repressors exhibit anticorrelation between motif activity and gene
249 expression[39], because chromatin bound by repressors is generally not accessible, and loss of
250 repressor expression would expose more motifs normally occupied by the repressor. In agreement
251 with these data, we find HIC2 motif to be enriched in cells where HIC2 expression is
252 downregulated with CRISPRi (Figure 3H and 3I). Having the paired RNA-seq and ATAC-seq data
253 from the same single cells, along with perturbation identities, was essential in establishing the anti-
254 correlative relationship between HIC2 expression and motif enrichment. We thus demonstrate that,
255 when applied in specific cellular contexts, CAT-ATAC can reveal novel context-dependent
256 activation and repressive behaviors of transcription factors.

257 **HIC2 indirectly activates genes potentially underlying dasatinib resistance.**

258 Since both our genome-wide CRISPRi screen and CAT-ATAC data identified HIC2 as a
259 significant hit we went on to further identify genes controlled by HIC2 that drive dasatinib
260 resistance. It was clear from the differential expression analysis that the HIC2 knockdown group
261 had broadly altered gene expression patterns under dasatinib treatment conditions (Figure 3F),
262 leading us to hypothesize that HIC2 targets' expression are altered upon treatment with dasatinib.
263 Considering that HIC2 downregulation sensitizes cells to dasatinib we hypothesized that some of
264 the genes differentially expressed upon HIC2 knockdown should also affect sensitivity to
265 dasatinib. Since HIC2 has repressor activity, the genes upregulated in the HIC2 knockdown cluster
266 (DASA-HIC2 cells) would potentially be direct targets of HIC2. De-repression of these genes in
267 K562 cells should sensitize them to dasatinib. In genes downregulated in DASA-HIC2 cells we
268 expected that downregulation of these genes in K562 would sensitize the cells to dasatinib and
269 their upregulation upon treatment with dasatinib could constitute the resistance response of the
270 cells to dasatinib. We plotted the mean normalized read counts for the differentially expressed
271 genes in the three cell groups, DMSO-treated cells expressing non-targeting control gRNAs
272 (DMSO_NT), dasatinib-treated cells expressing non-targeting control gRNAs (DASA_NT) and
273 dasatinib-treated cells expressing HIC2 gRNAs (DASA_HIC2) and clustered them using k-means
274 clustering to identify patterns of gene expression changes (Figure 4A). We performed GO analysis
275 for two sets of genes which follow Up-Down-Up (cluster 4) and Down-Up-Down (cluster 1)
276 patterns of gene expression in the three conditions DMSO_NT, DASA_NT and DASA_HIC2
277 respectively (Figure 4B and 4C). We observed that genes following the Up-Down-Up pattern

278 which potentially include direct targets of HIC2 are mostly involved with processes exhibiting
279 post-transcriptional control of gene expression (Figure 4B) while genes following a Down-Up-
280 Down pattern of gene expression, which may include dasatinib-resistance genes, were associated
281 with transcriptional regulation functions (Figure 4C). As CAT-ATAC can only help us infer
282 transcriptional control mechanisms we could not identify the genes controlled by direct targets of
283 HIC2. Hence, we focused on identifying how the genes downregulated in the DASA-HIC2 cluster
284 might contribute to dasatinib resistance via construction of gene regulatory networks.

285 **Gene Regulatory Networks identify novel pathways involved in dasatinib resistance**

286 Recent single-cell GRN inference methods use paired scRNA-seq and scATAC-seq data to infer
287 direct regulatory relationships between genes[40,41,42]. We used Pando[42] to construct a gene
288 regulatory network using scRNA and scATAC-seq data (Supplementary Figure 5). We queried
289 this network for the genes downregulated in the DASA-HIC2 cluster and identified a sub-network
290 comprising of 1st degree connections to the genes in the query list (Figure 4D). The orange-colored
291 connections show activating interactions while the grey connections represent repressive
292 interactions. Upon inspecting the expression levels of genes in the subnetwork we found that
293 majority of genes with activating interactions were upregulated in the DASA_NT cells as
294 compared to DMSO_NT and DASA_HIC2 cells (Supplementary Figure 6A) while genes with
295 repressive interactions did not follow this pattern of expression (Supplementary Figure 6B). We
296 hypothesized that the genes with activating interactions in this sub-network are necessary for
297 dasatinib resistance.

298 One such gene is ZFPM2, a transcription factor which is one of the central nodes in the sub-
299 network in Figure 4D. The majority of interactions exhibited by ZFPM2 in this sub-network are
300 activating interactions, leading us to hypothesize that ZFPM2 may be responsible for dasatinib
301 resistance. One of ZFPM2's activating interactors is the transcription factor TCF4. Upon
302 inspecting the TCF4 locus we identified an ATAC peak upstream of TCF4 gene which was less
303 pronounced in the DASA-HIC2 cells and harbored a predicted ZFPM2 binding site (Figure 4E and
304 4F). Decrease in accessibility of this peak correlated with a decrease in expression of the TCF4
305 gene in the DASA-HIC2 cell group (Figure 4E). We found that this peak overlaps with active
306 histone marks H3K27Ac and P300 in K562 cells and is a potential active enhancer for TCF4 in
307 K562 cells (Figure 4F). ZFPM2 could control TCF4 expression by binding an enhancer upstream
308 of TCF4, an effector for the WNT signaling pathway[43] which is known to be activated in the

309 context of TKI resistance[44,45,46]. Thus, ZFPM2 may activate WNT downstream targets by
310 controlling TCF4 expression, thereby conferring dasatinib resistance. Information about
311 perturbation status of the target gene, ATAC-seq and RNA-seq data from the same cell helped us
312 identify that HIC2 gene is necessary for ZFPM2 expression which in turn might contribute to
313 activating TCF4 gene as a WNT pathway effector necessary for cell survival when treated with
314 dasatinib. Paired RNA-seq and ATAC-seq data from CAT-ATAC allowed for the identification
315 of the cis-regulatory relationship between the TCF4 promoter and accessible DNA site upstream
316 of TCF4. CAT-ATAC thus provides the ability to build tesTable mechanistic hypothesis by
317 combining multiple data types from different aspects of cellular regulatory processes.

318 **ZFPM2 is necessary for increased dasatinib resistance.**

319 To determine whether ZFPM2 is necessary for dasatinib resistance, we knocked down HIC2
320 (positive control), and ZFPM2 using CRISPRi in K562 cells (guide sequences in Supplemental
321 Table 1). qRT-PCR analysis showed significant knockdown of gene expression for both genes
322 when compared to non-targeting controls (Figure 5A). We next performed flow cytometry analysis
323 post-dasatinib treatment to elucidate the effect of knockdowns on dasatinib sensitivity. The gating
324 strategy for the flow cytometry analysis is illustrated in Supplemental Figure 7. After treating the
325 cells with 8nM dasatinib for 72 hours, we observed that the cells with gene knockdowns showed
326 a significant decrease in survival compared to the non-targeting control expressing cells thereby
327 indicating these genes caused increased resistance to dasatinib (Figure 5B). We also performed a
328 longitudinal timecourse experiment quantifying % viability of cells over time and saw significant
329 decrease in viability for DASA-HIC2 and DASA-ZFPM2 cells across all the timepoints tested
330 (Figure 5C). Thus, we can conclude that ZFPM2 confers dasatinib resistance.

331 **Discussion**

332 We have shown that CAT-ATAC robustly captures guide RNAs, resulting in up to 77% of guide
333 assignment in tested cell lines, while producing high-quality scRNA- and scATAC-seq data.
334 The main application of this approach is to establish a novel platform for conducting CRISPR
335 screens, by combining small-scale to genome-wide gene perturbations with massively parallel
336 sequencing of RNA and chromatin accessibility at the single-cell level. CAT-ATAC generates
337 paired gene expression and open chromatin data without the need to perform separate pooled
338 screens for each of the modalities, thus reducing the avenues for technical and biological
339 confounders. The experiments in this study used a primer containing the CS1 direct capture

340 sequence to reverse transcribe guide RNAs (Figure 1B), but this assay is designed to be universally
341 compatible with any type of CRISPR perturbation or synthetic construct by including a splinted
342 reverse transcription primer against the targeted sequence. Two recent studies have demonstrated
343 alternative approaches to perform CRISPR screens with multimodal gene expression and open
344 chromatin readouts[47,48]. However, these methods would either require re-cloning most labs'
345 existing CRISPR guide libraries into a modified vector[47] or implementing complicated
346 experimental approaches with multiple rounds of barcoding[48]. CAT-ATAC offers a
347 straightforward and easy-to-perform workflow, incorporating minimal modifications to the 10X
348 Genomics Multiome assay and is compatible with existing designs of the CRISPR gRNA libraries
349 while reporting comparable capture efficiencies of up to 77%, which could likely be optimized
350 further by employing enrichment strategies using biotinylated PCR primers similar to that used for
351 MultiPerturb-seq[48]. Additionally, the ability to directly capture guide RNAs enables the use of
352 constructs that express multiple guide RNAs targeting one or several genes[23,49,50] which makes the
353 studies of combinatorial perturbations feasible. Our chemical-genetic screen identified genetic
354 modifiers that sensitize cells to dasatinib. These genetic modifiers included several genes with
355 transcriptional regulatory activity. We show for the first time that HIC2 expression drives cellular
356 resistance to dasatinib. Our results support an observation from a study comparing the gene
357 expression profiles of CD34+ cells from CML patients and controls which reported overexpression
358 of HIC2 in patient cells suggesting a role for HIC2 in disease progression[51]. When applied to
359 identifying cancer drug resistance mechanisms CAT-ATAC was able to identify a gene regulatory
360 network that is activated under dasatinib treatment and requires expression of HIC2 gene, and to
361 further identify ZFPM2 as a central transcription factor in this network. Using the scATAC and
362 scRNA-seq data from HIC2 knockdown cells we were able to predict a novel mechanism whereby
363 ZFPM2 putatively binds a cis-regulatory element upstream of TCF4, thereby controlling
364 expression of TCF4, a transcription factor known to be an effector for the WNT signaling pathway
365 that contributes to TKI resistance for CML and other myeloid leukemias.
366 Besides capturing guide RNAs in CRISPR screens, this method is further extendable to any other
367 nucleic acid target, such as a reporter DNA or RNA, BCR/TCR immunoglobulin sequences, or
368 DNA conjugated to antibodies and antigens which can be used to measure protein abundance along
369 with gene expression and chromatin accessibility. These types of multiparametric single-cell
370 readouts will be transformative to shape our understanding of fundamental biological processes,

cell fate decisions and disease mechanisms by harnessing the heterogeneity inherent to complex cell populations instead of reducing datasets to bulk analysis. By providing a better understanding of the link between DNA accessibility, gene transcription, and protein translation and processing, these approaches can be broadly applied to probe regulation by non-coding regions of the genome, allowing for more systematic studies of the links between gene regulatory elements and gene expression. In addition, these tools can be leveraged to gain biological and mechanistic insights into the relationship between recognition motifs, transcription factor binding, and downstream gene activity, enabling further exploration of the dynamics of transcription factor-driven cell fate changes. Furthermore, these types of multimodal perturbation datasets will prove invaluable for building foundational models in biology such as the virtual cell[52].

Data availability

Raw and processed CAT-ATAC sequencing data can be downloaded from Gene Expression Omnibus (GSE288996). Data from previously published studies were from the Sequence Read Archive or Gene Expression Omnibus: SHARE-seq[53] (GSE140203), SNARE-seq[54] (GSE126074), Paired-seq[55] (GSE130399), sci-CAR-seq[56] (GSE117089) and MultiPerturb-seq[48] (GSE277747). The human genome hg38 (GENCODE v32/Ensembl98) was from 10x Genomics (https://cf.10xgenomics.com/supp/cell-exp/refdata-gex-GRCh38-2020-A.tar.gz).

Code availability

All the code generated for analysis of the CAT-ATAC data including for counting guides, guide to cell assignment and downstream differential expression and gene regulatory analysis will be made available upon request to the corresponding authors and released upon publication of this manuscript.

Acknowledgements

We thank the UCSF Center for Advanced Technology (CAT) for NGS sequencing services. Sequencing was performed at the UCSF CAT, supported by UCSF PBBR, RRP IMIA, and NIH 1S10OD028511-01 grants. We would also like to thank LGR team members Greyson Lewis, Sailaja Peddada, and Adam Litterman for their advice and Lauren Enriquez for help in setting up automation protocols for guide cloning. Some Figure elements were created using BioRender.com.

Funding

This work is supported by the Laboratory for Genomics Research established by GSK, UCSF, and UC Berkeley.

402 **Author Contributions**

403 K.S, Y.A.Y, E.D.C and L.P contributed to the study's overall conception, design, interpretation
404 and co-wrote the manuscript. E.D.C conceptualized CAT-ATAC, Y.A.Y performed the CAT-
405 ATAC optimization and method development experiments. K.S designed the screening concept.
406 K.S and K.M worked on genome-wide dasatinib screen and CAT-ATAC follow-up screen. Y.A.Y,
407 K.S. and K.F performed data analysis and interpretation. V.S, K.F and S.F wrote the data analysis
408 pipeline. J.P wrote the code for guide RNA assignment. G.A helped in making the multiome
409 libraries. K.M and H.E worked on validation experiments. R.M cloned arrayed guide plasmids.
410 R.E, K.S, E.D.C and S.S helped Y.A.Y troubleshoot CAT-ATAC method development
411 experiments.

412 **Conflict of Interest**

413 Y.A.Y, V.S, J.P, R.E, G.A and S.S are employees of GSK. E.D.C. is a co-founder of Survey
414 Genomics.

415 **Methods**

416 **Cell lines**

417 The human biological samples were sourced ethically, and their research use was in accord with
418 the terms of the informed consent under an IRB/EC approved protocol. iPSC cell line (Proprietary
419 for GSK use generated by Takara) was engineered to express doxycycline inducible CRISPRa
420 machinery (dCas9-VP64-p65-Rta). The CRISPRa machinery was integrated into the AAVS1 safe
421 harbor locus to minimize silencing.

422 The K562-CRISPRi cell line was a gift from the Gilbert lab. K562-CRISPRi line was maintained
423 in the standard conditions of DMEM+10% FBS with antibiotics as recommended by ATCC.

424 **gRNA cloning**

425 For the pilot experiment top two guide RNAs from the Weissman V2 CRISPRi library targeting
426 the 3 transcription factors (GATA1, MEOX1 and NGN2) and were cloned in dual guide lentiviral
427 construct compatible with Perturb-seq (pLGR002 Addgene #188320) using Gibson assembly. For
428 the arrayed validation of hits from the dasatinib screen top two guides which exhibited the highest
429 phenotype score for sensitizing the cells to dasatinib were cloned individually in the dual guide
430 gRNA vector using Gibson assembly. Gibson assembly process was as follows:

431 Dual guide RNAs were cloned in a lentiviral vector backbone pLGR002 (Addgene #188320) using
432 Gibson assembly cloning. An insert DNA sequence with flanking BsmBI restriction sites and

433 containing a constant region with perturb-seq compatible capture sequence 1 (CS1:
434 GCTTTAAGGCCGGTCCTAGCAA) was cloned into an intermediate transfer vector and
435 transformed into bacteria for plasmid amplification (See insert sequence below). To prepare a
436 linear insert for cloning the amplified plasmid was digested using BsmBI restriction digestion and
437 the desired fragment was purified using gel purification. pLGR002 vector was also linearized using
438 BsmBI restriction digestion. ssDNA oligos for guides were ordered with overhangs which partially
439 overlap the constant region and mU6 promoter for guide 1 and constant region and hU6 promoter
440 for guide 2. Linearized vector, insert and guide oligos were then assembled into dual guide vectors
441 in a Gibson assembly reaction. . Colonies were validated using sanger sequencing post cloning.
442 Insert sequence:
443 gtttcAGAGCTAAGCACAAGAGTGCATAGCAAGTTGAAATAAGGCTAGTCCGTTTACA
444 ACTTGGCCGCTTTAAGGCCGGTCCTAGCAAGGCCAAGTGGCACCCGAGTCGGGTGC
445 TTTTTTTGCTCGAATCTACACTCAGCTATGGCGCGCCCCAAGGTCGGGCAGGAAGAG
446 GGCCTATTTCCCATGATTCCTTCATATTTGCATATACGATACAAGGCTGTTAGAGAG
447 ATAATTGGAATTAATTTGACTGTAAACACAAAGATATTAGTACAAAATACGTGACGT
448 AGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAATTATGTTTTAAAATGGAC
449 TATCATATGCTTACCGTAACTTGAAAGTATTTCGATTTCTTGGCTTTATATATCTTGT
450 GGAAAGCCAgaaacatg
451 Plasmids for the vectors targeting the 20 selected genes along with 6 NTCs required for the pooled
452 multiome screen were then mixed in equimolar amounts to generate a pooled plasmid library.
453 **Lentivirus production and transduction**
454 For all arrayed experiments lentivirus was generated in an arrayed format by transfecting LentiX
455 cells (Takara 632180) with individual gRNA vectors. To generate the lentivirus for transducing
456 iPSC cells LentiX cells were forward transfected with the gRNA vectors and lentiviral packaging
457 plasmids (dR8.91 and MD2G) using the Mirus transfection reagent (MIR 2700). The three
458 plasmids were mixed in 1:1:0.1 ratios for gRNA vector, MD2G and dR8.91 respectively. The
459 lentivirus was then harvested 48hr post transfection in MTeSR1 medium for transduction in iPSCs.
460 The lentivirus for arrayed validation experiment was generated in a 96 well format where the
461 LentiX cells were reverse transfected using TransIT-VirusGEN® Reagent (MIR 6704). The gRNA
462 and packaging plasmids were mixed in the same 1:1:0.1 ratio followed by lentivirus harvest in
463 media with 10% FBS. Pooled lentiviruses were generated in either 6 well or 15 cm dishes

464 depending on the amount of lentivirus required. LentiX cells were forward transduced using Mirus
465 transfection reagent and the plasmids were mixed in the ratios mentioned above. Lentivirus was
466 then harvested in media containing 10% FBS.

467 **Genome-wide CRISPRi screen**

468 K562 cells containing dCas9-KRAB CRISPRi machinery were infected with CRISPRi V2 library
469 containing the top5 guides (Addgene 83969) to achieve 1000-fold coverage at an MOI of 0.3.
470 Transduced cells were selected with puromycin until 90% of the cell population showed BFP
471 expression indicating guide expression. Cells were then treated with a single dose of 8nM dasatinib
472 (determined from a drug titration experiment) for 72hr. This was followed by recovery for 5 days.
473 An equivalent number of transduced cells were treated with DMSO as a vehicle control. Both the
474 dasatinib and DMSO treated cells were then collected and genomic DNA was isolated using the
475 Macherey-Nagel Nucleospin Blood XL kit (740950). The cassette encoding the gRNAs was then
476 amplified by PCR and gRNA abundance was determined by next generation sequencing (NGS) as
477 previously described[57,58,59]. The libraries were sequenced on an Illumina NextSeq550 sequencer
478 and the sequencing data was analyzed using the screen processing pipeline[57].

479 **Arrayed validation of the hits from CRISPRi screen**

480 A dual-guide lentiviral library targeting 60 genes and two non-targeting controls was prepared in-
481 house as described above and used for arrayed lentiviral generation in HEK 293T cells. K562-
482 CRISPRi cells were then "spinfected" at 1000 RCF for 1 hour at 22C with this lentivirus in an
483 arrayed fashion in the presence of 1 ug/mL of polybrene, left overnight at 37C, and then selected
484 via puromycin at a concentration of 2ug/mL for 48 hours. Following selection, guide-infected cell
485 lines were treated with 8nM Dasatinib and their viability was recorded using Propidium Iodide
486 nuclear stain (Invitrogen P1304MP) every 24 hours following treatment on an Invitrogen Attune
487 NxT Flow Cytometer.

488 **CAT-ATAC**

489 **CAT-ATAC pilot experiment in iPSCs**

490 iPSCs were transduced at MOI of 0.3 in the presence of 1 ug/mL of polybrene overnight in an
491 arrayed format, followed by puromycin selection for one week at a concentration of 2 ug/mL.
492 Three transcription factors were overexpressed (NEUROG2, GATA5, MEOX1) by supplementing
493 doxycycline (4uM) in the media (mTeSR1, StemCell Technologies) for 7 days prior to CAT-
494 ATAC analysis. As a control, a non-targeting scrambled gRNA sequence was used. After 7 days

495 of overexpression cells were dissociated into a single-cell suspension using Accutase
496 (ThermoFisher Scientific A1110501), counted and then equal numbers of cells per perturbation
497 were mixed to make a pool of cells prior to processing for CAT-ATAC.

498 **Pooled Lentivirus Generation and Dasatinib Treatment**

499 A pooled lentiviral library containing 26 dual-guide vectors targeting 20 previously identified
500 transciptional regulatory genes and 6 non-targeting genes was generated as described previously.
501 The resulting virus was then titered to determine the amount of virus required for desired MOI of
502 0.1 in K562 cells. Accordingly, 1.5*10^7 cells containing CRISPRi machinery were spun down
503 and resuspended directly into 2.4mL of this pooled virus and "spinfected" at 1000 RCF for 1 hour
504 at 22C in the presence of 1 ug/mL of polybrene and incubating overnight at 37C. Guide infected
505 cells were sorted via FACS based on BFP expression and allowed to recover for 5 days. 2.5*10^7
506 of these sorted cells were then aliquoted into 4 shaking flasks, two of which were treated with
507 Dasatinib and two of which were treated with DMSO for 48 hours before being collected and
508 aliquoted for CAT-ATAC.

509 **CAT-ATAC method**

510 Detailed CAT-ATAC protocol is in Supplemental Text 1. Briefly cells were harvested, nuclei were
511 extracted and tagmented according to the 10X Multiome protocol. During the GEM Generation
512 step, a duplex splint oligo is spiked into the barcoding reaction mix to allow reverse transcription
513 of gRNAs and ligation onto the capture oligo of 10X Multiome gel beads. The resulting cDNA
514 from captured guide RNAs is amplified to make an Illumina-compatible sequencing library
515 containing P5 and P7 adaptors as well as sample indexes. GEX libraries were sequenced on
516 NovaSeqX with paired-end sequencing, 28x10x10x90, at a minimum of 20,000 read pairs/cell.
517 ATAC and CRISPR libraries were pooled at a 5:1 ratio and were sequenced on NovaSeqX with
518 paired-end sequencing, 151x8x24x151 (since at least 75bp PE is needed to be able to sequence
519 through protospacer region for CRISPR libraries), at a minimum of 25,000 read pairs/cell for
520 ATAC libraries and 5,000 read pairs/cell for CRISPR libraries.

521 **CAT-ATAC data analysis**

522 **Processing FASTQ files after NGS**

523 We extracted the protospacer-cell barcode pairs from the guide capture FASTQ files using a
524 custom Bash script that first finds the reads containing both the protospacer and capture sequence
525 via BBDuk alignment (https://sourceforge.net/projects/bbmap/) at a Hamming distance=1. The

526 matching reads are then demultiplexing based on unique UMIs using a guide calling algorithm
527 described in the Guide Assignment section, and count tables are generated. GEX and ATAC
528 FASTQs were aligned to GRCh38 version of the human genome using prebuilt packages for Cell
529 Ranger ARC v.2.0.1. Post alignment filtering, barcode counting and counting of GEX and ATAC
530 molecules in the FASTQ files was performed using the cellranger-arc count function.

531 **Guide Assignment**

532 For the pilot CRISPRa experiment in iPSCs cells were assigned a guide identity using HTOdemux
533 function in Seurat which is based on a previously published algorithm that was developed for cell
534 hashing oligos[60]. We found that this method had low signal to noise ratio in our pooled CRISPRi
535 screen, so we developed a new algorithm to assign guides to cells. For the pooled CRISPRi CAT-
536 ATAC screen the UMIs per cell for each guide are modeled using a mixture of a zero-inflated
537 negative binomial distribution and a negative binomial distribution (without zero-inflation). To
538 maintain identifiability, the parameter space is constrained so that the negative binomial
539 distribution without zero inflation has a larger mean than the negative binomial component of the
540 zero-inflated distribution. Estimates for the negative binomial parameters and mixture proportions
541 were obtained by maximum likelihood and used to determine the posterior probability that each
542 observation arose from the negative binomial component with the larger mean. A cell was
543 classified as expressing a guide if the posterior probability that the corresponding UMI
544 measurement came from the negative binomial component with the larger mean was greater than
545 0.5.

546 **Data QC, filtering, integration and differential testing**

547 Seurat v.4.3.0[29] was used for subsequent analysis. Cells with mitochondrial RNA percentage >
548 20%, RNA count <1000 and >40000, nucleosome signal > 2, and TSS enrichment <1 were filtered
549 out. Seurat SCTransform v2.0 was used to normalize and scale the data as well as to find variable
550 features. Cell cycle-associated genes were also regressed out using Seurat's CellCycleScoring
551 SCTransform functions to mitigate their effects on downstream analyses. We then ran Mixscape
552 on each sample to calculate perturbation-specific signatures for each cell. This step included the
553 identification and exclusion of cells that evaded CRISPR-mediated perturbation. The method also
554 facilitated the visualization of similarities and differences among cells subjected to various
555 perturbations.

556 The top 3,000 highly variable genes were used for PCA and UMAP dimensionality reduction.
557 Libraries were integrated using Seurat's merge and IntegrateEmbeddings functions. Cells were
558 clustered using the FindClusters function and the Louvain algorithm. To identify significant
559 changes associated with each perturbation, differential expression testing was conducted using
560 Wilcoxon rank sum test through Seurat FindMarkers function and differential accessibility testing
561 was conducted using logistic regression in Signac.

562 **Module score analysis**

563 Genes associated with erythroid differentiation were identified from Velten et.al. [36]. Module scores
564 for these sets of genes were calculated using AddModuleScore function in Seurat. These module
565 scores were then plotted on the rnaUMAP to visualize the enrichment of the module in specific
566 treatment groups.

567 **H3K27Ac and P300 ChIPseq data**

568 Bigwig files for H3K27Ac (ENCSR000AKP) and P300 (ENCSR000EGE) ChIPseq datasets in
569 K562 were downloaded from ENCODE (The Encyclopedia of DNA Elements) consortium.

570 **Gene Ontology analysis**

571 Selected gene sets either identified as hits in the CRISPR screens or as significant differentially
572 expressed genes were used for gene ontology analysis as mentioned in the results section. Gene
573 ontology analysis was performed using the DAVID knowledgebase
574 (https://davidbioinformatics.nih.gov/).

575 **Motif enrichment and footprinting analysis**

576 Motif and footprinting analysis were performed using Signac's AddMotifs, FindMotifs, and
577 FootPrint functions. Motif activities were calculated using Signac's RunChromVAR function and
578 JASPAR2024 database[61].

579 **Gene Regulatory Network inference with Pando**

580 To perform gene regulatory network inference, we used the functions initiate_grn, find_motifs,
581 infer_grn, find_modules, and NetworkModules from the Pando package[42] . We first built a global
582 GRN including all the cells in the integrated multiome object. To identify HIC2 specific network
583 we subset this global GRN using a list of genes that were differentially downregulated in
584 DASA_HIC2 cells as compared to DASA_NT cells.

585 **Validation of ZFPM2**

586 The top two guide RNAs from the Weissman V2 CRISPRi library targeting ZFPM2 gene were
587 cloned into the LGR dual guide vector as described earlier. Lentivirus was generated as described
588 before and concentrated using Lenti-X™ Concentrator (cat # 631232). K562 cells containing
589 CRISPRi machinery were then transduced with guide lentiviruses for HIC2 (positive control), a
590 non-targeting control (NTC) and ZFPM2. The transduced cells were selected with puromycin and
591 then used for downstream validation experiments.

592 **Validation of target knockdown post CRISPRi**

593 Total RNA was extracted from K562 CRISPRi constructs targeted against NTC, HIC-2,
594 and ZFPM2 (*n* = 3/each group) by using RNeasy Plus Mini Kit (Cat # 74134) and converted to
595 cDNA using SuperScript Vilo cDNA synthesis kit (Cat # 11754050). A total of 1 µg RNA was
596 converted to cDNA. cDNA was amplified using a SYBR-green qPCR master mix (Cat #43-676-
597 59) and specific primers in Supplemental Table 4. GAPDH was used as the reference gene.
598 The PCR steps consisted of an initial activation 50°C for 2 mins, pre-soak step at 95 °C for 10
599 min, denaturation of at 95 °C for 15 sec, followed by annealing step at 60 °C for 1 min for 40
600 cycles. The last step is the melting curve (95°C for 15 sec, 60°C for 15 sec, and 95°C for 15 sec).
601 Detection, quantification, and data analysis were performed in the CFX manager real-time
602 detection system (Bio-Rad). The Ct values were calculated for each gene and were compared with
603 the reference gene GAPDH, followed by estimation of ΔΔCt and fold change ($2^{-\Delta\Delta Ct}$) to assess the
604 relative gene expression.

605 **Dasatinib treatment and cell viability analysis**

606 Dasatinib (Millipore Sigma, cat. # SML2589) was diluted with DMSO to final concentration of
607 8nM, 0.01% DMSO was used as a vehicle control. K562 cells expressing either NTC, HIC2 or
608 ZFPM2 guides were treated with 8nM dasatinib for 72 hr. Live-dead staining was performed using
609 2uM calcien-AM (cat # C1430) and 2.5uM ETHD1 EthD-1 (cat # E1169) and cells were analyzed
610 using the Attune NextGen flow cytometer (Thermo Scientific, USA) at 72h, 96h, 120h, 144h and
611 168h post dasatinib treatment. FSC and SSC dot plot was used to gate out the debris present in the
612 suspension and percentages of live (calcien-AM positive) and dead (EthD-1 positive) cells were
613 estimated.

614

615

616 **References**

617 1. Shalem, O., Sanjana, N. E. & Zhang, F. High-throughput functional genomics using
618 CRISPR–Cas9. *Nat. Rev. Genet.* **16**, 299–311 (2015).
619 2. Bock, C. *et al.* High-content CRISPR screening. *Nat. Rev. Methods Primer* **2**, 1–23 (2022).
620 3. Przybyla, L. & Gilbert, L. A. A new era in functional genomics screens. *Nat. Rev. Genet.*
621 **23**, 89–103 (2022).
622 4. Shevade, K., Peddada, S., Mader, K. & Przybyla, L. Functional genomics in stem cell
623 models: considerations and applications. *Front. Cell Dev. Biol.* **11**, 1236553 (2023).
624 5. Adamson, B. *et al.* A Multiplexed Single-Cell CRISPR Screening Platform Enables
625 Systematic Dissection of the Unfolded Protein Response. *Cell* **167**, 1867-1882.e21
626 (2016).
627 6. Dixit, A. *et al.* Perturb-Seq: Dissecting Molecular Circuits with Scalable Single-Cell RNA
628 Profiling of Pooled Genetic Screens. *Cell* **167**, 1853-1866.e17 (2016).
629 7. Jaitin, D. A. *et al.* Dissecting Immune Circuits by Linking CRISPR-Pooled Screens with
630 Single-Cell RNA-Seq. *Cell* **167**, 1883-1896.e15 (2016).
631 8. Datlinger, P. *et al.* Pooled CRISPR screening with single-cell transcriptome readout. *Nat.*
632 *Methods* **14**, 297–301 (2017).
633 9. Valencia, A. M. & Pașca, S. P. Chromatin dynamics in human brain development and
634 disease. *Trends Cell Biol.* **32**, 98–101 (2022).
635 10. Klemm, S. L., Shipony, Z. & Greenleaf, W. J. Chromatin accessibility and the regulatory
636 epigenome. *Nat. Rev. Genet.* **20**, 207–220 (2019).
637 11. Buenrostro, J. D. *et al.* Single-cell chromatin accessibility reveals principles of
638 regulatory variation. *Nature* **523**, 486 (2015).

639 12. Rubin, A. J. *et al.* Coupled Single-Cell CRISPR Screening and Epigenomic Profiling
640 Reveals Causal Gene Regulatory Networks. *Cell* **176**, 361-376.e17 (2019).
641 13. Liscovitch-Brauer, N. *et al.* Profiling the genetic determinants of chromatin
642 accessibility with scalable single-cell CRISPR screens. *Nat. Biotechnol.* **39**, 1270–1277
643 (2021).
644 14. Pierce, S. E., Granja, J. M. & Greenleaf, W. J. High-throughput single-cell chromatin
645 accessibility CRISPR screens enable unbiased identification of regulatory networks in
646 cancer. *Nat. Commun.* **12**, 2969 (2021).
647 15. Zhang, K. *et al.* A single-cell atlas of chromatin accessibility in the human genome. *Cell*
648 **184**, 5985-6001.e19 (2021).
649 16. Li, Y. E. *et al.* A comparative atlas of single-cell chromatin accessibility in the human
650 brain. *Science* **382**, eadf7044 (2023).
651 17. Bukowski, K., Kciuk, M. & Kontek, R. Mechanisms of Multidrug Resistance in Cancer
652 Chemotherapy. *Int. J. Mol. Sci.* **21**, 3233 (2020).
653 18. Olivieri, A. & Manzione, L. Dasatinib: a new step in molecular target therapy. *Ann.*
654 *Oncol.* **18**, vi42–vi46 (2007).
655 19. Hu, Y. *et al.* Magnitude and Temporal Trend of the Chronic Myeloid Leukemia: On the
656 Basis of the Global Burden of Disease Study 2019. *JCO Glob. Oncol.* **7**, GO.21.00194
657 (2021).
658 20. Wagle, M. *et al.* A role for FOXO1 in BCR–ABL1-independent tyrosine kinase inhibitor
659 resistance in chronic myeloid leukemia. *Leukemia* **30**, 1493–1501 (2016).

660 21. Singh, P., Kumar, V., Gupta, S. K., Kumari, G. & Verma, M. Combating TKI resistance in
661 CML by inhibiting the PI3K/Akt/mTOR pathway in combination with TKIs: a review. *Med.*
662 *Oncol. Northwood Lond. Engl.* **38**, 10 (2021).
663 22. Mu, H., Zhu, X., Jia, H., Zhou, L. & Liu, H. Combination Therapies in Chronic Myeloid
664 Leukemia for Potential Treatment-Free Remission: Focus on Leukemia Stem Cells and
665 Immune Modulation. *Front. Oncol.* **11**, 643382 (2021).
666 23. Replogle, J. M. *et al.* Maximizing CRISPRi efficacy and accessibility with dual-sgRNA
667 libraries and optimal effectors. *eLife* **11**, e81856 (2022).
668 24. Replogle, J. M. *et al.* Combinatorial single-cell CRISPR screens by direct guide RNA
669 capture and targeted sequencing. *Nat. Biotechnol.* **38**, 954–961 (2020).
670 25. Oh, Y. & Jang, J. Directed Differentiation of Pluripotent Stem Cells by Transcription
671 Factors. *Mol. Cells* **42**, 200–209 (2019).
672 26. Ng, A. H. M. *et al.* A comprehensive library of human transcription factors for cell fate
673 engineering. *Nat. Biotechnol.* **39**, 510–519 (2021).
674 27. Joung, J. *et al.* A transcription factor atlas of directed differentiation. *Cell* **186**, 209-
675 229.e26 (2023).
676 28. Hulme, A. J., Maksour, S., St-Clair Glover, M., Miellet, S. & Dottori, M. Making neurons,
677 made easy: The use of Neurogenin-2 in neuronal differentiation. *Stem Cell Rep.* **17**, 14–
678 34 (2021).
679 29. Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573-
680 3587.e29 (2021).

681 30. Hao, Y. *et al.* Dictionary learning for integrative, multimodal and scalable single-cell
682 analysis. *Nat. Biotechnol.* **42**, 293–304 (2024).
683 31. Stuart, T., Srivastava, A., Madad, S., Lareau, C. A. & Satija, R. Single-cell chromatin
684 state analysis with Signac. *Nat. Methods* **18**, 1333–1341 (2021).
685 32. Tzifi, F. *et al.* The Role of BCL2 Family of Apoptosis Regulator Proteins in Acute and
686 Chronic Leukemias. *Adv. Hematol.* **2012**, 524308 (2012).
687 33. Zhang, H. *et al.* LZTR1: A promising adaptor of the CUL3 family. *Oncol. Lett.* **22**, 564
688 (2021).
689 34. Hart, T. *et al.* Evaluation and Design of Genome-Wide CRISPR/SpCas9 Knockout
690 Screens. *G3 Bethesda Md* **7**, 2719–2727 (2017).
691 35. Kang, C.-D. *et al.* Role of Ras/ERK-dependent pathway in the erythroid differentiation of
692 K562 cells. *Exp. Mol. Med.* **31**, 76–82 (1999).
693 36. Velten, L. *et al.* Human haematopoietic stem cell lineage commitment is a continuous
694 process. *Nat. Cell Biol.* **19**, 271–281 (2017).
695 37. Schep, A. N., Wu, B., Buenrostro, J. D. & Greenleaf, W. J. chromVAR: inferring
696 transcription-factor-associated accessibility from single-cell epigenomic data. *Nat.*
697 *Methods* **14**, 975–978 (2017).
698 38. Huang, P. *et al.* HIC2 controls developmental hemoglobin switching by repressing
699 BCL11A transcription. *Nat. Genet.* **54**, 1417–1426 (2022).
700 39. Berest, I. *et al.* Quantification of Differential Transcription Factor Activity and
701 Multiomics-Based Classification into Activators and Repressors: diffTF. *Cell Rep.* **29**,
702 3147-3159.e12 (2019).

703 40. Kim, D. *et al.* Gene regulatory network reconstruction: harnessing the power of single-
704 cell multi-omic data. *NPJ Syst. Biol. Appl.* **9**, 51 (2023).
705 41. Bravo González-Blas, C. *et al.* SCENIC+: single-cell multiomic inference of enhancers
706 and gene regulatory networks. *Nat. Methods* **20**, 1355–1367 (2023).
707 42. Fleck, J. S. *et al.* Inferring and perturbing cell fate regulomes in human brain organoids.
708 *Nature* **621**, 365–372 (2023).
709 43. Hrckulak, D. *et al.* Wnt Effector TCF4 Is Dispensable for Wnt Signaling in Human
710 Cancer Cells. *Genes* **9**, 439 (2018).
711 44. Zhang, B. *et al.* Microenvironmental protection of CML stem and progenitor cells from
712 tyrosine kinase inhibitors through N-cadherin and Wnt–β-catenin signaling. *Blood* **121**,
713 1824–1838 (2013).
714 45. Grassi, S. *et al.* The WNT Pathway Is Relevant for the BCR-ABL1-Independent
715 Resistance in Chronic Myeloid Leukemia. *Front. Oncol.* **9**, (2019).
716 46. Patel, A. B., O'Hare, T. & Deininger, M. W. Mechanisms of resistance to ABL kinase
717 inhibition in CML and the development of next generation ABL kinase inhibitors.
718 *Hematol. Oncol. Clin. North Am.* **31**, 589–612 (2017).
719 47. Metzner, E., Southard, K. M. & Norman, T. M. Multiome Perturb-seq unlocks scalable
720 discovery of integrated perturbation effects on the transcriptome and epigenome.
721 2024.07.26.605307 Preprint at https://doi.org/10.1101/2024.07.26.605307 (2024).
722 48. Yan, R. E. *et al.* Pooled CRISPR screens with joint single-nucleus chromatin
723 accessibility and transcriptome profiling. *Nat. Biotechnol.* 1–7 (2024)
724 doi:10.1038/s41587-024-02475-x.

725 49. Campa, C. C., Weisbach, N. R., Santinha, A. J., Incarnato, D. & Platt, R. J. Multiplexed
726 genome engineering by Cas12a and CRISPR arrays encoded on single transcripts. *Nat.*
727 *Methods* **16**, 887–893 (2019).
728 50. Hsiung, C. C.-S. *et al.* Engineered CRISPR-Cas12a for higher-order combinatorial
729 chromatin perturbations. *Nat. Biotechnol.* 1–15 (2024) doi:10.1038/s41587-024-02224-
730 0.
731 51. Bruns, I. *et al.* The hematopoietic stem cell in chronic phase CML is characterized by a
732 transcriptional profile resembling normal myeloid progenitor cells and reflecting loss of
733 quiescence. *Leukemia* **23**, 892–899 (2009).
734 52. Bunne, C. *et al.* How to build the virtual cell with artificial intelligence: Priorities and
735 opportunities. *Cell* **187**, 7045–7063 (2024).
736 53. Ma, S. *et al.* Chromatin Potential Identified by Shared Single-Cell Profiling of RNA and
737 Chromatin. *Cell* **183**, 1103-1116.e20 (2020).
738 54. Chen, S., Lake, B. B. & Zhang, K. High-throughput sequencing of the transcriptome and
739 chromatin accessibility in the same cell. *Nat. Biotechnol.* **37**, 1452–1457 (2019).
740 55. Zhu, C. *et al.* An ultra high-throughput method for single-cell joint analysis of open
741 chromatin and transcriptome. *Nat. Struct. Mol. Biol.* **26**, 1063–1070 (2019).
742 56. Cao, J. *et al.* Joint profiling of chromatin accessibility and gene expression in thousands
743 of single cells. *Science* **361**, 1380–1385 (2018).
744 57. Gilbert, L. A. *et al.* Genome-Scale CRISPR-Mediated Control of Gene Repression and
745 Activation. *Cell* **159**, 647–661 (2014).

58. Bassik, M. C. *et al*. A Systematic Mammalian Genetic Interaction Map Reveals Pathways Underlying Ricin Susceptibility. *Cell* **152**, 909–922 (2013).

59. Kampmann, M., Bassik, M. C. & Weissman, J. S. Functional genomics platform for pooled screening and mammalian genetic interaction maps. *Nat. Protoc*. **9**, 1825–1847 (2014).

60. Stoeckius, M. *et al*. Cell Hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. *Genome Biol*. **19**, 224 (2018).

61. Rauluseviciute, I. *et al*. JASPAR 2024: 20th anniversary of the open-access database of transcription factor binding profiles. *Nucleic Acids Res*. **52**, D174–D182 (2024).

Figure 1

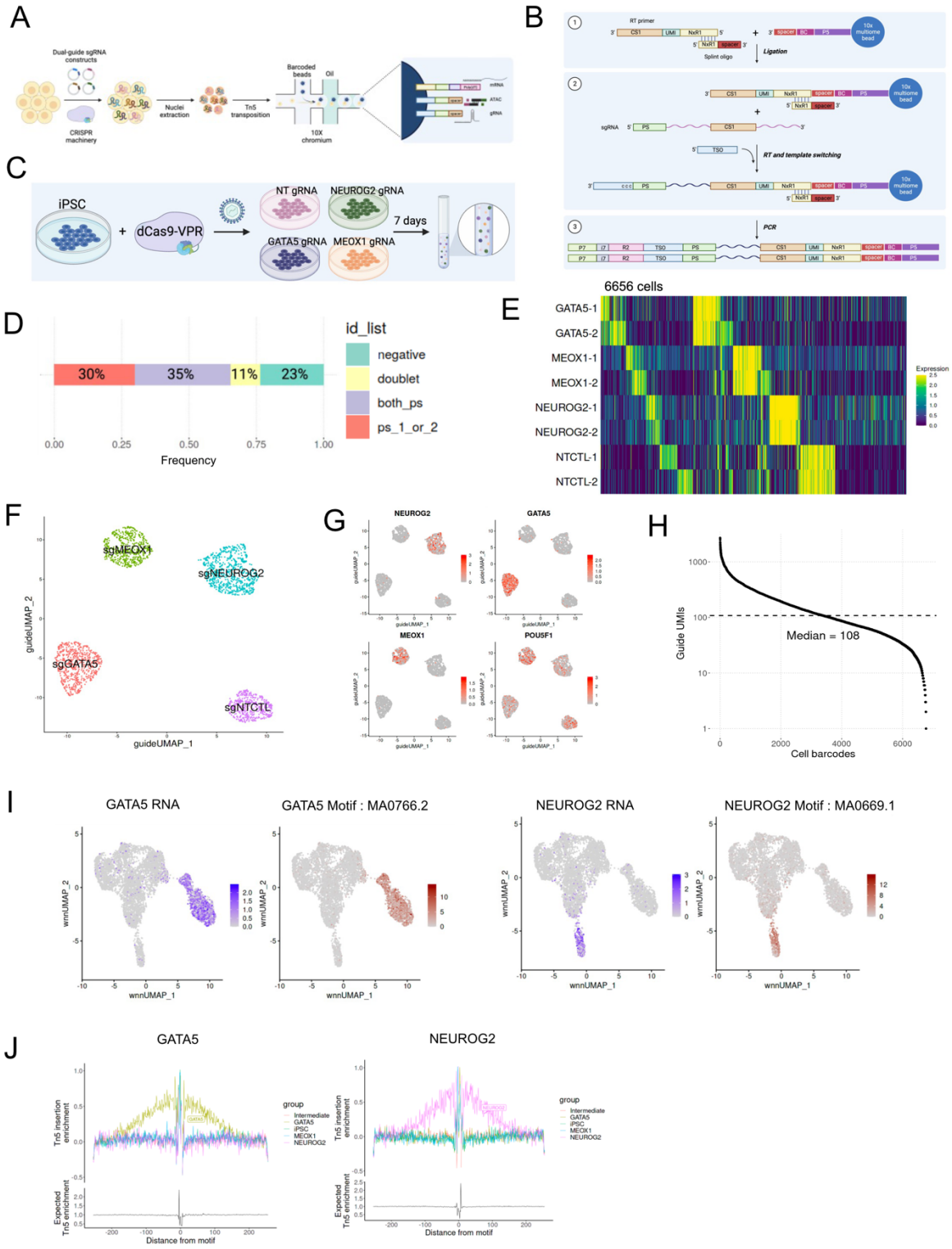

A
Dual-guide sgRNA constructs
CRISPR machinery
Nuclei extraction
Tn5 transposition
Barcoded beads
Oil
10X chromium
mRNA
ATAC
gRNA
B
RT primer
Splint oligo
Ligation
10x multiome bead
sgRNA
RT and template switching
PCR
C
iPSC
dCas9-VPR
NT gRNA
NEUROG2 gRNA
GATA5 gRNA
MEOX1 gRNA
7 days
D
id_list
negative
doublet
both_ps
ps_1_or_2
30%
35%
11%
23%
Frequency
E
6656 cells
GATA5-1
GATA5-2
MEOX1-1
MEOX1-2
NEUROG2-1
NEUROG2-2
NTCTL-1
NTCTL-2
Expression
F
sgMEOX1
sgNEUROG2
sgGATA5
sgNTCTL
guideUMAP_1
guideUMAP_2
G
NEUROG2
GATA5
MEOX1
POU5F1
H
Guide UMIs
Median = 108
Cell barcodes
I
GATA5 RNA
GATA5 Motif : MA0766.2
NEUROG2 RNA
NEUROG2 Motif : MA0669.1
wnnUMAP_1
wnnUMAP_2
J
GATA5
NEUROG2
Tn5 insertion enrichment
Expected Tn5 enrichment
Distance from motif
group
Intermediate
GATA5
iPSC
MEOX1
NEUROG2

Figure 2

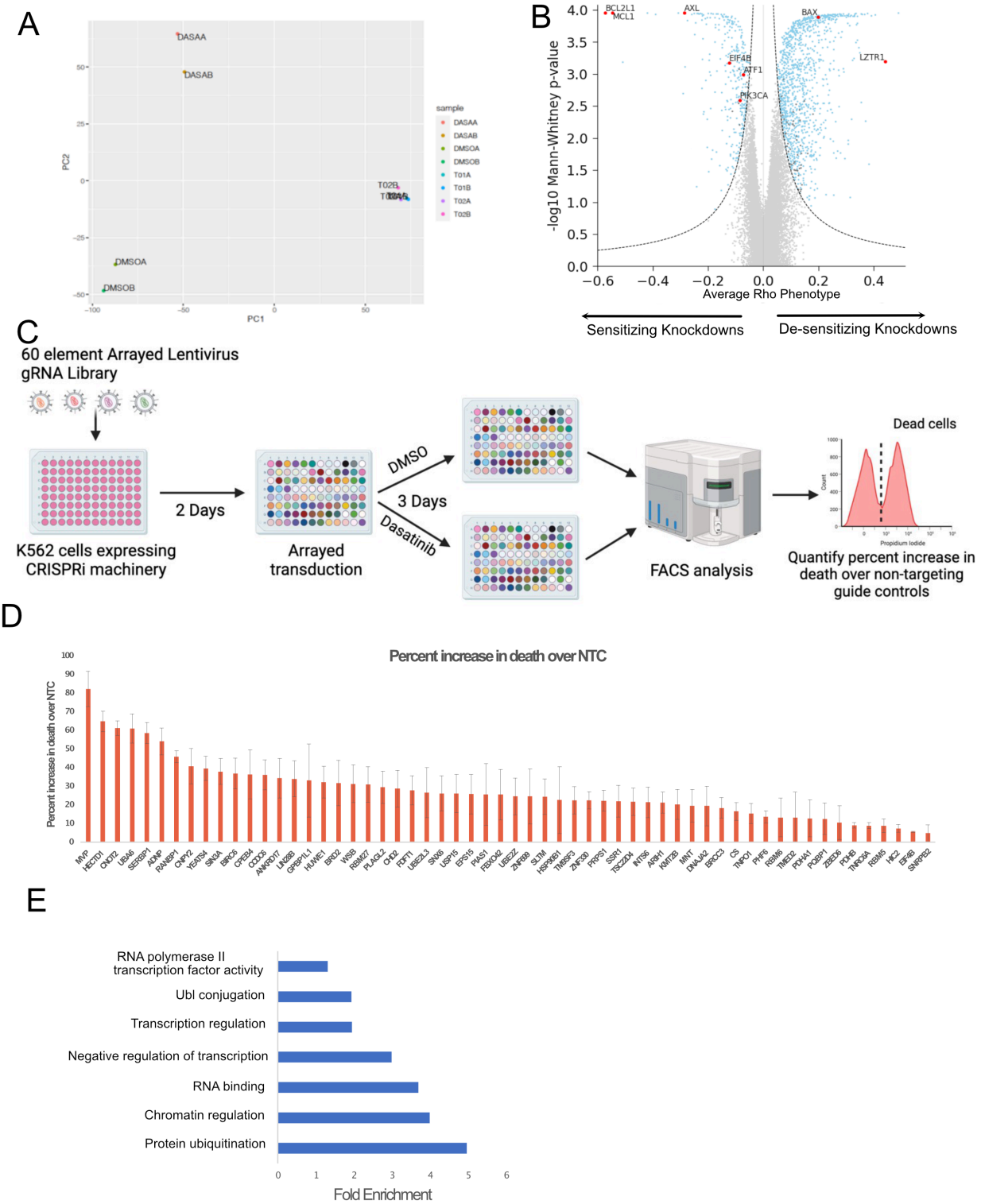

A
DASAA
DASAB
DMSOA
DMSOB
T02B
PC1
PC2
sample
DASAA
DASAB
DMSOA
DMSOB
T01A
T01B
T02A
T02B
B
BCL2L1
MCL1
AXL
BAX
LZTR1
EIF4B
ATF1
PIK3CA
-log10 Mann-Whitney p-value
Average Rho Phenotype
Sensitizing Knockdowns
De-sensitizing Knockdowns
C
60 element Arrayed Lentivirus gRNA Library
K562 cells expressing CRISPRi machinery
2 Days
Arrayed transduction
DMSO
3 Days
Dasatinib
FACS analysis
Dead cells
Quantify percent increase in death over non-targeting guide controls
D
Percent increase in death over NTC
E
RNA polymerase II transcription factor activity
Ubl conjugation
Transcription regulation
Negative regulation of transcription
RNA binding
Chromatin regulation
Protein ubiquitination
Fold Enrichment

Figure 3

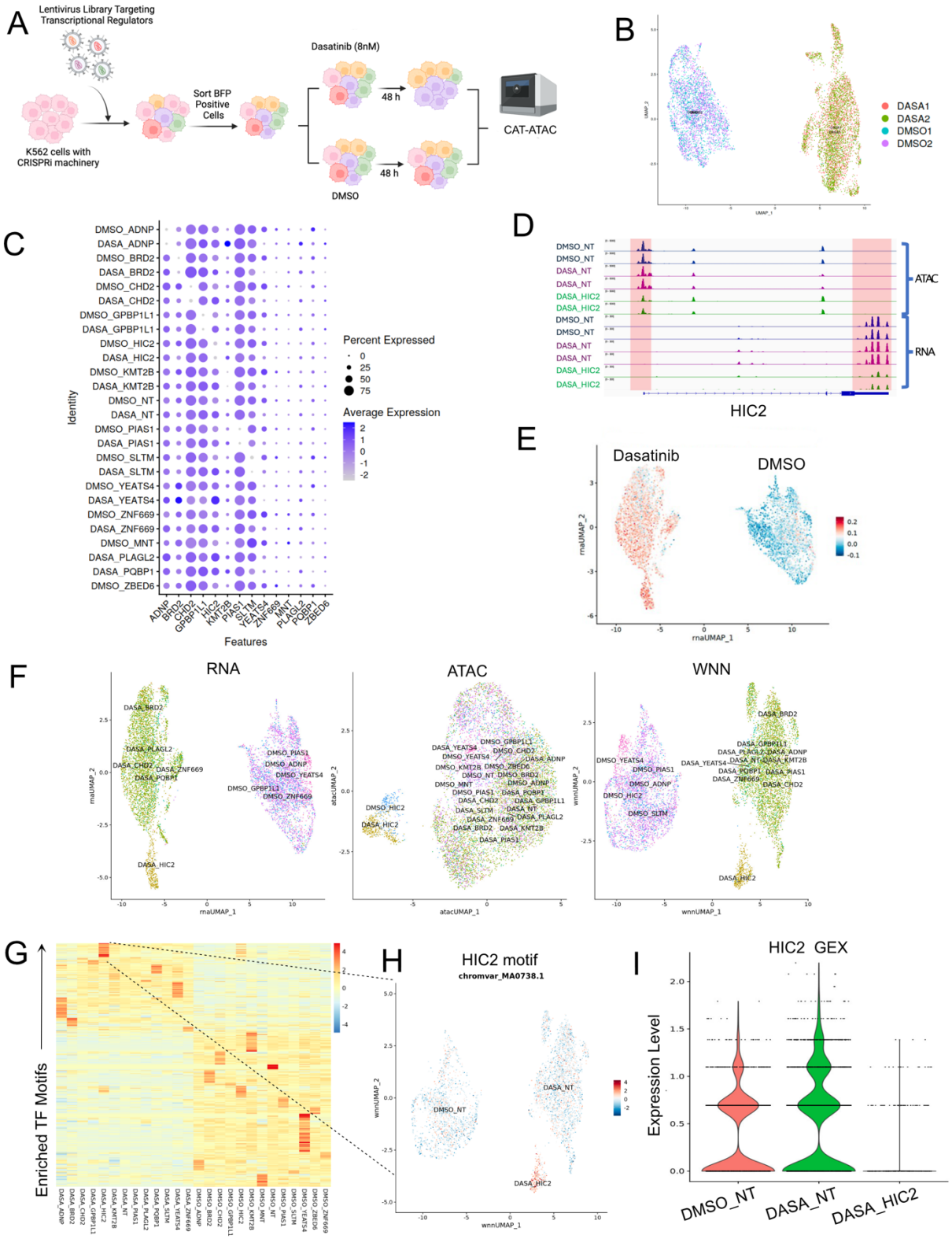

A
Lentivirus Library Targeting Transcriptional Regulators
K562 cells with CRISPRi machinery
Sort BFP Positive Cells
Dasatinib (8nM)
48 h
48 h
DMSO
CAT-ATAC
B
DASA1
DASA2
DMSO1
DMSO2
C
Identity
Features
Percent Expressed
Average Expression
D
ATAC
RNA
HIC2
E
Dasatinib
DMSO
F
RNA
ATAC
WNN
G
Enriched TF Motifs
H
HIC2 motif
chromvar_MA0738.1
I
HIC2 GEX
Expression Level
DMSO_NT
DASA_NT
DASA_HIC2

Figure 4

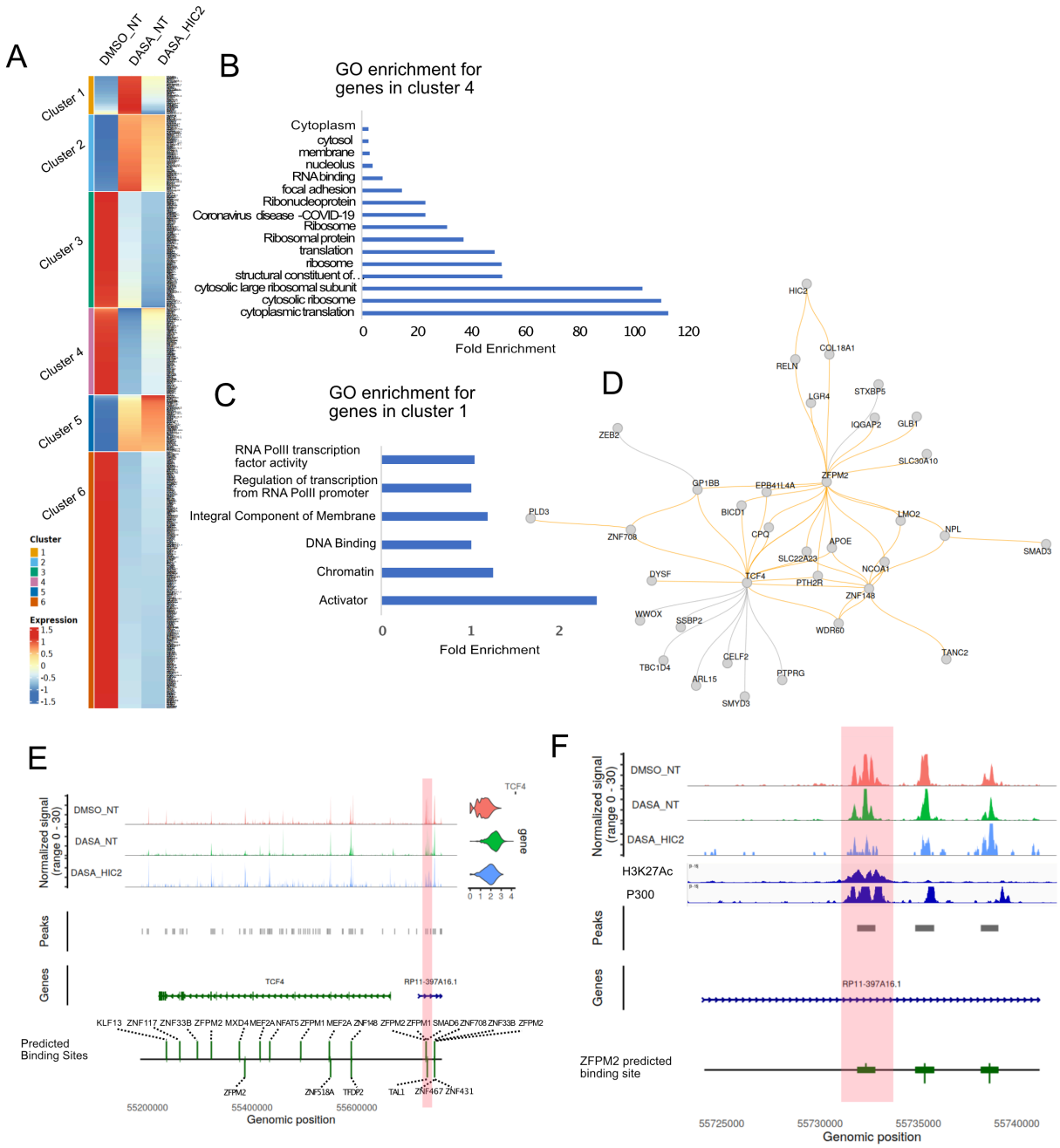


Figure 5

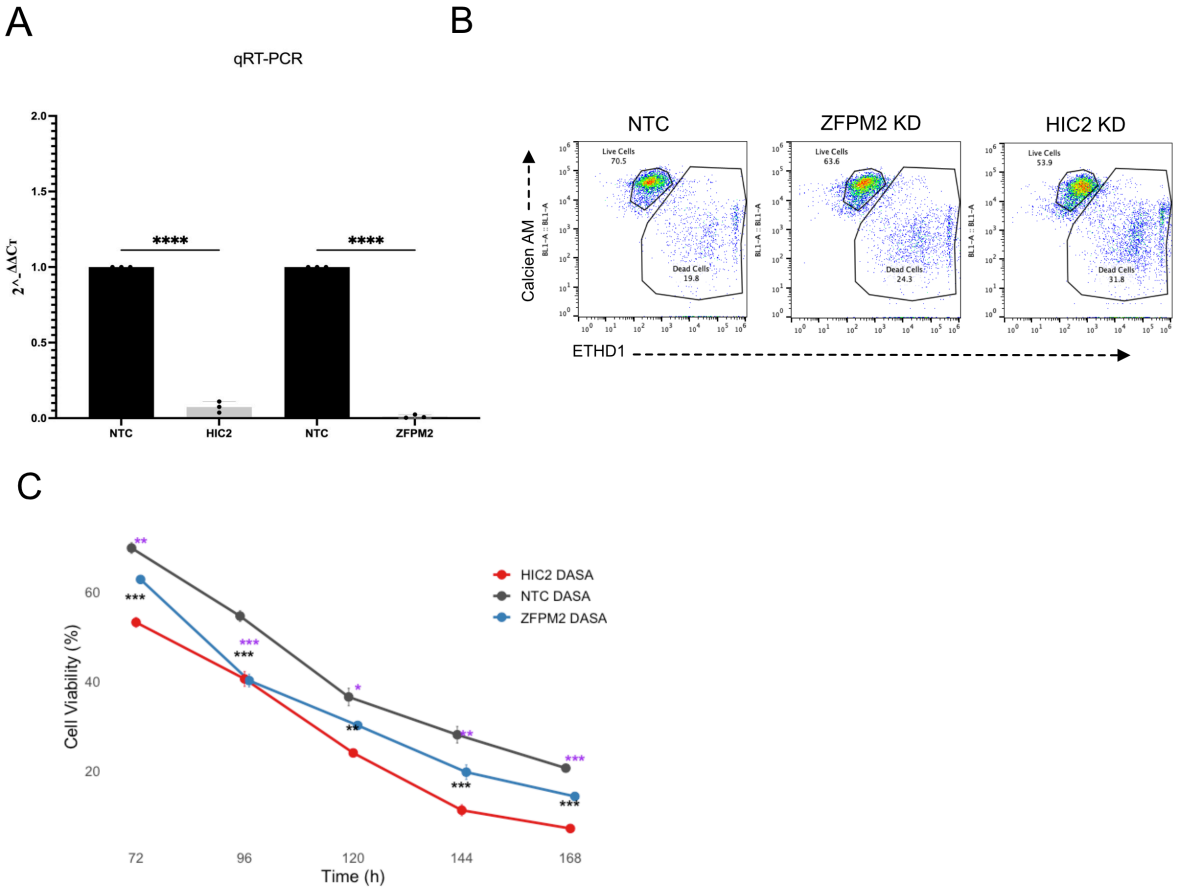


Figure Captions:

Figure 1: CAT-ATACmethod and validation

A. Illustration of CAT-ATAC workflow
B. Detailed schematic illustrating the guide capture process.
C. Pilot experiment overview
D. Overall guide assignment for the pilot experiment. 76% of the total cells could be assigned to guides, 11% of which were identified as doublets.
E. Heatmap showing guide assignment in all 6656 cells.
F. UMAP embedding of singlets expressing 2 protospacers (Singlet_2_PS) based on guide expression.
G. RNA expression plots of targeted genes in guide expression UMAP.
H. Barcode rank plot for sgRNA UMI counts, showing a median of 108 UMIs for captured guide sequences per cell.
I. Expression of the TF shown in purple, motif enrichment shown in brown.
J. TF footprinting for GATA5 and NEUROG2.

Figure 2: Pooled CRISPRi screen identifies gene knockdowns which modulate resistance to dasatinib

A. PCA plot showing reproducibility of replicate samples from the genome-wide screen.

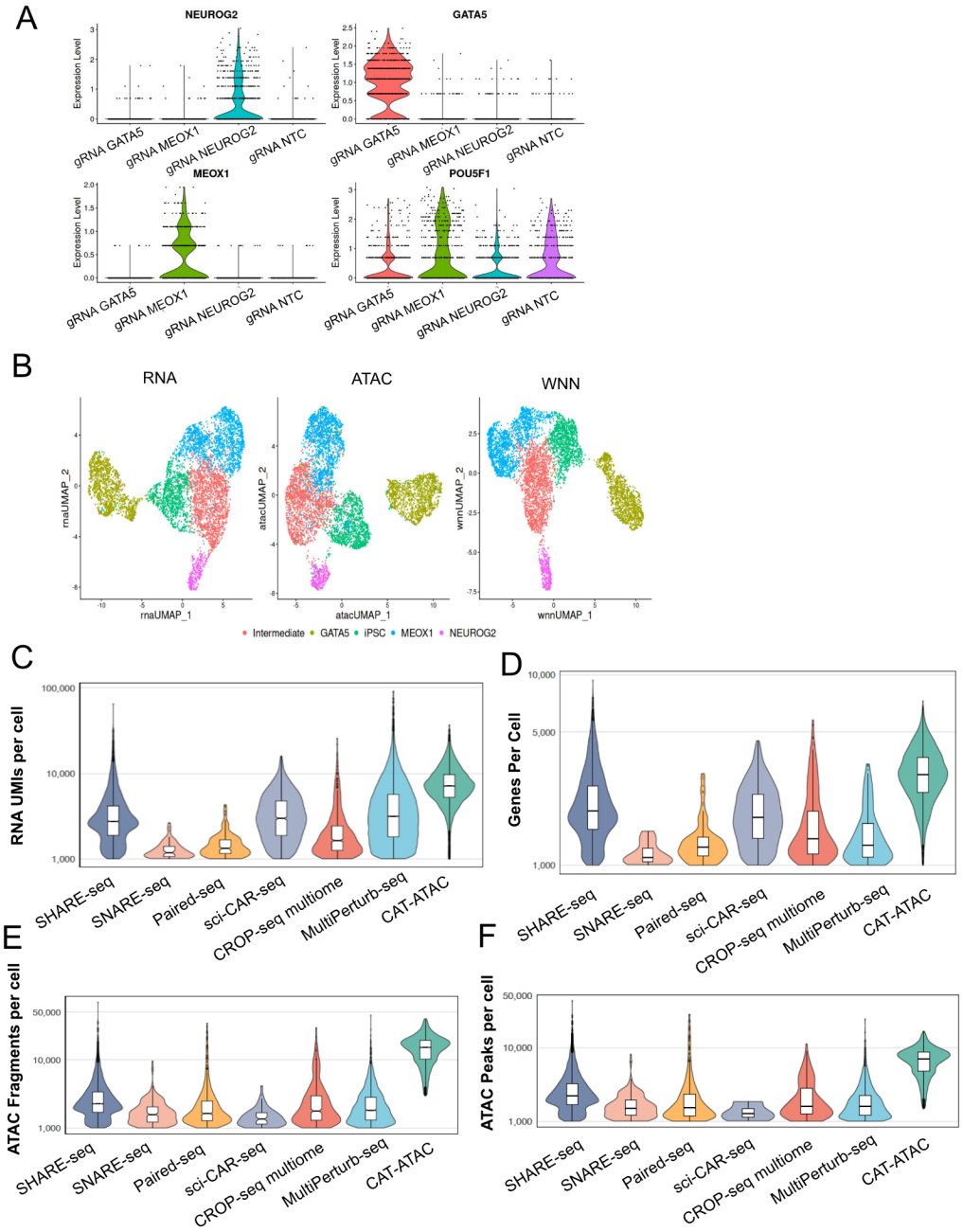
B. Volcano plot showing gene knockdowns exhibiting sensitizing and desensitizing phenotypes. Blue dots represent all significant hits. Genes highlighted in red were previously known to affect dasatinib sensitivity.
C. Schematic showing overview of the arrayed validation experiment conducted to validate hits from the primary genome-wide screen.
D. Bar plot displaying results from the arrayed validation experiment showing percent increase in cell death for each of the guide expressing cells over non-targeting control guide expressing cells. Error bars represent standard deviation.
E. GO analysis for the 60 genes included in the arrayed validation experiment.

Figure 3: CAT-ATAC elucidates the combinatorial effect of gene and drug perturbation on RNA and ATAC profiles.
A. Schematic showing the experimental outline for the CAT-ATAC experiment.
B. UMAPs after sample integration showing replicate reproducibility.
C. Dot plot showing target gene knockdown in cell populations which received the indicated guide in both dasatinib and DMSO treated conditions.
D. Genome browser tracks for ATACseq and RNAseq peaks for DMSO_NT, DASA_NT and DASA_HIC2 cell groups at the HIC2 gene locus. Highlighted areas mark the promoter and 3' end of the HIC2 gene.
E. Module score plot showing that dasatinib treatment results in upregulation of erythrocyte progenitors genes identified by Velten et.al.
F. UMAPs showing RNA, ATAC and weighted nearest neighbor profiles for cells expressing specific guideRNAs.
G. Heatmap showing drug treatment and knockdown specific changes in chromvar motif activities.
H. UMAP showing motif enrichment for HIC2 in the cell groups DMSO_NT, DASA_NT and DASA_HIC2.
I. Violin plots showing RNA expression of HIC2 in the cell groups DMSO_NT, DASA_NT and DASA_HIC2.

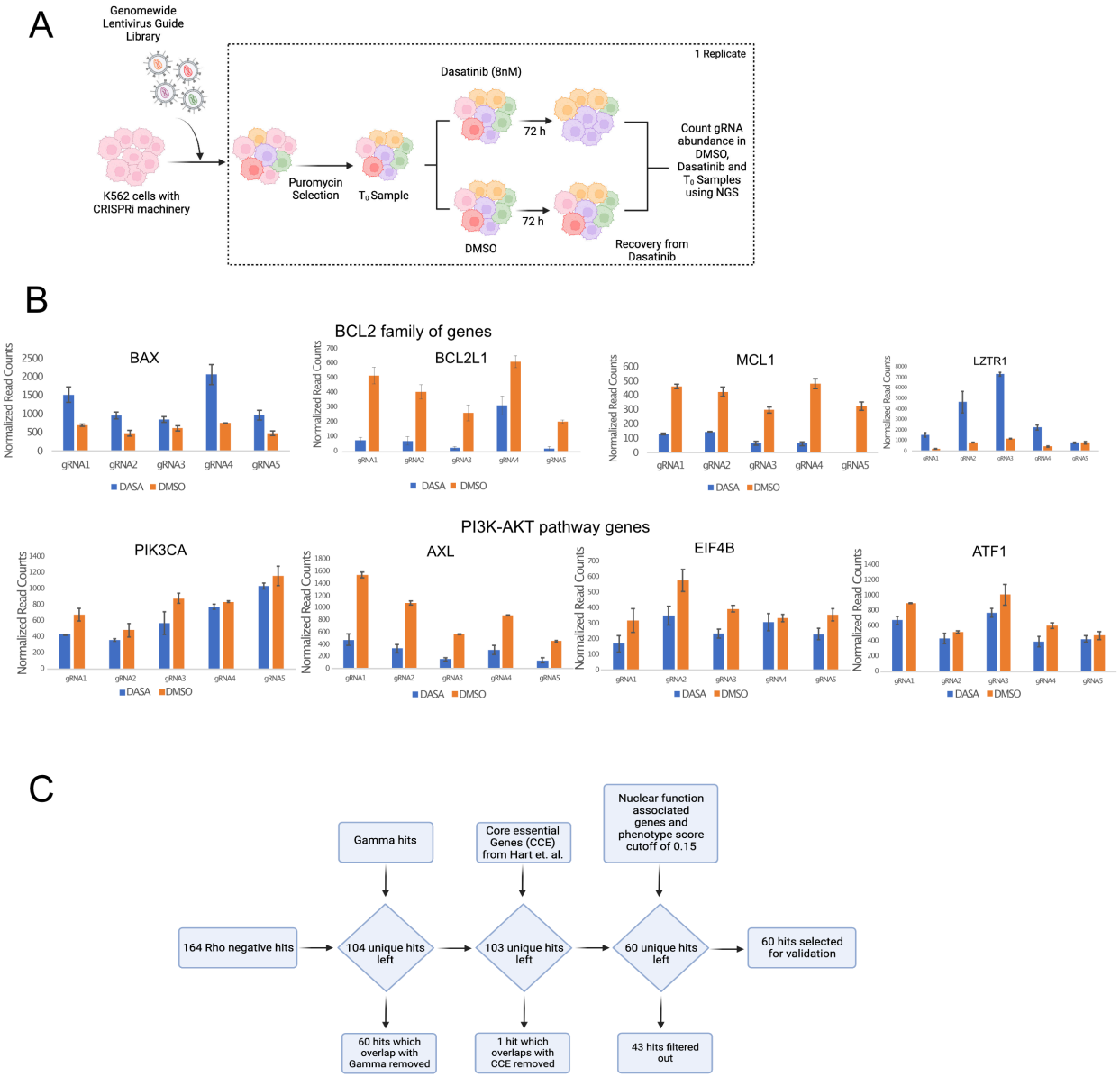
Figure 4: Identifying gene regulatory network underlying dasatinib resistance.
A. Heatmap showing gene expression patterns in DMSO treated cells expressing non-targeting guides (DMSO_NT), dasatinib treated cells expressing non-targeting guides (DASA_NT) and dasatinib treated cells expressing HIC2 targeting guides (DASA_HIC2).
B. GO analysis for genes in cluster 4 from panel A.
C. GO analysis for genes cluster 1 from panel A.
D. Subsetted gene regulatory network identified by querying the global network for genes differentially downregulated in dasatinib treated cells expressing HIC2 guides as compared to dasatinib treated cells expressing NT guides.
E. Genome browser tracks showing the TCF4 locus. The first three tracks show ATACseq signal for the cell groups DMSO_NT, DASA_NT and DASA_HIC2, violin plots show the gene expression of TCF4 in the same cell groups. The last track shows the transcription factor binding sites predicted by Pando. Highlighted region shows ATAC-seq peaks upstream of TCF4 with predicted ZFPM2 binding sites.
F. Zoomed in view of the highlighted region from panel E. The highlighted region in this track shows a ATACseq peak exhibiting reduced accessibility in DASA_HIC2 cell group. The

851 next two tracks showing H3K27Ac and P300 enrichment in K562 cells were downloaded
852 from ENCODE. The last track shows ZFPM2 binding sites as predicted by Pando.
853
854 Figure 5: ZFPM2 and PTH2R are necessary and sufficient for increased resistance to dasatinib.
855 A. qPCR data showing CRISPRi mediated knockdown of gene expression for the genes HIC2,
856 PTH2R and ZFPM2.
857 B. Flow-cytometry plots at 72h post dasatinib treatment showing the distribution of cells for
858 each of the three gene knockdown conditions along with control cells treated with dasatinib
859 and stained with Calcien-AM and EthD-1 to identify live and dead cells respectively.
860 C. Timecourse experiment showing flow-cytometry data from 72h to 168h post dasatinib
861 treatment quantifying % viable cells in HIC2 KD, ZFPM2 KD and control cells. Each data
862 point represents the mean ± SD of three biological replicates (n=3). Statistical significance
863 was assessed at each time point using unpaired two-tailed t-tests comparing HIC2 KD or
864 ZFPM2 KD vs. NTC. Black asterisks indicate significant differences for HIC2 KD vs. NTC,
865 while purple asterisks indicate significant differences for ZFPM2 KD vs. NTC (*
866 corresponds to $p < 0.05$, ** corresponds to $p < 0.01$ and *** corresponds to $p < 0.001$).
867
868
869
870
871
872 Supplementary Figures

Supplemental Figure 1

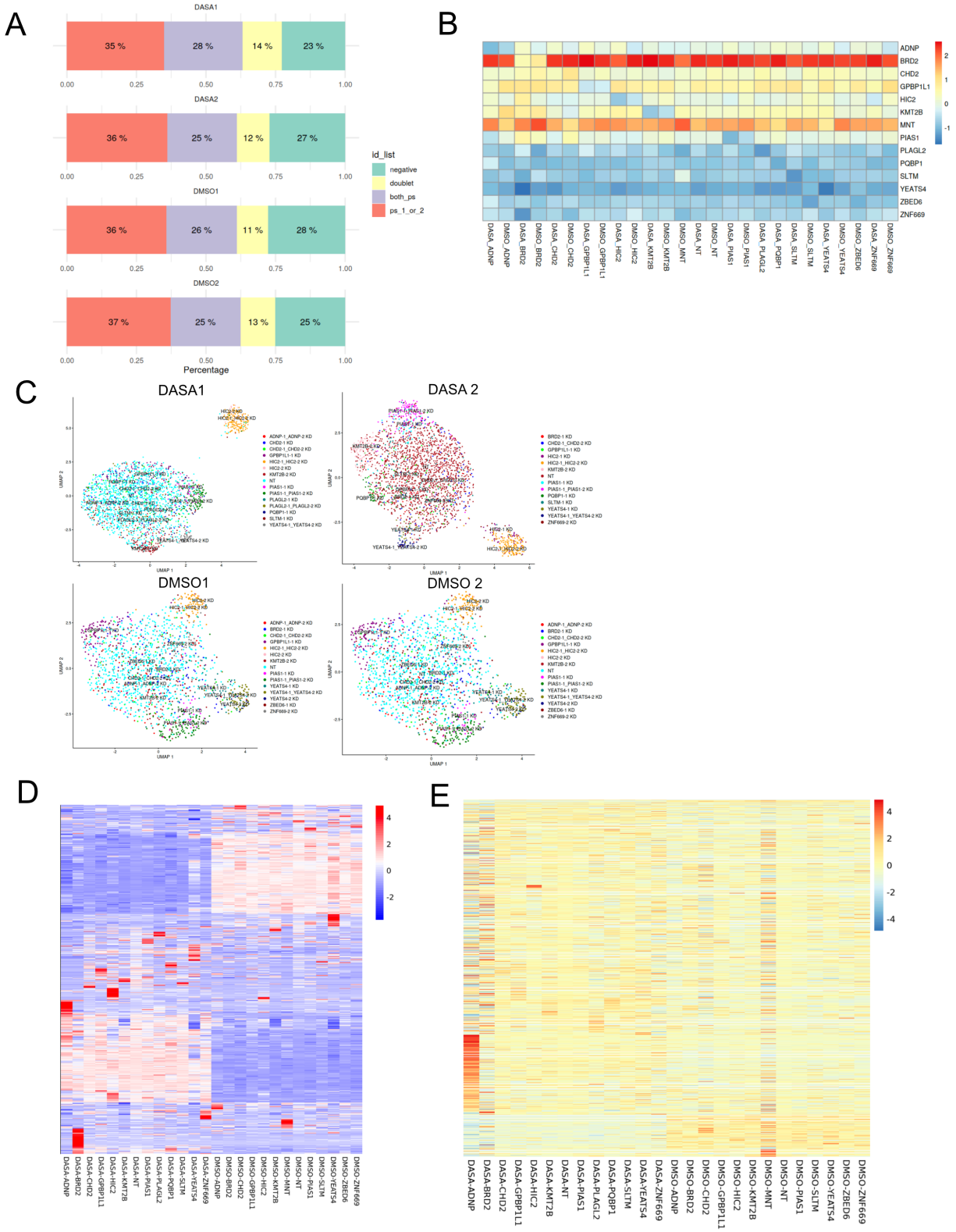


882 Supplemental Figure 2

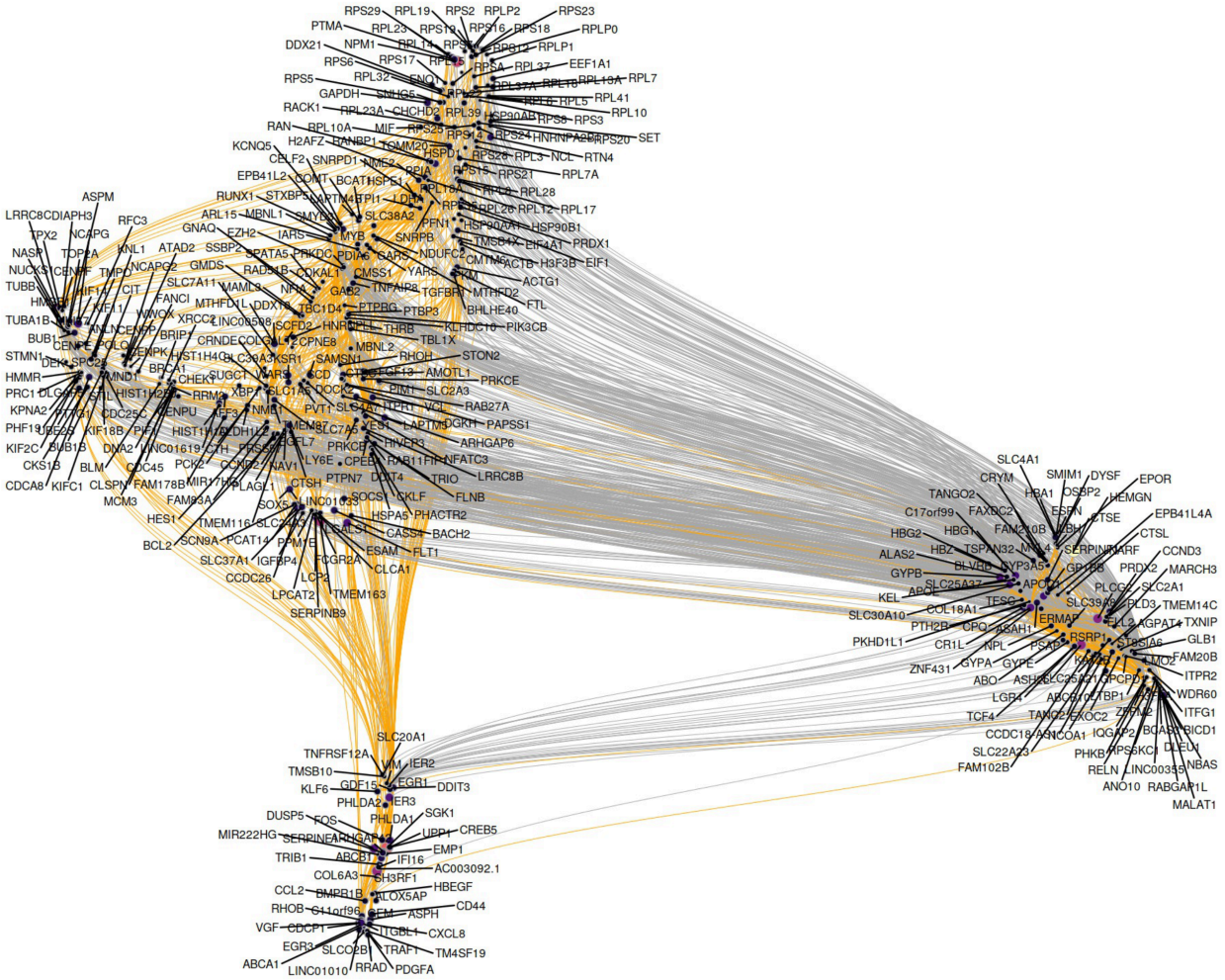


883
884

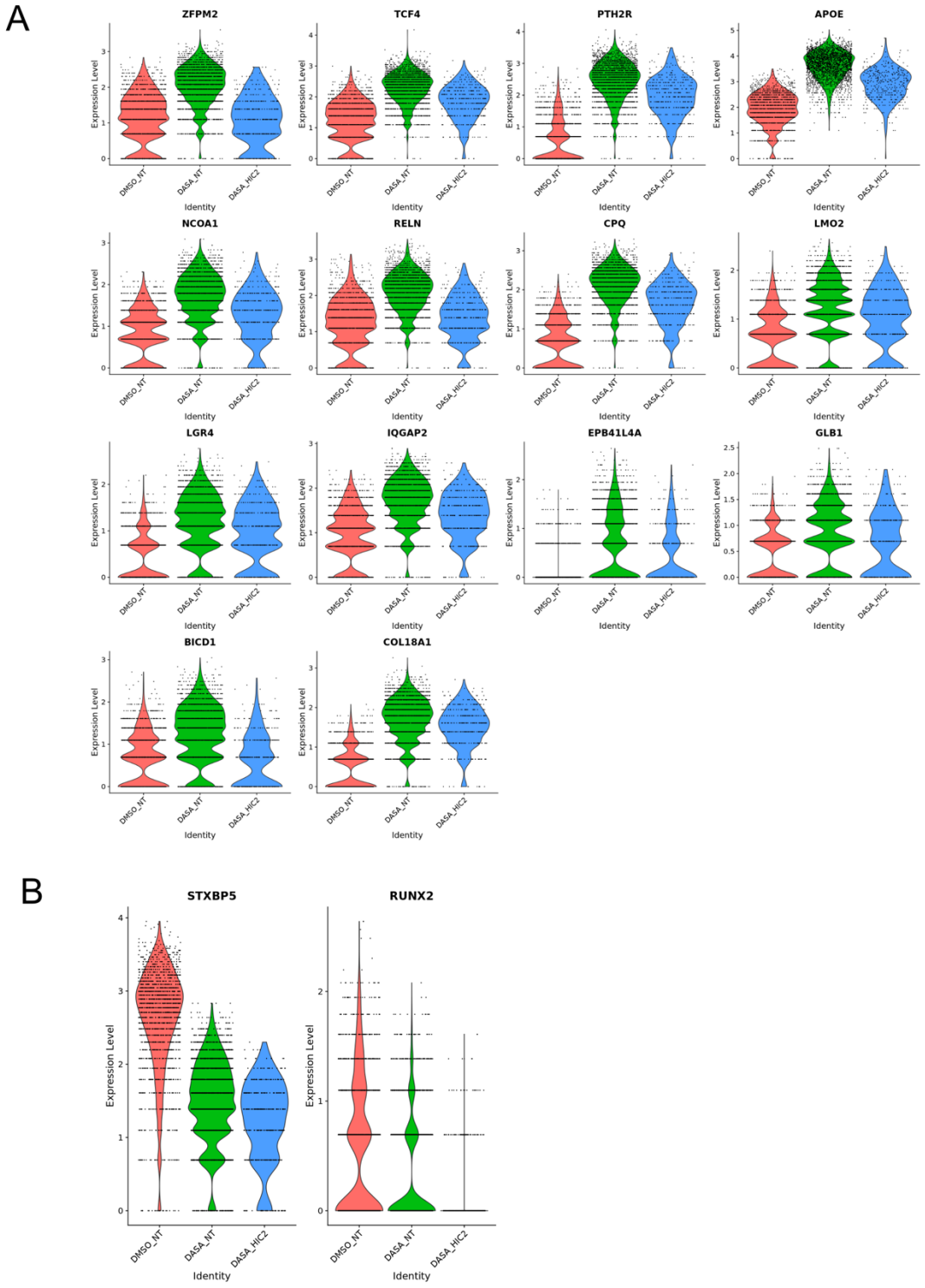
Supplemental Figure 3



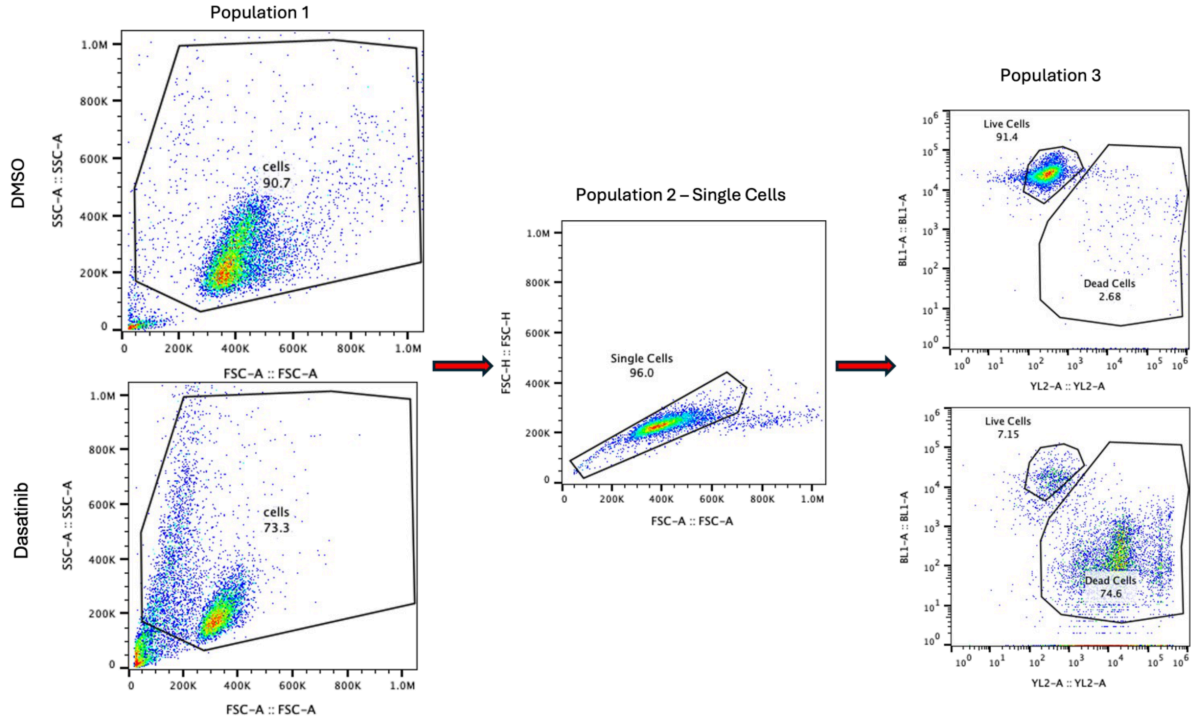
Supplemental Figure 4



Supplemental Figure 5



Supplemental Figure 6



Supplemental Figure Captions:

Supplementary Figure 1: CAT-ATAC QC and comparisons with other technologies

A. Violin plots showing protospacer expression in clusters shown in UMAP in Figure 1 panel F.
B. UMAPs based on RNA, ATAC and joint profiles for CAT-ATAC in iPSCs with 3 TFs (GATA5, NEUROG2, MEOX1) and 1 non-targeting control overexpressed by CRISPRa.
C. Violin plots comparing RNA UMIs per cell metric for various Multiome approaches with CAT-ATAC.
D. Violin plots comparing Genes per cell metric for various Multiome approaches with CAT-ATAC.
E. Violin plots comparing ATAC fragments per cell metric for various Multiome approaches with CAT-ATAC.
F. Violin plots comparing ATAC peaks per cell metric for various Multiome approaches with CAT-ATAC.

Supplementary Figure 2: Dasatinib genomewide screen design and QC

A. Design of the dasatinib survival screen.
B. Normalized read counts of guides targeting some of the genes that were identified as hits and were highlighted in Figure 2 panel B. Error bars represent standard deviation.
C. Strategy to filter the sensitizing hits.

Supplementary Figure 3: CAT-ATAC QC for dasatinib screen

A. Bar graph showing guide assignment for each of the 4 samples from the 10x Multiome screen.
B. Heatmap showing average accessibility at the promoters of targeted genes in both dasatinib and dmso treated samples.
C. UMAPs showing RNA expression profiles of cells for individual samples grouped by perturbation classes based on gRNA expression.
D. Heatmap showing top 1000 differentially expressed genes from each of the perturbation classes across the two treatment groups dasatinib and DMSO.
E. Heatmap showing differentially accessible ATAC peaks from each of the perturbation classes across the two treatment groups dasatinib and DMSO.

Supplementary Figure 4: Global GRN

Global GRN constructed using Pando

Supplementary Figure 5: Expression of genes in the subsetted GRN

A. Violin plots showing expression of genes having activating connections in the subsetted gene regulatory network.
B. Violin plots showing expression of genes having repressive connections in the subsetted gene regulatory network.

Supplementary Figure 6: Flow cytometry gating strategy

Gating strategy for flow cytometry data of k562/CRISPRi cells to study the effect of dasatinib on CRISPRi constructs targeted against HIC-2, PTH2R, and ZFPM2 stained with 1ug/ml EthD-1 and 2 μM calcein-AM analyzed by flow cytometry.

The first panel shows all events with a gate (population 1) for cells. The second panel shows population 1 with a gate for singlets (population 2). The third panel shows Quarter 1 = Green or BL1-A for live cells and Quarter 3 = Red or YL2-A for dead cells (population 3).