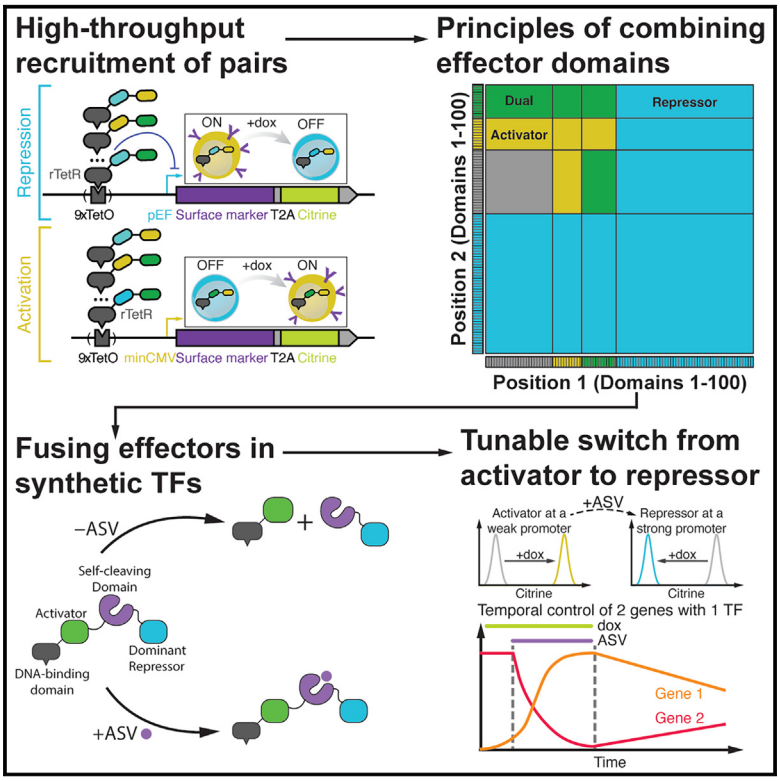
Article

Cell Systems

High-throughput functional characterization of combinations of transcriptional activators and repressors

Graphical abstract



Highlights

- We screen 8,400 effector domain pairs for transcriptional activation and repression
- We find synergy between weak activators and antagonism between strong activators
- We find that repressors tend to dominate activators when co-recruited
- We build a TF that changes from activator to repressor upon adding a small molecule

Authors

Adi X. Mukund, Josh Tycko, Sage J. Allen, ..., Kaitlyn Spees, Michael C. Bassik, Lacramioara Bintu

Correspondence

lbintu@stanford.edu

In brief

Mukund et al. measure transcriptional activation and repression for thousands of transcriptional effector domain pairs using a high-throughput reporter assay in human cells. This study helps outline the basic principles of how effector domains can be combined to build novel synthetic biology tools for precise and flexible gene regulation.

Mukund et al., 2023, Cell Systems *14*, 1–18
September 20, 2023
https://doi.org/10.1016/j.cels.2023.07.001

CellPress

Cell Systems

CellPress

Article

High-throughput functional characterization of combinations of transcriptional activators and repressors

Adi X. Mukund,[1] Josh Tycko,[2] Sage J. Allen,[3] Stephanie A. Robinson,[4] Cecelia Andrews,[5] Joydeb Sinha,[6] Connor H. Ludwig,[3] Kaitlyn Spees,[2] Michael C. Bassik,[2] and Lacramioara Bintu[3,7,*]

[1]Biophysics Program, Stanford University, Stanford, CA 94305, USA
[2]Department of Genetics, Stanford University, Stanford, CA 94305, USA
[3]Department of Bioengineering, Stanford University, Stanford, CA 94305, USA
[4]Department of Chemistry, Stanford University, Stanford, CA 94305, USA
[5]Department of Developmental Biology, Stanford University, Stanford, CA 94305, USA
[6]Department of Chemical and Systems Biology, Stanford University, Stanford, CA 94305, USA
[7]Lead contact
*Correspondence: lbintu@stanford.edu
https://doi.org/10.1016/j.cels.2023.07.001

SUMMARY

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

INTRODUCTION

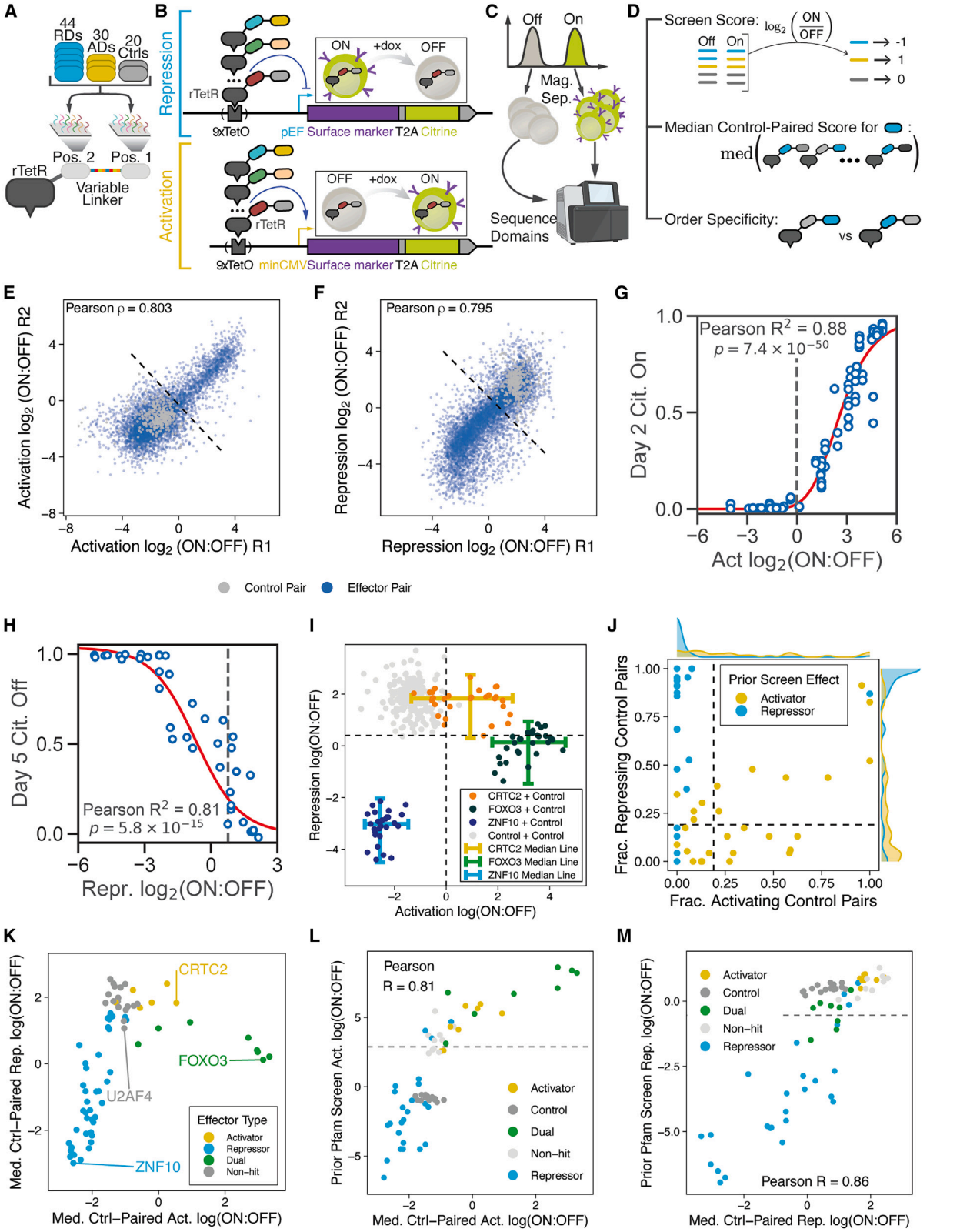
Transcription factors (TFs) and chromatin regulators (CRs) contain short effector domains that can act as repressors or activators when recruited at a target gene.[1–5] Site-specific recruitment assays of effector domains and full-length TFs at reporter genes have long been used to understand their effects on gene expression and develop better tools for gene regulation.[6–15]

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

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

In order to systematically test combinations of effector domains, we need high-throughput methods: even testing all possible combinations resulting from pairing 100 effector domains requires 10,000 measurements. Recently, pooled screens have been developed for the high-throughput characterization of individual TFs and effector domains in yeast,[30,31] drosophila,[32] and human cells.[20,21,33,34] Arrayed high-throughput measurements of combinations of CRs in concert with VP16 have been performed in yeast to test 223 combinations,[7,14] and

Please cite this article in press as: Mukund et al., High-throughput functional characterization of combinations of transcriptional activators and repressors, Cell Systems (2023), https://doi.org/10.1016/j.cels.2023.07.001



(legend on next page)

Please cite this article in press as: Mukund et al., High-throughput functional characterization of combinations of transcriptional activators and repressors, Cell Systems (2023), https://doi.org/10.1016/j.cels.2023.07.001

low-throughput measurements have been performed in human cells.[35] However, we are not aware of any systematic high-throughput studies to date in mammalian cells that measure how effector domains act in combination to regulate gene expression. As such, it remains unclear whether pairs of activator domains synergize when combined, whether pairs of repressor domains do the same, and how activator and repressor domains affect each other when recruited simultaneously. Here, we modify a recently developed pooled high-throughput method[21] to test thousands of combinations of previously characterized protein domains that can activate and repress gene expression in human cells to start addressing these questions.

RESULTS

Developing a workflow for combinatorial screening of transcriptional effector domains

We began by selecting a panel of effector domains and controls from our previous screen of single domains[21]: 44 repressors, 30 activators, and 20 control domains (Figure 1A; STAR Methods; Table S1). These effectors were chosen to span a wide range of individual activation and repression strengths when recruited at the same reporter gene (Figure S1A). The control domains were chosen to be either random sequences or fragments of the DMD protein that were shown to have no effect on gene expression when recruited to a reporter individually.[21] To avoid problems with protein stability, we chose only domains that were well-expressed on their own (i.e., above the detection threshold), as measured by FLAG staining.[21] We generated combinations of these individual domains using a 2-step cloning strategy to build a pool of domain-linker-domain concatenations that were then cloned into a lentiviral backbone vector as fusions to the reverse TetR (rTetR)-inducible DNA-binding domain (Figure 1A; Figure S1B; STAR Methods). We delivered this pool into K562 reporter cells using lentivirus with a low multiplicity of infection (MOI = 0.24), such that most cells expressed a single library element. We measured the ability of each pair in this library to activate or repress a reporter gene using a

Figure 1. Combinatorial screening of transcriptional effector domains

(A) Construction of the combinatorial library. A library consisting of approximately 100 effector domains was cloned into a lentiviral backbone in two positions (Pos. 1 and 2), connected by a 16-amino acid XTEN linker featuring a varying DNA but constant amino acid sequence, and fused to the DNA-binding domain rTetR (STAR Methods). This library consisted of domains that were previously identified as repressive domains (RDs), activator domains (ADs), and controls with no effect on transcription (Ctrls).

(B) Schematic of synthetic reporter system. rTetR-fused domain pairs are recruited to TetO sites upstream of a synthetic reporter expressing a surface marker (exposing an IgG epitope) and an mCitrine fluorophore. The reporter gene's promoter can be varied, with pEF being used to measure repression (top, blue arrow) and minCMV for activation (bottom, yellow arrow).

(C) Magnetic separation. Cells are bound with magnetic beads that recognize the reporter surface marker and are used to separate ON and OFF populations. The DNA encoding for domains expressed in cells from each population are then sequenced via next-generation sequencing.

(D) Computation of screen measurements. Domain pairs are assigned a screen score equal to $\log_2$(ON:OFF) where ON is the relative proportion of a domain in the ON population and OFF is the relative proportion of a domain in the OFF population. The median control-paired score for a given domain is computed as the median score of all pairs featuring that domain and a negative control domain. Order specificity was determined by measuring if the domain function significantly differently when positioned upstream or downstream of the negative control domain.

(E) Correlation between two replicates for the activation screen measurements for all domain pairs after 2 days of recruitment at the minCMV promoter, with domains colored based on whether they are composed of two negative control domains (gray) or at least one effector domain (blue). Dashed line is the threshold for a combination being labeled as activating, two standard deviations away from the mean of the negative control-control combinations.

(F) Correlation between two replicates for the repression screen measurements for all domain pairs after 5 days of recruitment at the pEF promoter, with domains colored based on whether they are composed of two negative control domains (gray) or at least one effector domain (blue). Dashed line is the threshold for a combination being labeled as repressing, two standard deviations away from the mean of the negative control-control combinations.

(G) Correlation between activation $\log_2$(ON:OFF) scores (x axis) and the fraction of cells activated after 2 days of doxycycline recruitment to the weak minCMV promoter (y axis) during individual low-throughput measurements of each pair's function via flow cytometry. The sigmoid curve represents the best-fit line of the form y = 1/(1 + (x/k)^n). Vertical dashed line corresponds to the activation threshold for effector domain combinations as defined above.

(H) Correlation between repression $\log_2$(ON:OFF) scores (x axis) and the fraction of cells repressed after 5 days of doxycycline recruitment to the strong pEF promoter (y axis) during individual low-throughput measurements of each pairs's function via flow cytometry. The sigmoid curve represents the best-fit line of the form y =1 - 1/(1 + (x/k)^n). Vertical dashed line corresponds to the repression threshold for effector domain combinations as defined above.

(I) Variation in control-paired measurements. The $\log_2$(ON:OFF) scores for activation (x axis) and repression (y axis) is shown for negative control-control domain pairs in any orientation (gray), CRTC2-negative control domain pairs in any orientation (yellow), FOXO3-negative control domain pairs in any orientation (green), and ZNF10 KRAB-negative control domain pairs in any orientation (blue). Solid lines represent the median activation/repression values for each domain, and the extent of each line corresponds to the min to max range of activation/repression scores. Dashed lines represent activation and repression thresholds as defined above.

(J) Number of control-including concatenations that exceed the threshold for activation (x axis) and repression (y axis) for all domains, colored by whether that domain acts as a repressor (blue) or activator (yellow) when recruited alone. Horizontal and vertical dashed lines indicate thresholds needed for a domain to be called as an activator or repressor.

(K) Median control-paired activation (x axis) and repression (y axis) scores for individual effector domains. Activators are colored yellow, repressors blue, dual-functional effectors green, and non-hits gray.

(L) Correlation between activation scores from the current domain combination screen and prior nuclear Pfam single-domain activation screen.[21] Domains present in this screen that were measured for activation in a prior screen are shown, with the median control-paired score in this screen shown on the x axis and the $\log_2$(ON:OFF) score from the prior screen shown on the y axis. The type of effector (activator, repressor, dual, negative control, or non-hit) is indicated by the color of each point. Dashed line indicates the threshold for activation in the prior screen.

(M) Correlation between repression scores from the current domain combination screen and prior nuclear Pfam single-domain repression screen.[21] 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 $\log_2$(ON:OFF) score from the prior screen shown on the y axis. The type of effector (activator, repressor, dual, negative control, or non-hit) is indicated by the color of each point. Dashed line indicates the threshold for repression in the prior screen.

Please cite this article in press as: Mukund et al., High-throughput functional characterization of combinations of transcriptional activators and repressors, Cell Systems (2023), https://doi.org/10.1016/j.cels.2023.07.001

high-throughput pooled method we recently developed: HT-recruit.[21] Briefly, by adding doxycycline (dox), we recruited the rTetR-concatenated fusions to either a minimal promoter to measure activation or to a highly expressed constitutive promoter to measure repression (Figure 1B). We separated cells into populations with high (ON) and low (OFF) reporter gene expressions using magnetic separation (STAR Methods; Figure 1C; Figure S1C) and computed the relative enrichment of individual effector combinations in each population (Figure 1D). We recovered the large majority of concatenations that were present in our combinatorial library with sufficient sequencing depth (Figure S1D) and found a good correlation across replicates for both our activation (Figure 1E; Pearson R = 0.803 and $p < 2.2 \times 10^{-16}$) and repression (Figure 1F; Pearson $\rho = 0.795$ and $p < 2.2 \times 10^{-16}$) screens. To identify effector combinations for which we could reliably measure activity, we used the distribution of negative control-only combinations: we considered a pair to be activating and/or repressing if it had a score of at least 2 standard deviations away from the mean of the negative control combinations.

To validate these high-throughput measurements and our detection threshold, we measured activation for 61 domain concatenations and repression for 38 domain concatenations by recruiting them individually to our synthetic reporters and measuring gene expression via flow cytometry. We measured the fraction of cells activated at the weak minCMV promoter after 2 days of recruitment and found strong correlations between our screen measurements and the fraction of cells activated (Figure 1G; $R^2 = 0.88$ and $p = 7.4 \times 10^{-50}$). We found a similarly strong correlation between repressive screen scores and the fraction of cells silenced after 5 days of recruitment to the strong pEF promoter (Figure 1H; $R^2 = 0.81$ and $p = 5.8 \times 10^{-15}$). We concluded that our screen accurately measured these concatenations' ability to either activate or repress gene expression.

We began our analysis of the screen data by examining the behavior of effector domains in our library in combination with the negative control domains. We identified 4 of 20 negative control domains that significantly affected either activation or repression scores across all effector domains and filtered them from downstream analyses (Figures S1E–S1F). We then sought to determine if the order of the domains within concatenations altered their effect on gene expression. Although prior literature has demonstrated that the ordering of effector domains in synthetic TFs can significantly affect the function of fusions,[24,25] we found that few of the domains in our library changed effect size significantly when switched from the first to the second position across all of their fusions with negative controls (Figure S1G). We excluded concatenations featuring any domain in the orientation that ablated its function from downstream analysis (Figure S1H).

To check to what extent effectors maintained their function when fused to a control sequence, we examined the distribution of scores for each effector domain when paired with all negative controls (Figure 1I; Figures S2A–S2B). Although for some domains, all scores clustered together in one quadrant (e.g., the strong KRAB repressor domain from ZNF10; Figure 1I, blue), we found that other domains dropped under the detection threshold when paired with certain negative controls (e.g., weak CRTC2 activation domain; Figure 1I, yellow). Some domains, such as the FOXO3 activation domain, have been previously reported to act as dual effectors that both activate a minimal reporter and repress a constitutive one when recruited individually.[36] We indeed found that most FOXO3-control fusions acted as dual activator-repressors, although a minority of fusions only acted as activators (Figure 1I, green). In order to classify each domain, we calculated the number of effector-control pairs that met our hit threshold for activation and repression (Figure 1J). We used these data, along with the magnitude of activation and repression scores for each effector when paired with negative controls, to label effectors as activators, repressors, dual-functional effectors (both activators and repressors), or non-hits (Figure 1K; STAR Methods). We labeled a domain as an activator if it exhibited at least minimal activation averaged over all negative controls or if it demonstrated significant activation with some negative controls. We did the same for repression and labeled a domain as dual-functional if it was both an activator and a repressor. The labeling of an effector domain as an activator, repressor, or dual is thus a summary label accounting for its behaviors when paired with any of a number of negative control domains; its median activation or repression score when paired with negative controls serves as a quantitative shorthand for this varied profile of behaviors. We found that the median activation or repression score of each effector domain when paired with negative control domains correlated well with prior screen measurements of activation (Figure 1L; Pearson R = 0.81 and $p < 2.2 \times 10^{-16}$) and repression (Figure 1M; Pearson R = 0.85 and $p < 2.2 \times 10^{-16}$). Convinced that we could appropriately characterize how individual effectors behave when paired with negative controls, we proceeded to analyze the behavior of effector domain combinations with each other.

Weak activator domain pairs synergistically drive gene expression from weak promoters

The high-throughput measurements for transcriptional activation identified a large number of activator-activator pairs whose activation scores exceeded the sum of each individual activator's scores when paired with negative controls (Figure 2A, left). In total, these measurements encompassed 63 activator-activator pairs, 72 activator-dual pairs, and 81 dual-dual pairs. We proceeded to validate individual examples of activator pairs to verify that the scores in the screen correlate well with individual flow cytometry measurements (Figure 1G; Figures S3A–S3B) and that synergy could be replicated in low-throughput (Figure 2B, top). Notable examples of synergy included the pair of ANM2's SH3 domain and KIBRA's WW-1 domain, as well as ANM2's SH3 domain and NOTC2's LNR-2 domain. Although the full role of the ANM2 SH3 domain is not yet clear,[37] it is known to regulate PRMT1 activity in a methylation-dependent fashion,[38] is required for the function of the actin nucleator Cobl,[39] and modulates alternative splicing of BCL-X.[40] KIBRA's WW-1 domain binds PPxY motifs in other proteins[41] and is essential for KIBRA-mediated regulation of Hippo signaling via interactions with LATS1/2,[42] whereas NOTC2's second Lin-12/Notch repeat (LNR) domain sits within its negative regulatory region[43] and is cleaved off during ligand binding.[44] It is important to note that this synergy is relative to domains paired with negative controls, not relative to individual activation domains on their own. When comparing the estimated fraction of

Please cite this article in press as: Mukund et al., High-throughput functional characterization of combinations of transcriptional activators and repressors, Cell Systems (2023), https://doi.org/10.1016/j.cels.2023.07.001

CellPress

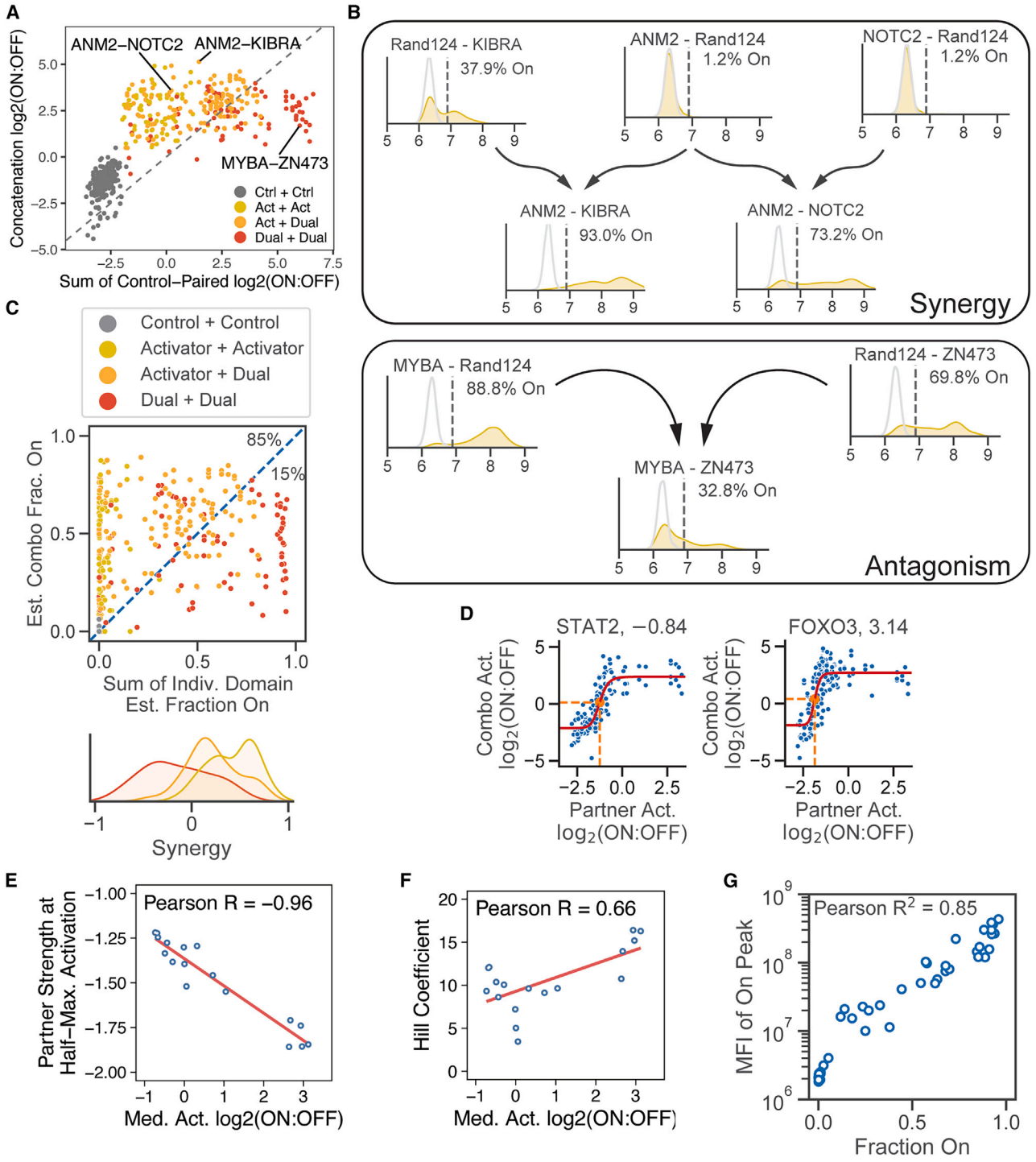


Figure 2. Activators work synergistically when paired together

(A) Evidence for activator-activator synergy in the high-throughput screen after 2 days of recruitment at the minCMV minimal promoter. For negative control-control (dark gray), activator-activator (yellow), activator-dual (orange), and dual-dual (red) domain pairs, the sum of their individual control-paired scores (x axis) is plotted against the average score of the combination (y axis). Individual examples of pairs highlighted in subsequent panels are labeled and indicated by lines.

(B) Example low-throughput validations of activator-negative control and activator-activator pairs at the minCMV promoter. No-doxycycline recruitment is shown in light gray, whereas the probability density of fluorescence after 2 days of dox recruitment is shown in yellow. Vertical dashed lines indicate the threshold for calling cells ON or OFF, with the fraction of cells ON in the dox population labeled within each histogram.

(C) Top: estimated fraction of cells activated at day 2 for every activator-activator (yellow), activator-dual (orange), and dual-dual (red) combination (y axis) versus the sum of the component domains' individual estimated fraction of cells activated based on control-paired scores (x axis). Bottom: density plot of degree of

(legend continued on next page)

Please cite this article in press as: Mukund et al., High-throughput functional characterization of combinations of transcriptional activators and repressors, Cell Systems (2023), https://doi.org/10.1016/j.cels.2023.07.001

cells in combination with the sum of each domain's individual estimates from our prior screen,[21] we find fewer synergistic domains (Figure S3C) because fusing random controls to effector domains systematically reduces their ability to activate target genes (Figure S3D), even after excluding negative control domains that reduce effector domain function across all activators and repressors (STAR Methods). Further experiments with larger panels of activation and random control domains could help clarify whether this interference is caused by the hindrance of multivalent interactions and phase separation,[45] destabilization or induced degradation of the activator domain,[46] or some other mechanism.

We found that the strongest activators at minCMV, which were generally dual-functional effectors that also repressed gene expression at pEF, resulted in lower activation scores at minCMV when paired together than when either activator was paired with negative controls (Figure 2A, upper right quadrant). To verify that the antagonism of dual-dual effector pairs was real and not simply a result of promoter saturation, we individually tested the MYBA-ZN473 pair and found that, indeed, the combination did not activate as many cells as either domain when paired with negative control (Figure 2B, bottom). We wished to determine if this antagonism could be explained by an overall reduction in cell health resulting in reduced rates of cell proliferation. We measured the division rate of cells expressing MYBA-ZN473 along with a number of other effector pairs over the course of 5 days (the timescale of our experiments) and found no significant differences in cell proliferation across all tested concatenations (Figures S4A and S4B). Since we did note a gradual silencing of the MYBA-ZN473-T2A-mcherry transgene over a period of weeks (Figure S4C, although less pronounced at the time point when we performed our activation measurements), we made sure only to examine effector-expressing cells in our analyses by gating for mCherry expression (STAR Methods; Figure S4C). Last, to test whether reduced expression of the effector pair could explain antagonism, we measured the expression of a number of domains via western blot (Figures S4D–S4G). We found that the average expression of the two domains as measured in prior screens[21] correlated reasonably well with the expression of the pairs we measured via western blot (Figure S4H) and found little correlation between the average expression of the two domains and the activity of the combination (Figure S4I). For example, cells express MYBA much lower than ZN473 despite both effectors activating nearly all cells[21] (Figure S4J), suggesting that reduced expression of the pair alone does not explain the antagonism between these dual-dual effector pairs. Although we cannot rule out additional pleiotropic effects causing the dual-dual antagonism we observed, our results do not seem explicable solely by underexpression of dual-dual concatenations or changes in the rate of cell division.

Using the previously computed functions connecting screen scores with individual flow cytometry measurements of gene activation (Figure 1G), we compared the estimated fraction of cells activated for each combination of activators with the sum of each activator's control-paired estimated fraction activated. We found that activator-activator pairs tended to synergize, activator-dual pairs were additive, and dual-dual pairs acted antagonistically with low activation scores (Figures 2A and 2C, top). We computed an estimated quantity of synergy for each combination by taking the difference between that combination's estimated fraction and the sum of its individual domains' fractions. Doing so, we found that activator-activator pairs tended to feature more synergy than activator-dual pairs, which in turn tended to feature more synergy than dual-dual pairs (Figure 2C, bottom).

For each activator or dual-functional domain, we then tracked the activity of all combinations containing that domain (including combinations with repressors and controls) and examined how the strength of the combination varied with the strength of the partner's activation score. We fit these data using a sigmoidal Hill function (Figure 2D; Figure S3E). We found that the strength of the partner domain at the half-maximal point of the Hill function decreased as the activation domain's strength increased, meaning that strong activators could reach half-maximal activation with less help from their partner domain than weak activators (Figure 2E). We also found that the Hill coefficient of these functions went up as the activation domain's strength increased, indicating that the increase from minimal to substantial gene activation happened more rapidly in stronger activation domains as the strength of their partners increased (Figure 2F).

Across all individual validations, we found a strong correlation between the fraction of cells activated and the mean fluorescence intensity (MFI) of activated cells for our effector domain pairs (Figure 2G). This correlation and the fluorescent distributions of reporter expression (Figures 2B and S3A) are consistent with activation domains modulating transcriptional bursting kinetics at the minCMV promoter: either increasing burst frequency or burst magnitude.[47–51] Since the degradation rates of our reporter mRNA and protein are slow, taking multiple days to dilute out,[10,21] reporter molecules in our cells persist long after the cessation of a transcriptional burst. As such, it would be difficult to determine whether reporter protein molecules were made all at once in one big burst or bit by bit (in many small bursts) over a longer period of time. Thus, in this regime of slow degradation rates, increasing either the duration of a transcriptional burst (burst size) or the frequency of such bursts (burst frequency) would lead to both a higher fraction of cells measured to be on

synergy (x axis) for activator-activator (yellow), activator-dual (orange), and dual-dual (red) combination pairs. Synergy is defined as the difference between the combination's estimated day 2 fraction off and the sum of the individual domains' day 2 fractions off.

(D) Sigmoidal shape of activator behavior in the high-throughput screen. For STAT2 (left) and FOXO3 (right), the control-paired score of its partner in any given pair (x axis) is plotted against the score of the combination of the domain and that partner (y axis) for all concatenations featuring the given domain.

(E) For all activator domains, the control-paired activation strength of the activator (x axis) is plotted against the half-maximal point of the corresponding sigmoid fit to screen data (y axis). Line shown is the line of best fit.

(F) For all activator domains, the control-paired activation strength of the activator (x axis) is plotted against the Hill coefficient of the corresponding sigmoid fit to screen data (y axis). Line shown is the line of best fit. (G) Correlation between the fraction of activated cells (x axis) and the mean fluorescence intensity (MFI) of activated cells (y axis) for each combination measured individually.

and a greater MFI of those on cells.[52] Our findings are consistent with recent work showing that changes in burst kinetics can combine to produce transcriptional synergy.[53]

Repressor domain combinations generate robust promoter silencing

Our measurements of transcriptional repression included 1,408 repressor-repressor pairs, 342 repressor-dual pairs, and 81 dual-dual pairs (Figure 3A). To determine how individual types of repressors behave in combinations with other domains on the screen, we plotted the strength of every combination featuring a particular repressor versus the strength of its partner (repressor, activator, dual, or control) in that combination. We found that the trendline for repressors was linear, with a lower slope for stronger repressors (Figure 3B; Figures S5A and S5B). For each repressor domain, we determined the slope and y-intercept of its corresponding trendline and found that stronger repressors featured both a flatter slope closer to 0 (Figure S5C; Pearson R = 0.89 and $p < 2.2 \times 10^{-16}$) and a lower y-intercept (Figure S5D; Pearson R = 0.93 and $p < 2.2 \times 10^{-16}$). These results suggest that weak repressors are tunable in that the strength of a pair featuring a weak repressor can vary over a wide range depending on the strength of the partner; however, pairs featuring a strong repressor will generally only act as strong repressors.

This analysis suggests that points that are above the diagonal and under −3 come from the saturation of repression at our reporter; these effector pairs are predicted to silence all cells according to our validation curve (Figure 1H). Indeed, when we tested a number of domain pairs in low throughput, we found that recruitment of pairs of strong repressors that each fully repressed the reporter when paired with controls (e.g., ZNF10 and CBX1) also leads to full repression (Figure 3C, top row; Figure S5E). In contrast, when we tested pairs of moderately strong repressors, we found that the strength of the pair was closer to the average of the two domains, rather than full silencing (Figure 3C, bottom two rows, Figures S5E and S5F).

Using the best-fit curve for low-throughput validations mapping the repression $\log_2$(ON:OFF) scores to the fraction of cells silenced at day 5 (Figure 1H), we estimated the fraction of cells silenced at day 5 for each repressor or dual effector in our screen based on its median control-paired repression $\log_2$(ON:OFF) score. We found that the strength of effector-control pairs that could repress spanned a wide range, and we categorized each repressor domain as weak, moderate, or strong based on its estimated fraction silenced (Figure 3D). We compared the estimated fraction of cells silenced at day 5 for each repressor-repressor combination versus the sum of its individual domains' respective fractions and found that although concatenations with at least one strong repressor were expected to and did silence virtually all cells after 5 days, combinations with only moderate and/or weak repressors generally featured a wider range of silent fractions that reached total silencing much less frequently (Figure 3E). In marked contrast to our data about activator-activator pairs, we found that although only 9% of repressor-repressor pairs were consistent with synergistic repression, approximately 90% were estimated to be similar to the average or sum of their individual domains (Figures S5G and S5H).

When examining the fluorescence distributions of our reporter throughout silencing in individual validations, we found that most domain combinations silenced either all or none of the cells by day 5 (e.g., Figure 3F; rTetR-only and ZNF10-CBX1), as expected from previous experiments with a small number of CRs.[10] However, a minority of combinations either silenced slowly, resulting in a fraction of cells remaining active after 5 days, (Figure 3F; SMCA4-U2AF4) or reduced gene expression without fully silencing the gene (Figure 3F; DPY30-HXA13). For DPY30-HXA13, the reporter MFI stabilized to a new intermediate level during days 2–5 of recruitment, indicating a new steady state (Figure 3F). This partial silencing is not likely due to the lack of expression of these effector combinations: their average expression scores based on previous measurements are high (Figure S4I), and the transgene encoding for these combinations has a high percentage of mCherry positive cells that we gated on (Figure S4C).

We wanted to see if these results are consistent with prior models of stochastic gene silencing by extending a mathematical model of gene expression we have used before to understand chromatin regulation.[52] We did so by incorporating parameters for background silenced cells at the beginning of the time course, basal gene expression from the pEF promoter, lag time prior to silencing, the rate of protein decay upon. silencer recruitment, and the rates of gene silencing and reactivation (Figure S6A; STAR Methods). After fitting this model to our experimental data, we found that it could accurately represent the dynamics of gene silencing for both strong silencers such as ZNF10-CBX1 (Figure S6B) and weaker silencers such as SMCA2-U2AF4 that only silenced the reporter in a fraction of cells (Figure S6C). For these domains, we found a good correlation between the rates of silencing predicted by this model and both the screen data (Figure S6D; $R^2 = 0.81$ and $p = 5.8 \times 10^{-12}$) and the fraction of cells silenced at day 5 (Figure S6E; $R^2 = 0.80$ and $p = 1.22 \times 10^{-11}$). Silencing rates predicted by the model suggested that the rate of silencing of a combination may be linear in the sum of the silencing rates of the individual domains (Figure S6F), although the correlation between screen data and silencing rates was not strong enough for a more definitive conclusion.

However, we found that our model of all-or-none gene silencing could not capture the dynamics of silencers that reduced but did not fully ablate gene expression by day 5, such as DPY30-HXA13 (Figures S6G–S6I), which featured a domain from DPY30 that worked as an activator on its own in prior screens and a repressor homeodomain from HXA13.[21] For these silencers, whose MFI stabilized to an intermediate level after 2–3 days of recruitment (Figures S6H and S6I), their effect on gene expression is better explained by a decrease in the production rate in the active state, rather than transitioning to a fully silent state. Altogether, these results suggest that although strong silencers in combinations tend to rapidly and fully silence gene expression, weaker repressors can produce more complex patterns of gene silencing.

High-throughput profiling of activator-repressor combinations

To identify principles underlying the behavior of activator-repressor combinations, we looked at all pairs featuring an

Please cite this article in press as: Mukund et al., High-throughput functional characterization of combinations of transcriptional activators and repressors, Cell Systems (2023), https://doi.org/10.1016/j.cels.2023.07.001

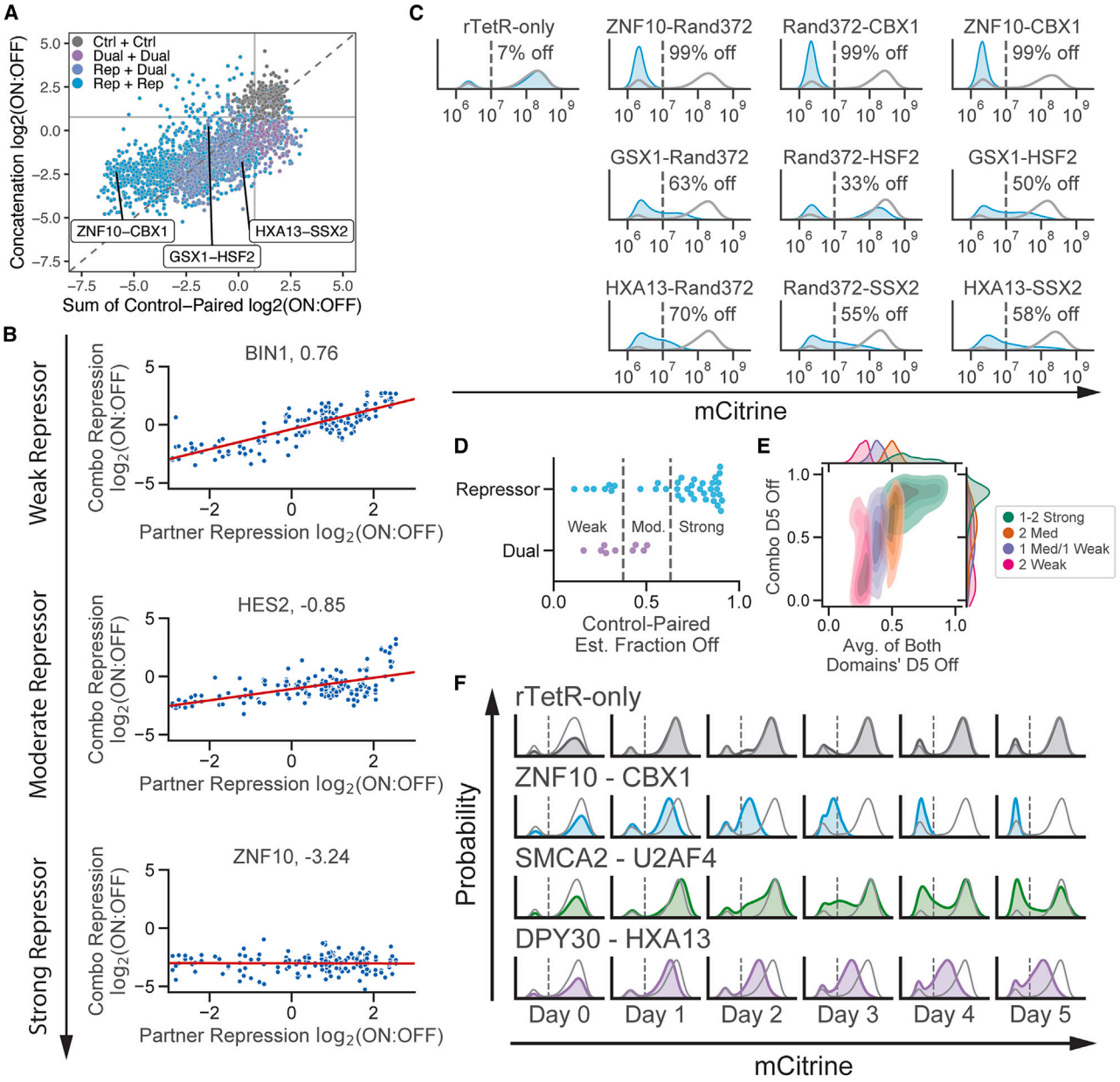


Figure 3. Repressors combine to drive gene silencing

(A) Repressor-repressor pair behavior in the high-throughput screen after 5 days of recruitment at the pEF promoter. For negative control-negative control (dark gray), repressor-repressor (light blue), repressor-dual (dark blue), and dual-dual (purple) domain pairs, the sum of their individual control-paired scores (x axis) is plotted against the average score of the combination (y axis). Lower ON:OFF scores denote more repression. Horizontal and vertical lines denote the repression threshold as defined in Figure 1F, whereas the diagonal dotted line corresponds to the line y = x.

(B) Linear shape of repressor behavior in the high-throughput screen. For a given domain (BIN1, top; HES2, middle; and ZNF10, bottom), the control-paired score of its partner in any given pair (x axis) is plotted against the score of the combination of the domain and that partner (y axis) for all concatenations featuring the given domain. Red lines indicate linear fits to the data.

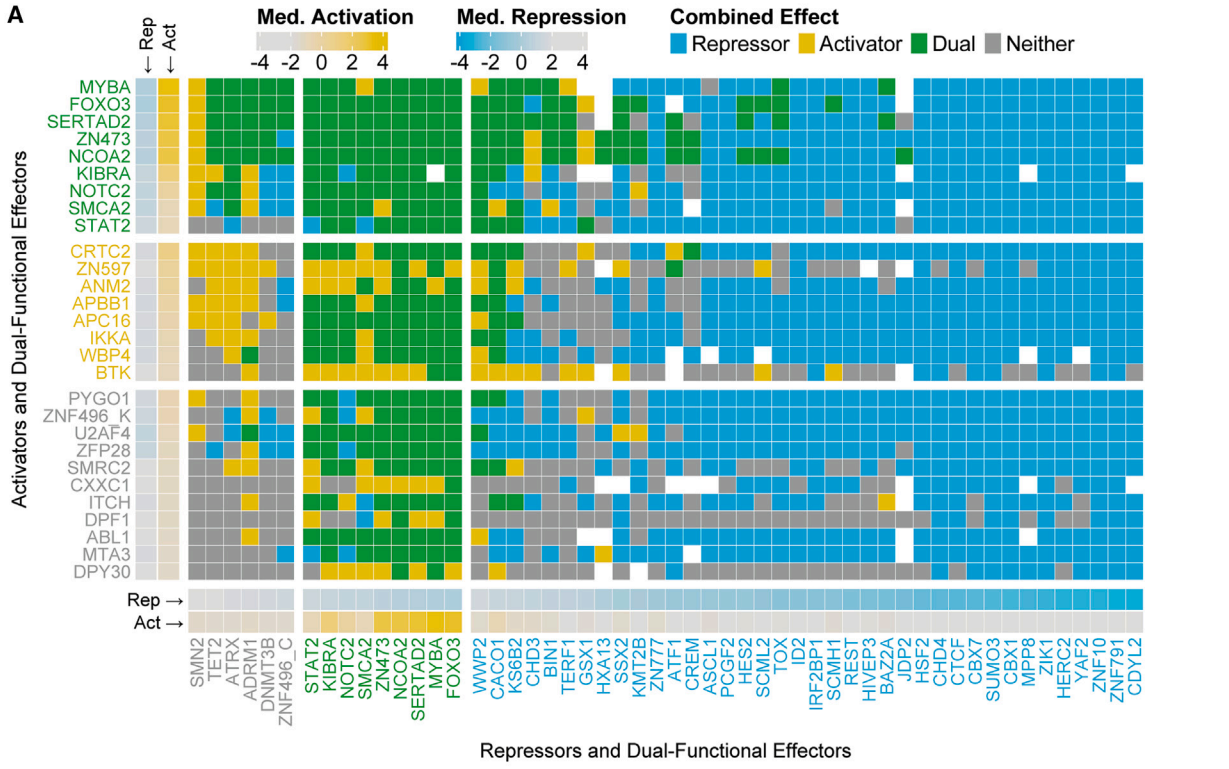
(C) Fluorescent distributions of the citrine reporter gene from low-throughput flow cytometry measurements of repressor function of various domain combinations after 5 days of recruitment at the pEF promoter, with Rand372 denoting a negative control domain. Blue, 1,000 ng/mL doxycycline; gray, no dox. Vertical dashed lines indicate the threshold for calling cells ON or OFF, with the fraction of cells OFF in the dox population labeled within each histogram.

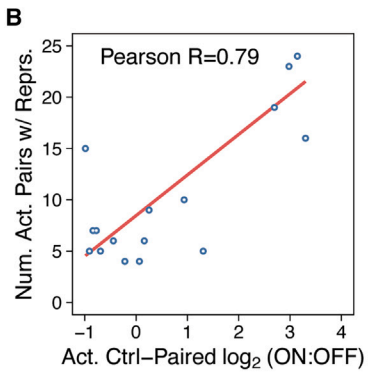
(D) Estimated fraction of cells silenced at day 5 (x axis) for repressors (blue) and dual-functional domains (purple) when paired with negative controls. Repressors were partitioned into three categories: weak (left), moderate (middle), and strong (right).

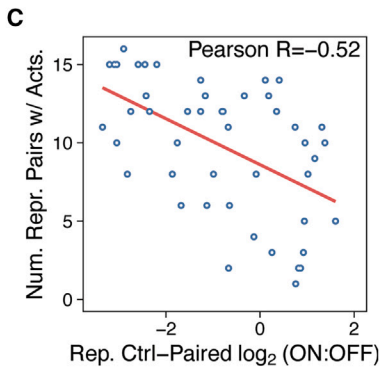
(E) For all repressor-repressor combinations, the sum of the two repressors' individual estimated control-paired fraction silenced at day 5 (x axis) is plotted against the estimated fraction of cells silenced at day 5 by the combination (y axis). Colors represent the composition of the repressor pairs in terms of strong, moderate, and weak repressors.

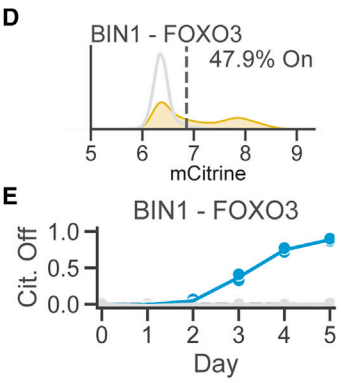
(F) Individual plots of transcriptional repression. Probability density (y axis) plots of mCitrine-A levels (x axis) presented for four different pairs of domains (rows) as measured by flow cytometry over 5 days of doxycycline recruitment (columns).

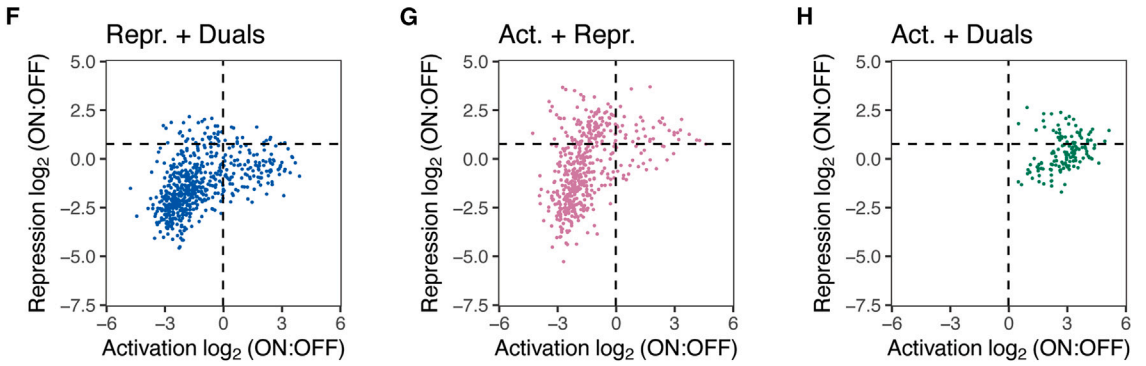
Please cite this article in press as: Mukund et al., High-throughput functional characterization of combinations of transcriptional activators and repressors, Cell Systems (2023), https://doi.org/10.1016/j.cels.2023.07.001











(legend on next page)

Please cite this article in press as: Mukund et al., High-throughput functional characterization of combinations of transcriptional activators and repressors, Cell Systems (2023), https://doi.org/10.1016/j.cels.2023.07.001

activator domain in combination with a repressor domain and determined if that pair functioned as an activator-only pair, repressor-only pair, dual-functional pair, or non-hit pair (Figure 4A; Figure S7A). We found that almost all pairs featuring moderate to strong repressors tended to behave exclusively as repressors, reducing gene expression at the strong pEF promoter while failing to drive transcription at the weak minCMV promoter (blue right region, Figure 4A). Domains that were dual-functional when paired with negative controls (e.g., MYBA, FOXO3, and SERTAD2) tended to be dual-functional in combinations with either weak repressors or other dual effectors (green region, Figure 4A), although they were dominated by strong repressors (e.g., ZNF10).

For most domains, their activator-repressor and repressor-activator orientations had similar strengths for both activation (Figure S7B; Pearson R = 0.80 and $p < 2.2 \times 10^{-16}$) and repression (Figure S7C; Pearson R = 0.79 and $p < 2.2 \times 10^{-16}$), indicating that these results were not the result of position-specific effects of the protein domains. Activator-repressor combinations with stronger activators were significantly more likely to be able to activate the minimal promoter than combinations with weaker activators (Figure 4B; Pearson R = 0.79 and $p = 1.6 \times 10^{-4}$). For repressors, this relationship was significantly less pronounced: weak repressors were only slightly less effective than strong repressors when paired with activator domains (Figure 4C; Pearson R = -0.52 and $p = 1.6 \times 10^{-4}$). Testing pairs in low throughput validated that the combination of the weak repressive SH3 domain from BIN1 and FOXO3's TAD could both activate gene expression at the weak minCMV promoter (Figure 4D) and repress gene expression when recruited to the strong pEF promoter (Figure 4E). As expected, adding the weak repressor domain from BIN1 decreased activation and increased repression compared with FOXO3 alone (Figures S2 and S3A).

Although in general, we found that strong repressors overpowered virtually any non-repressor effector domain they were paired with, we did find a few examples of activators that could prevent dual-functional domains from repressing strong promoters (Figure 4A). These activators include the activator KRAB from ZN597, a variant KRAB domain that functions as an activator; the SH3 domain from ANM2 previously described in this work; and the activating SH3 domain from BTK. Although BTK's SH3 domain is known to regulate BTK kinase activity upon autophosphorylation[54] and BTK can translocate to the nucleus,[55] we are unaware of a previously described role for the SH3 domain in regulating gene expression of BTK targets.

We found a number of non-hit domains that could not alter gene expression when paired with negative controls but did affect the function of other effector domains (gray labels, Figure 4A; Figure S7A). For example, we found that the aforementioned domain from DPY30, a core subunit of the SET1/MLL methyltransferase complex that interacts with ASH2L to establish H3K4me3,[56] could ablate repressor function when paired with not only dual-functional domains but even weak to moderate strength repressors. Additionally, the N-terminal domain of DPF1, which links the NF-κB RelA/p52 heterodimer with SWI/SNF complex subunits to drive transcription,[57] could prevent repressor function from strong repressors, such as HERC2's Cyt-b5 domain. We also found non-hit domains that prevented activation when paired with other effectors, including the tile from the N-terminal-disordered region of DNMT3B thought to be part of a broader region mediating interactions with the methyltransferases DNMT1 and DNMT3A.[58] The C2HR domain from the zinc finger ZNF496, which has been shown to overpower the variant activator KRAB present on the same transcription factor,[22] could similarly prevent all activators and all but the strongest dual-functional domains from driving gene expression. Altogether, these effectors did not themselves activate or repress transcription when paired with negative controls but did modulate the activity of partner effectors on the same molecule in a manner consistent with the function of their native proteins.

We wondered how the distribution of effector combinations along both repression and activation $\log_2$(ON:OFF) scores varied

Figure 4. Activator-repressor combinations

(A) Representation of activator-repressor behavior measured by the high-throughput screen. Domains that were identified as repressors when recruited alone are arrayed along the x axis separated into three categories based on their effect when paired with controls (non-hits on the left, dual-functional domains in the middle, and repressors on the right) with the stronger repressors within each category, as measured by their control-paired repression $\log_2$(ON:OFF), on the right. Boxes just above each repressor domain indicate the control-paired activation/repression score for each effector. Domains that were identified as activators are arrayed along the y axis, separated into three categories based on their effect when paired with controls (non-hits on the bottom, activators in the middle, and dual-functional domains on the top) with the stronger activators in each category, as measured by control-paired activation scores, on the top. Boxes just to the right of each activator domain indicate the control-paired activation/repression score for each effector. Boxes in the interior represent combinations of the corresponding activator and repressor, and are colored based on whether the combination acts as a repressor only (blue), activator only (yellow), both a repressor and an activator (green), or neither (gray).

(B) Correlation between the control-paired activation score for a set of activators (x axis) and the number of pairs including any repressor domain and that activator domain that were called as an activator pair hit (y axis). Red line represents the line of best fit.

(C) Correlation between the control-paired repression score for a set of repressors (x axis) and the number of pairs including any activator domain and that activator domain that were called as a repressor pair hit (y axis). Red line represents the line of best fit.

(D) Low-throughput measurement of the activation strength of the pair of BIN1's SH3 domain and FOXO3's TAD when recruited at the minCMV for 2 days. The gray line represents the no-doxycycline condition, and the yellow 2 days doxycycline recruitment. The vertical dashed line indicates the threshold for labeling a cell as ON or OFF.

(E) Fraction of cells with pEF promoter silenced by the fusion between BIN1's SH3 domain and FOXO3's TAD over 5 days of recruitment. The fraction of cells silenced was normalized to the no-dox condition (STAR Methods).

(F) Distribution of repression $\log_2$(ON:OFF) scores (y axis) versus activation $\log_2$(ON:OFF) scores (x axis) for combinations featuring one repressor domain and one dual-functional domain.

(G) Distribution of repression $\log_2$(ON:OFF) scores (y axis) versus activation $\log_2$(ON:OFF) scores (x axis) for combinations featuring one repressor domain and one activator domain.

(H) Distribution of repression $\log_2$(ON:OFF) scores (y axis) versus activation $\log_2$(ON:OFF) scores (x axis) for combinations featuring one activator domain and one dual-functional domain.

between repressor-dual, activator-dual, and repressor-activator combinations. We found that although most repressor-dual combinations functioned as pure repressors, some dual-functional domains could combine with repressors to produce overall dual-functional combinations (Figure 4F, bottom right quadrant). In contrast, combining activators with repressors produced a larger number of effector pairs that did not activate the weak promoter and relatively fewer dual-functional combinations (Figure 4G). Pairing pure activators with dual-functional domains produced combinations that mostly acted as dual activator-repressors, with a smaller fraction of combinations acting as pure activators without maintaining the repressor effect of the dual-functional domain (Figure 4H).

Systematic characterization of domains that influence KRAB-mediated repression

Our screen data so far indicated that virtually any concatenation including the ZNF10 KRAB domain functioned as a strong repressor (~180 KRAB-containing pairs). We were interested in investigating this KRAB domain in more detail, as it is widely used in the well-known CRISPRi system,[59] has been harnessed in conjunction with DNA methyltransferases to produce more durable epigenetic silencing,[9,25] and has been deeply characterized.[21] We decided to test whether this pattern of KRAB dominance held when using a larger panel of partner domains and generated a lentiviral library encoding a set of concatenations fusing ZNF10 KRAB with a library comprising ~5,000 80-amino acid proteins sequences consisting of Pfam-annotated domains from nuclearly localized proteins and a set of negative controls.[21] We transduced this library into K562 cells expressing a reporter gene driven by the strong pEF promoter (Figure 5A) and recruited the concatenations to the reporter gene for 5 days by adding doxycycline. We used a lower doxycycline concentration than that used in our previous screens (100 ng/mL vs. 1,000 ng/mL) to slow down KRAB silencing (Figure S8A) and thus allow for a wider dynamic range for measuring both the decrease and increase of function. We measured the enrichment of domain pairs in the ON versus OFF populations (Figure S8B) on day 2 of recruitment to determine the speed of silencing; we opted not to take a sequencing measurement on day 5 at the end of recruitment, as the population was virtually entirely silent (Figure 5B).

On day 2, the majority of pairs featuring well-expressed domains (as measured by FLAG staining before[21]) scored similarly to concatenations of KRAB with a negative control domain (Figure 5C), consistent with the flow cytometry results showing that the majority of the cells were silenced (Figure 5B). This matched our expectations from the prior screens, where concatenations featuring KRAB silenced similarly to KRAB on its own or with negative control domains. In this case, the negative controls consisted of a large set of random sequences and DMD fragments[21] and contained both well-expressed and poorly expressed proteins. We found that a number of domains were lowly expressed on their own ablated KRAB function when paired with it (Figures 5C and 5D). The FLAG(ON:OFF) scores for these low-expressed domains were low enough that we would not expect them to be visible via a western blot.[21] We hypothesize that these poorly expressed domains when fused to KRAB reduce their expression levels to a point where it impairs KRAB's repressive function. For example, we verified that the poorly expressed DHX16 OB_NTP and BAZ1A DDT domains inhibited KRAB function at 100 ng/mL dox (Figure 5E, top). Increasing the dox concentration to 1,000 ng/mL dox permitted some silencing for KRAB-DHX16 and full silencing for KRAB-BAZ1A (Figure 5E, bottom), consistent with the loss of function of these KRAB fusions coming from decreased protein abundance.

We also observed a loss of KRAB function when fused to certain well-expressed domains from proteins that are part of the basic transcriptional machinery, namely the 2nd WD40 domain from TAF5L and the fork domain of RPB2. Consistent with the high-throughput measurements, individual validations at 100 ng/mL dox show a complete loss of KRAB silencing for the fusions with these domains (Figure 5E, top row). The annotated RBP2 domain is smaller than the 80 aa sequence we used in our screen, and the 64 aa trimmed version had a lower capability of opposing KRAB, allowing for more silencing (Figure S8E). Some amount of silencing was also restored at saturating dox concentrations for both KRAB-TAF5L and KRAB-RBP2 (Figure 5E, bottom row), showing that KRAB can still dominate if enough of it is recruited at the locus. These results were corroborated by an overall increase in the rate of gene silencing when KRAB concatenations were recruited at 1,000 ng/mL dox compared with recruitment at 100 ng/mL dox (Figure S8F; $p = 0.0239$ via paired t test).

In the high-throughput measurements, we also identified a small number of domains that increased KRAB silencing (Figure 5C, bottom left). We validated that the library tile containing the homeodomain from GSX2 increases KRAB silencing at day 2 when recruited at 100 and 1,000 ng/mL doxycycline (Figure 5E), consistent with its behavior as a repressor when recruited on its own in our previous Pfam screen.[21] Interestingly, the trimmed version that contains only the annotated homeodomain does not enhance KRAB silencing (Figure S8C). In contrast, the NHR2 domain from MTG8R—on its own, a slightly weaker repressor than GSX2's homeodomain in our prior screens—could not modify KRAB-mediated silencing at both 100 ng/mL and 1,000 ng/mL dox (Figure 5E).

Composing effector domains to generate multifunctional synthetic transcription factors

We wanted to take advantage of the fact that the KRAB repressor is dominant over activators to build a more versatile transcription factor that can switch between repressor and activator based on the addition or removal of a second drug. We engineered a version of our rTetR-FOXO3-ZNF10 KRAB concatenation plasmid where the two effector domains were separated by a StaPL domain, which cleaves itself in the absence of asunaprevir (ASV)[60] (Figure 6A). Thus, in the absence of ASV, the KRAB domain would be cleaved, leaving behind the dual-functional rTetR-FOXO3, whereas upon addition of the ASV inhibitor, the KRAB domain would act as a dominant repressor (Figure 6B).

We first verified that the construct worked as expected when driving gene activation at the minCMV promoter (Figure 6C) and gene repression at the pEF promoter (Figure 6D; Figure S9A) at maximum recruitment with a saturating dose of dox (1,000 ng/mL). At minCMV, the rTetR-FOXO3-StaPL-KRAB drove strong gene activation in the absence of ASV (Figure 6C, bottom right green) to a similar level as rTetR-FOXO3 only (Figure 6C, top right). At the same promoter, addition of ASV reduced activation

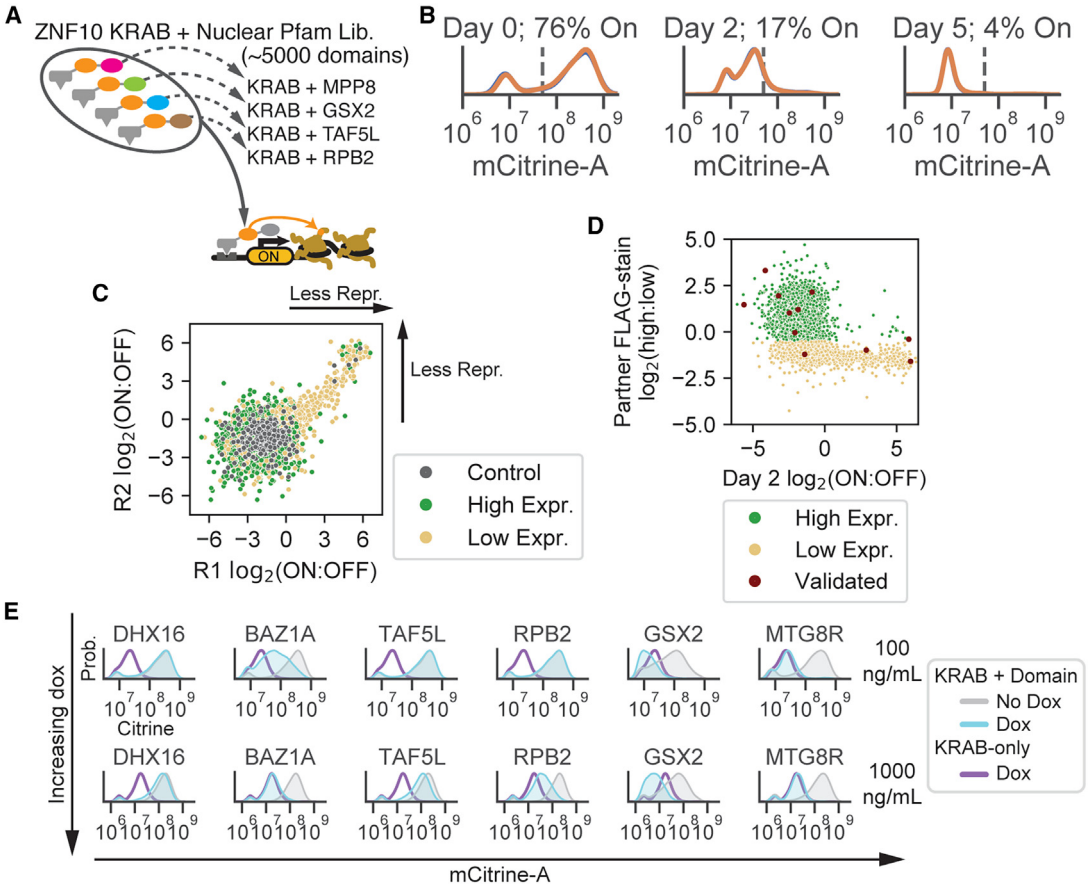


Figure 5. KRAB repression when used in combination with other protein domains

(A) Schematic of KRAB-Pfam screen. A library of ~5,000 Pfam-annotated nuclear-localized domains[21] was cloned into a backbone downstream of an rTetR DNA-binding domain fused to the ZNF10 KRAB repressor and recruited at the constitutive pEF promoter.

(B) Flow cytometry measurements of mCitrine fluorescence (x axis) at 0, 2, and 5 days after 100 ng/mL dox-mediated recruitment at the pEF promoter. Data shown consists of two replicates.

(C) Correlation between replicate 1 (x axis) and replicate 2 (y axis) of screen $\log_2$(ON:OFF) scores after 2 days of dox recruitment for KRAB-domain pairs where domain pairs were previously annotated as well-expressed (green), poorly expressed (yellow), or negative controls (gray).

(D) Correlation between screen $\log_2$(ON:OFF) scores (x axis) and prior FLAG-stain measurements of KRAB partner domain expression (y axis) for high-expressed domains (green) and low-expressed domains (yellow). Domains that were individually validated via flow cytometry are shown in red.

(E) Probability density curves of mCitrine fluorescence for select domains that were individually validated in combination with KRAB (blue) using flow cytometry. Data shown are after 2 days of recruitment at 100 ng/mL dox (top row) or 1,000 ng/mL dox (bottom row) at the pEF promoter. No-dox controls are shown in gray, whereas recruitment of KRAB alone is shown in purple.

by rTetR-FOXO3-StaPL-KRAB to a minimum (Figure 6C, bottom right blue), comparable with rTetR-KRAB recruitment and consistent with KRAB dominating over FOXO3. At pEF, the rTetR-FOXO3-StaPL-KRAB repressed virtually all cells upon ASV addition (Figure 6D, bottom right blue), the same as rTetR-KRAB alone. Without ASV, this fusion still repressed 50% of the cells (Figure 6D, bottom right green), consistent with rTetR-FOXO3 alone being able to silence pEF in about ~54%–57% of cells (Figure 6D, top right). The negative controls rTetR alone and rTetR-StaPL did not change gene expression at either promoter (Figures 6C and 6D). Altogether, the construct behaved as expected: adding ASV changed its behavior from FOXO3-like to KRAB-like.

The ability to toggle the FOXO3-StaPL-KRAB TF from activator to repressor and then back to activator again allowed us to test whether the activator behaves differently at a promoter before and after KRAB-induced repression. In order to do this, we added dox to cells containing this fusion for a period of 20 days to induce recruitment, and within this interval, we varied ASV to toggle between FOXO3 (−ASV) and KRAB dominant (+ASV) (Figures 6E and 6F). At the minCMV reporter, we saw rapid gene activation upon the addition of doxycycline that was lost when ASV was added and KRAB was recruited to the promoter (Figure 6E). Removal of ASV, leading to the recruitment of FOXO3 alone at this KRAB-silenced minCMV promoter produced much slower gene activation than at the unsilenced minCMV promoter (Figure 6E; days 7–20 vs. 0-2). The slow reactivation upon ASV washout suggests that although FOXO3 may be able to drive gene expression from a minimal promoter at a permissive locus, it is less efficient at reactivating that promoter after it is silenced with potentially methylated and/or compacted chromatin. At the pEF promoter (Figure 6F), we

Please cite this article in press as: Mukund et al., High-throughput functional characterization of combinations of transcriptional activators and repressors, Cell Systems (2023), https://doi.org/10.1016/j.cels.2023.07.001

CellPress

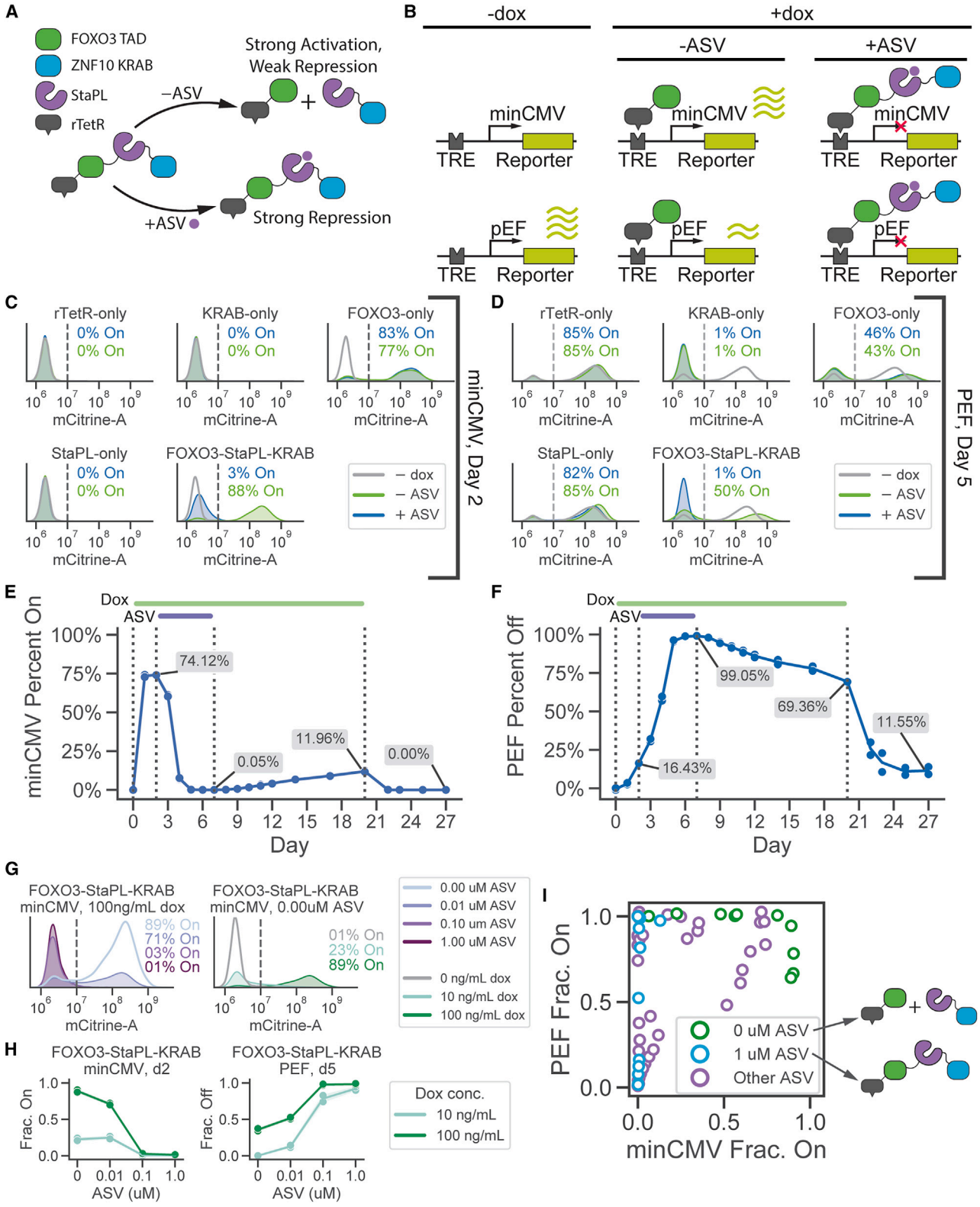


Figure 6. A synthetic TF that can switch from strong repressor to strong activator

(A) Schematic of synthetic transcription factor system. An ASV-inducible StaPL (purple) is fused in between FOXO3's TAD (green) and ZNF10's KRAB (blue) domains. In the absence of ASV, the StaPL and KRAB are cleaved off, resulting in only FOXO3's TAD being recruited to the reporter. Adding ASV stabilizes the StaPL domain, resulting in recruitment of FOXO3's TAD fused to ZNF10's KRAB domain.

(legend continued on next page)

Please cite this article in press as: Mukund et al., High-throughput functional characterization of combinations of transcriptional activators and repressors, Cell Systems (2023), https://doi.org/10.1016/j.cels.2023.07.001

observed a slow reduction in gene expression without ASV during days 0–2, consistent with the dual FOXO3 acting as a weak repressor at pEF. The percentage of cells silenced rapidly increased upon the addition of ASV and recruitment of KRAB culminating in complete silencing by day 7. After the removal of ASV, we saw slow but measurable reactivation that increased dramatically upon doxycycline removal (Figure 6F; days 7–20 vs. 20–27), suggesting that although FOXO3 can behave as an activator and drive gene expression at minCMV, its repressive capacity may inhibit proper reactivation of the pEF promoter.

In order to characterize the range of behaviors achievable with this inducible synthetic TF, we varied the dosing of both ASV and doxycycline while recruiting the rTetR-FOXO3-StaPL-KRAB construct at both promoters. At 100 ng/mL dox increasing doses of ASV led to decreased gene activation at the minCMV promoter (Figure 6G, left; Figure S9B), consistent with the expectation that increasing doses of ASV led to an increase in the species containing KRAB relative to FOXO3 only. We also found that increasing the dose of dox at 0 ASV led to increasing gene activation at the minCMV promoter, consistent with a model in which the absence of ASV leads to virtually minimal KRAB recruitment (Figure 6G, right; Figure S9B). Low to moderate doses of ASV, 0.01 and 0.1uM, produced intermediate profiles between FOXO3 and KRAB (Figures S9B and S9C). Altogether, these results suggest a model where increasing the dose of ASV titrates between FOXO3-like and KRAB-like behavior, whereas increasing the dose of dox increases the degree of gene activation and/or repression.

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

DISCUSSION

Despite considerable efforts to parse the combinatorial logic of TF binding and gene regulation,[19,61–63] our understanding of how multiple distinct functional transcriptional effector domains work together within a single TF remains limited. Improved characterization of the combined function of effector domains is important for understanding the function of natural TFs featuring multiple effector domains,[2] building complex synthetic biology tools to manipulate gene expression,[64] and building cell therapies to detect and treat diseases.[26,65–68] Here, we present the results of screening thousands of effector domain combinations, another step toward uncovering the basic principles of combinatorial gene regulation.

We found a number of negative control domains, including control domains composed of randomly chosen amino acids (“randomers”), that could tune the level of activation and repression by other effector domains. Although we have previously reported that random amino acid sequences could occasionally activate or repress target genes,[21] to the best of our knowledge, this marks the first report of randomers that do not alter gene expression on their own but still modulate the effect of partner domains, for example, by reducing domains that weakly activate transcription on their own to non-hits. Understanding the mechanism behind this phenomenon requires a more systematic interrogation of the sequence[13,30,69] and structural[70,71] motifs, as well as compositional biases,[20,34] present within these domains and other randomers that behave similarly.

We found that weak activators can synergize to drive robust gene expression even when the individual activators being paired were not particularly strong. This is in agreement with previous results showing that multiple activation domains can act synergistically in yeast when recruited at a synthetic reporter[7,14] and in human cells when recruited at reporters or endogenous genes using dCas9.[24,72] At the other extreme, we found that some of the strongest activators, which in our system acted as

(B) Expected dynamics of gene expression at the minCMV promoter (top row) or pEF promoter (bottom row) without dox (left), with dox and no ASV (middle), and dox and ASV (right).
(C) Histograms of transcriptional activation after 2 days of recruitment (with 1000 ng/ml dox) of indicated fusions to the synthetic reporter driven by the minCMV promoter in the presence (blue) or absence (green) of 1μM ASV, with the no dox control shown in gray.
(D) Histograms of transcriptional repression after 5 days of recruitment (with 1000 ng/ml dox) of indicated fusions to the synthetic reporter driven by the pEF promoter in the presence (blue) or absence (green) of 1μM ASV, with theno dox control shown in gray.
(E) Transcriptional activation after ASV washout. K562 cells expressing the minCMV-driven reporter and the FOXO3-StaPL-KRAB construct were treated with 1000 ng/ml doxycycline during days 0–20 (green line) and 1μM ASV during days 2–7 (purple line). Cells were washed 1× with PBS when removing ASV and doxycycline.
(F) Transcriptional repression after ASV washout. K562 cells expressing the pEF-driven reporter and the FOXO3-StaPL-KRAB construct were treated with 1000 ng/ml doxycycline during days 0–20 (green line) and 1μM ASV during days 2–7 (purple line). Cells were washed 1× with PBS when removing ASV and doxycycline.
(G) Histograms of transcriptional activation after 2 days of recruitment to the synthetic reporter driven by the minCMV promoter with either 100 ng/mL dox and varying doses of ASV (purple, left) or 0 uM ASV and varying doses of dox (green, right).
(H) (left) Fraction of cells activated after 2 days of recruitment to the synthetic reporter driven by the minCMV promoter (y axis) for different ASV doses (x axis), and for different doses of doxycycline (light vs. dark green). (right) Fraction of cells silenced after 5 days of recruitment to the pEF promoter.
(I) Scatter of minCMV-reporter fraction on (x axis) versus pEF-reporter fraction on (y axis) measured across all FOXO3-StaPL-KRAB experiments. Each point constitutes 1 time point, and the color of each point indicates the ASV dose.

Please cite this article in press as: Mukund et al., High-throughput functional characterization of combinations of transcriptional activators and repressors, Cell Systems (2023), https://doi.org/10.1016/j.cels.2023.07.001

dual-functional domains that could also repress a constitutive promoter, were often antagonistic: when paired with each other, they produced less gene activation from a minimal promoter than when recruited on their own. Antagonism between transcriptional activators has been reported in other contexts: TFs have been shown to interfere with each other's ability to activate genes via direct interactions with each other[73] or squelching mechanisms involving the sequestration of coactivator proteins.[74] Although our findings suggest a more general negative feedback mechanism triggered by high levels of activators at a promoter, more thorough confirmation of this phenomenon is still needed. For example, synergy between strong activator domains has been seen with dCas9-VPR.[24] However, since in those cases, the synthetic TF is recruited at a single site, there are fewer activator copies acting on the promoter compared with our system. Certain TFs, such as CRX, have been reported to be able to switch from activation to repression with an increase in the number of binding sites.[75] Moreover, further experiments with a larger panel of strong activator-strong activator pairs would help clarify how many strong activators are dual-functional domains, whether non-dual strong activator pairs act antagonistically as dual-dual pairs appear to do, and whether this phenomenon depends on the number of binding sites.

Our activation data showed a tight coupling between the fraction of actively transcribing cells and the degree of transcription in those cells, consistent with control of transcriptional bursting dynamics.[52,76] Recently, certain TFs were shown to increase only burst size or only burst frequency, depending on their molecular mechanism of action.[77] Although our long-lived protein reporter does not allow us to differentiate between changes in burst size or frequency for our combinations, it would be interesting to repeat these high-throughput measurements coupled with RNA FISH or a destabilized reporter that allows the extraction of burst parameters.

We found that repressors generally overpowered activators, although select activator domains could weaken the repressor function of dual-functional or weakly repressive domains. We confirmed this result more thoroughly via an in-depth screen measuring KRAB function when paired with a panel of 5,000 domains from nuclearly localized proteins and found that KRAB maintained its repressive function when paired with most domains except ones that had decreased expression and a few domains from general TAFs and RNA polymerase. We used these insights to build a synthetic transcription factor whose function could be switched from a dual-functional FOXO3-like profile (that activates a minimal promoter and weakly represses a constitutive one) to a repressive KRAB-like profile via the addition of the small molecule drug ASV. We used this switchable TF to show that both the activation and repression functions of the FOXO3 effector domain changed when the target promoter was first silenced by KRAB. Although activation from a minimal promoter was impaired by prior silencing as expected, recruitment of FOXO3 increased KRAB-mediated epigenetic memory and reduced reporter gene reactivation at the pEF promoter. Importantly, since ASV tuned the behavior of this synthetic TF independent of dox-mediated recruitment, we could generate gene expression profiles that were not accessible with either individual effector domain alone. Looking forward, this synthetic switchable TF can be used to split a cell population into well-defined percentages that express the desired combination of two target genes, e.g., 50% of the cells expressing gene 1 and 50% expressing gene 2. Such flexibility could allow for more sophisticated and complex gene circuits and the engineering of "higher-order" cell behaviors and programs.[78]

When we started this study, we had a limited number of validated activation and repressive domains to test in combination. Recently, however, we have identified a much larger set of effector domains from human TFs.[20] Prior efforts have characterized individual examples of multiple effector domains being combined within a larger TF[22,79]; it would be instructive to design a library that systematically tests combinations of domains that come from the same TF and compare these results with recent ORFeome screens measuring activation and repression of full-length TFs.[33] Moreover, emerging work has begun to connect TF and effector function with measurements of their affinities for transcriptional coactivators and corepressors.[13,33,80] Interactions with cofactors and hence the effect of each combination could vary with context such as target promoter, number of binding sites, and cell type.[9,12,81–84] Our work characterized effector combinations at two promoters in one human cell line, and our system for effector recruitment featured 9 copies of the TetO-binding site upstream of the promoter. Therefore, some of the synergistic or antagonistic effects that we have identified could be a result of binding stoichiometry, interactions between different transcription factor molecules, and the number of cofactors recruited to the promoter. We are excited about future efforts to broaden these investigations to include other DNA-binding domains beyond rTetR, other promoters, and other cell types. Such efforts will undoubtedly aid projects aiming to build robust tools for cell engineering that can function across multiple contexts.[85–87]

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead Contact
  - Materials Availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Cell lines and cell culture
- METHOD DETAILS
  - Library design
  - Co-recruit cloning
  - Co-recruit screen
  - Co-recruit library prep and sequencing
  - Computing co-recruit enrichment scores
  - Labeling domains as activators, repressors, and duals
  - KRAB-Nuclear Pfam screen cloning and cell culture
  - Magnetic Separation
  - Low-throughput validations
- MODEL OF TRANSCRIPTIONAL REPRESSION
  - Flow cytometry analysis
  - Western Blot
  - CFSE

- QUANTIFICATION AND STATISTICAL ANALYSIS
- ADDITIONAL INFORMATION

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cels.2023.07.001.

ACKNOWLEDGMENTS

We would like to thank Michaela Hinks, Jennifer Gucwa, Nicole DelRosso, Abby Thurm, David Yao, Priyanka Shrestha, Zeppelin Cat, Ophelia Hinks, Shivam Verma, Mary Frances Gallagher, Kaushik Ragunathan, Zara Weinberg, and all members of the Bintu lab for helpful conversations and assistance. We would like to thank Anshul Kundaje and Will Greenleaf for valuable feedback throughout the project. A.X.M. was supported by Stanford University Medical Scientist Training Program grants T32-GM007365 and T32-GM145402, and a Stanford Bio-X SIGF. J.T. is supported by the F99/K00 fellowship of the National Institutes of Health (NIH-1F99DK126120-01 and NIH-4K00DK126120-03). S.A.R. is supported by an NSF GRFP (DGE-1656518). J.S. is supported by NIH Training Program grant T32-GM113854. M.C.B. is supported by a grant from Stanford ChEM-H and an NIH Director's New Innovator Award (1DP2HD08406901). This work was supported by BWF CASI (L.B.), NIH-NIGMS R35M128947 (L.B.), and NIH-NIGRI R01HG011866 (L.B. and M.C.B.).

AUTHOR CONTRIBUTIONS

A.X.M. and L.B. designed the study with significant intellectual contributions from J.T. A.X.M. and J.T. designed the combinatorial recruitment library and performed the KRAB-Pfam screen. K.S. cloned the KRAB-Pfam library. A.X.M. performed the combinatorial recruitment screen with help from C.L. A.X.M., S.J.A., S.A.R., and C.A. all performed individual recruitment assay experiments. A.X.M. performed data analysis with input from L.B., J.T., and M.C.B. A.X.M. performed CellTrace proliferation experiments. A.X.M. and J.S. performed western blot experiments. A.X.M. and L.B. wrote the manuscript, with contributions and feedback from all authors. L.B. supervised the project.

DECLARATION OF INTERESTS

L.B., M.C.B., and J.T. acknowledge outside interest in Stylus Medicine.

INCLUSION AND DIVERSITY

We support inclusive, diverse, and equitable conduct of research. One or more of the authors of this paper self-identifies as a gender minority in their field of research. One or more of the authors of this paper self-identifies as a member of the LGBTQIA+ community. One or more of the authors of this paper self-identifies as living with a disability.

Received: December 20, 2022
Revised: June 26, 2023
Accepted: July 6, 2023
Published: August 4, 2023

REFERENCES

1. Frietze, S., and Farnham, P.J. (2011). Transcription factor effector domains. In A Handbook of Transcription Factors, T.R. Hughes, ed. (Springer Netherlands), pp. 261–277. https://doi.org/10.1007/978-90-481-9069-0_12.
2. Soto, L.F., Li, Z., Santoso, C.S., Berenson, A., Ho, I., Shen, V.X., Yuan, S., and Fuxman Bass, J.I. (2022). Compendium of human transcription factor effector domains. Mol. Cell *82*, 514–526. https://doi.org/10.1016/j.molcel.2021.11.007.
3. Lambert, S.A., Jolma, A., Campitelli, L.F., Das, P.K., Yin, Y., Albu, M., Chen, X., Taipale, J., Hughes, T.R., and Weirauch, M.T. (2018). The human transcription factors. Cell *172*, 650–665. https://doi.org/10.1016/j.cell.2018.01.029.
4. Stegmaier, P., Kel, A.E., and Wingender, E. (2004). Systematic DNA-binding domain classification of transcription factors. Genome Inform. *15*, 276–286. https://doi.org/10.11234/gi1990.15.2_276.
5. Harrison, S.C. (1991). A structural taxonomy of DNA-binding domains. Nature *353*, 715–719. https://doi.org/10.1038/353715a0.
6. Hathaway, N.A., Bell, O., Hodges, C., Miller, E.L., Neel, D.S., and Crabtree, G.R. (2012). Dynamics and memory of heterochromatin in living cells. Cell *149*, 1447–1460. https://doi.org/10.1016/j.cell.2012.03.052.
7. Keung, A.J., Bashor, C.J., Kiriakov, S., Collins, J.J., and Khalil, A.S. (2014). Using targeted chromatin regulators to engineer combinatorial and spatial transcriptional regulation. Cell *158*, 110–120. https://doi.org/10.1016/j.cell.2014.04.047.
8. Kungulovski, G., Nunna, S., Thomas, M., Zanger, U.M., Reinhardt, R., and Jeltsch, A. (2015). Targeted epigenome editing of an endogenous locus with chromatin modifiers is not stably maintained. Epigenetics Chromatin *8*, 12. https://doi.org/10.1186/s13072-015-0002-z.
9. Amabile, A., Migliara, A., Capasso, P., Biffi, M., Cittaro, D., Naldini, L., and Lombardo, A. (2016). Inheritable silencing of endogenous genes by hit-and-run targeted epigenetic editing. Cell *167*, 219–232.e14. https://doi.org/10.1016/j.cell.2016.09.006.
10. Bintu, L., Yong, J., Antebi, Y.E., McCue, K., Kazuki, Y., Uno, N., Oshimura, M., and Elowitz, M.B. (2016). Dynamics of epigenetic regulation at the single-cell level. Science *351*, 720–724. https://doi.org/10.1126/science.aab2956.
11. Tycko, J., Van, M.V., Elowitz, M.B., and Bintu, L. (2017). Advancing towards a global mammalian gene regulation model through single-cell analysis and synthetic biology. Eng. *4*, 174–193. https://doi.org/10.1016/j.cobme.2017.10.011.
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. (2017). DCas9-based epigenome editing suggests acquisition of histone methylation is not sufficient for target gene repression. Nucleic Acids Res. *45*, 9901–9916. https://doi.org/10.1093/nar/gkx578.
13. Sanborn, A.L., Yeh, B.T., Feigerle, J.T., Hao, C.V., Townshend, R.J., Lieberman Aiden, E., Dror, R.O., and Kornberg, R.D. (2021). Simple biochemical features underlie transcriptional activation domain diversity and dynamic, fuzzy binding to Mediator. eLife *10*, e68068. https://doi.org/10.7554/eLife.68068.
14. Lee, J.B., Caywood, L.M., Lo, J.Y., Levering, N., and Keung, A.J. (2021). Mapping the dynamic transfer functions of eukaryotic gene regulation. Cell Syst. *12*, 1079–1093.e6. https://doi.org/10.1016/j.cels.2021.08.003.
15. Chasman, D.I., Leatherwood, J., Carey, M., Ptashne, M., and Kornberg, R.D. (1989). Activation of yeast polymerase II transcription by herpesvirus VP16 and GAL4 derivatives in vitro. Mol. Cell. Biol. *9*, 4746–4749. https://doi.org/10.1128/mcb.9.11.4746-4749.1989.
16. Furth, N., and Shema, E. (2022). It's all in the combination: Decoding the epigenome for cancer research and diagnostics. Curr. Opin. Genet. Dev. *73*, 101899. https://doi.org/10.1016/j.gde.2022.101899.
17. Wolberger, C. (1998). Combinatorial transcription factors. Curr. Opin. Genet. Dev. *8*, 552–559. https://doi.org/10.1016/s0959-437x(98)80010-5.
18. Ravasi, T., Suzuki, H., Cannistraci, C.V., Katayama, S., Bajic, V.B., Tan, K., Akalin, A., Schmeier, S., Kanamori-Katayama, M., Bertin, N., et al. (2010). An atlas of combinatorial transcriptional regulation in mouse and man. Cell *140*, 744–752. https://doi.org/10.1016/j.cell.2010.01.044.
19. Reiter, F., Wienerroither, S., and Stark, A. (2017). Combinatorial function of transcription factors and cofactors. Curr. Opin. Genet. Dev. *43*, 73–81. https://doi.org/10.1016/j.gde.2016.12.007.
20. DelRosso, N., Tycko, J., Suzuki, P., Andrews, C., Aradhana, Mukund, A., Liongson, I., Ludwig, C., Spees, K., Fordyce, P., et al. (2023). Large-scale mapping and mutagenesis of human transcriptional effector domains. Nature *616*, 365–372. https://doi.org/10.1038/s41586-023-05906-y.

21. Tycko, J., DelRosso, N., Hess, G.T., Aradhana, Banerjee, A., Mukund, A., Van, M.V., Ego, B.K., Yao, D., Spees, K., et al. (2020). High-throughput discovery and characterization of human transcriptional effectors. Cell *183*, 2020–2035.e16. https://doi.org/10.1016/j.cell.2020.11.024.

22. Losson, R., and Nielsen, A.L. (2010). The NIZP1 KRAB and C2HR domains cross-talk for transcriptional regulation. Biochim. Biophys. Acta *1799*, 463–468. https://doi.org/10.1016/j.bbagrm.2010.02.003.

23. Beerli, R.R., Segal, D.J., Dreier, B., and Barbas, C.F., 3rd (1998). Toward controlling gene expression at will: Specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks. Proc. Natl. Acad. Sci. USA *95*, 14628–14633. https://doi.org/10.1073/pnas.95.25.14628.

24. Chavez, A., Scheiman, J., Vora, S., Pruitt, B.W., Tuttle, M., P R Iyer, E., Lin, S., Kiani, S., Guzman, C.D., Wiegand, D.J., et al. (2015). Highly efficient Cas9-mediated transcriptional programming. Nat. Methods *12*, 326–328. https://doi.org/10.1038/nmeth.3312.

25. Nuñez, J.K., Chen, J., Pommier, G.C., Cogan, J.Z., Replogle, J.M., Adriaens, C., Ramadoss, G.N., Shi, Q., Hung, K.L., Samelson, A.J., et al. (2021). Genome-wide programmable transcriptional memory by CRISPR-based epigenome editing. Cell *184*, 2503–2519.e17. https://doi.org/10.1016/j.cell.2021.03.025.

26. Thakore, P.I., Black, J.B., Hilton, I.B., and Gersbach, C.A. (2016). Editing the epigenome: Technologies for programmable transcription and epigenetic modulation. Nat. Methods *13*, 127–137. https://doi.org/10.1038/nmeth.3733.

27. Holtzman, L., and Gersbach, C.A. (2018). Editing the epigenome: Reshaping the genomic landscape. Annu. Rev. Genomics Hum. Genet. *19*, 43–71. https://doi.org/10.1146/annurev-genom-083117-021632.

28. Black, J.B., and Gersbach, C.A. (2018). Synthetic transcription factors for cell fate reprogramming. Curr. Opin. Genet. Dev. *52*, 13–21. https://doi.org/10.1016/j.gde.2018.05.001.

29. Elkon, R., and Agami, R. (2017). Characterization of noncoding regulatory DNA in the human genome. Nat. Biotechnol. *35*, 732–746. https://doi.org/10.1038/nbt.3863.

30. Erijman, A., Kozlowski, L., Sohrabi-Jahromi, S., Fishburn, J., Warfield, L., Schreiber, J., Noble, W.S., Söding, J., and Hahn, S. (2020). A high-throughput screen for transcription activation domains reveals their sequence features and permits prediction by deep learning. Mol. Cell *79*, 1066. https://doi.org/10.1016/j.molcel.2020.08.013.

31. Staller, M.V., Holehouse, A.S., Swain-Lenz, D., Das, R.K., Pappu, R.V., and Cohen, B.A. (2018). A high-throughput mutational scan of an intrinsically disordered acidic transcriptional activation domain. Cell Syst. *6*, 444–455.e6. https://doi.org/10.1016/j.cels.2018.01.015.

32. Arnold, C.D., Nemčko, F., Woodfin, A.R., Wienerroither, S., Vlasova, A., Schleiffer, A., Pagani, M., Rath, M., and Stark, A. (2018). A high-throughput method to identify trans-activation domains within transcription factor sequences. EMBO J. *37*, e98896. https://doi.org/10.15252/embj.201798896.

33. Alerasool, N., Leng, H., Lin, Z.-Y., Gingras, A.-C., and Taipale, M. (2022). Identification and functional characterization of transcriptional activators in human cells. Mol. Cell *82*, 677–695.e7. https://doi.org/10.1016/j.molcel.2021.12.008.

34. Staller, M.V., Ramirez, E., Kotha, S.R., Holehouse, A.S., Pappu, R.V., and Cohen, B.A. (2022). Directed mutational scanning reveals a balance between acidic and hydrophobic residues in strong human activation domains. Cell Syst. *13*, 334–345.e5. https://doi.org/10.1016/j.cels.2022.01.002.

35. Gao, Y., Xiong, X., Wong, S., Charles, E.J., Lim, W.A., and Qi, L.S. (2016). Complex transcriptional modulation with orthogonal and inducible dCas9 regulators. Nat. Methods *13*, 1043–1049. https://doi.org/10.1038/nmeth.4042.

36. DelRosso, N., Tycko, J., Suzuki, P., Andrews, C., Aradhana, Mukund, A., Liongson, I., Ludwig, C., Spees, K., Fordyce, P., et al. (2022). Large-scale mapping and systematic mutagenesis of human transcriptional effector domains. bioRxiv. https://doi.org/10.1101/2022.08.26.505496.

37. Cura, V., and Cavarelli, J. (2021). Structure, activity and function of the PRMT2 protein arginine methyltransferase. Life (Basel) *11*, 1263. https://doi.org/10.3390/life11111263.

38. Pak, M.L., Lakowski, T.M., Thomas, D., Vhuiyan, M.I., Hüsecken, K., and Frankel, A. (2011). A protein arginine N-methyltransferase 1 (PRMT1) and 2 heteromeric interaction increases PRMT1 enzymatic activity. Biochemistry *50*, 8226–8240. https://doi.org/10.1021/bi200644c.

39. Hou, W., Nemitz, S., Schopper, S., Nielsen, M.L., Kessels, M.M., and Qualmann, B. (2018). Arginine methylation by PRMT2 controls the functions of the actin nucleator Cobl. Dev. Cell *45*, 262–275.e8. https://doi.org/10.1016/j.devcel.2018.03.007.

40. Vhuiyan, M.I., Pak, M.L., Park, M.A., Thomas, D., Lakowski, T.M., Chalfant, C.E., and Frankel, A. (2017). PRMT2 interacts with splicing factors and regulates the alternative splicing of BCL-X. J. Biochem. *162*, 17–25. https://doi.org/10.1093/jb/mvw102.

41. Kremerskothen, J., Plaas, C., Büther, K., Finger, I., Veltel, S., Matanis, T., Liedtke, T., and Barnekow, A. (2003). Characterization of KIBRA, a novel WW domain-containing protein. Biochem. Biophys. Res. Commun. *300*, 862–867. https://doi.org/10.1016/s0006-291x(02)02945-5.

42. Xiao, L., Chen, Y., Ji, M., and Dong, J. (2011). KIBRA regulates Hippo signaling activity via interactions with large tumor suppressor kinases. J. Biol. Chem. *286*, 7788–7796. https://doi.org/10.1074/jbc.M110.173468.

43. Gordon, W.R., Vardar-Ulu, D., Histen, G., Sanchez-Irizarry, C., Aster, J.C., and Blacklow, S.C. (2007). Structural basis for autoinhibition of Notch. Nat. Struct. Mol. Biol. *14*, 295–300. https://doi.org/10.1038/nsmb1227.

44. Bray, S.J. (2016). Notch signalling in context. Nat. Rev. Mol. Cell Biol. *17*, 722–735. https://doi.org/10.1038/nrm.2016.94.

45. Trojanowski, J., Frank, L., Rademacher, A., Mücke, N., Grigaitis, P., and Rippe, K. (2022). Transcription activation is enhanced by multivalent interactions independent of phase separation. Mol. Cell *82*, 1878–1893.e10. https://doi.org/10.1016/j.molcel.2022.04.017.

46. Desterro, J.M., Rodriguez, M.S., and Hay, R.T. (2000). Regulation of transcription factors by protein degradation. Cell. Mol. Life Sci. *57*, 1207–1219. https://doi.org/10.1007/pl00000760.

47. Raj, A., Peskin, C.S., Tranchina, D., Vargas, D.Y., and Tyagi, S. (2006). Stochastic mRNA synthesis in mammalian cells. PLOS Biol. *4*, e309. https://doi.org/10.1371/journal.pbio.0040309.

48. Friedman, N., Cai, L., and Xie, X.S. (2006). Linking stochastic dynamics to population distribution: An analytical framework of gene expression. Phys. Rev. Lett. *97*, 168302. https://doi.org/10.1103/PhysRevLett.97.168302.

49. Dar, R.D., Razooky, B.S., Singh, A., Trimeloni, T.V., McCollum, J.M., Cox, C.D., Simpson, M.L., and Weinberger, L.S. (2012). Transcriptional burst frequency and burst size are equally modulated across the human genome. Proc. Natl. Acad. Sci. USA *109*, 17454–17459. https://doi.org/10.1073/pnas.1213530109.

50. Peccoud, J., and Ycart, B. (1995). Markovian modeling of gene-product synthesis. Theor. Popul. Biol. *48*, 222–234. https://doi.org/10.1006/tpbi.1995.1027.

51. Hao, N., and O'Shea, E.K. (2011). Signal-dependent dynamics of transcription factor translocation controls gene expression. Nat. Struct. Mol. Biol. *19*, 31–39. https://doi.org/10.1038/nsmb.2192.

52. Mukund, A., and Bintu, L. (2022). Temporal signaling, population control, and information processing through chromatin-mediated gene regulation. J. Theor. Biol. *535*, 110977. https://doi.org/10.1016/j.jtbi.2021.110977.

53. Martinez-Corral, R., Park, M., Biette, K., Friedrich, D., Scholes, C., Khalil, A.S., Gunawardena, J., and DePace, A.H. (2020). Transcriptional kinetic synergy: A complex landscape revealed by integrating modelling and synthetic biology. bioRxiv. https://doi.org/10.1101/2020.08.31.276261.

54. Park, H., Wahl, M.I., Afar, D.E., Turck, C.W., Rawlings, D.J., Tam, C., Scharenberg, A.M., Kinet, J.P., and Witte, O.N. (1996). Regulation of Btk function by a major autophosphorylation site within the SH3 domain. Immunity *4*, 515–525. https://doi.org/10.1016/s1074-7613(00)80417-3.

55. Mohamed, A.J., Vargas, L., Nore, B.F., Backesjo, C.M., Christensson, B., and Smith, C.I. (2000). Nucleocytoplasmic shuttling of Bruton's tyrosine kinase. J. Biol. Chem. *275*, 40614–40619. https://doi.org/10.1074/jbc.M006952200.

56. 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). Mechanism for DPY30 and ASH2L intrinsically disordered regions to modulate the MLL/SET1 activity on chromatin. Nat. Commun. *12*, 2953. https://doi.org/10.1038/s41467-021-23268-9.

57. Tando, T., Ishizaka, A., Watanabe, H., Ito, T., Iida, S., Haraguchi, T., Mizutani, T., Izumi, T., Isobe, T., Akiyama, T., et al. (2010). Requiem protein links RelB/p52 and the Brm-type SWI/SNF complex in a noncanonical NF-κB pathway. J. Biol. Chem. *285*, 21951–21960. https://doi.org/10.1074/jbc.M109.087783.

58. Kim, G.-D., Ni, J., Kelesoglu, N., Roberts, R.J., and Pradhan, S. (2002). Co-operation and communication between the human maintenance and de novo DNA (cytosine-5) methyltransferases. EMBO J. *21*, 4183–4195. https://doi.org/10.1093/emboj/cdf401.

59. Larson, M.H., Gilbert, L.A., Wang, X., Lim, W.A., Weissman, J.S., and Qi, L.S. (2013). CRISPR interference (CRISPRi) for sequence-specific control of gene expression. Nat. Protoc. *8*, 2180–2196. https://doi.org/10.1038/nprot.2013.132.

60. Jacobs, C.L., Badiee, R.K., and Lin, M.Z. (2018). StaPLs: Versatile genetically encoded modules for engineering drug-inducible proteins. Nat. Methods *15*, 523–526. https://doi.org/10.1038/s41592-018-0041-z.

61. Reményi, A., Schöler, H.R., and Wilmanns, M. (2004). Combinatorial control of gene expression. Nat. Struct. Mol. Biol. *11*, 812–815. https://doi.org/10.1038/nsmb820.

62. Zeitlinger, J. (2020). Seven myths of how transcription factors read the cis-regulatory code. Curr. Opin. Syst. Biol. *23*, 22–31. https://doi.org/10.1016/j.coisb.2020.08.002.

63. Charest, J., Daniele, T., Wang, J., Bykov, A., Mandlbauer, A., Asparuhova, M., Röhsner, J., Gutiérrez-Pérez, P., and Cochella, L. (2020). Combinatorial action of temporally segregated transcription factors. Dev. Cell *55*, 483–499.e7. https://doi.org/10.1016/j.devcel.2020.09.002.

64. Bashor, C.J., Patel, N., Choubey, S., Beyzavi, A., Kondev, J., Collins, J.J., and Khalil, A.S. (2019). Complex signal processing in synthetic gene circuits using cooperative regulatory assemblies. Science *364*, 593–597. https://doi.org/10.1126/science.aau8287.

65. Roybal, K.T., Rupp, L.J., Morsut, L., Walker, W.J., McNally, K.A., Park, J.S., and Lim, W.A. (2016). Precision tumor recognition by T cells with combinatorial antigen-sensing circuits. Cell *164*, 770–779. https://doi.org/10.1016/j.cell.2016.01.011.

66. Weinberg, Z.Y., Hilburger, C.E., Kim, M., Cao, L., Khalid, M., Elmes, S., Diwanji, D., Hernandez, E., Lopez, J., Schaefer, K., et al. (2021). Sentinel cells enable genetic detection of SARS-CoV-2 Spike protein. bioRxiv. https://doi.org/10.1101/2021.04.20.440678.

67. Zhu, I., Liu, R., Garcia, J.M., Hyrenius-Wittsten, A., Piraner, D.I., Alavi, J., Israni, D.V., Liu, B., Khalil, A.S., and Roybal, K.T. (2022). Modular design of synthetic receptors for programmed gene regulation in cell therapies. Cell *185*, 1431–1443.e16. https://doi.org/10.1016/j.cell.2022.03.023.

68. 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. https://doi.org/10.1016/j.cels.2022.09.006.

69. Ecco, G., Imbeault, M., and Trono, D. (2017). KRAB zinc finger proteins. Development *144*, 2719–2729. https://doi.org/10.1242/dev.132605.

70. Liu, J., Perumal, N.B., Oldfield, C.J., Su, E.W., Uversky, V.N., and Dunker, A.K. (2006). Intrinsic disorder in transcription factors. Biochemistry *45*, 6873–6888. https://doi.org/10.1021/bi0602718.

71. Ravarani, C.N., Erkina, T.Y., De Baets, G., Dudman, D.C., Erkine, A.M., and Babu, M.M. (2018). High-throughput discovery of functional disordered regions: Investigation of transactivation domains. Mol. Syst. Biol. *14*, e8190. https://doi.org/10.15252/msb.20188190.

72. 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. https://doi.org/10.1038/nature14136.

73. 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. https://doi.org/10.1016/0092-8674(90)90397-w.

74. Gill, G., and Ptashne, M. (1988). Negative effect of the transcriptional activator GAL4. Nature *334*, 721–724. https://doi.org/10.1038/334721a0.

75. White, M.A., Kwasnieski, J.C., Myers, C.A., Shen, S.Q., Corbo, J.C., and Cohen, B.A. (2016). A simple grammar defines activating and repressing cis-regulatory elements in photoreceptors. Cell Rep. *17*, 1247–1254. https://doi.org/10.1016/j.celrep.2016.09.066.

76. Ali Al-Radhawi, M., Del Vecchio, D., and Sontag, E.D. (2019). Multi-modality in gene regulatory networks with slow promoter kinetics. PLoS Comput. Biol. *15*, e1006784. https://doi.org/10.1371/journal.pcbi.1006784.

77. Mamrak, N.E., Alerasool, N., Griffith, D., Holehouse, A.S., Taipale, M., and Lionnet, T. (2022). The kinetic landscape of human transcription factors. bioRxiv. https://doi.org/10.1101/2022.06.01.494187.

78. Lim, W.A. (2022). The emerging era of cell engineering: Harnessing the modularity of cells to program complex biological function. Science *378*, 848–852. https://doi.org/10.1126/science.add9665.

79. 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. https://doi.org/10.1074/jbc.272.7.4021.

80. Jacobs, J., Pagani, M., Wenzl, C., and Stark, A. (2022). Widespread regulatory specificities between transcriptional corepressors and enhancers in *Drosophila*. bioRxiv. https://doi.org/10.1101/2022.11.07.515017.

81. O'Geen, H., Tomkova, M., and Combs, J.A. (2022). Determinants of heritable gene silencing for KRAB-dCas9+ DNMT3 and Ezh2-dCas9+ DNMT3 hit-and-run epigenome editing. Nucleic Acids Res. *50*, 3239–3253. https://doi.org/10.1093/nar/gkac123.

82. Zhao, S., Hong, C.K.Y., Myers, C.A., Granas, D.M., White, M.A., Corbo, J.C., and Cohen, B.A. (2023). A single-cell massively parallel reporter assay detects cell-type-specific gene regulation. Nat. Genet. *55*, 346–354. https://doi.org/10.1038/s41588-022-01278-7.

83. Hong, C.K.Y., and Cohen, B.A. (2022). Genomic environments scale the activities of diverse core promoters. Genome Res. *32*, 85–96. https://doi.org/10.1101/gr.276025.121.

84. Cano-Rodriguez, D., Gjaltema, R.A.F., Jilderda, L.J., Jellema, P., Dokter-Fokkens, J., Ruiters, M.H.J., and Rots, M.G. (2016). Writing of H3K4Me3 overcomes epigenetic silencing in a sustained but context-dependent manner. Nat. Commun. *7*, 12284. https://doi.org/10.1038/ncomms12284.

85. 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. https://doi.org/10.1016/j.cels.2021.05.011.

86. 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. https://doi.org/10.1101/2020.05.30.122077.

87. Baetica, A.-A., Westbrook, A., and El-Samad, H. (2019). Control theoretical concepts for synthetic and systems biology. Curr. Opin. Syst. Biol. *14*, 50–57. https://doi.org/10.1016/j.coisb.2019.02.010.

88. Teague, B. (2022). Cytoflow: A python toolbox for flow cytometry. bioRxiv. https://doi.org/10.1101/2022.07.22.501078.

89. Zulkower, V., and Rosser, S. (2020). DNA Chisel, a versatile sequence optimizer. Bioinformatics *36*, 4508–4509. https://doi.org/10.1093/bioinformatics/btaa558.

Please cite this article in press as: Mukund et al., High-throughput functional characterization of combinations of transcriptional activators and repressors, Cell Systems (2023), https://doi.org/10.1016/j.cels.2023.07.001

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
FLAG M2 monoclonal antibody, 1:1000, mouse	Sigma-Aldrich	F1804; RRID: AB_262044
Histone H3 antibody, 1:2000, rabbit	Abcam	AB1791; RRID: AB_302613
HRP-conjugated secondary antibody, 1:2000	Cytiva	NA934/931; RRID: AB_772206
Bacterial and Virus Strains		
Endura Electrocompetent Cells	Lucigen	60242-2
Chemicals, Peptides, and Recombinant Proteins		
Doxycycline hyclate	Tocris	4090
Blasticidin	Gibco	A1113903
Asunaprevir	Sigma-Aldrich	ADV465749194
Critical Commercial Assays		
Dynabeads M-280 Protein G	Thermo Fisher	10003D
NEBNext Ultra II	NEB	M0544
Deposited Data		
Raw sequencing files, Co-Recruit	This manuscript	GEO GSE235591
Raw sequencing files, KRAB-Pfam screen	This manuscript	GEO GSE235699
Experimental Models: Cell Lines		
Human: HEK293T Lenti-X	Takara	632180
Human: K562	ATCC	CCL-243
Human: K562 + pJT039	Tycko et al. 2020[21]	N/A
Human: K562 + pDY032	Tycko et al. 2020[21]	N/A
Oligonucleotides		
Co-recruit combinatorial library	Twist	N/A
Pfam nuclear domain library	Tycko et al. 2020[21]	N/A
Variable Linker	This manuscript	TCGGGNAGCGAGACNCCNGGNACN TCNGAGTCNGCNACNCCNGAGAGC
Recombinant DNA		
pJT039 AAVS1-PuroR-9xTetO-pEF-IGKleader-hIgG1_FC-Myc-PDGFRb-T2A-Citrine-PolyA	Tycko et al. 2020[21]	Addgene #161927
pDY32 AAVS1-PuroR-9xTetO-minCMV-IGKleader-hIgG1_FC-Myc-PDGFRb-T2A-Citrine-PolyA	Tycko et al. 2020[21]	Addgene #161928
pJT126 lenti pEF-rTetR(SE-G72P)-3XFLAG-LibCloneSite-T2A-mCherry-BSD-WPRE	Tycko et al. 2020[21]	Addgene #161926
pMD2.G	Didier Trono	Addgene #12259
pRSV-Rev	Didier Trono	Addgene #12253
pMDLg/pRRE	Didier Trono	Addgene #12251
Software and Algorithms		
Custom analyses	This manuscript	Zenodo https://doi.org/10.5281/zenodo.7453681
Cytoflow	Teague 2022[88]	Github: https://cytoflow.github.io
Other		
Miscellaneous data and resources	This manuscript	Zenodo https://doi.org/10.5281/zenodo.7453681

Please cite this article in press as: Mukund et al., High-throughput functional characterization of combinations of transcriptional activators and repressors, Cell Systems (2023), https://doi.org/10.1016/j.cels.2023.07.001

RESOURCE AVAILABILITY

Lead Contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Lacramioara Bintu (lbintu@stanford.edu).

Materials Availability

Reporter and effector plasmids were previously published and are listed in the key resources table.

Data and code availability

All data besides high-throughput sequencing data and all code used for analysis has been deposited at Zenodo with https://doi.org/10.5281/zenodo.7453681. High-throughput sequencing data has been deposited for the co-recruit screen in the NCBI GEO database with accession number GSE235591. High-throughput sequencing data has been deposited for the KRAB screen in the NCBI GEO database with accession number GSE235699. Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell lines and cell culture

Cell culture was performed as described in.[36] Briefly, all experiments presented here were carried out in K562 cells (ATCC, CCL-243, female). Cells were cultured in a controlled humidified incubator at 37C and 5% CO2, in RPMI 1640 (Gibco, 11-875-119) media supplemented with 10% FBS (Omega Scientific, 20014T), and 1% Penicillin-Streptomycin-Glutamine (Gibco, 10378016). HEK293T-LentiX (Takara Bio, 632180, female) cells, used to produce lentivirus, as described below, were grown in DMEM (Gibco, 10569069) media supplemented with 10% FBS (Omega Scientific, 20014T) and 1% Penicillin Streptomycin Glutamine (Gibco, 10378016). minCMV and pEF reporter cell line generation is described in.[21] pEF and minCMV promoter reporter cell lines were generated by TALEN-mediated homology-directed repair to integrate donor constructs (pEF promoter: Addgene #161927, minCMV promoter: Addgene #161928) into the *AAVS1* locus by electroporation of K562 cells with 1000 ng of reporter donor plasmid and 500 ng of each TALEN-L (Addgene #35431) and TALEN-R (Addgene #35432) plasmid (targeting upstream and downstream the intended DNA cleavage site, respectively). After 7 days, the cells were treated with 1000 ng/mL puromycin antibiotic for 5 days to select for a population where the donor was stably integrated in the intended locus. Fluorescent reporter expression was measured by flow cytometry.

METHOD DETAILS

Library design

Library members were chosen from the Nuclear Pfam library described in.[21] A total of 20 negative control domains, 10 randomers and 10 tiles of the DMD protein; 30 activator domains; and 50 repressor domains were chosen for cloning. Of the 50 repressors, 6 were discovered to have errors introduced during design, and were discarded from downstream analysis. All domains, including those eventually discarded, are listed in Table S1. After library assembly, DNAChisel[89] was used to optimize coding sequences by removing duplicates, 7xC homopolymers, BsmBI and BbsI restriction sites, and rare codons. Codon usage was matched to human codon prevalence and GC content was restricted to be between 20% and 75% in any 50-nucleotide window and between 25% and 65% globally.

Co-recruit cloning

Cloning for the combinatorial screen proceeded in two stages; in the first, domain-linker-domain concatenations were assembled, and in the second, concatenations were placed into a lentiviral backbone. Domains in the N-terminal and C-terminal position were synthesized as 2 separate oligonucleotide pools (Twist Biosciences). Each pool was PCR amplified in a clean PCR hood to avoid DNA contamination. Each pool was split into 6x 50 uL reactions that were PCR amplified for 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 10mM dNTPs, and 10 uL of 5x Herculase buffer (Agilent). Reaction mixes were thermocycled at 98C for 3m; then 21x cycles of 98C for 20s, 61C for 20s, and 72C for 30s; and, finally, 72C for 3m. Reaction products were pooled and gel extracted by loading a 1% TAE gel, excising the 300 bp band, and purifying using a Zymo Research gel extraction kit. Each of the 2 extraction products (N-terminal and C-terminal) were then amplified for 23x cycles each using the same protocol in order to generate sufficient DNA for downstream reactions. To generate concatenations, 5 uL of 15 ng/uL product from each product, 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 M0202M), and 1uL of an XTEN linker-encoding DNA amplicon. This linker was encoded with variable nucleotides in codon wobble positions to permit variable DNA sequence in between domains and prevent lentiviral recombination without varying the amino acid sequence of the linker. To complete the GoldenGate reaction, the mix was thermocycled for 65x cycles at 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 product was run on a 1% TAE gel and the 600bp band of domain pair concatenations was extracted with a Zymo gel

extraction kit. The concatenation product was then PCR amplified as before, using the outer primers, for 15x cycles. Amplification products were then split into 12x GoldenGate reactions, each consisting of 75ng (1.33 uL) Bsmbi-v2 digested pJT039 lentiviral backbone, 10 ng (0.64 uL) concatenation amplicon, 2 uL T4 buffer, 1 uL BsmbI-v2 Golden Gate Assembly Kit (NEB E1602L), and 15.03 uL nuclease free H2O. Reactions were thermocycled with 65x cycles at 42C for 5m and 16C for 5m, before a final incubation at 42C for 5m and 70C for 20m. The reactions were then pooled and purified with a QIAgen MinElute column, eluting in 6 uL ddH2O. 2 uL of eluent was then electroporated into each of 2 tubes of 50 uL Endura electrocompetent cells (Lucigen, Cat#60242-2) as per manufacturer's instructions. Cells were then plated onto 4x large 10'' x 10'' LB plates with carbenicillin. After overnight growth colonies were scraped into a collection bottle and plasmid pools were extracted using a plasmid maxiprep kit (Qiagen #12662). 2 smaller plates of LB with carbenicillin were also prepared with 1:20 diluted cells to count colonies and confirm transformation efficiency. Approximately 50 colonies were Sanger sequenced (Quintara) to estimate cloning efficiencies and the proportion of empty backbone plasmids in the pool.

Co-recruit screen

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

Co-recruit library prep and sequencing

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

Computing co-recruit enrichment scores

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

Labeling domains as activators, repressors, and duals

We labeled a domain as an activator if the following 2 conditions held for all concatenations including that domain and a negative control: 1) at least 19% functioned as activators (i.e. for each, activation $\log_2$(ON:OFF) > -0.0236); 2) the average activation $\log_2$(ON:OFF) for all pairs was greater than or equal to -0.5, or the activation $\log_2$(ON:OFF) for all pairs that functioned as activators was greater than or equal to 0.95. We labeled a domain as a repressor if the following 2 conditions held for all concatenations including that domain and a negative control: 1) at least 19% functioned as repressors (i.e. repression $\log_2$(ON:OFF) < 0.771); 2) the average repression $\log_2$(ON:OFF) for all pairs was less than or equal to -0.3, or the repression $\log_2$(ON:OFF) for all pairs that functioned as repressors was less than or equal to 0.20. Domains that functioned as both activators and repressors were labeled as duals. Domains that functioned as neither were labeled as non-hits.

KRAB-Nuclear Pfam screen cloning and cell culture

An initial backbone was generated by digesting pJT126 (AddGene #161926) with BsmbI, then ligating in a sequence corresponding to the ZNF10 KRAB domain. Then, the nuclear Pfam library was cloned into the downstream GoldenGate cloning site, and used to produce lentivirus identically as in.[21] K562 cells were then spinfected in 2 separate replicates with this KRAB+nuclear Pfam library lentivirus at an average MOI of 0.015, corresponding to an infection coverage of 85x per replicate. 72 hours post-infection 10 mg/mL blasticidin (Gibco) was added. 48 hours after blasticidin addition, 100 ng/mL doxycycline was added and cells were maintained in both blast + dox for 5 more days. Cells were sampled at day 2 of doxycycline recruitment for magnetic separation and downstream

sequencing. After 5 days of recruitment, cells were spun down, washed with PBS, and resuspended in dox-free/blast-free RPMI. 5 days after doxycycline release, cells were again sampled for magnetic separation and downstream sequencing. Cells were maintained at a maintenance coverage of 3,000-30,000 per library element on average. Library preparation and sequencing was performed identically as in to sequence the variable effector domain.[21]

Magnetic Separation

Cells to be separated were spun down at 300 x *g* for 5m and media was aspirated. Cells were then resuspended in PBS (Gibco), spun down again, and PBS was aspirated in order to wash the cells. Dynabeads Protein G were resuspended by vortexing for 30 seconds. 50mL of blocking buffer was prepared per 2 × 10^8 cells by adding 1 g biotin-free BSA (Sigma 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 keeping on ice. For all screens 60 uL of beads were used per 10 million cells. Magnetic separation was otherwise performed as in.[21]

Low-throughput validations

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

MODEL OF TRANSCRIPTIONAL REPRESSION

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

$$P(c \mid Active) = \sum_{mRNA\ levels} P(cit \mid mRNA) \cdot P(mRNA) = \sum_{n=0}^{1500} Norm_{6.5+(n/600),\sigma_{ON}}(c) \cdot Poisson_{\lambda}(n)$$

$$P(c \mid Silent) = Norm_{\mu_{OFF},\sigma_{OFF}}(c)$$

$$P(c) = P_{Active} \cdot P(c \mid Active) + P_{Silent} \cdot P(c \mid Silent)$$

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

To measure silencing, for every day of recruitment, 2000 cells were sampled to ensure equal sampling of $\log_{10}$ mCitrine-A fluorescence intensity across recruitment timepoints. The fraction of background silenced cells was directly computed using the rTetR-only recruitment data, with cells being labeled as Silent if their mCitrine-A fluorescence value was less than 10^7. Then, parameters μ_{OFF} and σ_{OFF} were directly estimated by examining the distribution of cells with log fluorescence values less than 7. λ was determined by estimating the mean amount of mRNA produced in ON cells at day 0, using the above mapping between mRNA counts and fluorescence values. k_s and t_{lag} were fit by examining the fraction of cells silenced over time; we assumed that the fraction of Active cells decreased with rate k_s after t_{lag} time had elapsed.[10,52] Finally, γ was determined by fitting a 2-state Gaussian Mixture model to fluorescence data for each day and fitting a line to the locations of the higher-fluorescence peak; we assumed λ decays linearly after silencing with slope γ.

Flow cytometry analysis

Flow cytometry analysis was performed in Python using Cytoflow.[88] For all flow experiments, live cells were gated using Cytoflow's DensityGateOp, keeping 90% of cells, and then thresholded for mCherry+ cells using the ThresholdGateOp. To determine the fraction of cells that were on, we computed the fraction of cells with $\log_{10}$ mCitrine-A fluorescence greater than 7 as measured on a BioRad ZE5 flow cytometer. Normalization of mCitrine levels to the no-dox condition was 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 any given condition. Cytometry analysis was otherwise performed identically as in.[21,36]

Western Blot

Western blots were performed similarly to.[20] Briefly, 20 million cells were pelleted and washed 1x with 5 mL of PBS. Pelleted cells were resuspended in 500 uL of ice cold RIPA buffer (EMD Millipore 20-188) containing protease inhibitor cocktail (Roche 11697498001) and were put on a rotator at 4C for 30 minutes. Next, the lysates were sonicated with a COVARIS E220 ultra-sonicator for 15 minutes (Peak power: 140-175, Duty factor: 10, Cycles/burst: 200). Lysates were spun down at 20,000 g for 5 minutes. Protein amounts were quantified using the Qubit protein broad range assay kit (Thermo Scientific, # A50668). 30 ug were boiled at 100C for 10 minutes in reducing sample loading buffer containing DTT (Thermo Scientific # 39000) and subsequently loaded onto a gel and transferred to a PVDF membrane. Membrane was first blocked with Pierce™ StartingBlock™ (TBS) Blocking Buffer (Thermo Scientific, Cat. No. 37543) for 1 hour at room temperature, then probed using FLAG M2 monoclonal antibody (1:1000, mouse, Sigma-Aldrich, F1804) and Histone H3 antibody (1:2000, rabbit, Abcam, AB1791) as primary antibodies overnight. Next, the membrane was washed with TBS-T 3x, 5 minutes each before being blotted again with secondary HRP-conjugated antibodies (Cytiva NA934 or NA931) at 1:2000 for one hour at room temperature. Blots were imaged on a iBright FL1500 imager. Band intensities were quantified using the analysis routine in iBright Analysis Software (see Figure S1 for regions of interest used).

CFSE

To make the staining solution, 1 vial of lyophilized CellTrace Violet dye (Thermo Fisher, Cat. C34557) was dissolved in 500uL DMSO and mixed into 20mL PBS (Gibco). For each sample, 1e6 cells were spun down at 300rcf for 5 min, washed 1x with PBS, and resuspended in 1mL staining solution. Samples were then incubated for 20m, then resuspended in RPMI for 5m to quench any dye remaining in solution. Cells were spun down, resuspended in RPMI, and flowed every 24h for 5 days.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analysis were performed in Python and R and are provided with all code at Zenodo https://doi.org/10.5281/zenodo.7453681. The n for each analysis is indicated in the main text or in figure legends of relevant analyses. Significance was set at $p < 0.05$, or $q < 0.05$ when correcting for multiple comparisons. No methods were used to determine whether the data met assumptions of the statistical approach.

ADDITIONAL INFORMATION

Supplementary information is available for this paper. Correspondence and requests for materials should be addressed to lbintu@stanford.edu.