

# 1 High-throughput functional characterization of combinations of transcriptional activators and 2 repressors

3

## 4 Authors

5 Adi X. Mukund<sup>1</sup>, Josh Tycko<sup>2</sup>, Sage J. Allen<sup>3</sup>, Stephanie A. Robinson<sup>4</sup>, Cecelia Andrews<sup>5</sup>, Connor H. Ludwig<sup>3</sup>,  
6 Kaitlyn Spees<sup>3</sup>, Michael C. Bassik<sup>2</sup>, Lacramioara Bintu<sup>3</sup>

7

- 8     1. Biophysics Program, Stanford University  
9     2. Department of Genetics, Stanford University  
10    3. Department of Bioengineering, Stanford University  
11    4. Department of Chemistry, Stanford University  
12    5. Department of Developmental Biology, Stanford University

13

## 14 Abstract

15 Despite growing knowledge of the functions of individual human transcriptional effector domains, much less is  
16 understood about how multiple effector domains within the same protein combine to regulate gene expression.  
17 Here, we measure transcriptional activity for 8,400 effector domain combinations by recruiting them to reporter  
18 genes in human cells. In our assay, weak and moderate activation domains synergize to drive strong gene  
19 expression, while combining strong activators often results in weaker activation. In contrast, repressors combine  
20 linearly and produce full gene silencing, and repressor domains often overpower activation domains. We use  
21 this information to build a synthetic transcription factor whose function can be tuned between repression and  
22 activation independent of recruitment to target genes by using a small molecule drug. Altogether, we outline the  
23 basic principles of how effector domains combine to regulate gene expression and demonstrate their value in  
24 building precise and flexible synthetic biology tools.

25 **Main Text**

26

27 **Introduction**

28

29 Transcription factors (TFs) and chromatin regulators (CRs) contain short effector domains that can act  
30 as repressors or activators when recruited at a target gene<sup>1,23–51,2</sup>. Site-specific recruitment assays of effector  
31 domains and full length TFs at reporter genes have long been used to understand their effects on gene  
32 expression and develop better tools for gene regulation<sup>6–15</sup>.

33 Although recruitment assays have historically focused on recruiting only one transcriptional effector per  
34 cell, combinatorial function is a key property of both chromatin-mediated gene regulation<sup>16</sup> and transcription  
35 factor-mediated gene regulation<sup>17–19</sup>. Transcription factors frequently contain multiple effector domains with  
36 potentially opposing functions, with studies reporting up to 40% of transcription factors having at least 2 distinct  
37 effector domains<sup>20</sup>. For example, the chromatin regulator MGA was recently shown to feature two repressive  
38 domains with different rates of silencing and amounts of memory<sup>21</sup>, while an earlier study showed that the  
39 transcription factor NIZP1 features an activating KRAB domain that is dominated by a repressive C2HR  
40 domain<sup>22</sup>. Combinations of effector domains have a long history in the context of synthetic biology, where the  
41 well-known transcriptional activators VP64<sup>23</sup> and VPR<sup>24</sup> were built via combining multiple known activation  
42 domains. Similarly, repressor combinations featuring the KRAB domain of ZNF10 and DNA methyltransferases  
43 have been shown to produce a robust combination of rapid gene silencing and long-term epigenetic memory<sup>9,25</sup>.

44 A systematic understanding of how combinations of transcriptional effector domains function in human  
45 cells would expand the range of compact tools available for epigenetic perturbations and therapy<sup>26,27</sup>.  
46 Additionally, composing effector domains can serve as a useful strategy to design synthetic TFs capable of  
47 implementing gene regulatory functions not achievable by fusing individual effectors to DNA-binding domains.  
48 These TFs could be used for many applications, including more efficient reprogramming of cell lineage  
49 specification or high-throughput screening of the noncoding genome<sup>28,29</sup>.

50 In order to systematically test combinations of effector domains, we need high-throughput methods: even  
51 testing all possible combinations resulting from pairing 100 effector domains requires 10,000 measurements.  
52 Recently, pooled screens have been developed for high-throughput characterization of individual transcription  
53 factors and effector domains in yeast<sup>30,31</sup>, drosophila<sup>32</sup>, and human cells<sup>20,21,33,34</sup>. Arrayed high-throughput  
54 measurements of combinations of chromatin regulators in concert with VP16 have been performed in yeast to  
55 test 223 combinations<sup>7,14</sup> and low-throughput measurements have been performed in human cells<sup>35</sup>. However,  
56 we are not aware of any systematic high-throughput studies to date in mammalian cells that measure how  
57 effector domains act in combination to regulate gene expression. As such, it remains unclear whether pairs of  
58 activator domains synergize when combined, whether pairs of repressor domains do the same, and how activator  
59 and repressor domains affect each other when recruited simultaneously. Here, we modify a recently developed  
60 pooled high-throughput method<sup>21</sup> to test thousands of combinations of previously characterized protein domains  
61 that can activate and repress gene expression in human cells in order to start addressing these questions.

62

63 **Developing a workflow for combinatorial screening of transcriptional effector domains**

64

65 We began by selecting a panel of effector domains and controls from our previous screen of single  
66 domains<sup>21</sup>: 44 repressors, 30 activators, and 20 control domains (**Fig. 1A, Materials & Methods,**  
67 **Supplementary Table 1**). These effectors were chosen to span a wide range of individual activation and  
68 repression strengths when recruited at the same reporter gene, (**Fig. S1A**). The control domains were chosen

69 to be either random sequences or fragments of the DMD protein that were shown to have no effect on gene  
70 expression when recruited to a reporter individually<sup>21</sup>. To avoid problems with protein stability, we chose only  
71 well-expressed, stable domains as measured by FLAG-staining<sup>21</sup>. We generated combinations of these  
72 individual domains using a 2-step cloning strategy to build a pool of domain-linker-domain concatenations that  
73 were then cloned into a lentiviral backbone vector as fusions to the reverse TetR (rTetR) inducible DNA binding  
74 domain (**Fig. 1A, Fig. S1B, Materials & Methods**). We delivered this pool into K562 reporter cells using lentivirus  
75 with a low multiplicity of infection (MOI=0.24), such that most cells expressed a single library element. We  
76 measured the ability of each pair in this library to activate or repress a reporter gene using a high-throughput  
77 pooled method we recently developed: HT-recruit<sup>21</sup>. Briefly, by adding doxycycline (dox) we recruited the rTetR-  
78 concatenated fusions to either a minimal promoter to measure activation, or to a highly expressed constitutive  
79 promoter to measure repression (**Fig. 1B**). We separated cells into populations with high (ON) and low (OFF)  
80 reporter gene expression using magnetic separation (**Materials & Methods, Fig. 1C, Fig. S1C**), and computed  
81 the relative enrichment of individual effector combinations in each population (**Fig. 1D**). We recovered the large  
82 majority of concatenations that were present in our combinatorial library with sufficient sequencing depth (**Fig.**  
83 **S1D**), and found good correlation across replicates for both our activation (**Fig. 1E**, Pearson R=0.803, p<2.2×10<sup>-16</sup>)  
84 and repression (**Fig. 1F**, Pearson p=0.795, p<2.2×10<sup>-16</sup>) screens. To identify effector combinations for which  
85 we could reliably measure activity, we used the distribution of negative control-only combinations: we considered  
86 a pair to be activating and/or repressing if it had a score at least 2 standard deviations away from the mean of  
87 the negative control combinations.

88 To validate these high-throughput measurements and our detection threshold, we measured activation  
89 for 61 domain concatenations and repression for 38 domain concatenations by recruiting them individually to our  
90 synthetic reporters and measuring gene expression via flow cytometry. We measured the fraction of cells  
91 activated at the weak minCMV promoter after 2 days of recruitment and found strong correlations between our  
92 screen measurements and the fraction of cells activated (**Fig. 1G**, R<sup>2</sup>=0.88, p=7.4×10<sup>-50</sup>). We found a similarly  
93 strong correlation between repressive screen scores and the fraction of cells silenced after 5 days of recruitment  
94 to the strong pEF promoter (**Fig. 1H**, R<sup>2</sup>=0.84, p=4.6×10<sup>-13</sup>). We concluded that our screen accurately measured  
95 these concatenations' ability to either activate or repress gene expression.

96 We began our analysis of the screen data by examining the behavior of effector domains in our library in  
97 combination with the negative control domains. We identified 4 out of 20 negative control domains that  
98 significantly affected either activation or repression scores across all effector domains and filtered them from  
99 downstream analyses (**Fig. S1E-F**). We then sought to determine if the order of the domains within  
100 concatenations altered their effect on gene expression. While prior literature has demonstrated that the ordering  
101 of effector domains in synthetic TFs can significantly affect the function of fusions<sup>24,25</sup>, we found that few of the  
102 domains in our library changed effect size significantly when switched from the first to the second position across  
103 all of their fusions with negative controls (**Fig. S1G**). We excluded concatenations featuring any domain in the  
104 orientation that ablated its function from downstream analysis (**Fig. S1H**).

105 To check to what extent effectors maintained their function when fused to a control sequence, we  
106 examined the distribution of scores for each effector domain when paired with all negative controls (**Fig. 1I**).  
107 While for some domains all scores clustered together in one quadrant (e.g., the strong KRAB repressor domain  
108 from ZNF10, **Fig. 1I**, blue), we found that other domains dropped under the detection threshold when paired with  
109 certain negative controls (e.g., weak CRTC2 activation domain, **Fig. 1I**, yellow). Some domains, such as the  
110 FOXO3 activation domain, have been previously reported to act as dual effectors that both activate a minimal  
111 reporter and repress a constitutive one when recruited individually<sup>20</sup>. We indeed found that most FOXO3-control

112 fusions acted as dual activator-repressors, though a minority of fusions only acted as activators (**Fig. 1I**, green).  
113 In order to classify each domain, we calculated the number of effector-control pairs that met our hit threshold for  
114 activation and for repression (**Fig. 1J**). We used these data, along with the magnitude of activation and  
115 repression scores for each effector when paired with negative controls, to label effectors as activators,  
116 repressors, dual-functional effectors (both activators and repressors), or non-hits (**Fig. 1K, Materials &**

117 Methods). We found that the median activation or repression score of each effector domain when paired with  
118 negative control domains correlated well with prior screen measurements of activation (**Fig. 1L**, Pearson R=0.81,  
119 p<2.2×10<sup>-16</sup>) and repression (**Fig. 1M**, Pearson R=0.85, p<2.2×10<sup>-16</sup>). Convinced that we could appropriately  
120 characterize how individual effectors behave when paired with negative controls, we proceeded to analyze the  
121 behavior of effector domain combinations with each other.  
122

## 123 Weak activator domain pairs synergistically drive gene expression from weak promoters

124

125 The high-throughput measurements for transcriptional activation identified a large number of activator-  
126 activator pairs whose activation scores exceeded the sum of each individual activator's scores when paired with  
127 negative controls (**Fig. 2A**, left). We proceeded to validate individual examples of activator pairs to verify that the  
128 scores in the screen correlate well with individual flow cytometry measurements (**Fig. S2A-B**, **Fig. 1G**), and that  
129 synergy could be replicated in low-throughput (**Fig. 2B** top). Notable examples of synergy included the pair of  
130 ANM2's SH3 domain and KIBRA's WW-1 domain, as well as ANM2's SH3 domain and NOTC2's LNR-2 domain.  
131 While the full role of the ANM2 SH3 domain is not yet clear<sup>36</sup>, it is known to regulate PRMT1 activity in a  
132 methylation-dependent fashion<sup>37</sup>, is required for the function of the actin nucleator Cobl<sup>38</sup>, and modulates  
133 alternative splicing of BCL-X<sup>39</sup>. KIBRA's WW-1 domain binds PPxY motifs in other proteins<sup>40</sup>, and is essential  
134 for KIBRA-mediated regulation of Hippo signaling via interactions with LATS1/2<sup>41</sup>, while NOTC2's second Lin-  
135 12/Notch repeat (LNR) domain sits within its negative regulatory region<sup>42</sup> and is cleaved off during ligand  
136 binding<sup>43</sup>.

137 Surprisingly, we found that the strongest activators at minCMV, which were generally dual-functional  
138 effectors that also repressed gene expression at pEF, resulted in lower activation scores at minCMV when paired  
139 together than when either activator was paired with negative controls (**Fig. 2A**, upper right quadrant). To verify  
140 that the antagonism of dual-dual effector pairs was real and not simply a result of promoter saturation, we  
141 individually tested the MYBA-ZN473 pair and found that, indeed, the combination did not activate as many cells  
142 as either domain when paired with a negative control (**Fig. 2B** bottom). Using the previously computed functions  
143 connecting screen scores with individual flow cytometry measurements of gene activation (**Fig. 1G**), we  
144 compared the estimated fraction of cells activated for each combination of activators with the sum of each  
145 activator's control-paired estimated fraction activated. We found that activator-activator pairs tended to  
146 synergize, activator-dual pairs were additive, and dual-dual pairs acted antagonistically with low activation scores  
147 (**Fig. 2A**, **Fig. 2C** top). We computed an estimated quantity of synergy for each combination by taking the  
148 difference between that combination's estimated fraction on and the sum of its individual domains' fractions on.  
149 Doing so, we found that activator-activator pairs tended to feature more synergy than activator-dual pairs, which  
150 in turn tended to feature more synergy than dual-dual pairs (**Fig. 2C** bottom).

151 For each activator or dual-functional domain we then tracked the activity of all combinations containing  
152 that domain (including combinations with repressors and controls) and examined how the strength of the  
153 combination varied with the strength of the partner's activation score. We fit these data using a sigmoidal Hill  
154 function (**Fig. 2D**, **Fig. S2C**). We found that the strength of the partner domain at the half-maximal point of the  
155 Hill function decreased as the activation domain's strength increased, meaning that strong activators were able

156 to reach half-maximal activation with less help from their partner domain than weak activators (**Fig. 2E**). We also  
157 found that the Hill coefficient of these functions went up as the activation domain's strength increased, indicating  
158 that the increase from minimal to substantial gene activation happened more rapidly in stronger activation  
159 domains as the strength of their partners increased (**Fig. 2F**).

160 Across all individual validations, we found a strong correlation between the fraction of cells activated and  
161 the mean fluorescence intensity (MFI) of activated cells for our effector domain pairs (**Fig. 2G**). This correlation  
162 and the fluorescent distributions of reporter expression (**Fig. 2C**, **Fig. S2A**) are consistent with activation domains  
163 modulating transcriptional bursting kinetics at the minCMV promoter: either increasing burst frequency or burst  
164 magnitude<sup>44–48</sup>. Since the degradation rates of our reporter mRNA and protein are slow, taking multiple days to  
165 dilute out<sup>10,21</sup>, reporter molecules in our cells persist long after the cessation of a transcriptional burst. As such,  
166 it would be difficult to determine whether reporter protein molecules were made all at once in one big burst or bit  
167 by bit (in many small bursts) over a longer period of time. Thus, in this regime of slow degradation rates,  
168 increasing either the duration of a transcriptional burst (burst size) or the frequency of such bursts (burst  
169 frequency) would lead to both a higher fraction of cells measured to be on and a greater MFI of those on cells<sup>49</sup>.  
170 Our findings are consistent with recent work showing that changes in burst kinetics can combine to produce  
171 transcriptional synergy<sup>50</sup>.

172

### 173 Repressor domain combinations generate robust promoter silencing

174

175 Our measurements of transcriptional repression suggested a linear relationship between the repressive  
176 strength of each domain and the repressive strength of the pair (**Fig. 3A**, points on the diagonal in the left bottom  
177 quadrant). In this case, the points that are above the diagonal come from saturation of repression at our reporter;  
178 effector pairs with scores  $\leq -3$  are predicted to silence all cells according to our validation curve (**Fig. 1H**, **Fig.**  
179 **S3A**). Indeed, when we tested a number of domain pairs in low throughput, we found that recruitment of pairs of  
180 strong repressors that each fully repressed the reporter when paired with controls (e.g. ZNF10 and CBX1) can  
181 also repress 100% of the cells when the pair is recruited. (**Fig. 3B**).

182 Using the best-fit curve for low-throughput validations mapping the repression  $\log_2(\text{ON:OFF})$  scores to  
183 the fraction of cells silenced at day 5 (**Fig. 1H**), we estimated the fraction of cells silenced at day 5 for each  
184 repressor or dual effector in our screen based on its median control-paired repression  $\log_2(\text{ON:OFF})$  score. We  
185 found that while the strength of effectors that could repress gene expression spanned a wide range, the majority  
186 of such domains were estimated to silence  $\geq 75\%$  of cells after 5 days of recruitment (**Fig. 3C**). As a result, when  
187 comparing the estimated fraction of cells silenced at day 5 for a combination versus the sum of the individual  
188 domains' respective fractions, we found that most concatenations were expected to and did silence virtually all  
189 cells after 5 days, with stronger repressors more consistently producing more repressive combinations than  
190 weaker repressors (**Fig. 3D**, **Fig. S3B-D**).

191 Our high-throughput measurements indicated that the repression  $\log_2(\text{ON:OFF})$  score of a repressor  
192 domain pair was a linear function of the two individual domains, as opposed to the nonlinear and sigmoidal  
193 behavior of activation domain pairs. To determine if this feature held true for individual repressor domains, we  
194 plotted the strength of every combination featuring each domain versus the strength of the domain's partner in  
195 that combination. We found that the trendline for repressors was linear, with lower slope for stronger repressors  
196 (**Fig. 3E**, **Fig. S3E**). For each repressor domain, we determined the slope and y-intercept of its corresponding

197 trendline, and found that stronger repressors featured both a flatter slope closer to 0 (**Fig. S3F**, Pearson R=0.87,  
198 p<2.2×10<sup>-16</sup>) and a lower y-intercept (**Fig. S3G**, Pearson R=0.92, p<2.2×10<sup>-16</sup>). These results suggest that weak  
199 repressors are tunable in that the strength of a pair featuring a weak repressor can vary over a wide range  
200 depending on the strength of the partner; however, pairs featuring a strong repressor will generally only act as  
201 strong repressors.

202 When examining the fluorescence distributions of our reporter throughout silencing, we found that most  
203 domain combinations silenced either all or none of the cells by day 5 (**Fig. 3F**, top 2 rows), as expected from  
204 previous experiments on a small number of chromatin regulators<sup>10</sup>. However, a minority of combinations either  
205 silenced only a fraction of cells or reduced gene expression without fully silencing cells (**Fig. 3F**, bottom 2 rows).  
206 We attempted to see if these results were consistent with prior models of stochastic gene silencing by extending  
207 a mathematical model of gene expression we have used before to understand chromatin regulation<sup>49</sup>. We did so  
208 by incorporating parameters for background silenced cells at the beginning of the timecourse, basal gene  
209 expression from the pEF promoter, lag time prior to silencing, the rate of protein decay upon silencer recruitment,  
210 and the rates of gene silencing and reactivation (**Fig. S4A, Materials & Methods**). After fitting this model to our  
211 experimental data, we found that it was able to accurately represent the dynamics of gene silencing for both  
212 strong silencers such as ZNF10-CBX1 (**Fig. S4B**) and weaker silencers such as SMCA2-U2AF4 that only  
213 silenced the reporter in a fraction of cells (**Fig. S4C**). For these types of domains, the rates of silencing predicted  
214 by this extended telegraph model for the individual validations where the model was able to accurately fit the  
215 data, and we found a good correlation with both the screen data (**Fig. S4D**, R<sup>2</sup>=0.73, p=1.20×10<sup>-9</sup>) and the fraction  
216 of cells silenced at day 5 (**Fig. S4E**, R<sup>2</sup>=0.79, p=2.94×10<sup>-11</sup>). Silencing rates predicted by the model suggested  
217 that the rate of silencing of a combination may be linear in the sum of the silencing rates of the individual domains  
218 (**Fig. S4F**), although the correlation between screen data and silencing rates was not strong enough for a more  
219 definitive conclusion.

220 However, we found that our model of all-or-none gene silencing could not capture the dynamics of  
221 silencers that reduced but did not fully ablate gene expression by day 5, such as DPY30-HXA13 (**Fig. S4G**),  
222 which featured a domain from DPY30 that worked as an activator on its own in prior screens and a repressor  
223 homeodomain from HXA13<sup>21</sup>. For these silencers, their effect on gene expression is better explained by a  
224 decrease in the production rate in the active state, rather than transitioning to a fully silent state. Altogether,  
225 these results suggest that while strong silencers in combinations will rapidly and fully silence gene expression in  
226 most cases, weaker repressors can produce more complex dynamic patterns of gene expression.  
227

## 228 High-throughput profiling of activator-repressor interactions

229

230 To identify principles underlying activator-repressor interactions, we looked at all pairs featuring an  
231 activator domain in combination with a repressor domain and determined if that pair functioned as an activator  
232 only pair, repressor only pair, dual-functional pair, or non-hit pair (**Fig. 4A**, **Fig. S5A**). We found that almost all  
233 pairs featuring moderate to strong repressors tended to behave exclusively as repressors, reducing gene  
234 expression at the strong pEF promoter while failing to drive transcription at the weak minCMV promoter (blue  
235 right region, **Fig. 4A**). Domains that were dual-functional when paired with negative controls (e.g., MYBA,  
236 FOXO3, and SERTAD2) tended to be dual-functional in combinations with either weak repressors or other dual  
237 effectors (green region, **Fig. 4A**), although they were dominated by strong repressors (e.g. ZNF10).

238 For most domains their activator-repressor and repressor-activator orientations had similar strengths for  
239 both activation (**Fig. S5B**, Pearson R=0.82, p<2.2×10<sup>-16</sup>) and repression (**Fig. S5C** Pearson R=0.79, p<2.2×10<sup>-16</sup>), indicating that these results were not the result of position-specific effects of the protein domains. Activator-

241 repressor combinations with stronger activators were significantly more likely to be able to activate the minimal  
242 promoter than combinations with weaker activators (**Fig. 4B**, Pearson R=0.79, p=1.6×10<sup>-4</sup>). For repressors, this  
243 relationship was significantly less pronounced: weak repressors were only slightly less effective than strong  
244 repressors when paired with activator domains (**Fig. 4C**, Pearson R=-0.52, p=1.6×10<sup>-4</sup>). Testing pairs in low  
245 throughput validated that the combination of the weak repressive SH3 domain from BIN1 and FOXO3's TAD  
246 could both activate gene expression at the weak minCMV promoter (**Fig. 4D**) and repress gene expression when  
247 recruited to the strong pEF promoter (**Fig. 4E**). As expected, adding the weak repressor domain from BIN1  
248 decreased activation and increased repression compared to FOXO3 alone (**Fig. S2A, S3A**).

249 While in general we found that strong repressors overpowered virtually any non-repressor effector  
250 domain they were paired with, we did find a few examples of activators that could prevent dual-functional  
251 domains from repressing strong promoters (**Fig. 4A**). These activators include the activator KRAB from ZN597,  
252 a variant KRAB domain that functions as an activator; the SH3 domain from ANM2 previously described in this  
253 work; and the activating SH3 domain from BTK. While BTK's SH3 domain is known to regulate BTK kinase  
254 activity upon autophosphorylation<sup>51</sup>, and while BTK can translocate to the nucleus<sup>52</sup>, we are unaware of a  
255 previously described role for the SH3 domain in regulating gene expression of BTK targets.

256 We were excited to find a number of non-hit domains that could not alter gene expression when paired  
257 with negative controls but did affect the function of other effector domains (gray labels, **Fig. 4A, Fig. S5A**). For  
258 example, we found that the aforementioned domain from DPY30, a core subunit of the SET1/MLL  
259 methyltransferase complex that interacts with ASH2L to establish H3K4me3<sup>53</sup>, was able to ablate repressor  
260 function when paired with not only dual-functional domains but even weak to moderate strength repressors.  
261 Additionally, the N-terminal domain of DPF1, which links the NF-κB RelA/p52 heterodimer with SWI/SNF  
262 complex subunits to drive transcription<sup>54</sup>, was able to prevent repressor function from strong repressors, such  
263 as HERC2's Cyt-b5 domain. We also found non-hit domains that prevented activation when paired with other  
264 effectors, including the a tile from the N-terminal disordered region of DNMT3B thought to be part of a broader  
265 region mediating interactions with the methyltransferases DNMT1 and DNMT3A<sup>55</sup>. The C2HR domain from the  
266 zinc finger ZNF496, which has been shown to overpower the variant activator KRAB present on the same  
267 transcription factor<sup>22</sup>, was similarly able to prevent all activators and all but the strongest dual-functional domains  
268 from driving gene expression. Altogether, these effectors did not themselves activate or repress transcription  
269 when paired with negative controls, but did modulate the activity of partner effectors on the same molecule in a  
270 manner consistent with the function of their native proteins.

271 We wondered how the distribution of effector combinations along both repression and activation  
272 log<sub>2</sub>(ON:OFF) scores varied between repressor-dual, activator-dual, and repressor-activator combinations. We  
273 found that while most repressor-dual combinations functioned as pure repressors, some dual-functional domains  
274 were able to combine with repressors to produce overall dual-functional combinations (**Fig. 4F**). In contrast,  
275 combining activators with repressors produced a larger number of effector pairs that neither activated the weak  
276 promoter nor repressed the strong promoter, with relatively fewer dual-functional combinations (**Fig. 4G**). Pairing  
277 pure activators with dual-functional domains produced combinations that mostly acted as dual activator-  
278 repressors, with a smaller fraction of combinations acting as pure activators without maintaining the repressor  
279 effect of the dual-functional domain (**Fig. 4H**).  
280

## 281 Systematic characterization of domains that influence KRAB-mediated repression

282

283 Our screen data so far indicated that virtually any concatenation including the ZNF10 KRAB domain  
284 functioned as a strong repressor (~180 KRAB-containing pairs). We were interested in investigating this KRAB

285 domain in more detail, as it is widely used in the well-known CRISPRi system<sup>56</sup>, has been harnessed in  
286 conjunction with DNA methyltransferases to produce more durable epigenetic silencing<sup>9,25</sup>, and has been deeply  
287 characterized<sup>21</sup>. We decided to test whether this pattern of KRAB dominance held when using a larger panel of  
288 partner domains, and generated a lentiviral library encoding a set of concatenations fusing ZNF10 KRAB with a  
289 library comprised of ~5000 80-amino acid proteins sequences consisting of Pfam annotated domains from  
290 nuclearly localized proteins and a set of negative controls<sup>21</sup>. We transduced this library into K562 cells expressing  
291 a reporter gene driven by the strong pEF promoter (**Fig. 5A**) and recruited the concatenations to the reporter  
292 gene for 5 days by adding doxycycline. We used a lower doxycycline concentration than in our previous screens  
293 (100 ng/ml vs 1000 ng/ml); to slow down KRAB silencing (**Fig. S6A**) and thus allow for a wider dynamic range  
294 for measuring both decrease and increase of function. We measured the enrichment of domain pairs in the ON  
295 versus OFF populations (**Fig. S6B**) at day 2 of recruitment to determine the speed of silencing; we opted not to  
296 take a measurement on day 5 at the end of recruitment, as the population was virtually entirely silent (**Fig. 5B**).  
297

On day 2, the majority of pairs featuring well-expressed domains (as measured by FLAG staining  
298 before<sup>21</sup>) scored similarly to concatenations of KRAB with a negative control domain (**Fig. 5C-D**), consistent with  
299 the flow cytometry results showing that the majority of the cells were silenced (**Fig. 5B**). This matched our  
300 expectations from the prior screens, where concatenations featuring KRAB silenced similarly to KRAB on its own  
301 or with negative control domains. In this case, the negative controls consisted of a large set of random sequences  
302 and DMD fragments<sup>21</sup>, and contained both well-expressed and poorly expressed proteins. We found that a  
303 number of domains that were lowly expressed on their own ablated KRAB function when paired with it (**Fig. 5C-D**).  
304 For example, we verified that the poorly expressed DHX16 OB\_NTP and BAZ1A DDT domains inhibited  
305 KRAB function at 100 ng/mL dox (**Fig. 5E**, top). Interestingly, increasing the dox concentration to 1000 ng/mL  
306 dox permitted some silencing for KRAB-DHX16 and full silencing for KRAB-BAZ1A (**Fig. 5E**, bottom), consistent  
307 with the loss of function of these KRAB fusions coming from decreased protein abundance.  
308

We also observed loss of KRAB function when fused to certain well-expressed domains from proteins  
309 that are part of the basic transcriptional machinery, namely the 2nd WD40 domain from TAF5L and the fork  
310 domain of RPB2. Consistent with the high-throughput measurements, individual validations at 100 ng/ml dox  
311 show a complete loss of KRAB silencing for the fusions with these domains (**Fig. 5E**, top row). The annotated  
312 RBP2 domain is smaller than the 80 aa sequence we used in our screen, and the 64 aa trimmed version had a  
313 lower capability of opposing KRAB, allowing for more silencing (**Fig. S6E**). Some amount of silencing was also  
314 restored at saturating dox concentrations for both KRAB-TAF5L and KRAB-RBP2 (**Fig. 5E**, bottom row), showing  
315 that KRAB can still dominate if enough of it is recruited at the locus. These results were corroborated by an  
316 overall increase in the rate of gene silencing when KRAB concatenations were recruited at 1000 ng/mL dox as  
317 compared to recruitment at 100 ng/mL dox (**Fig. S6F**,  $p=0.0239$ , paired t-test).  
318

In the high-throughput measurements we also identified a small number of domains that increased KRAB  
319 silencing (**Fig. 5C**, bottom left). We validated that the library tile containing the homeodomain from GSX2  
320 increases KRAB silencing at day 2 when recruited at 100 and 1000 ng/ml doxycycline (**Fig. 5E**, left), consistent  
321 with its behavior as a repressor when recruited on its own in our previous Pfam screen<sup>21</sup>. Interestingly, the  
322 trimmed version that contains only the annotated homeodomain does not enhance KRAB silencing (**Fig. S6C**).  
323 In contrast, the NHR2 domain from MTG8R, on its own a slightly weaker repressor than GSX2's homeodomain  
324 in our prior screens, was unable to modify KRAB-mediated silencing at both 100 ng/mL and 1000 ng/mL dox  
325 (**Fig. 5E**).  
326  
327

## 328 Composing effector domains to generate multifunctional synthetic transcription factors

329

330 We wanted to take advantage of the fact that the KRAB repressor is dominant over activators to build a  
331 more versatile transcription factor that can switch between repressor and activator based on addition or removal  
332 of a second drug. We engineered a version of our rTetR-FOXO3-ZNF10 KRAB concatenation plasmid where  
333 the two effector domains were separated by a StaPL domain, which cleaves itself in the absence of asunaprevir  
334 (ASV)<sup>57</sup> (**Fig. 6A**). Thus, in the absence of ASV, the KRAB domain would be cleaved, leaving behind the dual-  
335 functional rTetR-FOXO3, while upon addition of the ASV inhibitor, the KRAB domain would act as a dominant  
336 repressor (**Fig. 6B**).

337 We first verified that the construct worked as expected when driving gene activation at the minCMV  
338 promoter (**Fig. 6C**) and gene repression at the pEF promoter (**Fig. 6D**, **Fig. S7A**) at maximum recruitment with  
339 a saturating dose of dox (1000 ng/mL). At minCMV, the rTetR-FOXO3-StaPL-KRAB drove strong gene activation  
340 in the absence of ASV (**Fig. 6C**, bottom right green) to a similar level as rTetR-FOXO3 only (**Fig. 6C**, top right).  
341 At the same promoter, addition of ASV reduced activation by the rTetR-FOXO3-StaPL-KRAB to a minimum (**Fig.**  
342 **6C**, bottom right blue), comparable to rTetR-KRAB recruitment, and consistent with KRAB dominating over  
343 FOXO3. At pEF, the rTetR-FOXO3-StaPL-KRAB repressed virtually all cells upon ASV addition (**Fig. 6D**, bottom  
344 right blue), the same as rTetR-KRAB alone. Without ASV, this fusion still repressed 50% of the cells (**Fig. 6D**,  
345 bottom right green), consistent with rTetR-FOXO3 alone being able to silence pEF in about ~54-57% percent of  
346 cells (**Fig. 6D**, top right). As expected, the negative controls rTetR alone and rTetR-StaPL did not change gene  
347 expression at either promoter (**Fig. 6C-D**). Altogether, the construct behaved as expected: adding ASV changed  
348 its behavior from FOXO3-like to KRAB-like.

349 The ability to toggle the FOXO3-StaPL-KRAB TF from activator to repressor and then back to activator  
350 again allowed us to test whether the activator behaves differently at a promoter before and after KRAB-induced  
351 repression. In order to do this, we added dox to cells containing this fusion for a period of 20 days to induce  
352 recruitment, and within this interval we varied ASV to toggle between FOXO3 (-ASV) and KRAB dominant  
353 (+ASV) (**Fig. 6E-F**). At the minCMV reporter, we saw rapid gene activation upon addition of doxycycline that was  
354 lost when ASV was added and KRAB recruited to the promoter (**Fig. 6E**). Removal of ASV, leading to recruitment  
355 of FOXO3 alone at this KRAB-silenced minCMV promoter produced much slower gene activation than at the  
356 unsilenced minCMV promoter (**Fig. 6E**, days 0-2 vs. 7-20). The slow reactivation upon ASV washout suggests  
357 that while FOXO3 may be able to drive gene expression from a minimal promoter at a permissive locus, it is less  
358 efficient at reactivating that promoter after it is silenced with potentially methylated and/or compacted chromatin.  
359 At the pEF promoter (**Fig. 6F**), we observed a slow reduction in gene expression without ASV during days 0-2,  
360 consistent with the dual FOXO3 acting as a weak repressor at pEF. The percentage of cells silenced rapidly  
361 increased upon the addition of ASV and recruitment of KRAB culminating in complete silencing by day 7. After  
362 removal of ASV, we saw slow but measurable reactivation that increased dramatically upon doxycycline removal  
363 (**Fig. 6F**, days 7-20 vs. 20-27), suggesting that while FOXO3 can behave as an activator and drive gene  
364 expression at minCMV, its repressive capacity may inhibit proper reactivation of the pEF promoter.

365 In order to characterize the range of behaviors achievable with this inducible synthetic TF, we varied the  
366 dosing of both ASV and doxycycline while recruiting the rTetR-FOXO3-StaPL-KRAB construct at both promoters.  
367 At 100 ng/mL dox increasing doses of ASV led to decreased gene activation at the minCMV promoter (**Fig. 6G**  
368 left, **S7B**), consistent with the expectation that increasing doses of ASV led to an increase in the species  
369 containing KRAB relative to FOXO3 only. We also found that increasing the dose of dox at 0 ASV led to  
370 increasing gene activation at the minCMV promoter, consistent with a model in which the absence of ASV leads  
371 to virtually minimal KRAB recruitment (**Fig. 6G** right, **Fig. S7B**). Low to moderate doses of ASV, 0.01uM and

372 0.1uM, produced intermediate profiles between FOXO3 and KRAB (**Fig. S7B-C**). Altogether, these results  
373 suggest a model where increasing the dose of ASV titrates between FOXO3-like and KRAB-like behavior, while  
374 increasing the dose of dox increases the degree of gene activation and/or repression.

375 We wished to understand what new profiles of gene expression the combined FOXO3-StaPL-KRAB  
376 transcription factor could generate that neither FOXO3 nor KRAB could produce on their own. For each  
377 experiment performed with the transcription factor, we plotted for each timepoint the fraction of cells actively  
378 expressing the pEF reporter versus the fraction expressing the minCMV promoter (**Fig. 6I**). We found a large  
379 number of states that the synthetic inducible TF was able to access at intermediate ASV doses (**Fig. 6I**, purple)  
380 that could not be achieved at 0uM ASV (FOXO3-like behavior) or 1uM ASV (KRAB-like behavior). These states  
381 were principally composed of situations where both minCMV and pEF were expressed in only a fraction of cells;  
382 FOXO3-like behavior mainly produced high levels of cells with minCMV on (**Fig. 6I**, green), while KRAB-like  
383 behavior did not permit minCMV expression at all (**Fig. 6I**, blue). In conclusion, we found that composing FOXO3  
384 and KRAB in this manner and varying the dose of ASV produced distinct profiles of gene expression not  
385 achievable with either effector domain individually.

386

## 387 Discussion

388

389 Despite considerable efforts to parse the combinatorial logic of TF binding and gene regulation<sup>19,58-60</sup>, our  
390 understanding of how multiple distinct functional transcriptional effector domains work together within a single  
391 TF remains limited. Improved characterization of the combined function of effector domains is important for  
392 understanding the function of natural TFs featuring multiple effector domains<sup>2</sup>, for building complex synthetic  
393 biology tools to manipulate gene expression<sup>61</sup>, and for building cell therapies to detect and treat diseases<sup>26,62-65</sup>.  
394 Here, we present the results of screening thousands of effector domain combinations, another step towards  
395 uncovering the basic principles of combinatorial gene regulation.

396 We found that weak activators can synergize to drive robust gene expression even when the individual  
397 activators being paired were not particularly strong. This is in agreement with previous results showing that  
398 multiple activation domains can act synergistically in yeast when recruited at a synthetic reporter<sup>7,14</sup>, and in  
399 human cells when recruited at reporters or endogenous genes using dCas9<sup>24,66</sup>. At the other extreme, we found  
400 that some of the strongest activators, which in our system acted as dual-functional domains that could also  
401 repress a constitutive promoter, were often antagonistic: when paired with each other, they produced less gene  
402 activation from a minimal promoter than when recruited on their own. Antagonism between transcriptional  
403 activators has been reported in other contexts: TFs have been shown to interfere with each other's ability to  
404 activate genes via direct interactions with each other<sup>67</sup> or via squelching mechanisms involving the sequestration  
405 of coactivator proteins<sup>68</sup>. However, our findings suggest a more general negative feedback mechanism triggered  
406 by high levels of activators at a promoter; more thorough confirmation of this phenomenon and subsequent  
407 investigations into its molecular underpinnings are still needed.

408 Our activation data showed a tight coupling between the fraction of actively transcribing cells and the  
409 degree of transcription in those cells, consistent with manipulation of transcriptional bursting dynamics<sup>49,69</sup>.  
410 Recently, certain TFs were shown to increase only burst size or only burst frequency, depending on their  
411 molecular mechanism of action<sup>70</sup>. While our long-lived protein reporter does not allow us to differentiate between  
412 changes in burst size or frequency for our combinations, it would be interesting to repeat these high-throughput  
413 measurements coupled with RNA FISH or a destabilized reporter that allows extraction of burst parameters.

414 We found that repressors generally overpowered activators, although select activator domains could  
415 weaken the repressor function of dual-functional or weakly repressive domains. We confirmed this result more

416 thoroughly via an in-depth screen measuring KRAB function when paired with a panel of 5,000 domains from  
417 nuclearly localized proteins, and found that KRAB maintained its repressive function when paired with most  
418 domains except ones that had decreased expression and a few domains from general TAFs and RNA  
419 polymerase. We used these insights to build a synthetic transcription factor whose function could be switched  
420 from a dual-functional FOXO3-like profile (that activates a minimal promoter and weakly represses a constitutive  
421 one) to a repressive KRAB-like profile via the addition of the small molecule drug ASV. We used this switchable  
422 TF to show that both the activation and repression functions of the FOXO3 effector domain changed when the  
423 target promoter was first silenced by KRAB. While activation from a minimal promoter was impaired by prior  
424 silencing as expected, recruitment of FOXO3 increased KRAB-mediated epigenetic memory and reduced  
425 reporter gene reactivation at the pEF promoter. Importantly, since ASV tuned the behavior of this synthetic TF  
426 independent of dox-mediated recruitment, we could generate gene expression profiles that were not accessible  
427 with either individual effector domain alone. Looking forward, this synthetic switchable TF can be used to split a  
428 cell population into well-defined percentages that express the desired combination of two target genes, e.g., 50%  
429 of the cells expressing gene 1 and 50% expressing gene 2. Such flexibility could allow for more sophisticated  
430 and complex gene circuits and the engineering of “higher-order” cell behaviors and programs<sup>71</sup>.

431 When we started this study, we had a limited number of validated activation and repressive domains to  
432 test in combination. Recently, however, we have identified a much larger set of effector domains from human  
433 TFs<sup>20</sup>. Prior efforts have characterized individual examples of multiple effector domains being combined within  
434 a larger TF<sup>22,72</sup>; it would be instructive to design a library that systematically tests combinations of domains that  
435 come from the same TF and compare these results with recent ORFeome screens measuring activation and  
436 repression of full-length TFs<sup>33</sup>. Moreover, emerging work has begun to connect TF and effector function with  
437 measurements of their affinities for transcriptional coactivators and corepressors<sup>13,33,73</sup>. In addition, while our  
438 work characterized effector combinations at two promoters in one human cell line, we are excited for future  
439 efforts to broaden these settings to include other DNA-binding domains beyond rTetR, other promoters, and  
440 other cell types. Such efforts will undoubtedly aid projects aiming to build robust tools for cell engineering that  
441 can function across multiple contexts<sup>74–76</sup>.

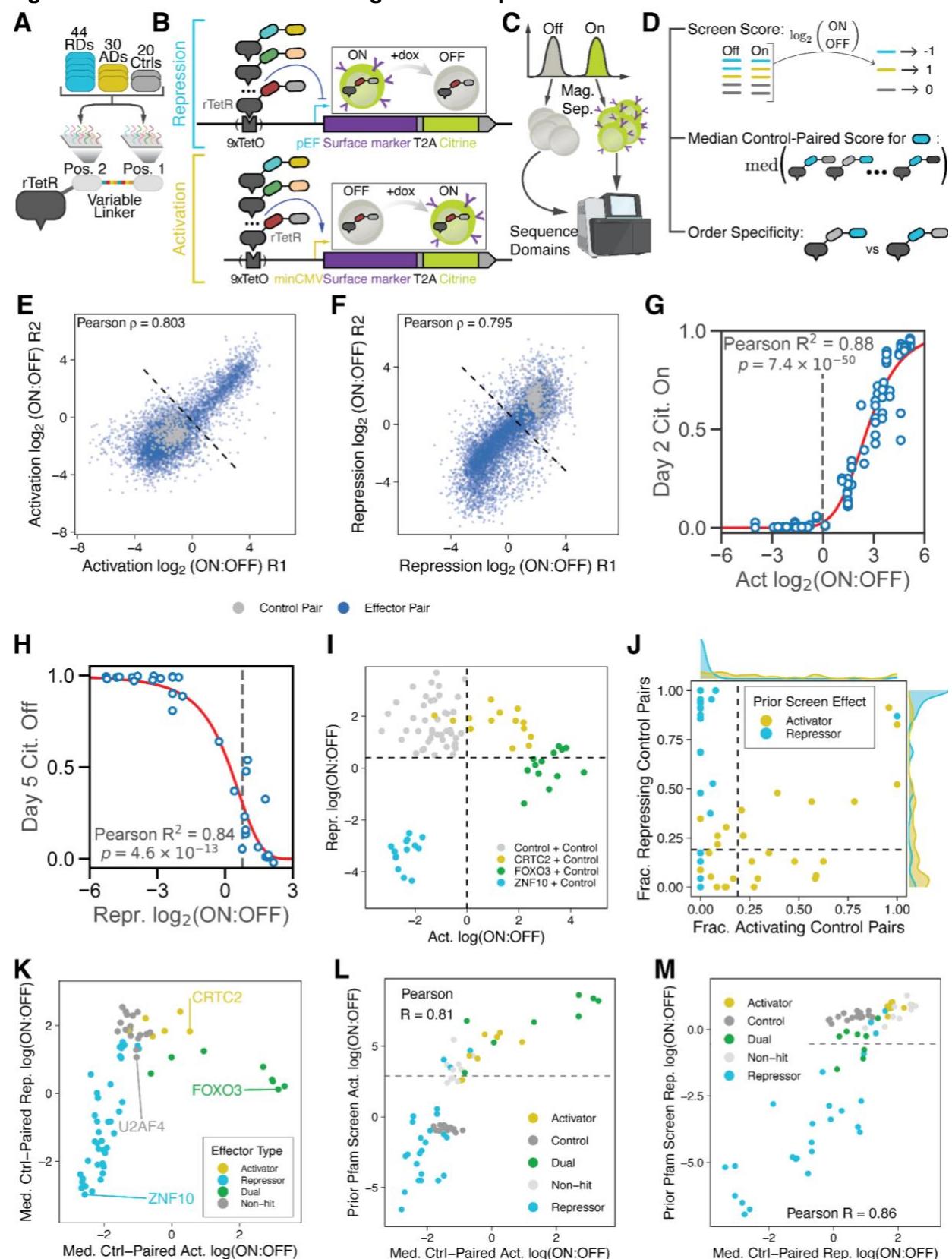
442

## Figures

443

444

### Figure 1 -- Combinatorial screening of transcriptional effector domains



445

- 446 A. Construction of the combinatorial library. A library consisting of approximately 100 effector domains was cloned into  
447 a lentiviral backbone in two positions (Pos. 1 and Pos. 2), connected by a 16-amino acid XTE linker featuring a  
448 varying DNA but constant amino acid sequence, and fused to the DNA binding domain rTetR (**Materials &**  
449 **Methods**). This library consisted of domains that were previously identified as repressive domains (RDs), activator  
450 domains (ADs) and controls with no effect on transcription (Ctrls).
- 451 B. Schematic of synthetic reporter system. rTetR-fused domain pairs are recruited to a synthetic reporter expressing  
452 a surface marker exposing an IgG epitope and an mCitrine fluorophore. The reporter gene's promoter can be varied,  
453 with pEF being used to measure repression (top, blue arrow) and minCMV for activation (bottom, yellow arrow)..
- 454 C. Magnetic separation. Cells are bound with magnetic beads that recognize the reporter surface marker and are used  
455 to separate ON and OFF populations. The DNA encoding for domains expressed in cells from each population are  
456 then sequenced via next-generation sequencing.
- 457 D. Computation of screen measurements. Domain pairs are assigned a screen score equal to  $\log_2(\text{ON:OFF})$  where  
458 ON is the relative proportion of a domain in the ON population and OFF is the relative proportion of a domain in the  
459 OFF population. The median control-paired score for a given domain is computed as the median score of all pairs  
460 featuring that domain and a negative control domain. Order specificity was determined by determining if the domain  
461 function significantly differently when positioned upstream or downstream of the negative control domain.
- 462 E. Correlation between 2 replicates for the activation screen measurements for all domain pairs after 2 days of  
463 recruitment at the minCMV promoter, with domains colored based on whether they are composed of 2 negative  
464 control domains (gray) or at least 1 effector domain (blue). Dashed line is the threshold for a combination being  
465 labeled as activating, 2 standard deviations away from the mean of the negative control-negative control  
466 combinations.
- 467 F. Correlation between 2 replicates for the repression screen measurements for all domain pairs after 5 days of  
468 recruitment at the pEF promoter, with domains colored based on whether they are composed of 2 negative control  
469 domains (gray) or at least 1 effector domain (blue). Dashed line is the threshold for a combination being labeled as  
470 repressing, 2 standard deviations away from the mean of the negative control-negative control combinations.
- 471 G. Correlation between activation  $\log_2(\text{ON:OFF})$  scores (x-axis) and the fraction of cells activated after 2 days of  
472 doxycycline recruitment to the weak minCMV promoter (y-axis) during individual low-throughput measurements of  
473 each pair's function via flow cytometry. The sigmoid curve represents the best-fit line of the form  $y = 1/(1 + (x/k)^n)$ .  
474 Vertical dashed line corresponds to the activation threshold for effector domain combinations as defined above.
- 475 H. Correlation between repression  $\log_2(\text{ON:OFF})$  scores (x-axis) and the fraction of cells repressed after 5 days of  
476 doxycycline recruitment to the strong pEF promoter (y-axis) during individual low-throughput measurements of each  
477 pair's function via flow cytometry. The sigmoid curve represents the best-fit line of the form  $y = 1 - 1/(1 + (x/k)^n)$ .  
478 Vertical dashed line corresponds to the repression threshold for effector domain combinations as defined above.
- 479 I. Variation in control-paired measurements. The  $\log_2(\text{ON:OFF})$  scores for activation (x-axis) and repression (y-axis)  
480 is shown for negative control-negative control domain pairs in any orientation (gray), CRTC2-negative control  
481 domain pairs in any orientation (yellow), FOXO3-negative control domain pairs in any orientation (green), and  
482 ZNF10 KRAB-negative control domain pairs in any orientation (blue). Dashed lines represent activation and  
483 repression thresholds as defined above.
- 484 J. Number of control-including concatenations that exceed the threshold for activation (x-axis) and repression (y-axis)  
485 for all domains, colored by whether that domain acts as a repressor (blue) or activator (yellow) when recruited alone.  
486 Horizontal and vertical dashed lines indicate thresholds needed for a domain to be called as an activator or  
487 repressor.
- 488 K. Median control-paired activation (x-axis) and repression (y-axis) scores for individual effector domains. Activators  
489 are colored yellow, repressors blue, dual-functional effectors green, and non-hits gray.
- 490 L. Correlation between activation scores from the current domain combination screen and prior nuclear Pfam single-  
491 domain activation screen. Domains present in this screen that were measured for activation in a prior screen are  
492 shown, with the median control-paired score in this screen shown on the x-axis and the  $\log_2(\text{ON:OFF})$  score from  
493 the prior activation screen shown on the y-axis. The type of effector (activator, repressor, dual, negative control,

494  
495

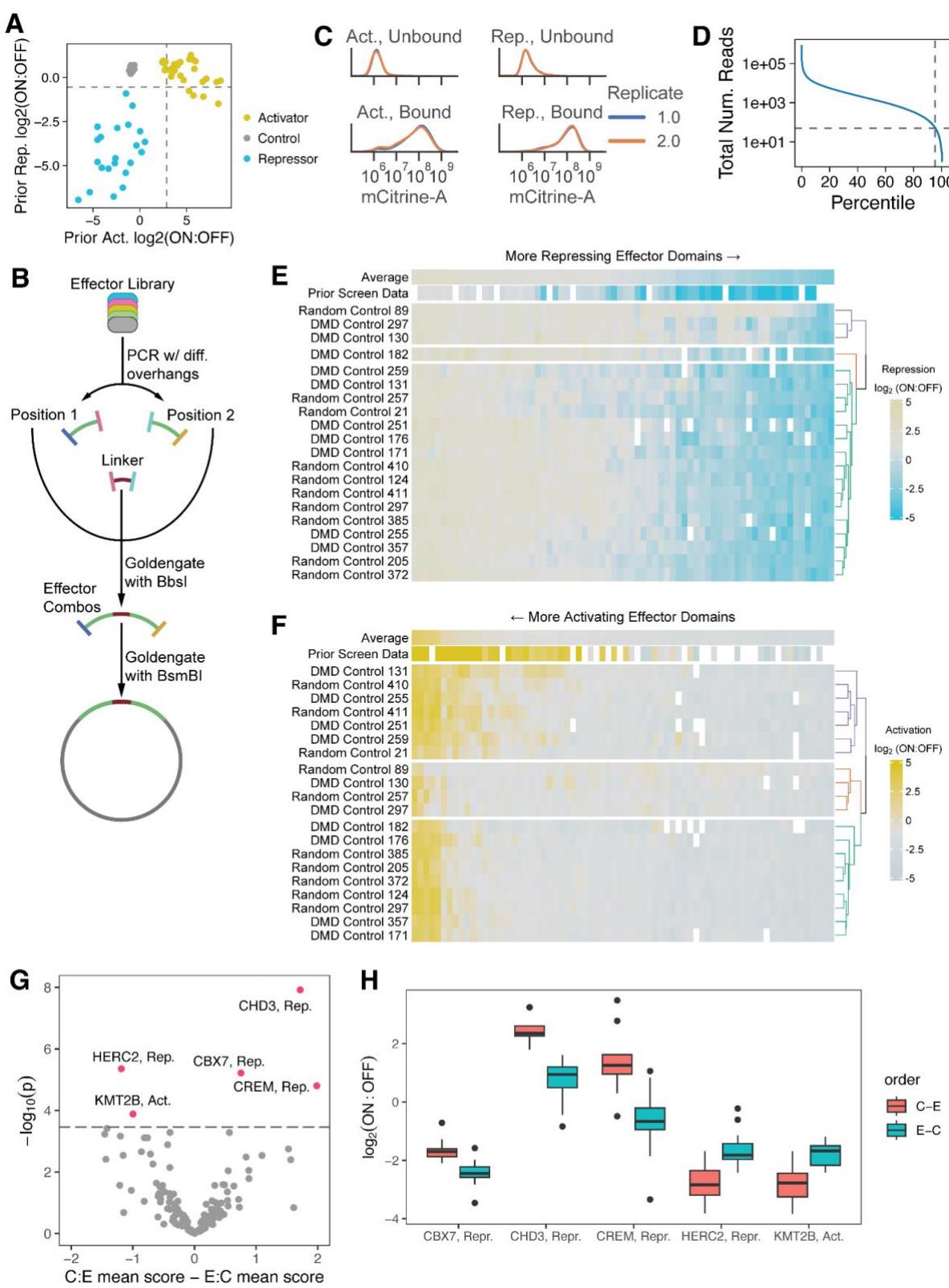
non-hit) is indicated by the color of each point. Dashed line indicates the threshold for activation in the prior activation screen.

496  
497  
498  
499  
500  
501

M. Correlation between repression scores from the current domain combination screen and prior nuclear Pfam single-domain repression screen. Domains present in this screen that were measured for repression in a prior screen are shown, with the median control-paired score in this screen shown on the x-axis and the  $\log_2(\text{ON:OFF})$  score from the prior activation screen shown on the y-axis. The type of effector (activator, repressor, dual, negative control, non-hit) is indicated by the color of each point. Dashed line indicates the threshold for repression in the prior activation screen.

502  
503

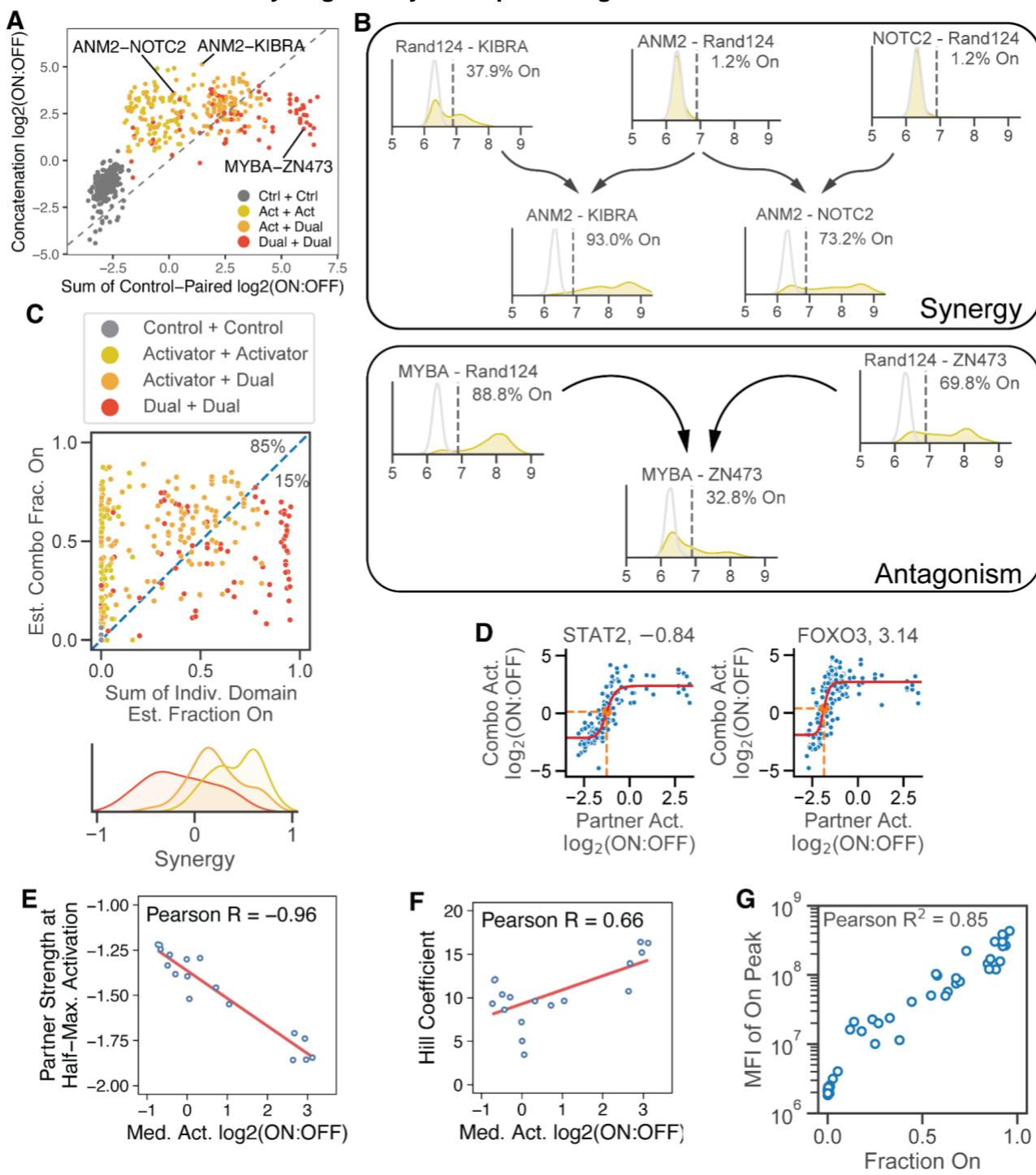
**Figure S1 – Screen Quality Control and Filtering**



504  
505

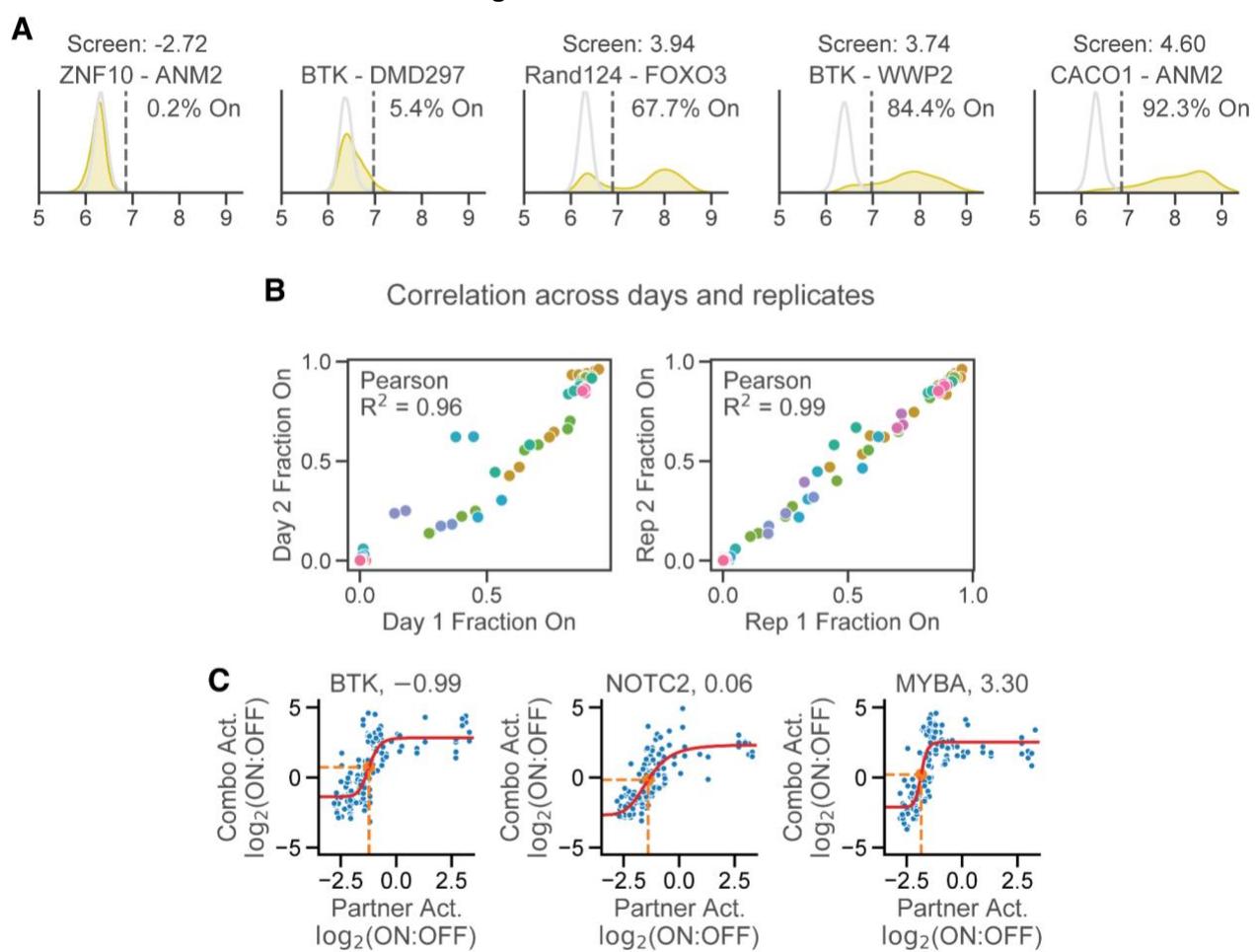
- 506 A. Scatterplot of prior data<sup>21</sup> for both activation at minCMV (x-axis) and repression at pEF (y-axis) for individual  
507 domains chosen in the screen, colored based on their prior behavior. Controls are random 80AA sequences or tiles  
508 from the DMD protein. Dashed lines represent hit thresholds in the prior screen.  
509 B. Schematic of co-recruit cloning strategy. Domain library is PCRed with distinct overhangs to produce position 1 and  
510 position 2 libraries, which are then joined to the XTEN linker in one Goldengate reaction to produce a library of  
511 effector concatenations. Concatenations are then ligated into a lentiviral backbone via a second Goldengate  
512 reaction.  
513 C. Probability density (y-axis) of mCitrine fluorescence levels (x-axis) for bound (top) and bound (bottom) cells post-  
514 magnetic separation for activation (left) and repression (right) screens, with each replicate shown in a different color.  
515 D. Total number of reads for each domain pair (y-axis); domain pairs are ordered from most to least reads, with x-axis  
516 marks indicating what percentile of pairs had more reads than those at that x-value.  
517 E. Clustermap of controls (rows) when paired with each repressing effector domain (columns), with average effector  
518 domain profiles shown along the top. Color indicates the repression log<sub>2</sub>(ON:OFF) score for a particular control-  
519 effector pair. Domains filtered out were random control 89 and DMD controls 130, 182, and 297.  
520 F. Clustermap of controls (rows) when paired with each activating effector domain (columns), with average effector  
521 domain profiles shown along the top. Color indicates the activation log<sub>2</sub>(ON:OFF) score for a particular control-  
522 effector pair. Domains filtered out were random controls 89 and 257 and DMD controls 130 and 297.  
523 G. Comparison of orientations for each domain. Difference between the average score with the effector C-terminal to  
524 the control domain (C:E) and the average score with the effector N terminal to the control domain (E:C) (x-axis) is  
525 plotted against the negative log of the p-value as determined by a Welch's t-test. Dashed line indicates the  
526 Bonferroni-corrected significance threshold, and domains with significant differences between orientations are  
527 marked in red.  
528 H. For domains with statistically significant orientation-dependence, box-and-whisker plots of their log<sub>2</sub>(ON:OFF)  
529 scores are shown for each of the 2 possible orientations. Orientations that weakened the domains' effects were  
530 removed from downstream analysis.

531 **Figure 2 -- Activators work synergistically when paired together**



- 533 A. Evidence for activator-activator synergy in the high-throughput screen after two days of recruitment at the minCMV  
534 minimal promoter. For negative control-negative control (dark gray), activator-activator (yellow), activator-dual  
535 (orange), and dual-dual (red) domain pairs, the sum of their individual control-paired scores (x-axis) is plotted  
536 against the average score of the combination (y-axis). Individual examples of pairs highlighted in subsequent panels  
537 are labeled and indicated by lines.
- 538 B. Example low-throughput validations of activator-negative control and activator-activator pairs at the minCMV  
539 promoter. No doxycycline recruitment is shown in light gray, while the probability density of fluorescence after 2  
540 days of dox recruitment is shown in yellow. Vertical dashed lines indicate the threshold for calling cells ON or OFF,  
541 with the fraction of cells ON in the dox population labeled within each histogram.
- 542 C. Top: estimated fraction of cells activated at day 2 for every activator-activator (yellow), activator-dual (orange), and  
543 dual-dual (red) combination (y-axis) versus the sum of the component domains' individual estimated fraction of cells  
544 activated based on control-paired scores (x-axis). Bottom: density plot of degree of synergy (x-axis) for activator-  
545 activator (yellow), activator-dual (orange), and dual-dual (red) combination pairs. Synergy is defined as the  
546 difference between the combination's estimated day 2 fraction off and the sum of the individual domains' day 2  
547 fractions off.
- 548 D. Sigmoidal shape of activator behavior in the high-throughput screen. For STAT2 (left) and FOXO3 (right) , the  
549 control-paired score of its partner in any given pair (x-axis) is plotted against the score of the combination of the  
550 domain and that partner (y-axis) for all concatenations featuring the given domain.
- 551 E. For all activator domains, the control-paired activation strength of the activator (x-axis) is plotted against the half-  
552 maximal point of the corresponding sigmoid fit to screen data (y-axis). Line shown is the line of best fit.
- 553 F. For all activator domains, the control-paired activation strength of the activator (x-axis) is plotted against the Hill  
554 coefficient of the corresponding sigmoid fit to screen data (y-axis). Line shown is the line of best fit.
- 555 G. Correlation between the fraction of activated cells (x-axis) and the mean fluorescence intensity (MFI) of activated  
556 cells (y-axis) for each combination measured individually.

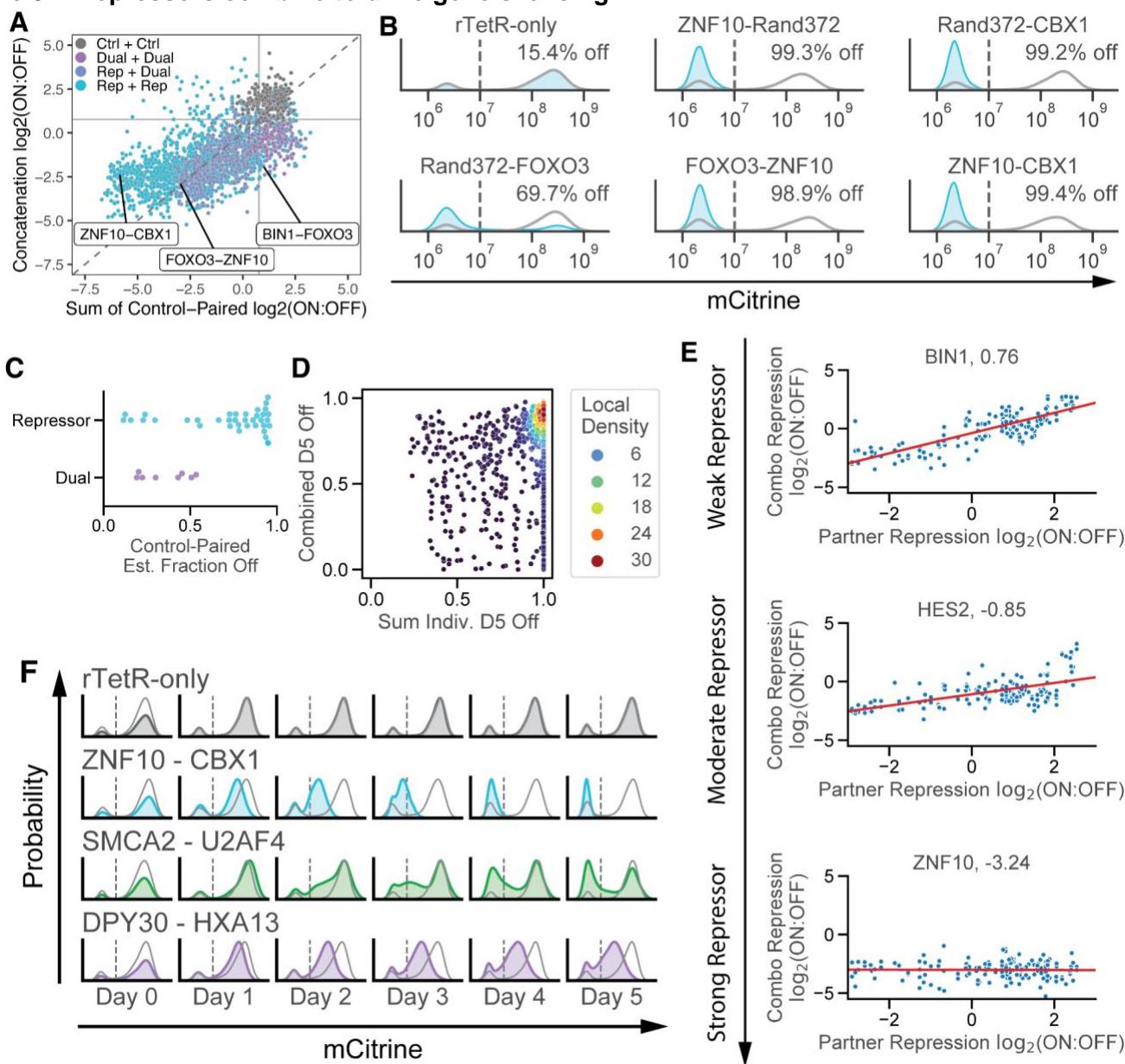
557 **Figure S2 – Activator Validations and Modeling**



558

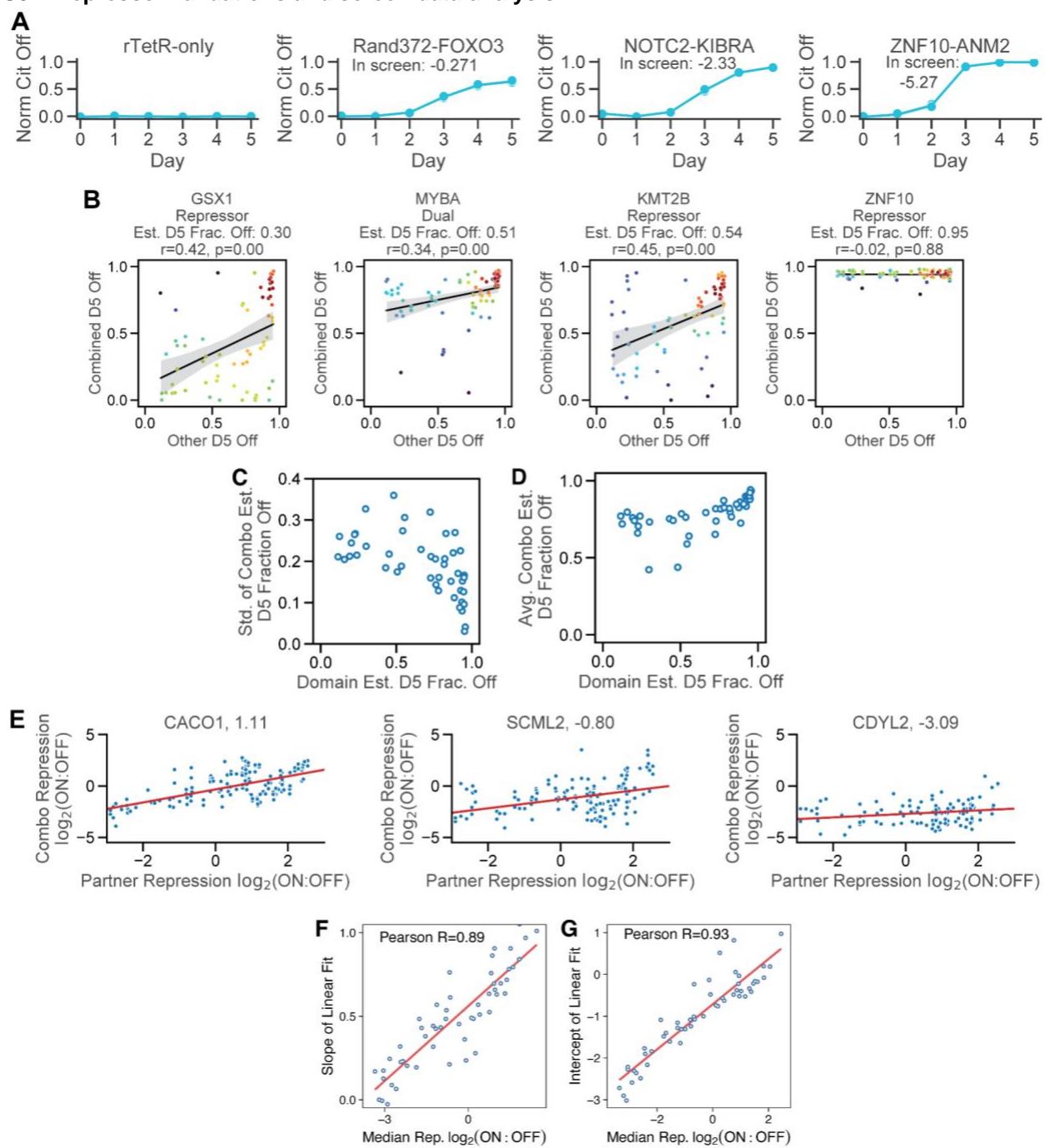
- 559 A. Example low-throughput validations of activator-negative control and activator-activator pairs. No doxycycline  
560 recruitment is shown in light gray, while the histogram of fluorescence after 2 days of dox recruitment is shown in  
561 yellow. Dashed lines represent the fluorescence threshold for calling cells On in reporter expression.  
562 B. Correlations between days 1 and 2 (left) and replicates 1 and 2 (right) for measuring the activation strength of  
563 domain concatenations. Colors indicate different constructs.  
564 C. Sigmoidal shape of activator behavior in the high-throughput screen. For BTK (left), NOTC2 (middle), and FOXO3  
565 (right) , the control-paired score of its partner in any given pair (x-axis) is plotted against the score of the combination  
566 of the domain and that partner (y-axis) for all concatenations featuring the given domain.

567 **Figure 3 -- Repressors combine to drive gene silencing**



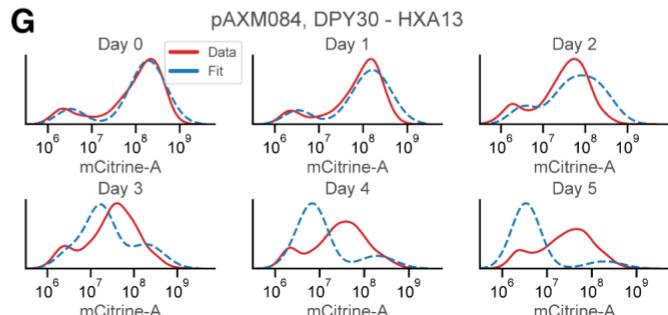
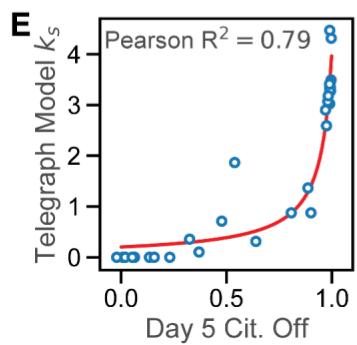
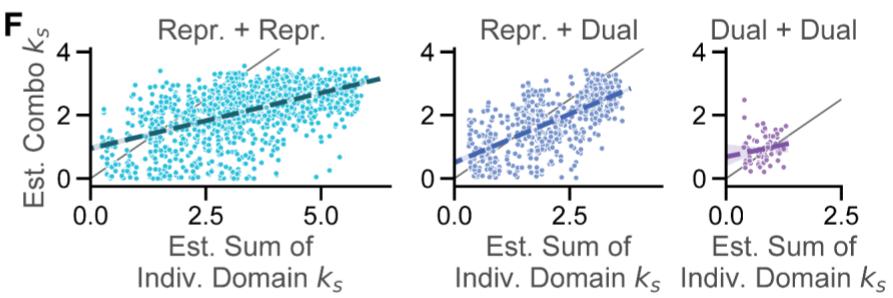
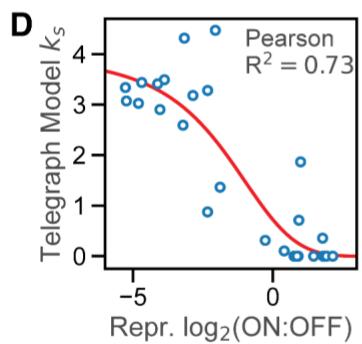
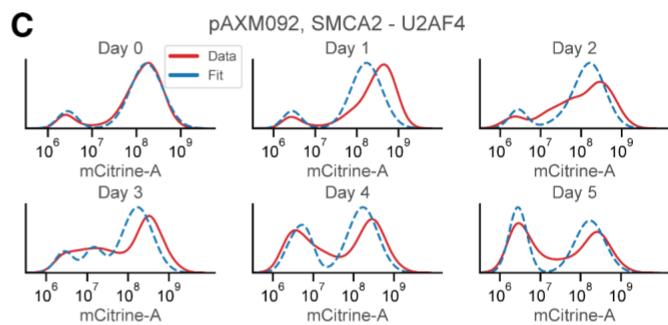
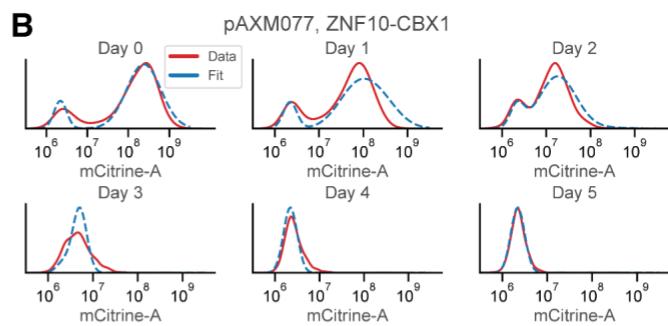
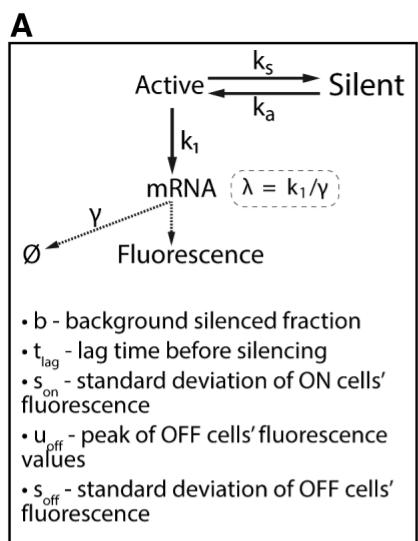
- 569 A. Repressor-repressor pair behavior in the high-throughput screen after five days of recruitment at the pEF promoter.  
570 For negative control-negative control (dark gray), repressor-repressor (light blue), repressor-dual (dark blue), and  
571 dual-dual (purple) domain pairs, the sum of their individual control-paired scores (x-axis) is plotted against the  
572 average score of the combination (y-axis). Horizontal and vertical lines denote the repression threshold as defined  
573 in **Fig. 1F**, while the diagonal dotted line corresponds to the line  $y=x$ .  
574 B. Fluorescent distributions of the citrine reporter gene from low-throughput flow cytometry measurements of repressor  
575 function of various domain combinations after 5 days of recruitment at the pEF promoter, with Rand372 denoting a  
576 negative control domain. Blue: 1000 ng/mL doxycycline; gray: no dox. Vertical dashed lines indicate the threshold  
577 for calling cells ON or OFF, with the fraction of cells OFF in the dox population labeled within each histogram.  
578 C. Estimated fraction of cells silenced at day 5 (x-axis) for repressors (blue) and dual-functional domains (purple) when  
579 paired with negative controls.  
580 D. For repressor-repressor, repressor-dual, and dual-dual domain pairs, the sum of the estimated day 5 fraction off for  
581 both individual domains (x-axis) is plotted against the estimated day 5 fraction off for the pair (y-axis). Color indicates  
582 density as determined by a Gaussian kernel density estimate.  
583 E. Linear shape of repressor behavior in the high-throughput screen. For a given domain (BIN1 - left, HES2 - middle,  
584 ZNF10 - right), the control-paired score of its partner in any given pair (x-axis) is plotted against the score of the  
585 combination of the domain and that partner (y-axis) for all concatenations featuring the given domain. Red lines  
586 indicate linear fits to the data.  
587 F. Individual plots of transcriptional repression. Probability density (y-axis) plots of mCitrine-A levels (x-axis) presented  
588 for four 4 different pairs of domains (rows) as measured by flow cytometry over 5 days of doxycycline recruitment  
589 (columns).

590 Figure S3 – Repressor validations and screen data analysis



- 592 A. Low-throughput measurement of repressor function of various domain combinations, with Rand372 denoting a  
593 negative control domain. Cells were labeled as Off when their mCitrine-A fluorescence dropped below  $10^7$  (**Fig.**  
594 **3B**).  
595 B. Effect of partner estimated day 5 fraction off on combination day 5 fraction off for repressor domains. For each  
596 repressor or dual domain (GSX1, MYBA, KMT2B, or ZNF10) the estimated day 5 fraction off of a given partner of  
597 that domain (x-axis) is plotted against the estimated day 5 fraction off for the concatenation (y-axis). Color denotes  
598 density as computed by a kernel density estimator, with red constituting denser, and blue less dense.  
599 C. Estimated day 5 fraction off (x-axis) for each domain versus the standard deviation of the estimated day 5 fractions  
600 off for all combinations including that domain (y-axis) for all repressor-repressor, repressor-dual, and dual-dual  
601 pairs.  
602 D. Estimated day 5 fraction off (x-axis) for each domain versus the average of the estimated day 5 fractions off for all  
603 combinations including that domain (y-axis) for all repressor-repressor, repressor-dual, and dual-dual pairs.  
604 E. Linear shape of repressor behavior in the high-throughput screen. For a given domain (CACO1 - left, SCML2 -  
605 middle, CDYL2 - right), the control-paired score of its partner in any given pair (x-axis) is plotted against the score  
606 of the combination of the domain and that partner (y-axis) for all concatenations featuring the given domain. Red  
607 lines indicate linear fits to the data.  
608 F. For all repressor domains, the control-paired strength of the repressor (x-axis) is plotted against the slope of the  
609 corresponding linear fit to screen data (y-axis). Line shown is the line of best fit.  
610 G. For all repressor domains, the control-paired strength of the repressor (x-axis) is plotted against the y-intercept of  
611 the corresponding linear fit to screen data (y-axis). Line shown is the line of best fit.

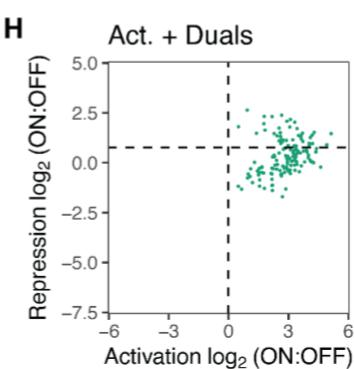
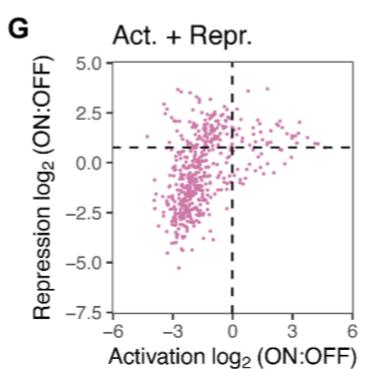
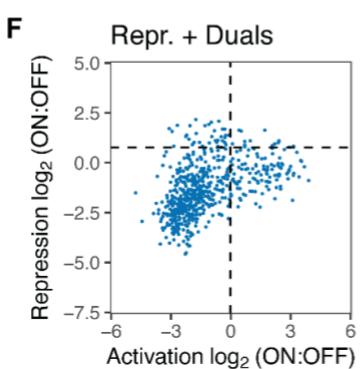
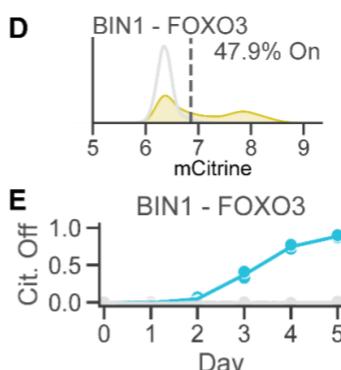
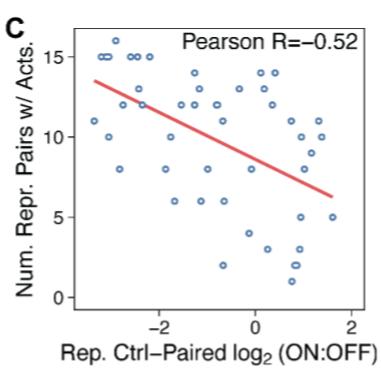
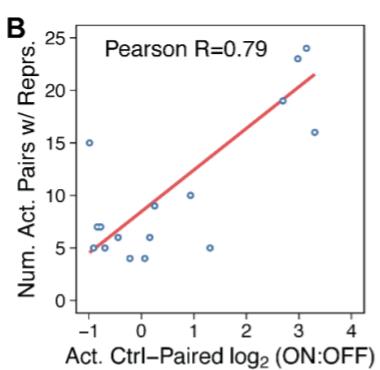
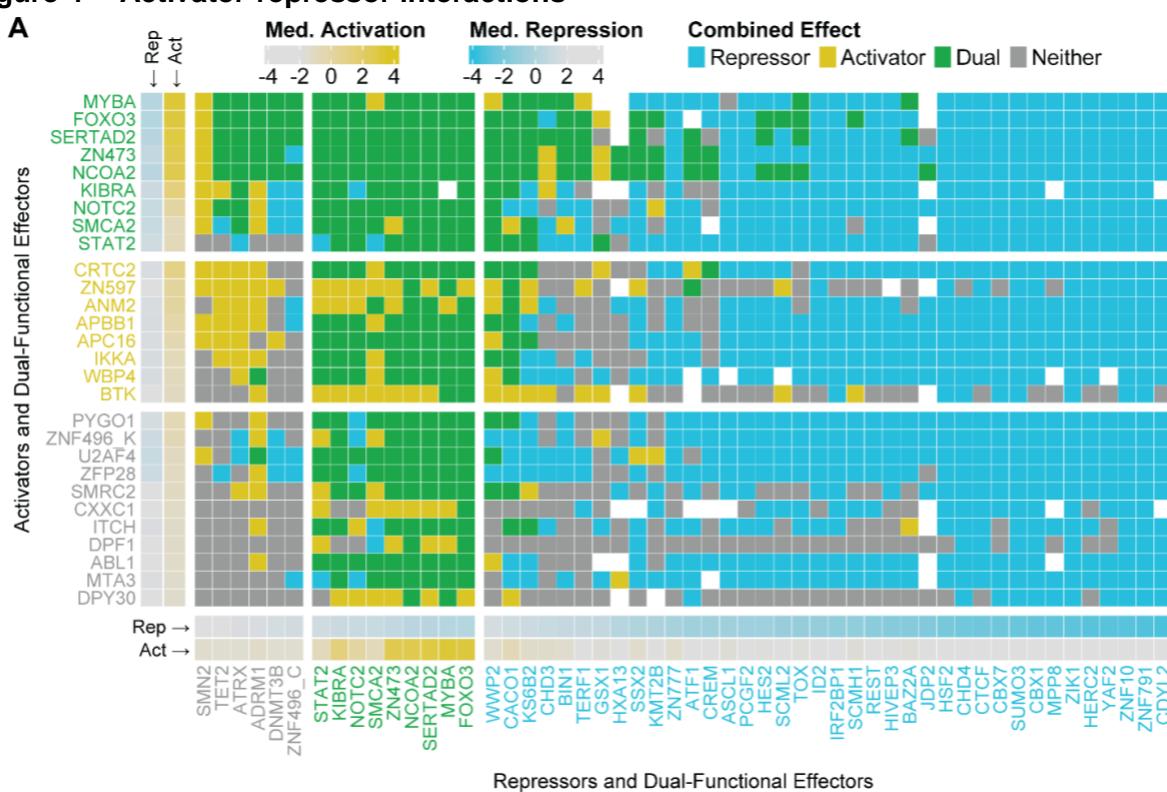
## 612 Figure S4 – Repressor modeling



- 615 A. Brief description of the silencing model (**Materials & Methods**). Cells in the Active state produce mRNA at rate  $k_1$ ,  
616 which decays with rate  $\gamma$ . We define  $\lambda = k_1/\gamma$  to be the average amount of mRNA in cells at steady state when Active.  
617 Cells in the Active state are silenced with rate  $k_s$  and reactivate at rate  $k_a$  upon recruitment. The model incorporates  
618 a permanently silent population  $b$  of background silenced cells. We assume a lag time  $t_{lag}$  before silencing. The  
619 standard deviation of the Active cells fluorescence is denoted  $s_{on}$  while the standard deviation of the Silent cells  
620 fluorescence is  $s_{off}$ . The fluorescence peak of Silent cells is defined as  $u_{off}$ .
- 621 B. Example fit of the silencing model for ZNF10-CBX1, with fluorescence indicated on the x-axis and density on the y-  
622 axis. Data gathered from flow cytometry measurements is shown in red, while the fitted model is shown in blue for  
623 each day (0-5).
- 624 C. Example fit of the silencing model for SMCA2-U2AF4, with fluorescence indicated on the x-axis and density on the  
625 y-axis. Data gathered from flow cytometry measurements is shown in red, while the fitted model is shown in blue  
626 for each day (0-5).
- 627 D. Correlation between repression  $\log_2(\text{ON:OFF})$  scores for validated concatenations (x-axis) and  $k_s$  values as  
628 estimated by the silencing model (y-axis), with best-fit sigmoid shown in red.
- 629 E. Correlation between day 5 fractions of cells silenced for validated concatenations (x-axis) and  $k_s$  values as  
630 estimated by the silencing model (y-axis), with best-fit asymptotically increasing curve shown in red.
- 631 F. For repressor-repressor (left), repressor-dual (middle), and dual-dual (right) pairs, the sum of the estimated  $k_s$  for  
632 both individual domains (x-axis) is plotted against the estimated  $k_s$  for the combination (y-axis). Line shown is a  
633 best-fit linear regression with a 95% confidence interval shaded in. Gray line indicates the  $y=x$  line.
- 634 G. Example fit of the silencing model for DPY30-HXA13, with fluorescence indicated on the x-axis and density on the  
635 y-axis. Data gathered from flow cytometry measurements is shown in red, while the fitted model is shown in blue  
636 for each day (0-5).

637

### Figure 4 -- Activator-repressor interactions

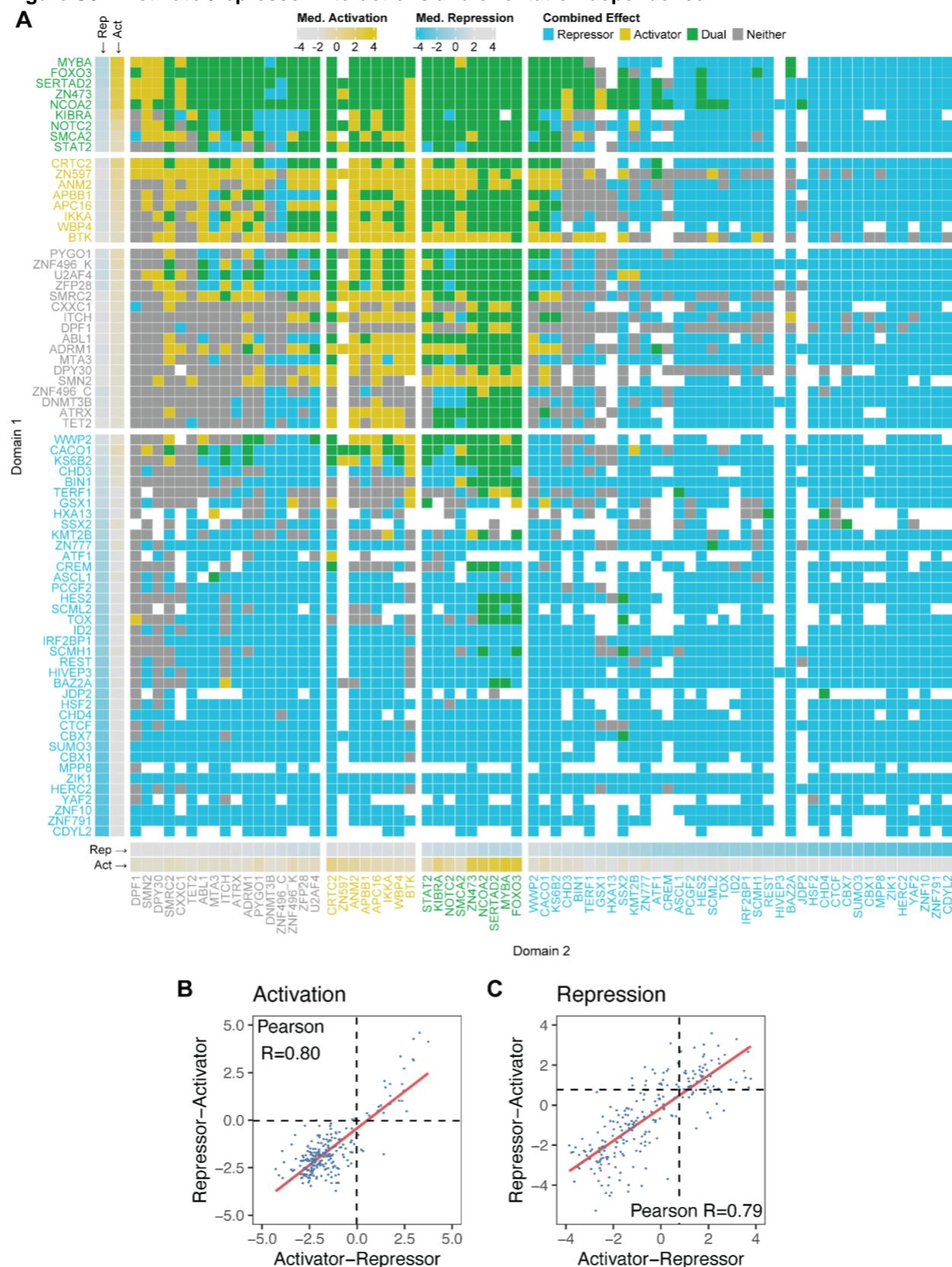


638

- 639 A. Representation of activator-repressor interactions measured by the high-throughput screen. Domains that were  
640 identified as repressors when recruited alone are arrayed along the x-axis separated into 3 categories based on  
641 their effect when paired with controls (non-hits on the left, dual-functional domains in the middle, and repressors on  
642 the right) with the stronger repressors within each category, as measured by their control-paired repression  
643  $\log_2(\text{ON:OFF})$ , on the right. Boxes just above each repressor domain indicate the control-paired  
644 activation/repression score for each effector. Domains that were identified as activators are arrayed along the y-  
645 axis, separated into 3 categories based on their effect when paired with controls (non-hits on the bottom, activators  
646 in the middle, and dual-functional domains on the top) with the stronger activators in each category, as measured  
647 by control-paired activation scores, on the top. Boxes just to the right of each activator domain indicate the control-  
648 paired activation/repression score for each effector. Boxes in the interior represent combinations of the  
649 corresponding activator and repressor, and are colored based on whether the combination acts as a repressor only  
650 (blue), activator only (yellow), both a repressor and an activator (green), or neither (gray).
- 651 B. Correlation between the control-paired activation score for a set of activators (x-axis) and the number of pairs  
652 including any repressor domain and that activator domain that were called as an activator pair hit (y-axis). Red line  
653 represents the line of best fit.
- 654 C. Correlation between the control-paired repression score for a set of repressors (x-axis) and the number of pairs  
655 including any activator domain and that activator domain that were called as a repressor pair hit (y-axis). Red line  
656 represents the line of best fit.
- 657 D. Low-throughput measurement of the activation strength of the pair of BIN1's SH3 domain and FOXO3's TAD when  
658 recruited at the minCMV for 2 days. The gray line represents the no-doxycycline condition, and the yellow 2 days  
659 doxycycline recruitment. The vertical dashed line indicates the threshold for labeling a cell as ON or OFF.
- 660 E. Fraction of cells with pEF promoter silenced by the fusion between BIN1's SH3 domain and FOXO3's TAD over 5  
661 days of recruitment. The fraction of cells silenced was normalized to a no-dox condition (**Materials & Methods**).
- 662 F. Distribution of repression  $\log_2(\text{ON:OFF})$  scores (y-axis) versus activation  $\log_2(\text{ON:OFF})$  scores (x-axis) for  
663 combinations featuring 1 repressor domain and 1 dual-functional domain.
- 664 G. Distribution of repression  $\log_2(\text{ON:OFF})$  scores (y-axis) versus activation  $\log_2(\text{ON:OFF})$  scores (x-axis) for  
665 combinations featuring 1 repressor domain and 1 activator domain.
- 666 H. Distribution of repression  $\log_2(\text{ON:OFF})$  scores (y-axis) versus activation  $\log_2(\text{ON:OFF})$  scores (x-axis) for  
667 combinations featuring 1 activator domain and 1 dual-functional domain.

668

### Figure S5 – Activator/repressor interactions and orientation dependence

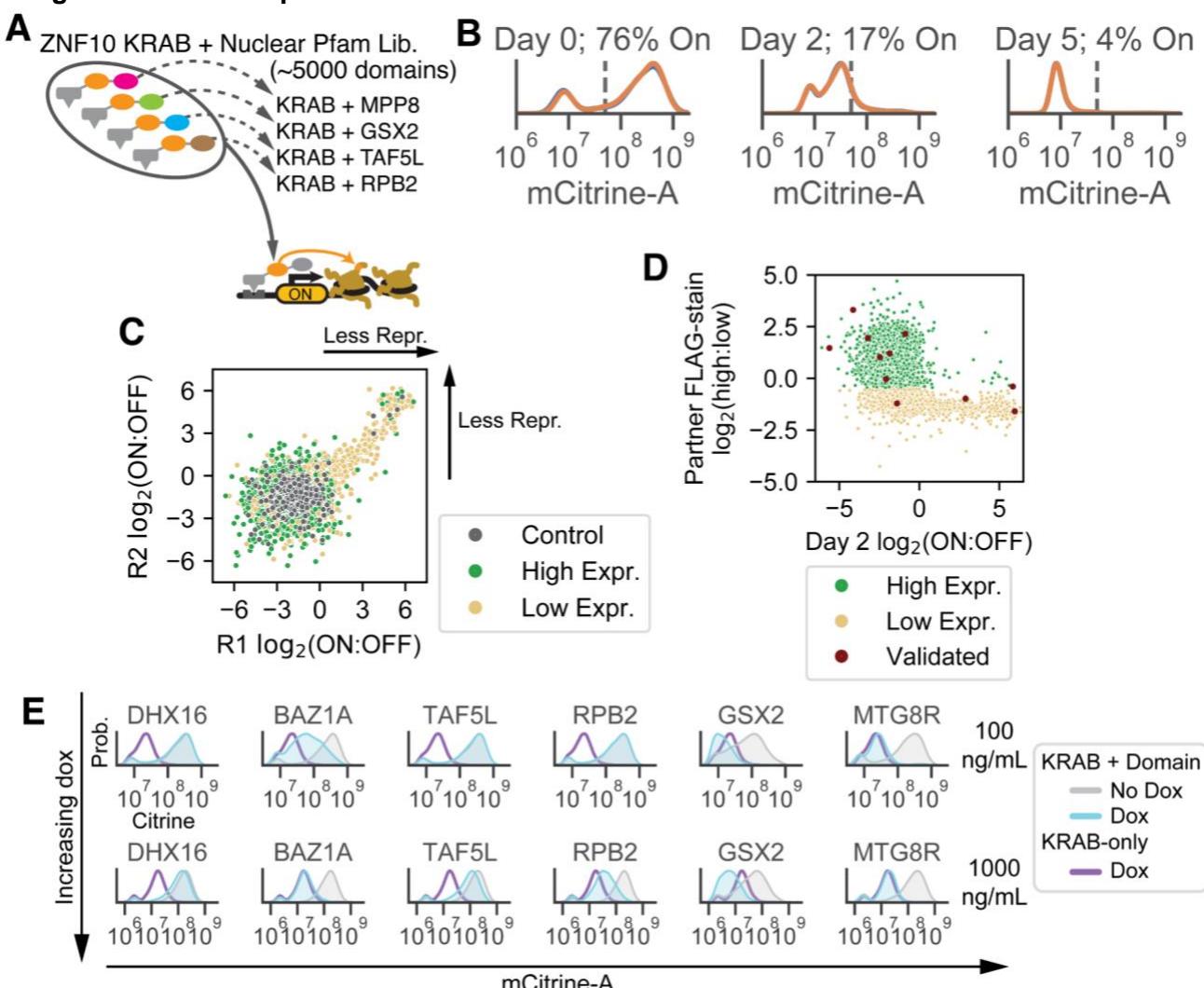


669

- 670 A. Representation of effector-effector interactions measured by the high-throughput screen. Domains in position 1 are  
671 arrayed along the y-axis separated into 4 categories (from bottom to top: repressors, non-hits, activators, and dual-  
672 functional domains), with the stronger duals, activators, and non-hits as measured by control-paired activation  
673 scores on the top of their respective groups. Repressors, in contrast, are ordered by their repression score, with the  
674 strongest repressors at the bottom. Boxes just to the right of each activator domain indicate the control-paired  
675 activation/repression score for each effector. Domains in position 2 are arrayed along the x-axis separated into 4  
676 categories (from left to right: non-hits, activators, dual-functional domains, and repressors), with the stronger  
677 repressors within each category as measured by control-paired activation scores on the right. Boxes just above  
678 each repressor domain indicate the control-paired activation/repression score for each effector. Boxes in the interior  
679 represent combinations of the corresponding activator and repressor, and are colored based on whether the  
680 combination acts as a repressor only (blue), activator only (yellow), both a repressor and an activator (green), or  
681 neither (gray).  
682 B. Correlation of activator-repressor pair activation  $\log_2(\text{ON:OFF})$  scores when recruited at minCMV (to test activation)  
683 between an orientation where the activator is N-terminal to the repressor (x-axis) and an orientation where the  
684 activator is C-terminal to the repressor (y-axis). Red line represents the line of best fit. Dashed lines indicate the  
685 threshold for an effector domain combination to be labeled an activator from **Fig. 1E**.  
686 C. Correlation of activator-repressor pair repression  $\log_2(\text{ON:OFF})$  scores when recruited at pEF (to test repression)  
687 between an orientation where the activator is N-terminal to the repressor (x-axis) and an orientation where the  
688 activator is C-terminal to the repressor (y-axis). Red line represents the line of best fit. Dashed lines indicate the  
689 threshold for an effector domain combination to be labeled an activator from **Fig. 1F**.

690

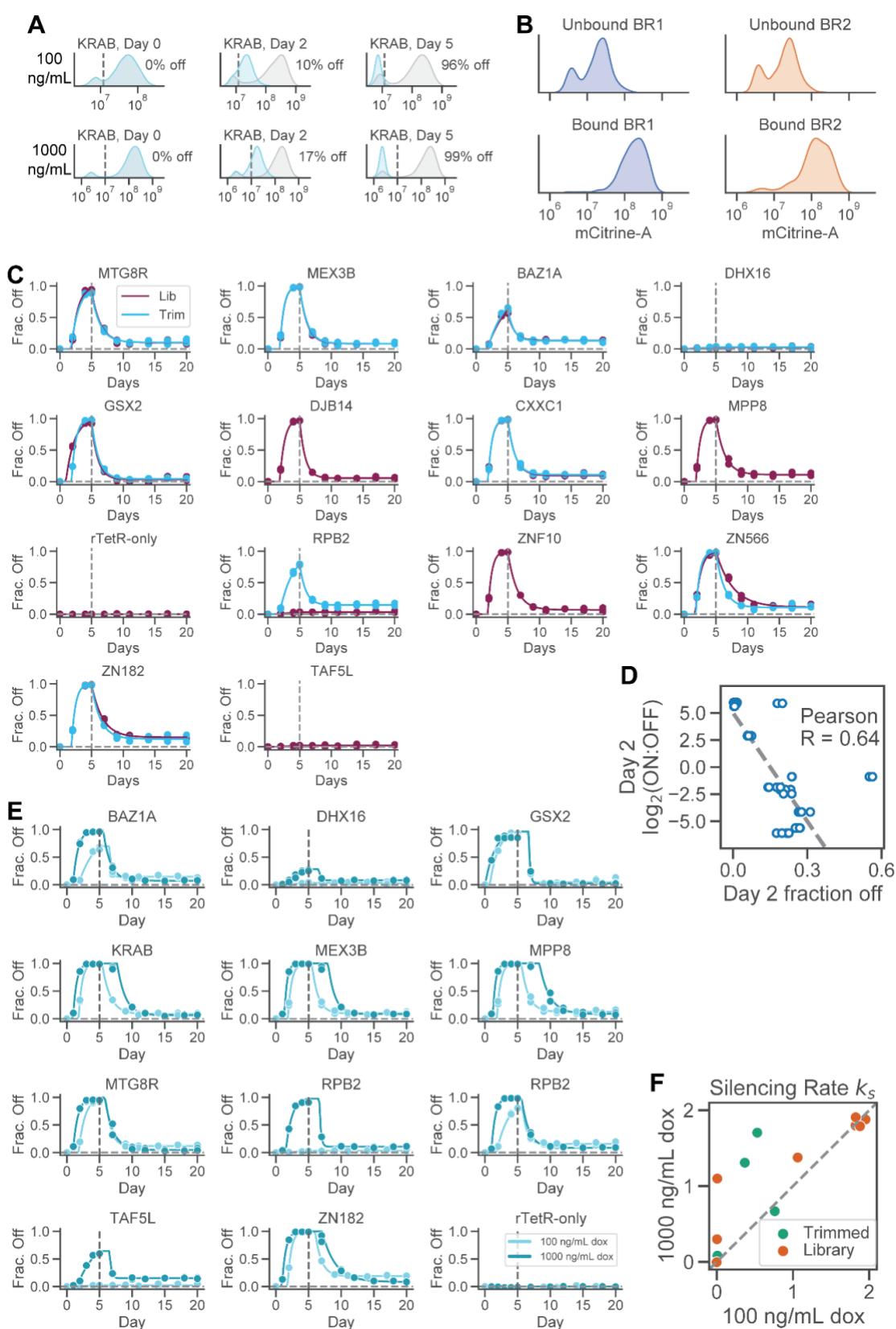
### Figure 5 -- KRAB repression



691

- 692 A. Schematic of KRAB-Pfam screen. A library of ~5000 Pfam-annotated nuclear-localized domains<sup>21</sup> was cloned into  
693 a backbone downstream of an rTetR DNA-binding domain fused to the ZNF10 KRAB repressor and recruited at the  
694 constitutive pEF promoter.  
695 B. Flow cytometry measurements of mCitrine fluorescence (x-axis) at 0, 2, and 5 days after 100ng/ml dox-mediated  
696 recruitment at the pEF promoter. Data shown consists of 2 replicates.  
697 C. Correlation between replicate 1 (x-axis) and replicate 2 (y-axis) of screen log<sub>2</sub>(ON:OFF) scores after 2 days of dox  
698 recruitment for KRAB-domain pairs where domain pairs were previously annotated as stable (green), unstable  
699 (yellow), or negative controls (gray).  
700 D. Correlation between screen log<sub>2</sub>(ON:OFF) scores (x-axis) and prior FLAG-stain measurements of KRAB partner  
701 domain expression (y-axis) for high-expressed domains (green), low-expressed domains (yellow). Domains that  
702 were individually validated via flow cytometry are shown in red.  
703 E. Probability density curves of mCitrine fluorescence for select domains that were individually validated using flow  
704 cytometry. Data shown is after 2 days of recruitment at 100 ng/mL dox (top row) or 1000 ng/mL dox (bottom row)  
705 at the pEF promoter. No dox controls are shown in gray, while recruitment of KRAB alone is shown in purple.

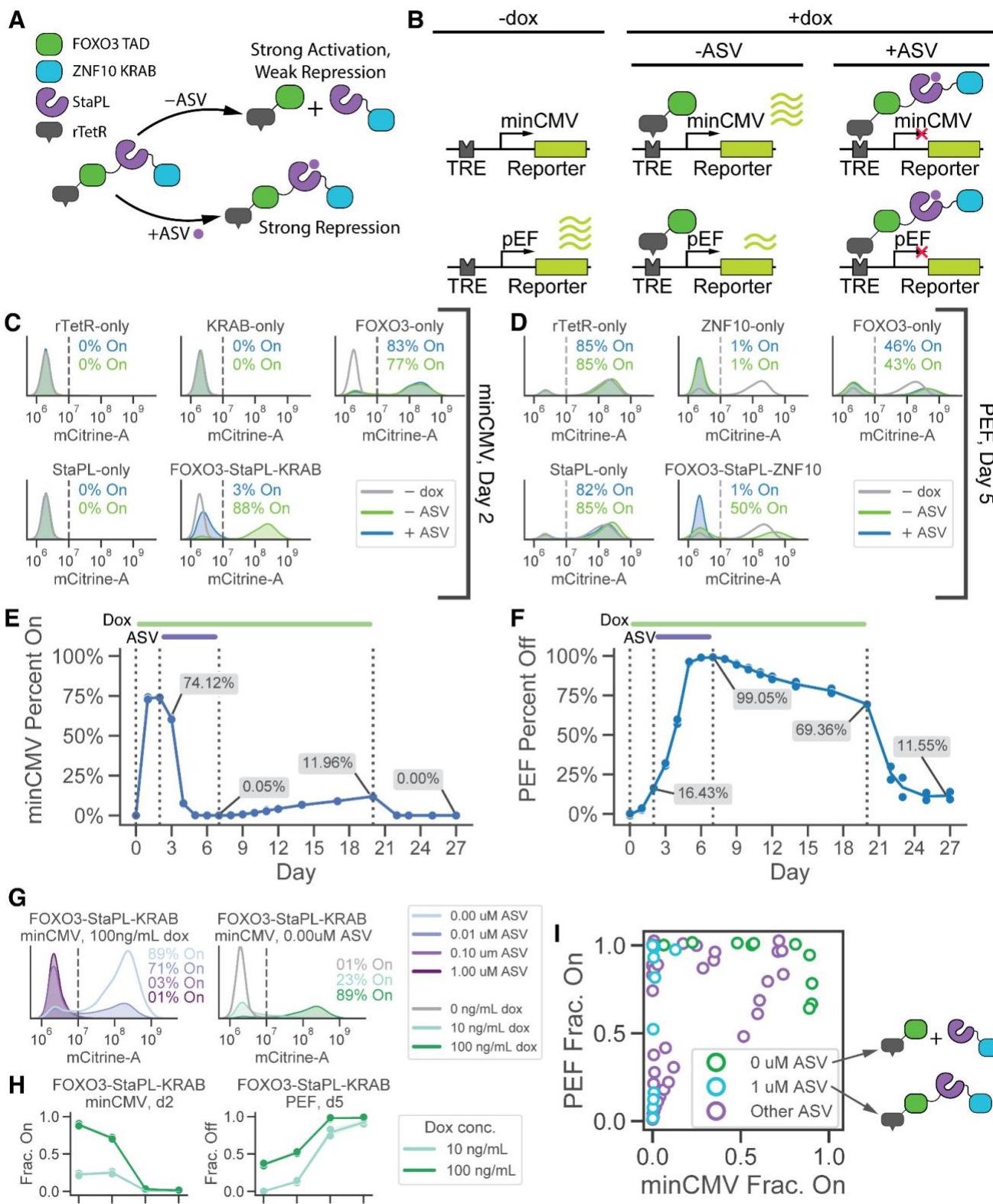
706 Figure S6 – KRAB screen quality control and filtering



- 708 A. Probability density (y-axis) of mCitrine fluorescence levels (x-axis) for KRAB at days 0, 2, and 5 when recruited at  
709 either 100 ng/mL dox (top) or 1000 ng/mL dox (bottom). Vertical dashed lines indicate the threshold for labeling  
710 cells as ON or OFF.
- 711 B. Probability density (y-axis) of mCitrine fluorescence levels (x-axis) for unbound (top) and bound (bottom) fractions  
712 post-magnetic separation for screen replicates 1 (left) and 2 (right) at day 2 after dox addition
- 713 C. Individual recruitment of domains in low throughput to a reporter gene driven by the strong pEF promoter. Effectors  
714 in the library are composed of 80 amino acids surrounding the Pfam-annotated domain (lib, purple line) but were  
715 also tested in trimmed formats comprising only the annotated domain (trim, blue line). Doxycycline was added to  
716 cells for 5 days, and then cells were monitored for gene reactivation 20 days post-dox. Fractions of cells silenced  
717 were normalized with respect to a no-doxycycline control (**Materials and Methods**). Datapoints were fit using the  
718 3-state phenomenological model<sup>10,49</sup>.
- 719 D. Correlation between the fraction of cells silenced at day 2 of recruitment during low-throughput measurements (x-  
720 axis) and the corresponding log<sub>2</sub>(ON:OFF) scores from the KRAB-Pfam screen data (y-axis). Gray line shown is  
721 the line of best fit.
- 722 E. Recruitment of domains at varying strength. KRAB-domain pairs were recruited to the pEF reporter in K562 cells at  
723 a dox concentration of either 100 ng/mL (light blue) or 1000 ng/mL (dark blue) for 5 days. Cells were monitored for  
724 20 days after doxycycline release to measure gene reactivation. Fractions of cells silenced were normalized with  
725 respect to a no-doxycycline control.
- 726 F. Stronger silencing at higher doxycycline doses. Rates of gene silencing computed using the 3-state model are  
727 compared for each domain measured in low-throughput at 100 ng/mL dox (x-axis) and 1000 ng/mL dox (y-axis).  
728 Gray line shown is y=x.

729  
730

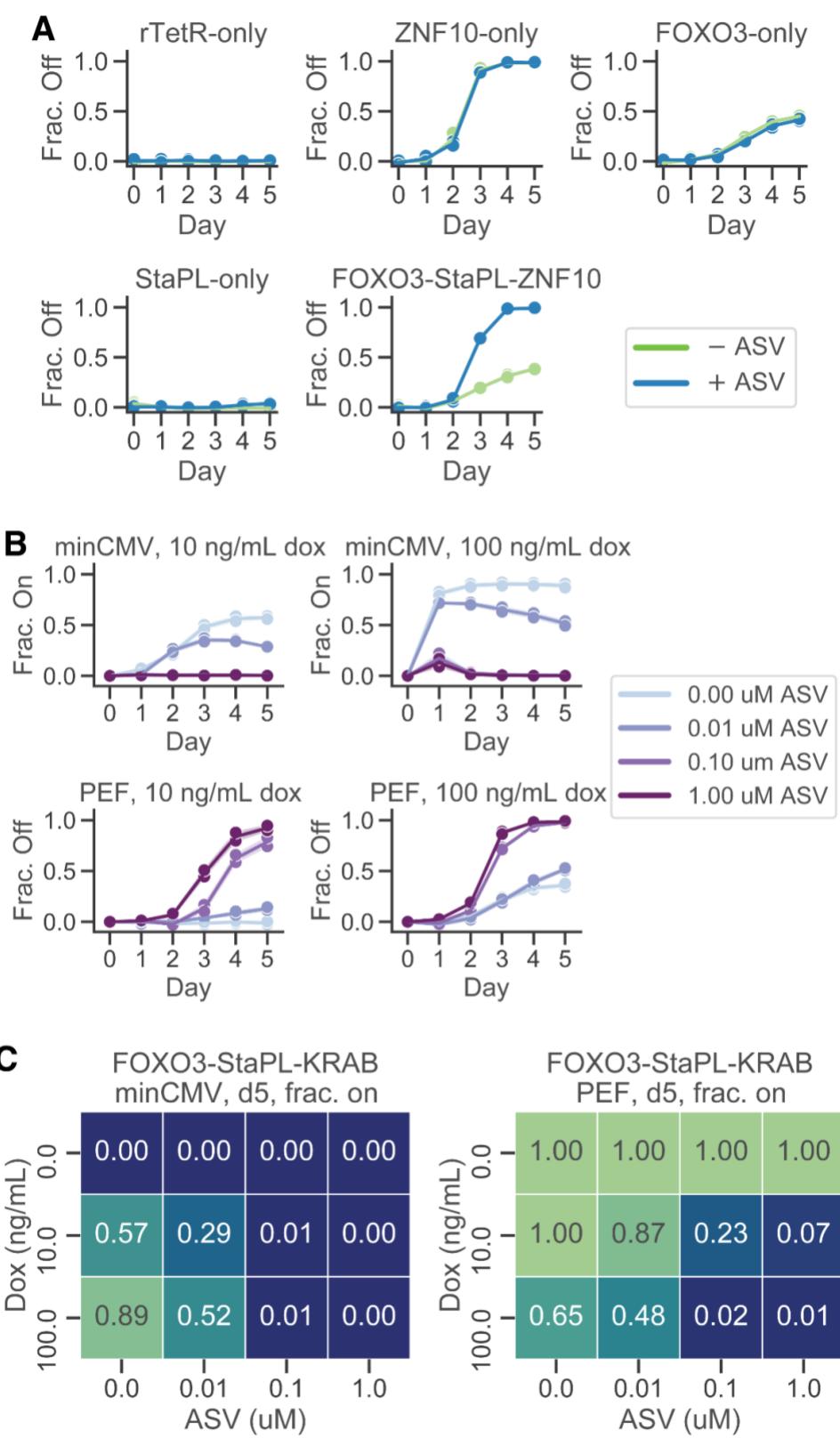
**Figure 6 -- A synthetic TF that can switch from strong repressor to strong activator**



731

- 732 A. Schematic of synthetic transcription factor system. An ASV-inducible StaPL (purple) is fused in between FOXO3's  
733 TAD (green) and ZNF10's KRAB (blue) domains. In the absence of ASV, the StaPL and KRAB are cleaved off,  
734 resulting in only FOXO3's TAD being recruited to the reporter. Adding ASV stabilizes the StaPL domain, resulting  
735 in recruitment of FOXO3's TAD fused to ZNF10's KRAB domain.
- 736 B. Expected dynamics of gene expression at the minCMV promoter (top row) or pEF promoter (bottom row) without  
737 dox (left), with dox and no ASV (middle), and dox and ASV (right).
- 738 C. Histograms of transcriptional activation after 2 days of recruitment to the synthetic reporter driven by the minCMV  
739 promoter in the presence (blue) or absence (green) of ASV, with the 0 ng/mL dox control shown in gray.
- 740 D. Histograms of transcriptional repression after 5 days of recruitment to the synthetic reporter driven by the pEF  
741 promoter in the presence (blue) or absence (green) of ASV, with the 0 ng/mL dox control shown in gray.
- 742 E. Transcriptional activation after ASV washout. K562 cells expressing the minCMV-driven reporter and the FOXO3-  
743 StaPL-KRAB construct were treated with doxycycline during days 0-20 (green line) and ASV during days 2-7 (purple  
744 line). Cells were washed 1x with PBS when removing ASV and doxycycline.
- 745 F. Transcriptional repression after ASV washout. K562 cells expressing the pEF-driven reporter and the FOXO3-  
746 StaPL-KRAB construct were treated with doxycycline during days 0-20 (green line) and ASV during days 2-7 (purple  
747 line). Cells were washed 1x with PBS when removing ASV and doxycycline.
- 748 G. Histograms of transcriptional activation after 2 days of recruitment to the synthetic reporter driven by the minCMV  
749 promoter with either 100 ng/mL dox and varying doses of ASV (purple, left) or 0 uM ASV and varying doses of dox  
750 (green, right).
- 751 H. Fraction of cells activated after 2 days of recruitment to the synthetic reporter driven by the minCMV promoter (y-  
752 axis) for different ASV doses (x-axis), and for varying doses of doxycycline (light vs dark green).
- 753 I. Scatter of minCMV-reporter fraction on (x-axis) versus pEF-reporter fraction on (y-axis) measured across all  
754 FOXO3-StaPL-KRAB experiments. Each point constitutes 1 timepoint, and the color of each point indicates the  
755 ASV dose.

756 Figure S7 – Functional characterization of FOXO3-StaPL-KRAB



- 758 A. Recruitment of rTetR-only, ZNF10-only, FOXO3-only, StaPL-only, and FOXO3-StaPL-KRAB to the pEF promoter  
759 at 1000 ng/mL dox for 5 days. Fraction of silent cells normalized to background (y-axis) is plotted against days of  
760 recruitment (x-axis). Data shown is the average of 2 replicates.  
761 B. Recruitment of FOXO3-StaPL-KRAB to the minCMV (top) or pEF (bottom) promoters for 5 days at different dox  
762 doses (100 ng/mL - left, 1000 ng/mL - right) and 4 different ASV doses. Each line shown is the average of 2  
763 replicates.  
764 C. Fraction of cells activated at minCMV after 5 days of recruitment (left) and repressed at pEF after 5 days of  
765 recruitment (right) at 3 different doxycycline doses (rows) and 4 different ASV doses (columns). Data shown is the  
766 average of 2 replicates.

767 **Supplementary Tables**

768

769 **Supplementary Table 1 – Base Oligo Library**

770 Sequences for domains used in the combinatorial recruitment screen related to Figs. 1-4 are attached in a CSV file.

771

772 **Supplementary Table 2 – Corecruit screen CSV**

773 Domain combinations and their corresponding measurements in the combinatorial recruitment screens for activation are

774 attached in a CSV file.

775

776 **Supplementary Table 3 – KRAB+Pfam screen CSV**

777 Domains fused to KRAB and corresponding measurements in the repression screen related to Fig. 5 are attached in a

778 CSV file.

## 779 Materials & Methods

780

### 781 Library design

782 Library members were chosen from the Nuclear Pfam library described in<sup>21</sup>. A total of 20 negative control domains, 10  
783 randomers and 10 tiles of the DMD protein; 30 activator domains; and 50 repressor domains were chosen for cloning. Of  
784 the 50 repressors, 6 were discovered to have errors introduced during design, and were discarded from downstream  
785 analysis. All domains, including those eventually discarded, are listed in **Supplementary Table 1**. After library assembly,  
786 DNACHisel<sup>77</sup> was used to optimize coding sequences by removing duplicates, 7xC homopolymers, BsmBI and BbsI  
787 restriction sites, and rare codons. Codon usage was matched to human codon prevalence and GC content was restricted  
788 to be between 20% and 75% in any 50-nucleotide window and between 25% and 65% globally.  
789

### 790 Cell culture

791 Cell culture was performed as described in<sup>20</sup>. Briefly, all experiments presented here were carried out in K562 cells (ATCC,  
792 CCL-243, female). Cells were cultured in a controlled humidified incubator at 37C and 5% CO<sub>2</sub>, in RPMI 1640 (Gibco, 11-  
793 875-119) media supplemented with 10% FBS (Omega Scientific, 20014T), and 1% Penicillin-Streptomycin-Glutamine  
794 (Gibco, 10378016). HEK293T-LentiX (Takara Bio, 632180, female) cells, used to produce lentivirus, as described below,  
795 were grown in DMEM (Gibco, 10569069) media supplemented with 10% FBS (Omega Scientific, 20014T) and 1% Penicillin  
796 Streptomycin Glutamine (Gibco, 10378016). minCMV and pEF reporter cell line generation is described in<sup>21</sup>. pEF and  
797 minCMV promoter reporter cell lines were generated by TALEN-mediated homology-directed repair to integrate donor  
798 constructs (pEF promoter: Addgene #161927, minCMV promoter: Addgene #161928) into the AAVS1 locus by  
799 electroporation of K562 cells with 1000 ng of reporter donor plasmid and 500 ng of each TALEN-L (Addgene #35431) and  
800 TALEN-R (Addgene #35432) plasmid (targeting upstream and downstream the intended DNA cleavage site, respectively).  
801 After 7 days, the cells were treated with 1000 ng/mL puromycin antibiotic for 5 days to select for a population where the  
802 donor was stably integrated in the intended locus. Fluorescent reporter expression was measured by flow cytometry.  
803

### 804 Co-recruit cloning

805 Cloning for the combinatorial screen proceeded in two stages; in the first, domain-linker-domain concatenations were  
806 assembled, and in the second, concatenations were placed into a lentiviral backbone. Domains in the N-terminal and C-  
807 terminal position were synthesized as 2 separate oligonucleotide pools (Twist Biosciences). Each pool was PCR amplified  
808 in a clean PCR hood to avoid DNA contamination. Each pool was split into 6x 50 uL reactions that were PCR amplified for  
809 21 cycles with 5 ng template, 1 uL of each 10mM primer, 1uL of Herculase II polymerase (Agilent), 1 uL of DMSO, 1 uL of  
810 10mM dNTPs, and 10 uL of 5x Herculase buffer (Agilent). Reaction mixes were thermocycled at 98C for 3m; then 21x cycles  
811 of 98C for 20s, 61C for 20s, and 72C for 30s; and, finally, 72C for 3m. Reaction products were pooled and gel extracted by  
812 loading a 1% TAE gel, excising the 300 bp band, and purifying using a Zymo Research gel extraction kit. Each of the 2  
813 extraction products (N-terminal and C-terminal) were then amplified for 23x cycles each using the same protocol in order to  
814 generate sufficient DNA for downstream reactions. To generate concatenations, 5 uL of 15 ng/uL product from each product,  
815 N-terminal and C-terminal, were mixed with 4 uL T4 buffer (NEB B0202S), 4 uL BbsI-HF (NEB R3539L, 1 uL T4 ligase (NEB  
816 M0202M), and 1uL of an XTEN linker-encoding DNA amplicon. This linker was encoded with variable nucleotides in codon  
817 wobble positions to permit variable DNA sequence in between domains and prevent lentiviral recombination without varying  
818 the amino acid sequence of the linker. To complete the GoldenGate reaction, the mix was thermocycled for 65x cycles at  
819 37C for 5m, then 16C for 5m, before a series of incubations at 37C for 1h, 65C for 20m, and 16C for 1h. The reaction  
820 product was run on a 1% TAE gel and the 600bp band of domain pair concatenations was extracted with a Zymo gel  
821 extraction kit. The concatenation product was then PCR amplified as before, using the outer primers, for 15x cycles.  
822 Amplification products were then split into 12x GoldenGate reactions, each consisting of 75ng (1.33 uL) BsmBI-v2 digested  
823 pJT039 lentiviral backbone, 10 ng (0.64 uL) concatenation amplicon, 2 uL T4 buffer, 1 uL BsmBI-v2 Golden Gate Assembly  
824 Kit (NEB E1602L), and 15.03 uL nuclease free H2O. Reactions were thermocycled with 65x cycles at 42C for 5m and 16C  
825 for 5m, before a final incubation at 42C for 5m and 70C for 20m. The reactions were then pooled and purified with a QIAgen  
826 MinElute column, eluting in 6 uL ddH2O. 2 uL of eluent was then electroporated into each of 2 tubes of 50 uL Endura

827 electrocompetent cells (Lucigen, Cat#60242-2) as per manufacturer's instructions. Cells were then plated onto 4x large 10"  
828 x 10" LB plates with carbenicillin. After overnight growth colonies were scraped into a collection bottle and plasmid pools  
829 were extracted using a plasmid maxiprep kit (Qiagen #12662). 2 smaller plates of LB with carbenicillin were also prepared  
830 with 1:20 diluted cells to count colonies and confirm transformation efficiency. Approximately 50 colonies were Sanger  
831 sequenced (Quintara) to estimate cloning efficiencies and the proportion of empty backbone plasmids in the pool.  
832

### 833 Co-recruit screen

834 We plated  $15 \times 10^6$  HEK293T cells on each of 10x 15-cm tissue culture plates in 30 mL DMEM, grew them overnight, and  
835 then transfected each with 8 ug of an equimolar mixture of three third-generation packaging plasmids (pMD2.G, psPAX2,  
836 pMDLg/pRRE) and 8 ug of rTetR-concatenation library vectors using 50 mL of polyethylenimine (PEI, Polysciences #23966).  
837 Packaging plasmids were gifts from Didier Trono. Lentivirus was harvested after 48 and 72 hours of incubation and filtered  
838 through a 0.45 mm PVDF filter (Millipore) to remove debris. K562 cells were infected with lentivirus via spinfection for 2  
839 hours, with 2 replicates each spinfected separately. After 48 hours of growth, infected cells were selected with 10 mg/mL  
840 blasticidin (Gibco). Infection and selection efficiency were monitored each day with flow cytometry using a Biorad ZE5 flow  
841 cytometer. Infection coverage was approximately 600x for each replicate, while maintenance coverage was maintained  
842 between 10,000 and 20,000x cells per library element. On day 8 post-infection, cells were treated with 1000 ng/mL  
843 doxycycline (Fisher Scientific) for 2 days for activation or 5 days for repression.  
844

### 845 Co-recruit library prep and sequencing

846 Genomic DNA was extracted with the QIAgen Blood Maxi Kit following the manufacturer's instructions with up to  $1 \times 10^8$   
847 cells per column. Domain sequences were amplified by PCR with primers containing Illumina adapters as extensions. 27-  
848 55x 100uL PCRs were set up on ice, with the genomic DNA available for each experiment dictating the number of reactions.  
849 10 ug of genomic DNA, 0.5 uL of each 100M primer, and 50uL of NEBnext Ultra 2x Master Mix (NEB) was used in each  
850 reaction. Samples were thermocycled at 98C for 3m; 33 cycles of 98C for 10s, 63C for 1m, 72C for 30s; and 72C for 2m.  
851 PCR reactions were pooled and run on a 1% TAE gel, the 600bp band was excised, and DNA was purified using QIAquick  
852 gel extraction kit (Qiagen) and eluted into nonstick tubes (Ambion). Samples were sequenced on an Illumina HiSeq  
853 (2x150bp, Admera Health).  
854

### 855 Computing co-recruit enrichment scores

856 Demultiplexed sequencing reads were provided by Admera. Domain 1 and domain 2 reads were aligned separately using  
857 bowtie2, trimming 30 base pairs from both 5' and 3' ends, and were then each paired read was identified as a particular  
858 library member using a python script iterating over each read pair. The total number of reads for each library member was  
859 then summed up using another python script. Library members were required to have at least 5 ON reads or 5 OFF reads  
860 as well as 50 total reads across both ON and OFF subpopulations to be considered for downstream analysis. For every  
861 qualifying library member, the ON or OFF read count was then set to 5 if it was less than 5. Then, ON read counts were  
862 normalized by dividing each library member's read count by the total number of ON reads, and the same was done for OFF  
863 counts. Overall enrichment scores were then computed by taking the  $\log_2$  of the ratio of normalized ON read counts to  
864 normalized OFF read counts.  
865

### 866 Labeling domains as activators, repressors, and duals

867 We labeled a domain as an activator if the following 2 conditions held for all concatenations including that domain and a  
868 negative control: 1) at least 19% functioned as activators (i.e. for each, activation  $\log_2(\text{ON:OFF}) > -0.0236$ ); 2) the average  
869 activation  $\log_2(\text{ON:OFF})$  for all pairs was greater than or equal to -0.5, or the activation  $\log_2(\text{ON:OFF})$  for all pairs that  
870 functioned as activators was greater than or equal to 0.95. We labeled a domain as a repressor if the following 2 conditions  
871 held for all concatenations including that domain and a negative control: 1) at least 19% functioned as repressors (i.e.  
872 repression  $\log_2(\text{ON:OFF}) < 0.771$ ); 2) the average repression  $\log_2(\text{ON:OFF})$  for all pairs was less than or equal to -0.3, or  
873 the repression  $\log_2(\text{ON:OFF})$  for all pairs that functioned as repressors was less than or equal to 0.20. Domains that

874 functioned as both activators and repressors were labeled as duals. Domains that functioned as neither were labeled as  
875 non-hits.  
876

## 877 KRAB-Nuclear Pfam screen cloning and cell culture

878 An initial backbone was generated by digesting pJT126 (AddGene #161926) with Bsml, then ligating in a sequence  
879 corresponding to the ZNF10 KRAB domain. Then, the nuclear Pfam library was cloned into the downstream GoldenGate  
880 cloning site, and used to produce lentivirus identically as in<sup>21</sup>. K562 cells were then spinfected in 2 separate replicates with  
881 this KRAB+nuclear Pfam library lentivirus at an average MOI of 0.015, corresponding to an infection coverage of 85x per  
882 replicate. 72 hours post-infection 10 mg/mL blasticidin (Gibco) was added. 48 hours after blasticidin addition, 100 ng/mL  
883 doxycycline was added and cells were maintained in both blast + dox for 5 more days. Cells were sampled at day 2 of  
884 doxycycline recruitment for magnetic separation and downstream sequencing. After 5 days of recruitment, cells were spun  
885 down, washed with PBS, and resuspended in dox-free/blast-free RPMI. 5 days after doxycycline release, cells were again  
886 sampled for magnetic separation and downstream sequencing. Cells were maintained at a maintenance coverage of 3,000-  
887 30,000 per library element on average. Library preparation and sequencing was performed identically as in to sequence the  
888 variable effector domain<sup>21</sup>  
889

## 890 Magnetic Separation

891 Cells to be separated were spun down at 300 x g for 5m and media was aspirated. Cells were then resuspended in PBS  
892 (Gibco), spun down again, and PBS was aspirated in order to wash the cells. Dynabeads Protein G were resuspended by  
893 vortexing for 30 seconds. 50mL of blocking buffer was prepared per 2 x 10<sup>8</sup> cells by adding 1 g biotin-free BSA (Sigma  
894 Aldrich) and 200 mL of 0.5 M pH 8.0 EDTA into DPBS (GIBCO), vacuum filtering with a 0.22-mm filter (Millipore), and  
895 keeping on ice. For all screens 60 uL of beads were used per 10 million cells. Magnetic separation was otherwise performed  
896 as in<sup>21</sup>.  
897

## 898 Low-throughput validations

899 Plasmids for each validation were produced using GoldenGate cloning as described in<sup>21</sup>. Then, 750ng plasmid and 750ng  
900 pack mix were mixed in 200uL serum-free OptiMEM (Gibco) along with 2-8uL PEI. Meanwhile, HEK293T cells were plated  
901 in 5mL DMEM supplemented with 10% FBS and 1% PSG in one well of a 6-well plate (Fisher Scientific) for each validation.  
902 The plasmid-pack mix-optiMEM mixture was added to the HEK293T cells to initiate lentivirus production, and virus was  
903 harvested 48 and 72 hours post incubation and filtered through a 0.45mm PDMS filter. For each validation, 2.50 x 10<sup>5</sup> K562  
904 cells were spinfected with 1mL virus for 2 hours, and 10 mg/mL blasticidin was added after 48 hours. After 7 days of  
905 blasticidin selection, cells were spun down and resuspended in RPMI with doxycycline for recruitment.  
906

## 907 Model of transcriptional repression

908 In our model,  $\lambda$  defines the amount of mRNA in Active (ON) cells. We used the experimental data to estimate these  
909 parameters for our synthetic reporter. We directly fit the mean and standard deviation of fluorescence of Silent (OFF) cells  
910 from experimental data. We assumed that a population of cells each containing  $m$  mRNA molecules produced a distribution  
911 of  $\log_{10}$  mCitrine-A fluorescence values centered at  $(m/600) + 6.5$ , and directly fit the standard deviation of this population  
912 from experimental data. We constrained mRNA levels to range between 0 and 1500 molecules per cell. Thus, the probability  
913 of seeing a cell produce a  $\log_{10}$  mCitrine-A fluorescence level of  $c$  given parameters  $\sigma_{ON}$ ,  $\mu_{OFF}$ ,  $\sigma_{OFF}$ , and  $\lambda$  is given by the  
914 equations:

$$915 P(c | Active) = \sum_{mRNA\ levels} P(cit | mRNA) \cdot P(mRNA) = \sum_{n=0}^{1500} Norm_{6.5 + (n/600), \sigma_{ON}}(c) \cdot Poisson_{\lambda}(n)$$
$$916 P(c | Silent) = Norm_{\mu_{OFF}, \sigma_{OFF}}(c)$$
$$917 P(c) = P_{Active} \cdot P(c | Active) + P_{Silent} \cdot P(c | Silent)$$

918 where  $P_{Active}$  and  $P_{Silent}$  are the probability that a given cell is active or silent, respectively. Parameters were fit directly to  
919 data using `scipy.optimize.curve_fit`. Note that  $P_{Active} + P_{Silent} = 1$  and that when  $t \leq t_{lag}$ ,  $P_{Silent} = 0$ .

920

921 To measure silencing, for every day of recruitment, 2000 cells were sampled to ensure equal sampling of  $\log_{10}$  mCitrine-A  
922 fluorescence intensity across recruitment timepoints. The fraction of background silenced cells was directly computed using  
923 the rTetR-only recruitment data, with cells being labeled as Silent if their mCitrine-A fluorescence value was less than  $10^7$ .  
924 Then, parameters  $\mu_{OFF}$  and  $\sigma_{OFF}$  were directly estimated by examining the distribution of cells with log fluorescence values  
925 less than 7.  $\lambda$  was determined by estimating the mean amount of mRNA produced in ON cells at day 0, using the above  
926 mapping between mRNA counts and fluorescence values.  $k_s$  and  $t_{lag}$  were fit by examining the fraction of cells silenced over  
927 time; we assumed that the fraction of Active cells decreased with rate  $k_s$  after  $t_{lag}$  time had elapsed<sup>10,49</sup>. Finally,  $\gamma$  was  
928 determined by fitting a 2-state Gaussian Mixture model to fluorescence data for each day and fitting a line to the locations  
929 of the higher-fluorescence peak; we assumed  $\lambda$  decays linearly after silencing with slope  $\gamma$ .

930

### 931 **Flow cytometry analysis**

932 Flow cytometry analysis was performed in Python using Cytoflow<sup>78</sup>. For all flow experiments, live cells were gated using  
933 Cytoflow's DensityGateOp, keeping 90% of cells, and then thresholded for mCherry+ cells using the ThresholdGateOp. To  
934 determine the fraction of cells that were on, we computed the fraction of cells with  $\log_{10}$  mCitrine-A fluorescence greater  
935 than 7 as measured on a BioRad ZE5 flow cytometer. Normalization of mCitrine levels to the no-dox condition was  
936 performed as follows:  $f_{off,norm} = (f_{off,dox} - f_{off,no\ dox}) \div (1 - f_{off,no\ dox})$  where  $f_{off}$  denotes the fraction of cells off for  
937 any given condition. Cytometry analysis was otherwise performed identically as in<sup>20,21</sup>.

938

### 939 **Data and code availability**

940 All raw NGS data and associated processed data generated in this study will be deposited in the NCBI GEO database  
941 upon publication. All data besides raw NGS data used for this manuscript, and all code, are available at Zenodo at DOI  
942 10.5281/zenodo.7453682.

## 943 Acknowledgements

944 We would like to thank Michaela Hinks, Jennifer Gucwa, Nicole DelRosso, Abby Thurm, David Yao, Priyanka Shrestha,  
945 Zeppelin Cat, Ophelia Hinks, Shivam Verma, Ben Doughty, Mary Frances Gallagher, Kaushik Ragunathan, Zara Weinberg,  
946 and all members of the Bintu lab for helpful conversations and assistance. We would like to thank Anshul Kundaje and Will  
947 Greenleaf for valuable feedback throughout the project. AXM was supported by Stanford University Medical Scientist  
948 Training Program grants T32-GM007365 and T32-GM145402, and a Stanford Bio-X SIGF. JT is supported by the F99/K00  
949 fellowship of the National Institutes of Health (NIH-1F99DK126120-01; NIH-4K00DK126120-03). SAR is supported by an  
950 NSF GRFP (DGE-1656518). MCB is supported by a grant from Stanford ChEM-H and an NIH Director's New Innovator  
951 Award (1DP2HD08406901). This work was supported by BWF CASI (LB), NIH-NIGMS R35M128947 (LB), and NIH-NIGRI  
952 R01HG011866 (LB and MCB).

953

## 954 Author Contributions

955 AXM and LB designed the study with significant intellectual contributions from JT. AXM and JT designed the combinatorial  
956 recruitment library and performed the KRAB-Pfam screen. KS cloned the KRAB-Pfam library. AXM performed the  
957 combinatorial recruitment screen with help from CL. AXM, SA, SAR, and CA all performed individual recruitment assay  
958 experiments. AXM performed data analysis with input from LB, JT, and MCB. AXM and LB wrote the manuscript, with  
959 contributions and feedback from all authors. LB and MCB supervised the project.

960

## 961 CRediT Statement

962 AXM: conceptualization, methodology, software, formal analysis, investigation, data curation, writing - original draft, writing  
963 - review & editing, visualization. JT: conceptualization, methodology, investigation, writing - review & editing. SA:  
964 investigation, writing - review & editing. SR: investigation, writing - review & editing. CA: investigation, writing - review &  
965 editing. CL: investigation, writing - review & editing. MCB: conceptualization, resources, writing - review & editing,  
966 supervision, project administration. LB: conceptualization, methodology, resources, writing - review & editing, supervision,  
967 project administration, funding acquisition.

968

## 969 Competing Interests

970 LB, MCB, and JT acknowledge outside interest in Stylus Medicine. All other authors declare that they have no known  
971 competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

972

## 973 Additional Information

974 Supplementary information is available for this paper. Correspondence and requests for materials should be addressed to  
975 [lbintu@stanford.edu](mailto:lbintu@stanford.edu).

## 976 References

- 977 1. Frieteze, S., and Farnham, P.J. (2011). Transcription Factor Effector Domains. In A Handbook of Transcription  
978 Factors, T. R. Hughes, ed. (Springer Netherlands), pp. 261–277. 10.1007/978-90-481-9069-0\_12.
- 979 2. Soto, L.F., Li, Z., Santoso, C.S., Berenson, A., Ho, I., Shen, V.X., Yuan, S., and Fuxman Bass, J.I. (2021).  
980 Compendium of human transcription factor effector domains. *Mol. Cell.* 10.1016/j.molcel.2021.11.007.
- 981 3. Lambert, S.A., Jolma, A., Campitelli, L.F., Das, P.K., Yin, Y., Albu, M., Chen, X., Taipale, J., Hughes, T.R., and  
982 Weirauch, M.T. (2018). The Human Transcription Factors. *Cell* 172, 650–665. 10.1016/j.cell.2018.01.029.
- 983 4. Stegmaier, P., Kel, A.E., and Wingender, E. (2004). Systematic DNA-binding domain classification of transcription  
984 factors. *Genome Inform.* 15, 276–286.
- 985 5. Harrison, S.C. (1991). A structural taxonomy of DNA-binding domains. *Nature* 353, 715–719. 10.1038/353715a0.
- 986 6. Hathaway, N.A., Bell, O., Hodges, C., Miller, E.L., Neel, D.S., and Crabtree, G.R. (2012). Dynamics and memory of  
987 heterochromatin in living cells. *Cell* 149, 1447–1460. 10.1016/j.cell.2012.03.052.
- 988 7. Keung, A.J., Bashor, C.J., Kiriakov, S., Collins, J.J., and Khalil, A.S. (2014). Using targeted chromatin regulators to  
989 engineer combinatorial and spatial transcriptional regulation. *Cell* 158, 110–120. 10.1016/j.cell.2014.04.047.
- 990 8. Kungulovski, G., Nunna, S., Thomas, M., Zanger, U.M., Reinhardt, R., and Jeltsch, A. (2015). Targeted epigenome  
991 editing of an endogenous locus with chromatin modifiers is not stably maintained. *Epigenetics Chromatin* 8, 12.  
992 10.1186/s13072-015-0002-z.
- 993 9. Amabile, A., Migliara, A., Capasso, P., Biffi, M., Cittaro, D., Naldini, L., and Lombardo, A. (2016). Inheritable  
994 Silencing of Endogenous Genes by Hit-and-Run Targeted Epigenetic Editing. *Cell* 167, 219–232.e14.  
995 10.1016/j.cell.2016.09.006.
- 996 10. Bintu, L., Yong, J., Antebi, Y.E., McCue, K., Kazuki, Y., Uno, N., Oshimura, M., and Elowitz, M.B. (2016). Dynamics  
997 of epigenetic regulation at the single cell level. *Science* 351, 720–724.
- 998 11. Tycko, J., Van, M.V., Elowitz, M.B., and Bintu, L. (2017). Advancing towards a global mammalian gene regulation  
999 model through single-cell analysis and synthetic biology. *Current Opinion in Biomedical Engineering* 4, 174–193.  
1000 10.1016/j.cobme.2017.10.011.
- 1001 12. O'Geen, H., Ren, C., Nicolet, C.M., Perez, A.A., Halmai, J., Le, V.M., MacKay, J.P., Farnham, P.J., and Segal, D.J.  
1002 (2017). DCas9-based epigenome editing suggests acquisition of histone methylation is not sufficient for target gene  
1003 repression. *Nucleic Acids Res.* 45, 9901–9916. 10.1093/nar/gkx578.
- 1004 13. Sanborn, A.L., Yeh, B.T., Feigerle, J.T., Hao, C.V., Townshend, R.J., Lieberman Aiden, E., Dror, R.O., and  
1005 Kornberg, R.D. (2021). Simple biochemical features underlie transcriptional activation domain diversity and dynamic,  
1006 fuzzy binding to Mediator. *eLife* 10. 10.7554/eLife.68068.
- 1007 14. Lee, J.B., Caywood, L.M., Lo, J.Y., Levering, N., and Keung, A.J. (2021). Mapping the dynamic transfer functions of  
1008 eukaryotic gene regulation. *Cell Syst* 12, 1079–1093.e6. 10.1016/j.cels.2021.08.003.
- 1009 15. Chasman, D.I., Leatherwood, J., Carey, M., Ptashne, M., and Kornberg, R.D. (1989). Activation of yeast polymerase  
1010 II transcription by herpesvirus VP16 and GAL4 derivatives in vitro. *Mol. Cell. Biol.* 9, 4746–4749.  
1011 10.1128/mcb.9.11.4746-4749.1989.
- 1012 16. Furth, N., and Shema, E. (2022). It's all in the combination: decoding the epigenome for cancer research and  
1013 diagnostics. *Curr. Opin. Genet. Dev.* 73, 101899. 10.1016/j.gde.2022.101899.
- 1014 17. Wolberger, C. (1998). Combinatorial transcription factors. *Curr. Opin. Genet. Dev.* 8, 552–559. 10.1016/s0959-  
1015 437x(98)80010-5.
- 1016 18. Ravasi, T., Suzuki, H., Cannistraci, C.V., Katayama, S., Bajic, V.B., Tan, K., Akalin, A., Schmeier, S., Kanamori-  
1017 Katayama, M., Bertin, N., et al. (2010). An atlas of combinatorial transcriptional regulation in mouse and man. *Cell*  
1018 140, 744–752. 10.1016/j.cell.2010.01.044.
- 1019 19. Reiter, F., Wienerroither, S., and Stark, A. (2017). Combinatorial function of transcription factors and cofactors. *Curr.*  
1020 *Opin. Genet. Dev.* 43, 73–81. 10.1016/j.gde.2016.12.007.
- 1021 20. DelRosso, N., Tycko, J., Suzuki, P., Andrews, C., Aradhana, Mukund, A., Liogson, I., Ludwig, C., Spees, K.,  
1022 Fordyce, P., et al. (2022). Large-scale mapping and systematic mutagenesis of human transcriptional effector  
1023 domains. *bioRxiv*, 2022.08.26.505496. 10.1101/2022.08.26.505496.
- 1024 21. Tycko, J., DelRosso, N., Hess, G.T., Aradhana, Banerjee, A., Mukund, A., Van, M.V., Ego, B.K., Yao, D., Spees, K.,  
1025 et al. (2020). High-Throughput Discovery and Characterization of Human Transcriptional Effectors. *Cell* 183, 2020–  
1026 2035.e16. 10.1016/j.cell.2020.11.024.
- 1027 22. Losson, R., and Nielsen, A.L. (2010). The NIZP1 KRAB and C2HR domains cross-talk for transcriptional regulation.

- 1028 1 Biochimica et Biophysica Acta - Gene Regulatory Mechanisms 1799, 463–468. 10.1016/j.bbagr.2010.02.003.
- 1029 23 Beerli, R.R., Segal, D.J., Dreier, B., and Barbas, C.F., 3rd (1998). Toward controlling gene expression at will: 1030 specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular 1031 building blocks. Proc. Natl. Acad. Sci. U. S. A. 95, 14628–14633. 10.1073/pnas.95.25.14628.
- 1032 24 Chavez, A., Scheiman, J., Vora, S., Pruitt, B.W., Tuttle, M., P R Iyer, E., Lin, S., Kiani, S., Guzman, C.D., Wiegand, 1033 D.J., et al. (2015). Highly efficient Cas9-mediated transcriptional programming. Nat. Methods 12, 326–328. 10.1038/nmeth.3312.
- 1035 25 Nuñez, J.K., Chen, J., Pommier, G.C., Cogan, J.Z., Replogle, J.M., Adriaens, C., Ramadoss, G.N., Shi, Q., Hung, 1036 K.L., Samelson, A.J., et al. (2021). Genome-wide programmable transcriptional memory by CRISPR-based 1037 epigenome editing. Cell 184, 2503–2519.e17. 10.1016/j.cell.2021.03.025.
- 1038 26 Thakore, P.I., Black, J.B., Hilton, I.B., and Gersbach, C.A. (2016). Editing the epigenome: technologies for 1039 programmable transcription and epigenetic modulation. Nature Methods. 10.1038/nmeth.3733.
- 1040 27 Holtzman, L., and Gersbach, C.A. (2018). Editing the Epigenome: Reshaping the Genomic Landscape. Annu. Rev. 1041 Genomics Hum. Genet. 19, 43–71. 10.1146/annurev-genom-083117-021632.
- 1042 28 Black, J.B., and Gersbach, C.A. (2018). Synthetic transcription factors for cell fate reprogramming. Curr. Opin. 1043 Genet. Dev. 52, 13–21. 10.1016/j.gde.2018.05.001.
- 1044 29 Elkon, R., and Agami, R. (2017). Characterization of noncoding regulatory DNA in the human genome. Nat. 1045 Biotechnol. 35, 732–746. 10.1038/nbt.3863.
- 1046 30 Erijman, A., Kozlowski, L., Sohrabi-Jahromi, S., Fishburn, J., Warfield, L., Schreiber, J., Noble, W.S., Söding, J., and 1047 Hahn, S. (2020). A High-Throughput Screen for Transcription Activation Domains Reveals Their Sequence Features 1048 and Permits Prediction by Deep Learning. Molecular Cell 79, 1066. 10.1016/j.molcel.2020.08.013.
- 1049 31 Staller, M.V., Holehouse, A.S., Swain-Lenz, D., Das, R.K., Pappu, R.V., and Cohen, B.A. (2018). A High-Throughput 1050 Mutational Scan of an Intrinsically Disordered Acidic Transcriptional Activation Domain. Cell Syst 6, 444–455.e6. 10.1016/j.cels.2018.01.015.
- 1052 32 Arnold, C.D., Nemčko, F., Woodfin, A.R., Wienerroither, S., Vlasova, A., Schleiffer, A., Pagani, M., Rath, M., and 1053 Stark, A. (2018). A high-throughput method to identify trans-activation domains within transcription factor sequences. 1054 EMBO J. 37, e98896. 10.15252/embj.201798896.
- 1055 33 Alerasool, N., Leng, H., Lin, Z.-Y., Gingras, A.-C., and Taipale, M. (2022). Identification and functional 1056 characterization of transcriptional activators in human cells. Mol. Cell 82, 677–695.e7. 10.1016/j.molcel.2021.12.008.
- 1057 34 Staller, M.V., Ramirez, E., Kotha, S.R., Holehouse, A.S., Pappu, R.V., and Cohen, B.A. (2022). Directed mutational 1058 scanning reveals a balance between acidic and hydrophobic residues in strong human activation domains. Cell Syst 1059 13, 334–345.e5. 10.1016/j.cels.2022.01.002.
- 1060 35 Gao, Y., Xiong, X., Wong, S., Charles, E.J., Lim, W.A., and Qi, L.S. (2016). Complex transcriptional modulation with 1061 orthogonal and inducible dCas9 regulators. Nat. Methods 13, 1043–1049. 10.1038/nmeth.4042.
- 1062 36 Cura, V., and Cavarelli, J. (2021). Structure, Activity and Function of the PRMT2 Protein Arginine Methyltransferase. 1063 Life 11. 10.3390/life11111263.
- 1064 37 Pak, M.L., Lakowski, T.M., Thomas, D., Vhuiyan, M.I., Hüsecken, K., and Frankel, A. (2011). A protein arginine N- 1065 methyltransferase 1 (PRMT1) and 2 heteromeric interaction increases PRMT1 enzymatic activity. Biochemistry 50, 1066 8226–8240. 10.1021/bi200644c.
- 1067 38 Hou, W., Nemitz, S., Schopper, S., Nielsen, M.L., Kessels, M.M., and Qualmann, B. (2018). Arginine Methylation by 1068 PRMT2 Controls the Functions of the Actin Nucleator Cobl. Dev. Cell 45, 262–275.e8. 10.1016/j.devcel.2018.03.007.
- 1069 39 Vhuiyan, M.I., Pak, M.L., Park, M.A., Thomas, D., Lakowski, T.M., Chalfant, C.E., and Frankel, A. (2017). PRMT2 1070 interacts with splicing factors and regulates the alternative splicing of BCL-X. J. Biochem. 162, 17–25. 10.1093/jb/mvw102.
- 1071 40 Kremerskothen, J., Plaas, C., Büther, K., Finger, I., Veltel, S., Matanis, T., Liedtke, T., and Barnekow, A. (2003). 1072 Characterization of KIBRA, a novel WW domain-containing protein. Biochem. Biophys. Res. Commun. 300, 862– 1073 867. 10.1016/s0006-291x(02)02945-5.
- 1074 41 Xiao, L., Chen, Y., Ji, M., and Dong, J. (2011). KIBRA regulates Hippo signaling activity via interactions with large 1075 tumor suppressor kinases. J. Biol. Chem. 286, 7788–7796. 10.1074/jbc.M110.173468.
- 1076 42 Gordon, W.R., Vardar-Ulu, D., Histen, G., Sanchez-Irizarry, C., Aster, J.C., and Blacklow, S.C. (2007). Structural 1077 basis for autoinhibition of Notch. Nat. Struct. Mol. Biol. 14, 295–300. 10.1038/nsmb1227.
- 1078 43
- 1079 44

- 1080 43. Bray, S.J. (2016). Notch signalling in context. *Nat. Rev. Mol. Cell Biol.* 17, 722–735. 10.1038/nrm.2016.94.
- 1081 44. Raj, A., Peskin, C.S., Tranchina, D., Vargas, D.Y., and Tyagi, S. (2006). Stochastic mRNA synthesis in mammalian  
1082 cells. *PLoS Biol.* 4, e309. 10.1371/journal.pbio.0040309.
- 1083 45. Friedman, N., Cai, L., and Xie, X.S. (2006). Linking stochastic dynamics to population distribution: an analytical  
1084 framework of gene expression. *Phys. Rev. Lett.* 97, 168302. 10.1103/PhysRevLett.97.168302.
- 1085 46. Dar, R.D., Razooky, B.S., Singh, A., Trimeloni, T.V., McCollum, J.M., Cox, C.D., Simpson, M.L., and Weinberger,  
1086 L.S. (2012). Transcriptional burst frequency and burst size are equally modulated across the human genome. *Proc.  
1087 Natl. Acad. Sci. U. S. A.* 109, 17454–17459. 10.1073/pnas.1213530109.
- 1088 47. Peccoud, J., and Ycart, B. (1995). Markovian modeling of gene-product synthesis. *Theoretical Population Biology*  
1089 48, 222–234. 10.1006/tpbi.1995.1027.
- 1090 49. Hao, N., and O'Shea, E.K. (2012). Signal-dependent dynamics of transcription factor translocation controls gene  
1091 expression. *Nat. Struct. Mol. Biol.* 19, 31–40. 10.1038/nsmb.2192.
- 1092 50. Mukund, A., and Bintu, L. (2022). Temporal signaling, population control, and information processing through  
1093 chromatin-mediated gene regulation. *J. Theor. Biol.* 535, 110977. 10.1016/j.jtbi.2021.110977.
- 1094 51. Martinez-Corral, R., Park, M., Biette, K., Friedrich, D., Scholes, C., Khalil, A.S., Gunawardena, J., and DePace, A.H.  
1095 (2021). Transcriptional kinetic synergy: a complex landscape revealed by integrating modelling and synthetic  
1096 biology. *bioRxiv*, 2020.08.31.276261. 10.1101/2020.08.31.276261.
- 1097 52. Park, H., Wahl, M.I., Afar, D.E., Turck, C.W., Rawlings, D.J., Tam, C., Scharenberg, A.M., Kinet, J.P., and Witte,  
1098 O.N. (1996). Regulation of Btk function by a major autophosphorylation site within the SH3 domain. *Immunity* 4,  
1099 515–525. 10.1016/s1074-7613(00)80417-3.
- 1100 53. Mohamed, A.J., Vargas, L., Nore, B.F., Backesjo, C.M., Christensson, B., and Smith, C.I. (2000). Nucleocytoplasmic  
1101 shuttling of Bruton's tyrosine kinase. *J. Biol. Chem.* 275, 40614–40619. 10.1074/jbc.M006952200.
- 1102 54. Lee, Y.-T., Ayoub, A., Park, S.-H., Sha, L., Xu, J., Mao, F., Zheng, W., Zhang, Y., Cho, U.-S., and Dou, Y. (2021).  
1103 Mechanism for DPY30 and ASH2L intrinsically disordered regions to modulate the MLL/SET1 activity on chromatin.  
1104 *Nat. Commun.* 12, 2953. 10.1038/s41467-021-23268-9.
- 1105 55. Tando, T., Ishizaka, A., Watanabe, H., Ito, T., Iida, S., Haraguchi, T., Mizutani, T., Izumi, T., Isobe, T., Akiyama, T.,  
1106 et al. (2010). Requiem protein links RelB/p52 and the Brm-type SWI/SNF complex in a noncanonical NF-κB  
1107 pathway. *J. Biol. Chem.* 285, 21951–21960.
- 1108 56. Kim, G.-D., Ni, J., Kelesoglu, N., Roberts, R.J., and Pradhan, S. (2002). Co-operation and communication between  
1109 the human maintenance and de novo DNA (cytosine-5) methyltransferases. *EMBO J.* 21, 4183–4195.  
1110 10.1093/emboj/cdf401.
- 1111 57. Larson, M.H., Gilbert, L.A., Wang, X., Lim, W.A., Weissman, J.S., and Qi, L.S. (2013). CRISPR interference  
1112 (CRISPRi) for sequence-specific control of gene expression. *Nat. Protoc.* 8, 2180–2196. 10.1038/nprot.2013.132.
- 1113 58. Jacobs, C.L., Badiee, R.K., and Lin, M.Z. (2018). StaPLs: versatile genetically encoded modules for engineering  
1114 drug-inducible proteins. *Nat. Methods* 15, 523–526. 10.1038/s41592-018-0041-z.
- 1115 59. Reményi, A., Schöler, H.R., and Wilmanns, M. (2004). Combinatorial control of gene expression. *Nat. Struct. Mol.  
1116 Biol.* 11, 812–815. 10.1038/nsmb820.
- 1117 60. Zeitlinger, J. (2020). Seven myths of how transcription factors read the cis-regulatory code. *Curr Opin Syst Biol* 23,  
1118 22–31. 10.1016/j.coisb.2020.08.002.
- 1119 61. Charest, J., Daniele, T., Wang, J., Bykov, A., Mandlbauer, A., Asparuhova, M., Röhssner, J., Gutiérrez-Pérez, P., and  
1120 Cochella, L. (2020). Combinatorial Action of Temporally Segregated Transcription Factors. *Dev. Cell* 55, 483–  
1121 499.e7. 10.1016/j.devcel.2020.09.002.
- 1122 62. Bashor, C.J., Patel, N., Choubey, S., Beyzavi, A., Kondev, J., Collins, J.J., and Khalil, A.S. (2019). Complex signal  
1123 processing in synthetic gene circuits using cooperative regulatory assemblies. *Science* 364, 593–597.  
1124 10.1126/science.aau8287.
- 1125 63. Weinberg, Z.Y., Hilburger, C.E., Kim, M., Cao, L., Khalid, M., Elmes, S., Diwanji, D., Hernandez, E., Lopez, J.,  
1126 Schaefer, K., et al. (2021). Sentinel cells enable genetic detection of SARS-CoV-2 Spike protein. *bioRxiv*.  
1127 10.1101/2021.04.20.440678.
- 1128 64. Zhu, I., Liu, R., Garcia, J.M., Hyrenius-Wittsten, A., Piraner, D.I., Alavi, J., Israni, D.V., Liu, B., Khalil, A.S., and  
1129 Roybal, K.T. (2022). Modular design of synthetic receptors for programmed gene regulation in cell therapies. *Cell*  
1130 1131

- 1132 185, 1431–1443.e16. 10.1016/j.cell.2022.03.023.
- 1133 65. Lee, S., Khalil, A.S., and Wong, W.W. (2022). Recent progress of gene circuit designs in immune cell therapies. *Cell Syst* 13, 864–873. 10.1016/j.cels.2022.09.006.
- 1134 66. Konermann, S., Brigham, M.D., Trevino, A.E., Joung, J., Abudayyeh, O.O., Barcena, C., Hsu, P.D., Habib, N., Gootenberg, J.S., Nishimasu, H., et al. (2015). Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* 517, 583–588. 10.1038/nature14136.
- 1135 67. Schüle, R., Rangarajan, P., Kliewer, S., Ransone, L.J., Bolado, J., Yang, N., Verma, I.M., and Evans, R.M. (1990). Functional antagonism between oncoprotein c-Jun and the glucocorticoid receptor. *Cell* 62, 1217–1226. 10.1016/0092-8674(90)90397-w.
- 1136 68. Gill, G., and Ptashne, M. (1988). Negative effect of the transcriptional activator GAL4. *Nature* 334, 721–724. 10.1038/334721a0.
- 1137 69. Al-Radhawi, M.A., Ali Al-Radhawi, M., Del Vecchio, D., and Sontag, E.D. (2019). Multi-modality in gene regulatory networks with slow promoter kinetics. *PLOS Computational Biology* 15, e1006784. 10.1371/journal.pcbi.1006784.
- 1138 70. Mamrak, N.E., Alerasool, N., Griffith, D., Holehouse, A.S., Taipale, M., and Lionnet, T. (2022). The kinetic landscape of human transcription factors. *bioRxiv*, 2022.06.01.494187. 10.1101/2022.06.01.494187.
- 1139 71. Lim, W.A. (2022). The emerging era of cell engineering: Harnessing the modularity of cells to program complex biological function. *Science* 378, 848–852. 10.1126/science.add9665.
- 1140 72. Majello, B., De Luca, P., and Lania, L. (1997). Sp3 is a bifunctional transcription regulator with modular independent activation and repression domains. *J. Biol. Chem.* 272, 4021–4026. 10.1074/jbc.272.7.4021.
- 1141 73. Jacobs, J., Pagani, M., Wenzl, C., and Stark, A. (2022). Widespread regulatory specificities between transcriptional corepressors and enhancers in *Drosophila*. *bioRxiv*, 2022.11.07.515017. 10.1101/2022.11.07.515017.
- 1142 74. Shakiba, N., Jones, R.D., Weiss, R., and Del Vecchio, D. (2021). Context-aware synthetic biology by controller design: Engineering the mammalian cell. *Cell Syst* 12, 561–592. 10.1016/j.cels.2021.05.011.
- 1143 75. Dods, G., Gómez-Schiavon, M., El-Samad, H., and Ng, A.H. (2020). Accurate prediction of genetic circuit behavior requires multidimensional characterization of parts. *bioRxiv*, 2020.05.30.122077. 10.1101/2020.05.30.122077.
- 1144 76. Baetica, A.-A., Westbrook, A., and El-Samad, H. (2019). Control theoretical concepts for synthetic and systems biology. *Current Opinion in Systems Biology* 14, 50–57. 10.1016/j.coisb.2019.02.010.
- 1145 77. Zulkower, V., and Rosser, S. (2020). DNA Chisel, a versatile sequence optimizer. *Bioinformatics* 36, 4508–4509. 10.1093/bioinformatics/btaa558.
- 1146 78. Teague, B. (2022). Cytoflow: A Python Toolbox for Flow Cytometry. *bioRxiv*, 2022.07.22.501078. 10.1101/2022.07.22.501078.
- 1147 1161
- 1148 1162