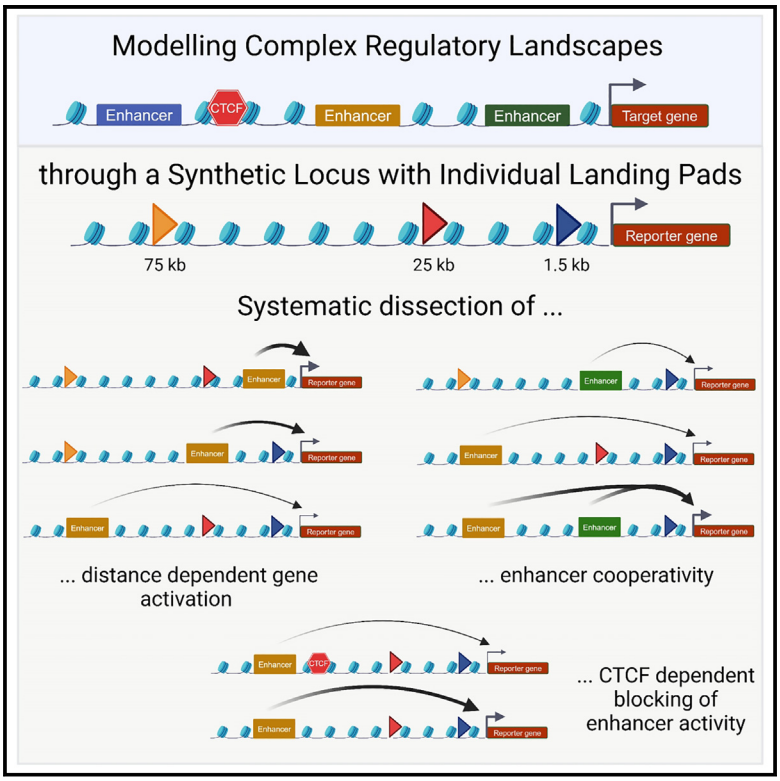
Molecular Cell

Article

Enhancer cooperativity can compensate for loss of activity over large genomic distances

Graphical abstract



Highlights

- Synthetic platform with multiple landing pads to build complex regulatory landscapes
- Distance-dependent activation is specific to each enhancer
- Position-dependent cooperation occurs between weak and strong enhancers
- CTCF sites can lower enhancer-dependent activity but do not interfere with synergy

Authors

Henry F. Thomas, Songjie Feng, Felix Haslhofer, ..., Mattia Pitasi, Alexander Stark, Christa Buecker

Correspondence

henry.fabian.thomas@univie.ac.at (H.F.T.), christa.buecker@maxperutzlabs.ac.at (C.B.)

In brief

Thomas, Feng, et al. introduce a synthetic platform that allows the building of complex regulatory landscapes. Integrating the same enhancer at different distances from a promoter uncovered that activation from a distance is an individual characteristic of each enhancer. Combining multiple enhancers revealed extensive cooperativity among individual elements.

Thomas et al., 2025, Molecular Cell *85*, 1–14
January 16, 2025
https://doi.org/10.1016/j.molcel.2024.11.008

CellPress

Molecular Cell

CellPress
OPEN ACCESS

Article

Enhancer cooperativity can compensate for loss of activity over large genomic distances

Henry F. Thomas,[1,2,3,6,*] Songjie Feng,[1,2,3,6] Felix Haslhofer,[1,2] Marie Huber,[1,2] María García Gallardo,[1,2,3] Vincent Loubiere,[4] Daria Vanina,[1,2] Mattia Pitasi,[1,2,3] Alexander Stark,[4,5] and Christa Buecker[1,2,7,*]

[1]Max Perutz Laboratories, Vienna BioCenter Campus (VBC), Dr.-Bohr-Gasse 9, 1030 Vienna, Austria
[2]University of Vienna, Center for Molecular Biology, Department of Microbiology, Immunobiology, and Genetics, Dr.-Bohr-Gasse 9, 1030 Vienna, Austria
[3]Vienna BioCenter PhD Program, a Doctoral School of the University of Vienna and the Medical University of Vienna, 1030 Vienna, Austria
[4]Research Institute of Molecular Pathology (IMP), Vienna BioCenter (VBC), 1030 Vienna, Austria
[5]Medical University of Vienna, Vienna BioCenter (VBC), Vienna, Austria
[6]These authors contributed equally
[7]Lead contact
*Correspondence: henry.fabian.thomas@univie.ac.at (H.F.T.), christa.buecker@maxperutzlabs.ac.at (C.B.)
https://doi.org/10.1016/j.molcel.2024.11.008

SUMMARY

Enhancers are short DNA sequences that activate their target promoter from a distance; however, increasing the genomic distance between the enhancer and the promoter decreases expression levels. Many genes are controlled by combinations of multiple enhancers, yet the interaction and cooperation of individual enhancer elements are not well understood. Here, we developed a synthetic platform in mouse embryonic stem cells that allows building complex regulatory landscapes from the bottom up. We tested the system by integrating individual enhancers at different distances and confirmed that the strength of an enhancer contributes to how strongly it is affected by increased genomic distance. Furthermore, synergy between two enhancer elements depends on the distance at which the two elements are integrated: introducing a weak enhancer between a strong enhancer and the promoter strongly increases reporter gene expression, allowing enhancers to activate from increased genomic distances.

INTRODUCTION

Transcription needs to be tightly controlled to ensure correct expression levels. In higher eukaryotes, *cis*-regulatory elements such as enhancers control when and in which cells a promoter is active.[1,2] Enhancers are short DNA sequences that consist of multiple transcription factor (TF) binding sites and control the expression of their target genes from a distance.[2] With increasing genomic distance, individual enhancers activate lower levels of expression from the promoter,[3–5] raising the question of how enhancers can bridge the distance to their promoter.

Enhancers are typically studied either in their native genomic context or in reporter assays.[2,6] Dissection of enhancer loci has, e.g., determined the contribution of tissue-specific TFs,[1,7–9] identified a role for 3D genome organization in enhancer-promoter communication,[2,10] and described various modes of enhancer cooperativity.[11–20] However, the complexity of endogenous regulatory landscapes can obstruct their analysis. A recent dissection of the α-globin locus used advances in synthesis of entire genomic loci to remove all enhancers and reintroduce them individually.[20] This approach revealed the super-additive behavior of several enhancers, contrasting previous conclusions of additive behavior at the same locus.[12,20] This underlines the importance of studying individual elements of multi-enhancer clusters both in the presence and absence of other *cis*-regulatory elements. Genetic modification of endogenous loci remains laborious and is limited to few selected enhancer clusters, making it challenging to derive general rules of enhancer cooperativity.

Conversely, reporter assays eliminate regulatory complexity and analyze the activity of individual enhancers and even combinations. Besides testing tissue specificity and enhancer strength,[6,21] reporter assays can be massively parallelized (massively parallel reporter assays, MPRAs) to measure activity of millions of fragments in one experiment and identify enhancer activity in the cell type of interest.[22,23] MPRAs are limited by the size of the analyzed DNA fragment. Therefore, enhancer-promoter communication over large genomic distances cannot be investigated.[24] Furthermore, most genes are controlled by multiple enhancers,[25] and elements with low or undetectable enhancer activity in reporter assays can exert critical regulatory functions at their endogenous loci.[18,20,26,27] While the first studies combined two enhancers to study cooperativity,[28,29]

Please cite this article in press as: Thomas et al., Enhancer cooperativity can compensate for loss of activity over large genomic distances, Molecular Cell (2024), https://doi.org/10.1016/j.molcel.2024.11.008

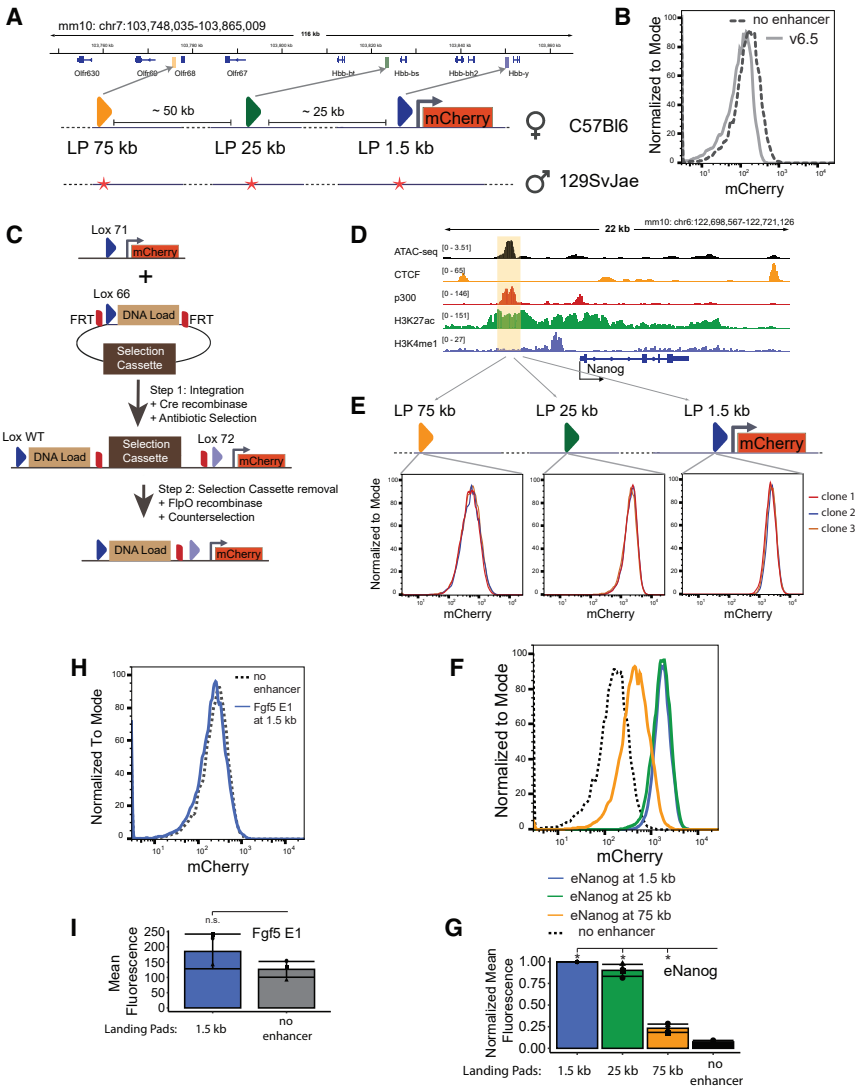


Figure 1. Generation and validation of a synthetic locus to systematically test enhancer activity

(A) Schematic representation of the different components of the synthetic locus generated. Top: UCSC browser track depicting the β-globin locus. The position of the different landing pads (LPs) is indicated. Below: different parts of the synthetic locus including three LP at different distances to the minimal promoter (arrow) and the mCherry reporter gene. All components are integrated on the C57BL6 allele. Bottom: red stars depict single nucleotide polymorphisms (SNPs).

(B) FACS plot showing a selected example of clone 1 no enhancer and the parental cell line v6.5.

(C) Schematic representation of the integration strategy (see text for details).

(D) IGV browser track depicting ATAC-seq,[33] CTCF,[35] p300,[34] H3K27ac,[34] and H3K4me1[34] ChIP-seq tracks for the *Nanog* locus.

(E) Representative FACS plots showing three individual clones with the *Nanog* enhancer (*eNanog*) integrated at the three LPs.

(F) FACS plots of individual examples for mCherry expression of *eNanog* integrated at each distance.

(G) Quantification of mean mCherry expression of *eNanog* integrated at the indicated LPs. All mean expressions were normalized to the 1.5 kb integration. Statistically significant different signals ($p < 0.05$, one-sided paired t tests) are marked by stars.

(H) Representative FACS plots showing an individual clonal cell line with the *Fgf5 E1* enhancer integrated at the 1.5 kb LP.

(I) Quantification of mean mCherry expression of *Fgf5 E1* integrated at the 1.5 kb LP compared with the no-enhancer control. Statistical significance was tested ($p < 0.05$, one-sided paired t tests).

See also Figures S1–S3.

the genomic distance between the elements was negligible. Therefore, new experimental systems are required to study enhancers in a native environment with higher throughput.

Here, we developed a versatile platform for testing enhancer activities and their distance dependencies: a cell line that allows for efficient integration of enhancer sequences at three different distances to a fluorescent reporter gene. This cell line allows building and analyzing complex regulatory landscapes from the bottom up. We compared multiple enhancers selected from different genomic contexts of the mouse genome at different distances to the same promoter. We demonstrate that all enhancers show reduced ability to activate a promoter with increasing distance but to different degrees. Furthermore, combinations of weak and strong enhancers strongly increase expression levels when the weak enhancer is placed between the strong enhancer and the promoter. Thus, synergy between enhancers depends on the individual enhancer activity and on the relative genomic distance between the enhancers and the promoter. Finally, we demonstrate that the interplay of genomic distance, enhancer cooperativity, and CTCF insulation determines gene expression levels.

RESULTS

Design of the reporter system

We aimed to generate a flexible, quantitative reporter system to analyze individual *cis*-regulatory elements and their combinations at different genomic distances from a reporter gene. We introduced the system into the β-globin locus, an inert environment in mouse embryonic stem cells (mESCs)[30] devoid of activating and low on repressive histone marks (Figure S1A[30–35]). We limited the reporter locus to one allele by targeting the C57BL/6 allele of v6.5 mESCs,[36] a C57BL/6 X 129/sv cross with single nucleotide polymorphisms (SNPs) distinguishing the parental alleles (Figure 1A). We inserted an mCherry reporter gene under the control of a minimal TK promoter[37,38] 1.5 kb

Please cite this article in press as: Thomas et al., Enhancer cooperativity can compensate for loss of activity over large genomic distances, Molecular Cell (2024), https://doi.org/10.1016/j.molcel.2024.11.008

from the *Hbb-γ* gene along with three landing pads (LPs) 1.5, 25, and 75 kb upstream of the reporter gene (Figure 1A). After the initial introduction of the reporter gene, we chose two independent clones for the integration of the additional LPs, reaching comparable conclusions with both clones.

mCherry allows for readout of expression levels by fluorescence-activated cell sorting (FACS), and we can obtain information on average expression values and the distribution of expression levels within a population of cells. The final reporter cell lines with the integration of the reporter gene and all three LPs, but without any enhancers, had slightly elevated levels of mCherry fluorescence compared with wild-type (WT) cells (Figures 1B and S1B).

For integrating enhancers at our locus, we used a variation of the Cre/lox recombination system. Recombination between identical lox sites within the genome and a donor plasmid leads to the plasmid integration into the genome. The efficiency is low due to its reversibility but can be increased by using lox66 and lox71 variants,[39,40] which contain mutations that reduce recognition by Cre recombinase after recombination, thereby minimizing excision. We integrated a lox71 site 1.5 kb 5′ of the reporter gene and designed plasmids containing lox66 sites for enhancer integration. Upon recombination, the entire plasmid is integrated into the genome (Figure 1C). To remove the plasmid backbone, we included compatible FRT sites[40] (Figure 1C). Expression of FlpO recombinase excises the entire backbone (Figure 1C), leaving only the enhancer and a single FRT site. For the LPs at 25 and 75 kb, we used different lox variants (here: lox2272-71 and loxm2-71) and designed specific targeting vectors for each LP. These vectors included a compatible lox66 site, a positive selection cassette fused to ΔThymidine Kinase for negative selection, and LP-specific FRT sites distinct from those used in other vectors (see Table S5). We validated that all lox sites and integrations occurred as expected for the two independent clones and multiple cell lines with integrated enhancers using Cas9-seq in combination with Nanopore sequencing.[41]

The *eNanog* activates gene expression in a distance-dependent manner

We amplified all selected enhancers from the *Mus musculus castaneus* strain to distinguish the integrated from the endogenous enhancer. First, we integrated the proximal enhancer from the *Nanog* locus (Figure 1D) into the targeting plasmid for the 1.5 kb LP. After co-transfection with a plasmid expressing Cre recombinase, roughly 25% of selected colonies had the desired plasmid integration. mCherry expression was slightly reduced upon plasmid integration (Figure S1C, right). We then transfected a plasmid expressing FlpO recombinase and selected for excision of the plasmid backbone with ganciclovir. Now, mCherry expression was strongly activated, and several independently derived cell lines showed almost indistinguishable mCherry levels (Figure 1E).

Interestingly, some clones initially contained few mCherry-negative cells that were quickly lost upon passaging (Figure S1D). We sorted mCherry-positive and -negative cells for one clone and measured mCherry fluorescence by FACS over several passages. While the mCherry-positive population remained stable, the negative population quickly gained mCherry expression (Figure S1D). Initiation of mCherry expression after backbone removal can be delayed but completes within a few passages. Therefore, all FACS analyses in this study were carried out after repeated passaging to ensure full activation.

Next, we integrated the *Nanog* enhancer (*eNanog*) at 25 and 75 kb. Integration efficiencies were similar to integration at 1.5 kb (see Table S5). At both distances, mCherry expression was not increased after initial plasmid integration (Figure S1C, middle and left). Different clones with integrations at 25 or 75 kb showed consistent behavior after backbone excision (Figure 1E, middle and left). We did not observe a fraction of *mCherry*-negative cells at 25 or 75 kb, and expression levels remained stable over several passages. Next, we compared mCherry expression across different e*Nanog* integration distances (Figures 1F and 1G). In a previous study, reporter gene fluorescence was linearly correlated with the number of mRNAs.[3] We performed Flow-fluorescence *in situ* hybridization (Flow-FISH)[42] for mCherry mRNA in cell lines with *eNanog* integrated at 1.5, 25, or 75 kb and the cell line without enhancers (Figure S1E). Both mRNA and protein levels decreased with increasing enhancer-promoter distance. We calculated the mean mCherry fluorescence, subtracting the autofluorescence of unmodified WT cells (Figure 1B), and used this as a direct measure of transcriptional activity. We normalized fluorescence at each LP to the 1.5 kb level to compare activity across distances (Figures 1G and S1F). mCherry expression was slightly reduced upon integrating e*Nanog* at 25 kb compared with 1.5 kb. However, when the enhancer is moved to 75 kb, mCherry expression strongly decreased across the population, though it remained higher than the control (Figures 1G and S1F). In summary, our reporter system allows for the efficient integration of enhancer sequences at three different distances from a reporter gene and confirms the critical role of genomic enhancer-promoter distance in transcriptional regulation.[3,4]

Next, we tested whether increased mCherry expression is due to the integration of active enhancers. *Fgf5 E1* is only activated during the differentiation of mESCs into so-called Epiblast-like cells (EpiLCs), is not actively repressed in mESCs, and has low intrinsic strength in luciferase assays.[18] As expected, integration of the *Fgf5 E1* enhancer at 1.5 kb did not increase mCherry expression beyond the no-enhancer control (Figures 1H, 1I, and S1G).

The β-globin locus is inactive in mESCs; however, individual promoters may be activated by integrating active enhancers. We performed ATAC-seq on selected cell lines with different enhancers and aligned the sequencing data against the relevant custom genomes. First, we analyzed the open chromatin landscape at the *β-globin* and *Nanog* loci, comparing cell lines with and without enhancers to the parental cell lines (Figures S2A and S2B). The profiles at the *Nanog* locus and other naive pluripotency genes were unchanged between v6.5 and the reporter cell lines (Figure S2B). At the *β-globin* locus, integration of the reporter and enhancers caused minimal changes in chromatin accessibility: peak calling identified open chromatin surrounding the integrated reporter and at only one additional peak about 14 kb 3′ of the reporter gene present in all cell lines with the synthetic locus independent of enhancer integration (Figure S2A). Importantly, the reporter

Please cite this article in press as: Thomas et al., Enhancer cooperativity can compensate for loss of activity over large genomic distances, Molecular Cell (2024), https://doi.org/10.1016/j.molcel.2024.11.008

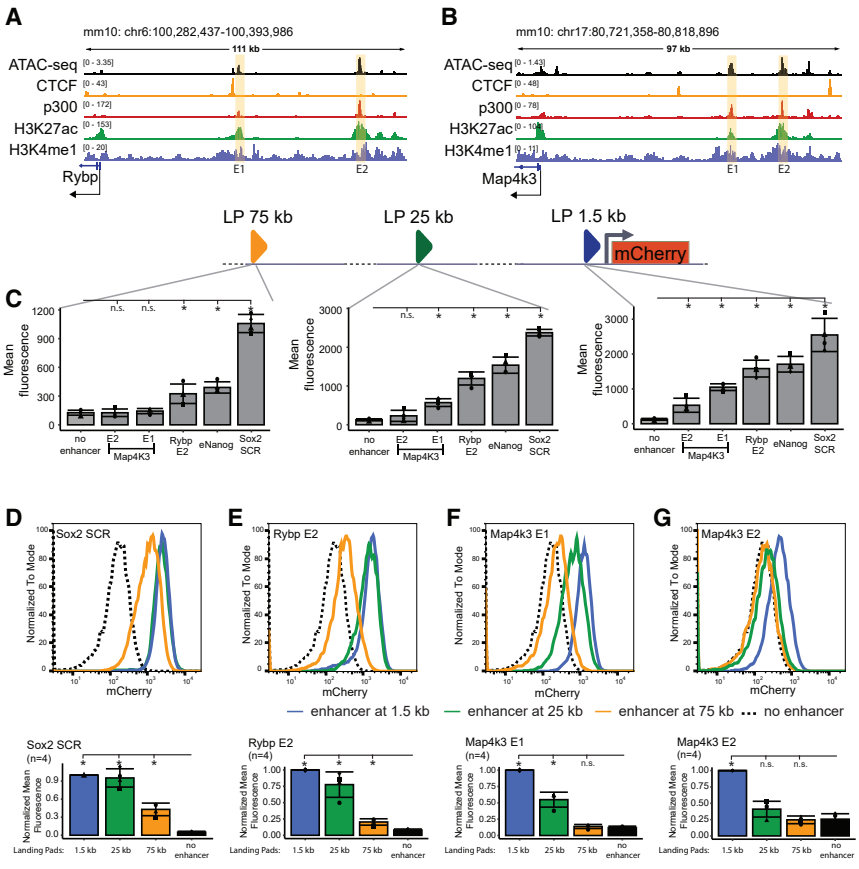


Figure 2. Enhancer strength determines dependency on enhancer-promoter distance

(A) IGV browser track depicting ATAC-seq,[33] CTCF,[35] p300,[34] H3K27ac,[34] and H3K4me1[34] ChIP-seq tracks for the *Rybp* locus.

(B) IGV browser track depicting ATAC-seq,[33] CTCF,[35] p300,[34] H3K27ac,[34] and H3K4me1[34] ChIP-seq tracks for the *Map4k3* locus.

(C) Quantification of mCherry expression of the five indicated enhancers compared with the no-enhancer control at the three different LPs. Left: integration at 75 kb; middle: integration at 25 kb; right: integration at 1.5 kb. Statistically significant signals are marked by stars compared with the no-enhancer control ($p < 0.05$, one-sided paired t tests). n.s.: not significant, $p>0.05$.

(D–G) Quantifying mean mCherry expression of indicated enhancers integrated at the indicated LPs. Top: individual examples; bottom: normalized mCherry expression. All mean expressions were normalized to the 1.5 kb integration. Statistically significant different signals ($p < 0.05$, one-sided paired t tests) are marked.

See also Figure S4.

gene gained accessibility while the endogenous genes remained unaffected. We aligned all sequencing results to custom genomes reflecting the specific enhancer integrations (Figures S2C, S2D, S3C, and S3D). All integrated enhancers gained accessibility at their respective integration sites. Finally, we analyzed the expression of *Hbb-γ* and *Olfr67* in cell lines with a strong enhancer integrated at 1.5, 25, and 75 kb before and after backbone removal (Figures S3A and S3B). While *Olfr67* is very lowly expressed and showed slight increase in expression upon integration of a strong enhancer at 25 kb, *Hbb-γ* levels increased with integration of a strong enhancer, correlating with activation of the reporter gene. However, the results were not statistically significant due to low and variable expression of *Hbb-γ*. In summary, while we observed minor differences after integrating the reporter and enhancers, the manipulations did not alter cell identity, and the minimal TK promoter is the main promoter activated at the synthetic locus.

Taken together, we have generated a highly versatile synthetic locus that allows for the reproducible interrogation of enhancer activity at defined distances to a reporter gene. The locus itself is mostly inert, and reporter gene expression is only activated by active enhancers.

Enhancer strength determines dependency on enhancer-promoter distance

Recently, both the *Sox2* control region (SCR) from the *Sox2* locus in mESCs and the human β-globin micro-LCR in erythroleukemia K562 cells have been demonstrated to depend on the genomic distance to the promoter in model loci.[3,4] Both studies used highly active enhancers, so it remains unclear if all enhancers respond similarly to genomic distance. The high efficiency of enhancer integration into our reporter system enabled us to systematically compare how different enhancers are affected by genomic distance. We selected enhancers from multi-enhancer clusters with varying strengths.[43] We included e*Nanog* (Figure 1[44]) the E2 enhancer element located 90 kb upstream of the *Rybp* locus (Figure 2A), and two enhancers (*E1* and *E2*) from the *Map4k3* locus (60 and 75 kb upstream of the transcription start site [TSS], Figure 2B). All enhancers activated luciferase in a plasmid-based assay (Figure S4A), albeit to different degrees: e*Nanog* and *Rybp E2* were the strongest activators, followed by *Map4k3 E2* and *Map4k3 E1*. We also included the well-studied SCR, a 6 kb multi-enhancer cluster roughly 100 kb upstream of the *Sox2* gene[3,45] (see also Figure S5A).

Next, we integrated each enhancer into each LP and measured mCherry levels using FACS (Figures 2C and S4B). At 1.5 kb, all enhancers increased *mCherry* expression, with the SCR being the strongest activator, followed by e*Nanog, Rybp* E2, *Map4k3* E1, and *Map4k3* E2 (Figures 2C, right; and S4B, right). The expression levels of mCherry were reproducible between different clones and replicates, though we observed more variability in lower-expressing cell lines with *Map4k3* E1 or *Map4k3* E2 integration. Despite higher activity in luciferase assays, *mCherry* levels were lower with *Map4k3* E2 integration compared to *Map4k3* E1 (Figure 2C). We then assessed mCherry expression when the same enhancers were integrated at either 25 kb (Figures 2C, middle; and S4B, middle) or 75 kb (Figures 2C, left; and S4B, left). At all distances, the order of relative strength of the enhancers remained consistent:

Please cite this article in press as: Thomas et al., Enhancer cooperativity can compensate for loss of activity over large genomic distances, Molecular Cell (2024), https://doi.org/10.1016/j.molcel.2024.11.008

CellPress
OPEN ACCESS

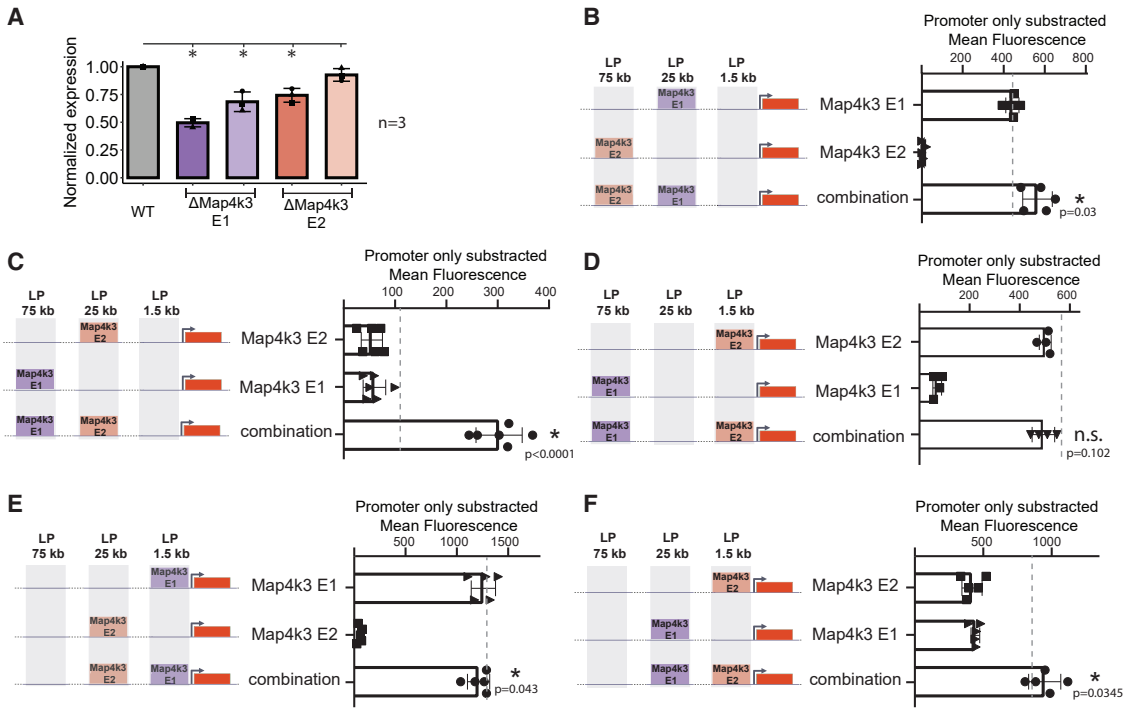


Figure 3. Weak enhancers activate transcription synergistically from a distance

(A) Quantification of *Map4k3* expression after deletion of the two different individual enhancers compared with the WT expression.

(B–F) Comparison of combinations of enhancers to the individual integrations and the expected additive behavior. Left: schematic of individual and combinations of integrations at the different LPs. The mean mCherry expression was calculated, and then the mean expression of the no-enhancer control was subtracted from all individual and combination mean mCherry expressions. The gray dashed line indicates the expected values under an additive model. Statistically significant signals compared with the expected additive model ($p < 0.05$, one-sided paired t tests) are marked by stars or n.s. (not significant). (B) *Map4k3 E2* at 75 kb combined with *Map4k3 E1* at 25 kb. (C) *Map4k3 E1* at 75 kb combined with *Map4k3 E2* at 25 kb. (D) *Map4k3 E1* at 75 kb combined with *Map4k3 E2* at 1.5 kb. (E) *Map4k3 E2* at 25 kb combined with *Map4k3 E1* at 1.5 kb. (F) *Map4k3 E1* at 25 kb combined with *Map4k3 E2* at 1.5 kb.

See also Figure S4.

SCR was the strongest, followed by e*Nanog*, *Rybp* E2, *Map4k3* E1, and finally *Map4k3* E2.

While the order in enhancer strength was preserved, mCherry expression decreased with increasing distances (Figures 2C and S4B). The strong enhancers *SCR* (Figures 2D and S4C), *eNanog* (Figures 1G and S1D), and *Rybp E2* (Figures 2E and S4D) showed strong activation at 1.5 and 25 kb, with no (*SCR*) or only slight loss (*eNanog* and *Rybp E2*) of activity at the intermediate distance 25 kb. In comparison, at 75 kb, mCherry expression is substantially reduced but still detectable. For the weaker enhancers, *Map4k3 E1* (Figures 2F and S4E) and *Map4k3 E2* (Figures 2G and S4F), activation from 25 kb was greatly reduced and almost undetectable at 75 kb. Of note, neither *Map4k3 E1* at 75 kb (Figures 2C and 2F) nor *Map4k3 E2* at 25 kb (Figures 2C and 2G) in clone 1 reached statistical significance. However, with increasing replicates (see Figures 3 and 4), both enhancers reached statistical significance at these loci. Nevertheless, the measured activity was very low, and the relative loss of activity compared with 1.5 kb was more pronounced for the weaker enhancers (Figures 2C and S4B). In summary, increasing distance to the minimal promoter decreases the ability of an enhancer to activate that promoter. Still, stronger enhancers can activate this promoter from larger distances.

The cooperation of weak enhancers can activate transcription from a distance

Map4k3 E2 could not elicit considerable mCherry expression from 25 kb. At its endogenous locus, this element is located about 75 kb from its potential target gene, *Map4k3*. We tested whether *Map4k3 E2* contributed to *Map4k3* expression. We deleted both *Map4k3* enhancers individually at the endogenous locus using CRISPR-Cas9, selected two clones each, and analyzed *Map4k3* expression using qPCR. Deletion of *E1* clearly reduced *Map4k3* expression in both analyzed clones (Figure 3A). Deletion of the *E2* enhancer also reduced the expression of the target gene (Figure 3A), albeit not as strongly as the *E1* enhancer and only reaching significance in one clone.

As the weak enhancer *Map4k3 E2* affects expression levels at its endogenous locus from a distance where it is inactive in our synthetic locus, we hypothesized that the presence of additional enhancers might influence how a given enhancer is affected by genomic distance. Therefore, we integrated both *Map4k3 E1* and *E2* elements step-wise at different LPs into our synthetic locus: first integrating one enhancer, selecting individual clones, and then integrating the second enhancer, followed by backbone removal at both integration sites. We used clone 1 for all dual enhancer experiments but validated

Please cite this article in press as: Thomas et al., Enhancer cooperativity can compensate for loss of activity over large genomic distances, Molecular Cell (2024), https://doi.org/10.1016/j.molcel.2024.11.008

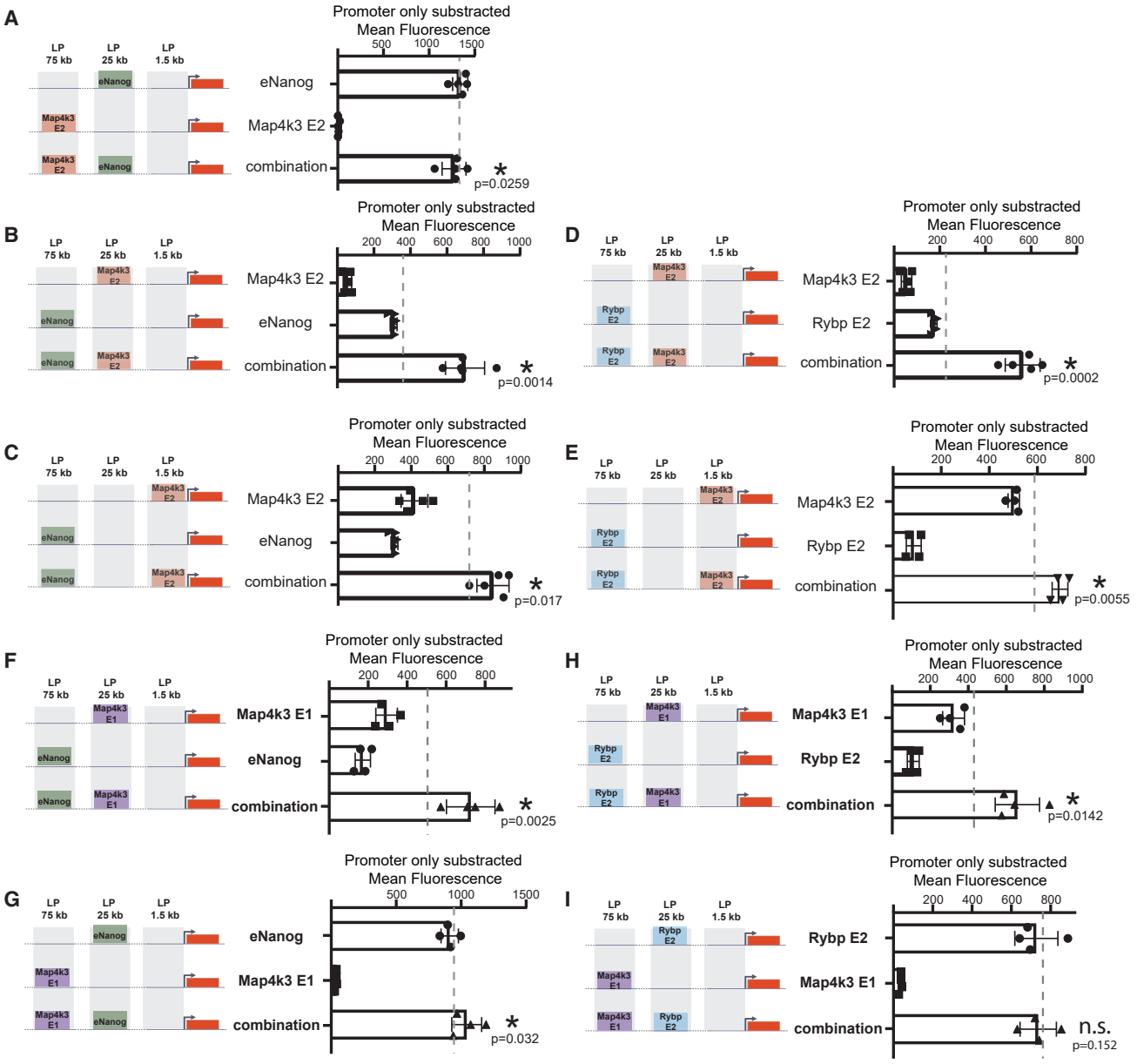


Figure 4. *Map4k3 E1 and E2* can cooperate with strong enhancers to activate transcription from a distance

Comparison of combinations of enhancers (third row) to the individual integrations (top two rows) and the expected additive behavior. Left: schematic of individual and combinations of integrations at the different LPs. The mean mCherry expression was calculated, and then the mean expression of the no-enhancer control was subtracted from all individual and combination mean mCherry expressions. The gray dashed line indicates the expected values under an additive model. Statistically significant signals compared with the expected additive model ($p < 0.05$, one-sided paired t tests) are marked by stars or n.s. (not significant). (A) *Map4k3 E2* at 75 kb combined with *eNanog* at 25 kb. (B) *eNanog* at 75 kb combined with *Map4k3 E2* at 25 kb. (C) *eNanog* at 75 kb combined with *Map4k3 E2* at 1.5 kb. (D) *Rybp E2* at 75 kb combined with *Map4k3 E2* at 25 kb. (E) *Rybp E2* at 75 kb combined with *Map4k3 E2* at 1.5 kb. (F) *eNanog* at 75 kb combined with *Map4K3 E1* at 25 kb. (G) *Map4k3 E1* at 75 kb combined with *eNanog* at 25 kb. (H) *Rybp E2* at 75 kb combined with *Map4k3 E1* at 25 kb. (I) *Map4K3 E1* at 75 kb combined with *Rybp E2* at 25 kb.

See also Figure S4.

selected combinations in clone 2. To compare dual to individual integrations to assess enhancer cooperativity, we calculated the expected activity for enhancer combinations based on an additive model (gray dotted lines), summing the individual enhancer contributions to predict expression. We then tested if the observed expression was significantly higher than the expected value. We subtracted the baseline *mCherry* expression from the no-enhancer control to avoid counting basic promoter activity twice when calculating the expected activity for dual integrations.

Please cite this article in press as: Thomas et al., Enhancer cooperativity can compensate for loss of activity over large genomic distances, Molecular Cell (2024), https://doi.org/10.1016/j.molcel.2024.11.008

First, we tested whether *Map4k3 E2* at 75 kb increased expression levels when combined with *Map4k3 E1* at 25 kb. Both elements together led to increased mCherry expression compared with the individual integrations (Figure 3B). Next, we swapped their positions: placing *Map4k3 E1* at 75 kb, and *Map4k3 E2* at 25 kb. While both enhancers showed minimal activity at these distances, their combination increased *mCherry* expression super-additively (Figures 3C and S4G). Hence, weak enhancers can synergize to activate transcription from distances where they are individually hardly active. Next, we placed *Map4k3 E2* closer to the promoter. Integrating *Map4k3* E2 at 1.5 kb and *Map4k3* E1 at 75 kb, super-additivity was not observed, and *mCherry* expression was similar to the level of *Map4k3* E2 alone at 1.5 kb (Figure 3D).

Next, we moved both elements closer to the promoter. Placing the stronger *Map4k3* E1 at 1.5 kb and the weaker *Map4k3* E2 at 25 kb did not further increase *mCherry* expression (Figure 3E). When we reversed the positions, with *Map4k3* E2 at 1.5 kb and *Map4k3* E1 at 25 kb, their combination slightly exceeded the expected additive behavior (Figure 3F).

Cooperation between strong and weak enhancers depends on their relative position

We next explored whether *Map4k3* E2 could cooperate with enhancers from other loci and combined *Map4k3 E2* with e*Nanog*. When *eNanog* was integrated at 25 kb, *Map4k3 E2* at 75 kb did not increase the expression of mCherry further (Figure 4A). However, placing e*Nanog* at 75 kb and *Map4k3 E2* at 25 kb resulted in strong super-additive mCherry expression (Figures 4B and S4H). Moving *Map4k3 E2* to 1.5 kb with e*Nanog* at 75 kb increased mCherry expression slightly above the combined individual enhancer levels (Figure 4C). The cooperative effect of *Map4k3 E2* was not limited to e*Nanog*: combining *Map4k3* E2 at 25 kb with *Rybp* E2 at 75 kb also produced super-additive *mCherry* expression (Figures 4D and S4I). Furthermore, placing *Map4k3* E2 at 1.5 kb and *Rybp* E2 at 75 kb yielded only minor increases above additive levels (Figure 4E).

Next, we combined *Map4k3 E1* at 25 kb with a stronger enhancer and integrated either *eNanog* or *Rybp E2* at 75 kb. In both cases, the enhancer combinations yielded stronger mCherry expression than the sum of their individual effects (Figures 4F and 4H). Similar to *Map4k3 E2*, the order of enhancers mattered: placing the weaker *Map4k3* E1 at 75 kb and the stronger enhancer at 25 kb did not yield a strong synergistic effect (Figures 4G and 4I). In summary, weak enhancers such as *Map4k3 E2 or Map4k3 E1* can synergize with strong enhancers, partially alleviating the distance-dependent drop in activation of the strong enhancer. This behavior somewhat depends on the weaker enhancer's position: closer integrations to the promoter produce additive-like expressions, while more distal combinations yield synergistic levels.

Map4k3 E1 and *E2* exemplify two weak enhancers with strong synergistic interactions, both with each other and with other enhancers. However, it remains unclear if all weak enhancers can cooperate with strong enhancers from a distance. Therefore, we selected four additional weak enhancers: *Rybp E1* (Figure 2A) and three well-described enhancers from the *Sox2* locus: *Sox2 E15*, *Sox2 E19*, and *Sox2 E20* (Figure S5A). *Sox2 E15* is a weak enhancer with a described role in transcriptional bursting of *Sox2*.[46] *Sox2 E19* and *E20* are potential facilitator elements, lacking intrinsic enhancer strength but capable of supporting *Sox2* expression.[19] First, we tested each enhancer's activity at 1.5 kb and at 25 kb (Figure 5A). *Sox2 E19* showed very low activity even at 1.5 kb. All other elements showed reduced mCherry expression at 25 kb compared with 1.5 kb. Notably, *Rybp E1* and *Sox2 E20* showed interesting behavior: both enhancers show lower mCherry expression at 1.5 kb than *Map4k3 E2* (Figure 5A), but the mCherry expression at 25 kb only drops by about 25% in the case of *Rybp E1* (Figure 5B), whereas for *Map4k3 E2* the drop is around 75% on average (Figure 2G). Therefore, the decrease in expression for *Rybp E1* and *Sox2 E20* from 25 kb compared with 1.5 kb was much shallower than expected for their respective enhancer strength at 1.5 kb (Figure 5B). The distance-dependent decrease in expression may be influenced not only by intrinsic enhancer strength but also by specific characteristics of each genomic element.

We next tested the cooperative behavior of these additional weak enhancers by integrating each element at 25 kb in cell lines with *eNanog* at 75 kb (Figures 5C–5F). In all these cell lines, we observed synergistic expression levels. Finally, we examined whether synergism is limited to weak enhancers by combining *eNanog* at 75 kb with either the *Rybp E2* or a second e*Nanog* at 25 kb (Figures S5B and S5C). *eNanog* at 25 kb alone produced higher levels of mCherry expression than *Rybp E2* at the same position (compare Figures S5B, S5C, and 2C, middle). Interestingly, the *mCherry* expression from *Rybp* E2 and *eNanog* combined exceeded the expected additive levels (Figure S5D), whereas combining two e*Nanog* at 75 and 25 kb did not produce a synergistic effect (Figure S5C). To test the limits of our minimal TK promoter, we combined *eNanog* and *Rybp E2* at 1.5 and 25 kb. The resulting gene expression was sub-additive, regardless of the order of the enhancers (Figures S5D and S5E). Although this combination produced stronger expression than two distal e*Nanog*, it was still lower than the *SCR* at 1.5 or 25 kb (Figure 2C). Therefore, promoter saturation may not function as a simple fixed threshold when combining strong enhancers (see discussion).

In summary, all tested weak enhancers can synergize with a distal strong enhancer, increasing expression at the target promoter beyond the additive levels of individual enhancer activities.

Two recent publications examined enhancer synergies using episomal plasmids but reached different conclusions[28,29]: in mESCs, enhancers worked predominantly additive,[29] while in *Drosophila* S2 cells most developmental enhancers were super-additive following a simple multiplicative model, with the caveat that strong enhancers might saturate limited promoter capacity.[28] We compared these models of enhancer cooperativity to explain our experimental results.

Multiple independently derived clones of individual and combinations of enhancers exhibited consistent expression changes. To assess variability, we first compared $\log_2$ fold changes to the no-enhancer control across clones derived after integration (step 1) and after backbone removal (step 2). Clones with the same enhancer integration showed exceptional reproducibility (Figure S5F), allowing us to pool data from all individual

Please cite this article in press as: Thomas et al., Enhancer cooperativity can compensate for loss of activity over large genomic distances, Molecular Cell (2024), https://doi.org/10.1016/j.molcel.2024.11.008

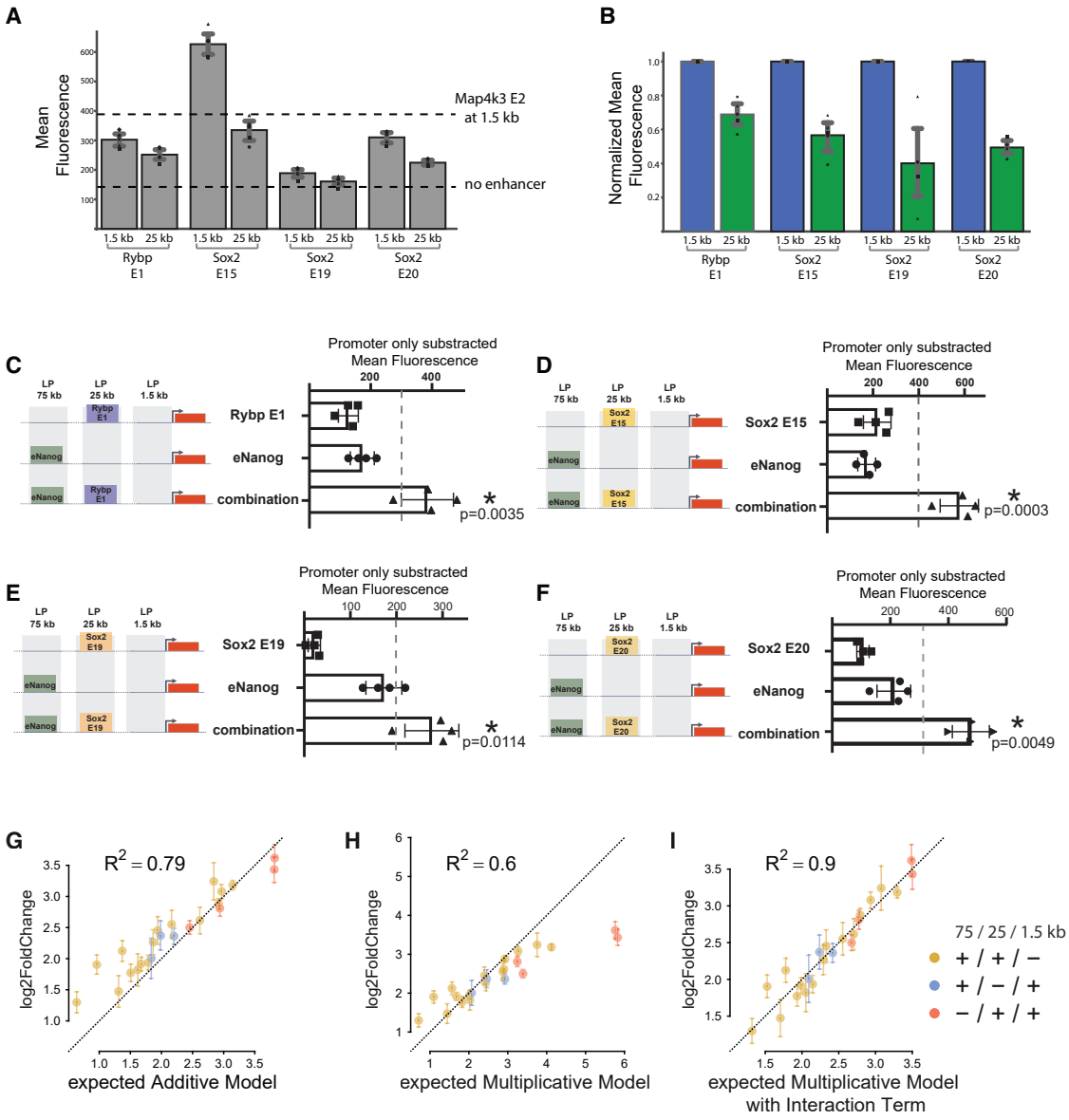


Figure 5. Weak enhancers cooperate with strong enhancers at distal positions

(A) Quantification of mCherry expression of the four indicated enhancers at the two indicated different LPs 1.5 and 25 kb. Both no enhancer and *Map4k3 E2* at 1.5 kb mCherry levels are indicated with dashed gray lines.

(B) Normalized mean mCherry expression of indicated enhancers integrated at the indicated LPs. All mean expressions were normalized to the 1.5 kb integration.

(C–F) Comparison of combinations of enhancers (third row) to the individual integrations (top two rows) and the expected additive behavior. Left: schematic of individual and combinations of integrations at the different LPs. The mean mCherry expression was calculated, and then the mean expression of the no-enhancer control was subtracted from all individual and combination mean mCherry expressions. The gray dashed line indicates the expected values under an additive model. Statistically significant signals compared with the expected additive model ($p < 0.05$, one-sided paired t tests) are marked by stars or n.s. (not significant).

(C) *eNanog* at 75 kb combined with *Rybp E1* at 25 kb. (D) *eNanog* at 75 kb combined with *Sox2 E15* at 25 kb. (E) *eNanog* at 75 kb combined with *Sox2 E19* at 25 kb. (F) *eNanog* at 75 kb combined with *Sox2 E20* at 25 kb. (G) Additive model of enhancer activity. (H) Multiplicative model of enhancer activity. (I) Multiplicative model with interaction term to predict enhancer activity.

See also Figure S5.

clones. We then tested if additive or multiplicative models could explain our results. The simple additive model predicts combined activity by summing individual enhancer activities (in linear gene expression space) and accurately described combinations when one enhancer is positioned proximal to the promoter (Figure 5G). However, many distal enhancer combinations deviated from this pattern, showing super-additive behavior (Figures 5G, 3, 4, and 5C–5F). A simple multiplicative model revealed a poor

Please cite this article in press as: Thomas et al., Enhancer cooperativity can compensate for loss of activity over large genomic distances, Molecular Cell (2024), https://doi.org/10.1016/j.molcel.2024.11.008

CellPress
OPEN ACCESS

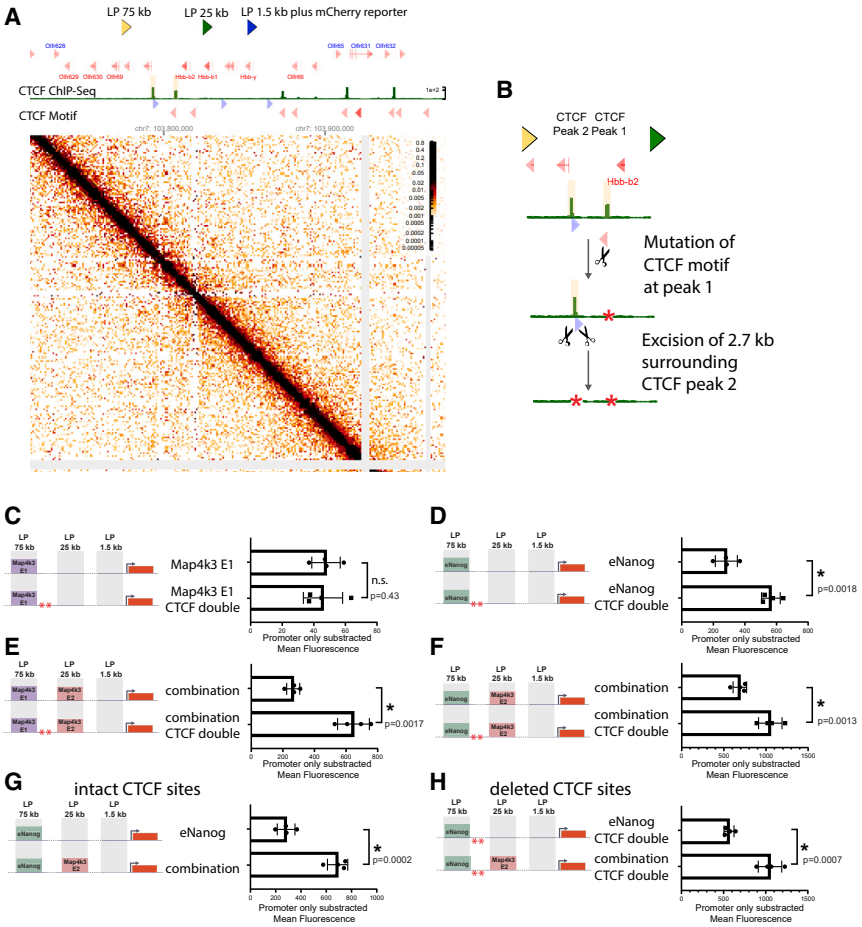


Figure 6. CTCF sites between the most distal LP and the promoter reduce activity of strong enhancers and enhancer combinations

(A) Top: positions of the LPs at the β-globin locus. Middle: ChIP-seq against CTCF[35] and direction of CTCF motifs; Bottom: MicroC data[48] across the β-globin locus.

(B) Strategy to eliminate both CTCF sites. Guide RNA against peak 1 binds in the CTCF motif of peak 1. In a second step, peak 2 was deleted with two guide RNAs binding about 3 kb apart from each other.

(C–H) Comparison of mCherry expression before and after deletion of both CTCF sites. Left: schematic of individual or combinations of integrations at the different LPs and the deletion of the CTCF sites. The mean mCherry expression was calculated, and then the mean expression of the no-enhancer control was subtracted from all individual and combination mean mCherry expressions. Statistically significant signals ($p < 0.05$, one-sided paired t tests) are marked by stars or n.s. (not significant). (C) *Map4k3 E1* integrated at 75 kb with and without CTCF motifs. (D) *eNanog* integrated at 75 kb with and without CTCF motifs. (E) *Map4k3 E1* integrated at 75 kb and *Map4k3 E2* integrated at 25 kb with and without CTCF motifs.(F) *eNanog* integrated at 75 kb and *Map4k3 E2* integrated at 25 kb with and without CTCF motifs. (G) Comparison of mCherry expression with *eNanog* at 75 kb without and with *Map4k3 E2* at 25 kb. (H) Comparison of mCherry expression with *eNanog* at 75 kb without and with *Map4k3 E2* at 25 kb after deletion of the CTCF motifs.

See also Figure S6.

fit: distal enhancer combinations showed higher activity than expected for multiplicative behavior, while combinations with a proximal enhancer yielded lower than predicted activity (Figure 5H). Neither the simple additive nor the simple multiplicative models alone accurately predicted combined activities. However, incorporating interaction terms into the multiplicative model significantly improved accuracy ($R^2 = 0.90$, Figure 5I). In this refined model, the individual activity of the two integrated enhancers is weighted with an interaction term, which considers how the activity of one enhancer modifies or scales the activity of the other. This model highlighted a linear relationship between individual and combined activities, allowing predictions based on individual activities at different integration sites. While each integration site contributed positively to the overall activity, the model revealed a significant negative interaction for enhancers placed closest to the promoter (−0.126), indicating that placing an enhancer at these proximal sites can reduce the combined effect (see Table S6).

In conclusion, how two enhancers cooperate depends strongly on their genomic distance to the promoter and their individual activity from that location, but strong individual activity close to the promoter is detrimental for strong synergistic behavior, suggesting that promoter saturation may limit stronger super-additive outcomes.[28]

CTCF sites between the most distal LP and the promoter reduce activity of strong enhancers and enhancer combinations

CTCF sites between enhancer and promoter can reduce target gene expression.[3,10,47] At the *β-globin* locus, two CTCF sites are positioned between the *Hbb-bt* and the *Olfr67* gene, located near the 75 and the 25 kb LPs (Figure 6A). In mESCs, this locus is inactive with no notable 3D genome structure surrounding the integration sites in the WT cells (Figure 6A, compare with the *Nanog* locus, Figure S6A). However, these CTCF sites might influence enhancer-driven gene expression at the locus. We devised a two-step strategy to remove the CTCF sites from existing enhancer cell lines (Figure 6B). We used a single guide RNA to target the CTCF motif near the 25 kb LP (peak 1). The second CTCF site (peak 2), located in a highly repetitive region, required deletion of a 3 kb genomic region. First, we examined whether removing CTCF sites would restore the activity of a weak enhancer, *Map4k3* E1, at 75 kb. Neither deletion of peak 1 alone (Figure S6B) nor removal of both peaks increased *mCherry* expression through *Map4k3 E1* (Figure 6C). We then tested the strong *eNanog* at 75 kb, observing a slight increase in *mCherry* expression upon single CTCF deletion (Figure S6C), which increased after removing both peaks (Figure 6D). We next examined enhancer combinations: both *Map4k3 E1* or

Please cite this article in press as: Thomas et al., Enhancer cooperativity can compensate for loss of activity over large genomic distances, Molecular Cell (2024), https://doi.org/10.1016/j.molcel.2024.11.008

eNanog at 75 kb showed synergy with *Map4k3 E2* at 25 kb. In both cases, overall expression of mCherry is slightly increased by mutation of peak 1 (Figures S6D and S6E), but much stronger increased upon removal of both CTCF peaks (Figures 6E and 6F). Finally, we compared the synergy between *eNanog* and *Map4k3 E2* with intact and disrupted CTCF sites. In both cases, the integration of the *Map4k3 E2* enhancer in combination with *eNanog* at 75 kb leads to increased mCherry expression. Finally, we compared the synergy between *eNanog* and *Map4k3 E2* with intact and disrupted CTCF sites. In both cases, the integration of the *Map4k3 E2* enhancer in combination with *eNanog* at 75 kb leads to increased mCherry expression. Taken together, the CTCF peak 1 has a minor negative effect on gene expression; however, removal of both CTCF peaks leads to context-dependent increases in gene expression at the promoter.

DISCUSSION

We developed a platform to study distance-dependent enhancer-enhancer cooperativity in the mammalian genome. Our flexible platform can accommodate different sequences at distinct positions, testing enhancers in their native chromatinized setting. As the locus contains no active regulatory elements besides the ones introduced, it provides a controlled environment for building and analyzing complex regulatory landscapes *de novo*.

Our system uncovered key principles of enhancer cooperativity and the impact of genomic distance on individual enhancers and their combinations. We found that increased genomic distance does not equally affect all enhancers: stronger enhancers generally retain more activation potential at greater distances. However, intrinsic strength alone cannot fully predict the activity drop when enhancers are integrated distally, suggesting additional mechanisms at play. Our findings also illuminate how genomic distance influences enhancer cooperativity. Combining a strong distal enhancer with a weaker enhancer at 25 kb consistently led to super-additive activation, while placing the strong enhancer closer to the promoter rarely produced this effect. Notably, *Map4k3 E1* can act as either a strong distal enhancer with *Map4k3* E2 or as a weaker enhancer at 25 kb with *Nanog* or *Rybp* E2, highlighting that relative strength between enhancers might matter more than absolute activity. Moving a weak enhancer closer to the promoter diminishes or eliminates super-additivity. This indicates that enhancer cooperation depends on both intrinsic strength and relative activity at specific distances, which may explain the variability in enhancer cooperativity modes at endogenous loci reported previously. Finally, how enhancers are affected by CTCF sites depends on the enhancer strength, as well as on the overall enhancer landscape.

Our results align with and expand on previous studies of distance-dependent regulation.[3,4] While increasing the enhancer-promoter distance generally reduces activation, many enhancers are positioned at distances where they should not activate a promoter individually. Interrogating multiple weak enhancers from three different genomic loci (*Map4k3*, *Rybp*, and *Sox2*) suggests that addition of weak enhancers can partially overcome decreased enhancer activity from increased genomic distances. Individually, weak enhancers cannot activate transcription strongly even from intermediate distances. However, placed between a strong enhancer and a promoter, they significantly boost promoter activation. This might explain recent findings from the *α-globin*[20] and the *Sox2* loci,[19] where weak enhancers help strong enhancers communicate with the promoter. However, we cannot rule out that these facilitators apply distinct mechanisms of enhancer cooperation.

Why can stronger enhancers bridge more considerable genomic distance than weaker ones, and how do weak enhancers synergize with distal strong enhancers? Increased genomic distance reduces enhancer-promoter contact frequency,[49,50] but contacts do not directly correlate with transcriptional activation,[3] and it remains unclear whether and how close enhancers need to come to their promoter to activate transcription.[51–53] Regulatory elements form large "transcription hubs," aggregating TFs, co-factors, and RNA polymerase II (RNA Pol II), which may bridge enhancers to promoters.[54,55] Actual 3D distance might matter less if the transcription hub still "touches" the promoter.[56] Stronger enhancers may form more substantial hubs, allowing them to activate from larger distances and making them less sensitive to small increases in 3D distance. This model could explain why the SCR maintains high expression from both 1.5 and 25 kb, unlike weaker enhancers. At large distances, strong enhancers may only intermittently contact the promoter via their hub, reducing gene expression. Combining strong distal enhancers with weak ones at intermediate distances could create larger hubs that increase promoter contact frequency, leading to super-additivity. Conversely, moving the weak enhancer closer to the promoter may allow its hub to contact the promoter independently, reducing the benefit of the strong distal enhancer and lowering super-additivity. However, intrinsic strength alone does not fully explain distance-dependent activation loss. Our comparison of weak enhancers suggests that distance sensitivity is characteristic of each enhancer, likely influenced by sequence composition and associated factor recruitment.

An alternative explanation could be provided by how promoters integrate enhancer input. Previous studies suggest that expression levels follow a sigmoidal curve relative to enhancer contact frequency[3,57]: at low contact frequency and expression levels, even large increases in contact produce only small increases in expression, whereas at intermediate contact frequencies, the same increase results in large jumps in expression levels. At high frequencies, further contact increases yield minimal expression gains. This behavior has been attributed to multiple transitions that a promoter has to undergo while in contact with the enhancer in order to be activated.[3,57] Sigmoidal integration of enhancer input into expression levels by promoters might be a general feature of transcriptional regulation that can explain our findings[58]: stronger enhancers might increase the rate of promoter transitions more than weaker enhancers. This could explain why strong enhancers are less affected by increased genomic distance than weaker ones. At the same time, introducing a weak enhancer alone at intermediate distances would have no effect since the promoter would be in the regime of the curve where big changes are required for increasing gene expression, but

adding a weak enhancer to a strong distal enhancer—where the promoter is in a more responsive, intermediate regime due to the presence of the distal enhancer—could lead to strong activation of gene expression. Such a model describes our results reasonably well and might explain why the strongest cooperativity arises when combining enhancers at distances where they have rather low individual activity. The sigmoidal model also suggests that high expression levels require increasing amounts of enhancer input, which might explain reduced or sub-additive effects observed for close enhancer combinations or for strong enhancers at 1.5 kb, similar to observations in the fly.[11] The response curve might vary by promoter; hence, some of the enhancer combinations might show different behavior when combined with a stronger promoter. Here, the promoter might be rate limiting or "saturated," even though drastically increasing enhancer input (e.g., by taking the strong SCR) can still push expression levels further. To fully assess whether a sigmoidal response curve can explain distance-dependent enhancer interactions, further studies with diverse enhancers, promoters, precise contact measurements, and mathematical modeling will be needed.

Gene expression regulation at the target locus depends on the interplay of promoters, enhancers, CTCF sites, and other regulatory elements. While enhancer-promoter distance remains the primary factor, multiple CTCF sites at our locus reduce gene expression without completely blocking it. CTCF may act as a boundary or insulator, hindering communication with the target gene. Interestingly, these CTCF sites do not form a detectable structure in WT lacking integrated reporters or enhancers, suggesting minimal loop extrusion in this region without active regulatory elements. Future studies on 3D contact changes will help clarify the emergence of potential new loops.

Previously, we demonstrated that elements lacking intrinsic activity can still contribute to gene expression.[18] Here, we expand on this by showing that weak enhancers positioned between a strong enhancer and the promoter can significantly enhance target gene expression. A recent MPRA study in human cell lines categorized *cis*-regulatory elements as classical enhancers or chromatin-dependent enhancers,[59] the latter marked by strong epigenomic signatures but low intrinsic activity. Both types often coexist within enhancer clusters, where they may serve distinct roles. Multiple weak enhancers could provide a buffering effect against the loss of individual elements. Additionally, combining weak and strong enhancers may increase the genomic distance range at which enhancer clusters can activate their target genes.[19,20] Moreover, we speculate that enhancer cooperativity could improve efficiency in target promoter selection: promoters associated with an additional weak enhancer may exhibit an increased responsiveness to a distal strong enhancer compared with other promoters without such an element.

Limitations

While mCherry levels were similar between independently generated cell lines, protein levels are not a direct readout of transcription. For selected enhancers the change in mRNA levels corresponded well with protein levels, however, other *cis*-regulatory elements could in principle also affect mRNA processing or half-life, in which case protein and transcription levels will not correlate well.

We have built the synthetic locus around the minimal TK promoter that might be rate limiting: most strong enhancers and combinations at close distances showed similar mean expression levels. In addition, strong activity from close distances penalized against a strong synergistic behavior in our linear model. The incorporation of different promoters in the future will test the limitations of the system and expand our toolbox. Similarly, we used the β-globin locus as our neutral environment and limited the tested distance to 75 kb.

In this proof-of-concept study, we analyzed a limited set of individual enhancers that showed very similar behavior; therefore, it is tempting to speculate that the super-additive cooperation of different enhancer elements is a general feature in transcriptional regulation. However, other, thus far untested elements could not follow this same trend, might antagonize each other, or simply not be compatible with each other.

Finally, we are currently limited to mere speculation regarding the molecular nature of enhancer cooperativity.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Dr. Christa Buecker (christa.buecker@maxperutzlabs.ac.at).

Materials availability

All unique/stable reagents generated in this study are available from the lead contact with a completed Materials Transfer Agreement.

Data and code availability

- All ATAC-seq datasets generated in this study have been deposited at ENA with the accession number PRJEB80727. FACS data have been deposited to Mendeley Data, and the DOI is reported in the key resources table. They are publicly available as of the date of publication.
- This study does not report any original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

ACKNOWLEDGMENTS

We would like to thank all members of the Buecker lab, Martin Leeb, and the members of his lab. In addition, we would like to thank Patricia Rothe and Mustafa Alaboo for technical help and Anton Goloborodko and Barak Cohen for insightful discussions. The BioOptics-FACS facility at the Max Perutz labs was instrumental in the success of this project. Cas9-seq and ATAC-seq were performed by the Next Generation Sequencing facility at the Vienna BioCenter Core Facilities (VBCF), a member of the Vienna BioCenter (VBC). The research was supported by the Austrian Science Fund FWF (P30599, P34123, and PAT9017923 to C.B., SMICH). Research at the IMP is supported by Boehringer Ingelheim GmbH and the Austrian Research Promotion Agency (FFG, FO999902549).

AUTHOR CONTRIBUTIONS

H.T. and C.B. designed experiments, analyzed results, and wrote the manuscript with the help from all authors. H.T. and S.F. performed the experimental work and analyzed results with the support of M.H. and D.V. F.H. generated and analyzed additional weak enhancer cell lines. M.G.G. performed

Flow-FISH. M.P. performed Cas9-seq analysis. V.L. and A.S. performed the modeling.

DECLARATION OF INTERESTS

The authors declare no competing interests.

STAR★METHODS

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

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - General culture conditions of mouse embryonic stem cells
  - Generation of reporter cell line
- METHOD DETAILS
  - Design of plasmids for enhancer integration
  - Integration of enhancer sequences
  - Cas9-targeted sequencing
  - FACS analysis of mCherry expression
  - FACS sorting of mCherry-populations
  - Luciferase assays
  - Enhancer KO and RT-qPCR
  - CTCF binding sites KO
  - RNA Flow-FISH
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Quantification and statistical analysis of FACS data
  - Modelling of expected combined activities
  - RT-qPCR of endogenous genes in the synthetic locus
  - ATAC-seq analysis

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.molcel.2024.11.008.

Received: January 2, 2024
Revised: October 1, 2024
Accepted: November 7, 2024
Published: December 2, 2024

REFERENCES

1. Kim, S., and Wysocka, J. (2023). Deciphering the multi-scale, quantitative cis-regulatory code. Mol. Cell *83*, 373–392. https://doi.org/10.1016/j.molcel.2022.12.032.
2. Long, H.K., Prescott, S.L., and Wysocka, J. (2016). Ever-changing landscapes: transcriptional enhancers in development and evolution. Cell *167*, 1170–1187. https://doi.org/10.1016/j.cell.2016.09.018.
3. Zuin, J., Roth, G., Zhan, Y., Cramard, J., Redolfi, J., Piskadlo, E., Mach, P., Kryzhanovska, M., Tihanyi, G., Kohler, H., et al. (2022). Nonlinear control of transcription through enhancer–promoter interactions. Nature *604*, 571–577. https://doi.org/10.1038/s41586-022-04570-y.
4. Rinzema, N.J., Sofiadis, K., Tjalsma, S.J.D., Verstegen, M.J.A.M., Oz, Y., Valdes-Quezada, C., Felder, A.-K., Filipovska, T., van der Elst, S., de Andrade dos Ramos, Z., et al. (2022). Building regulatory landscapes reveals that an enhancer can recruit cohesin to create contact domains, engage CTCF sites and activate distant genes. Nat. Struct. Mol. Biol. *29*, 563–574. https://doi.org/10.1038/s41594-022-00787-7.
5. Jensen, C.L., Chen, L.-F., Swigut, T., Crocker, O.J., Yao, D., Bassik, M.C., Ferrell, J.E., Boettiger, A.N., and Wysocka, J. (2024). Long range regulation of transcription scales with genomic distance in a gene specific manner. Preprint at bioRxiv. https://doi.org/10.1101/2024.07.19.604327.
6. Catarino, R.R., and Stark, A. (2018). Assessing sufficiency and necessity of enhancer activities for gene expression and the mechanisms of transcription activation. Genes Dev. *32*, 202–223. https://doi.org/10.1101/gad.310367.117.
7. Thanos, D., and Maniatis, T. (1995). Virus induction of human IFN beta gene expression requires the assembly of an enhanceosome. Cell *83*, 1091–1100. https://doi.org/10.1016/0092-8674(95)90136-1.
8. Farley, E.K., Olson, K.M., Zhang, W., Brandt, A.J., Rokhsar, D.S., and Levine, M.S. (2015). Suboptimization of developmental enhancers. Science *350*, 325–328. https://doi.org/10.1126/science.aac6948.
9. Goto, T., Macdonald, P., and Maniatis, T. (1989). Early and late periodic patterns of even skipped expression are controlled by distinct regulatory elements that respond to different spatial cues. Cell *57*, 413–422. https://doi.org/10.1016/0092-8674(89)90916-1.
10. Pachano, T., Haro, E., and Rada-Iglesias, A. (2022). Enhancer-gene specificity in development and disease. Development *149*, dev186536. https://doi.org/10.1242/dev.186536.
11. Bothma, J.P., Garcia, H.G., Ng, S., Perry, M.W., Gregor, T., and Levine, M. (2015). Enhancer additivity and non-additivity are determined by enhancer strength in the Drosophila embryo. eLife *4*, e07956. https://doi.org/10.7554/eLife.07956.
12. Hay, D., Hughes, J.R., Babbs, C., Davies, J.O.J., Graham, B.J., Hanssen, L.L.P., Kassouf, M.T., Marieke Oudelaar, A.M., Sharpe, J.A., Suciu, M.C., et al. (2016). Genetic dissection of the α-globin super-enhancer in vivo. Nat. Genet. *48*, 895–903. https://doi.org/10.1038/ng.3605.
13. Carleton, J.B., Berrett, K.C., and Gertz, J. (2017). Multiplex enhancer interference reveals collaborative control of gene regulation by estrogen receptor α-bound enhancers. Cell Syst. *5*, 333–344.e5. https://doi.org/10.1016/j.cels.2017.08.011.
14. Huang, J., Liu, X., Li, D., Shao, Z., Cao, H., Zhang, Y., Trompouki, E., Bowman, T.V., Zon, L.I., Yuan, G.-C., et al. (2016). Dynamic control of enhancer repertoires drives lineage and stage-specific transcription during hematopoiesis. Dev. Cell *36*, 9–23. https://doi.org/10.1016/j.devcel.2015.12.014.
15. Shin, H.Y., Willi, M., HyunYoo, K.H., Zeng, X., Wang, C., Metser, G., and Hennighausen, L. (2016). Hierarchy within the mammary STAT5-driven Wap super-enhancer. Nat. Genet. *48*, 904–911. https://doi.org/10.1038/ng.3606.
16. Hnisz, D., Schuijers, J., Lin, C.Y., Weintraub, A.S., Abraham, B.J., Lee, T.I., Bradner, J.E., and Young, R.A. (2015). Convergence of developmental and oncogenic signaling pathways at transcriptional super-enhancers. Mol. Cell *58*, 362–370. https://doi.org/10.1016/j.molcel.2015.02.014.
17. Moorthy, S.D., Davidson, S., Shchuka, V.M., Singh, G., Malek-Gilani, N., Langroudi, L., Martchenko, A., So, V., Macpherson, N.N., and Mitchell, J.A. (2017). Enhancers and super-enhancers have an equivalent regulatory role in embryonic stem cells through regulation of single or multiple genes. Genome Res. *27*, 246–258. https://doi.org/10.1101/gr.210930.116.
18. Thomas, H.F., Kotova, E., Jayaram, S., Pilz, A., Romeike, M., Lackner, A., Penz, T., Bock, C., Leeb, M., Halbritter, F., et al. (2021). Temporal dissection of an enhancer cluster reveals distinct temporal and functional contributions of individual elements. Mol. Cell *81*, 969–982.e13. https://doi.org/10.1016/j.molcel.2020.12.047.
19. Brosh, R., Coelho, C., Ribeiro-dos-Santos, A.M., Ellis, G., Hogan, M.S., Ashe, H.J., Somogyi, N., Ordoñez, R., Luther, R.D., Huang, E., et al. (2023). Synthetic regulatory genomics uncovers enhancer context dependence at the Sox2 locus. Mol. Cell *83*, 1140–1152.e7. https://doi.org/10.1016/j.molcel.2023.02.027.
20. Blayney, J.W., Francis, H., Rampasekova, A., Camellato, B., Mitchell, L., Stolper, R., Cornell, L., Babbs, C., Boeke, J.D., Higgs, D.R., et al. (2023). Super-enhancers include classical enhancers and facilitators to fully activate gene expression. Cell *186*, 5826–5839.e18. https://doi.org/10.1016/j.cell.2023.11.030.

21. Gasperini, M., Tome, J.M., and Shendure, J. (2020). Towards a comprehensive catalogue of validated and target-linked human enhancers. Nat. Rev. Genet. *21*, 292–310. https://doi.org/10.1038/s41576-019-0209-0.

22. Arnold, C.D., Gerlach, D., Stelzer, C., Boryń, Ł.M., Rath, M., and Stark, A. (2013). Genome-wide quantitative enhancer activity maps identified by STARR-seq. Science *339*, 1074–1077. https://doi.org/10.1126/science.1232542.

23. Smith, R.P., Taher, L., Patwardhan, R.P., Kim, M.J., Inoue, F., Shendure, J., Ovcharenko, I., and Ahituv, N. (2013). Massively parallel decoding of mammalian regulatory sequences supports a flexible organizational model. Nat. Genet. *45*, 1021–1028. https://doi.org/10.1038/ng.2713.

24. Muerdter, F., Boryń, Ł.M., Woodfin, A.R., Neumayr, C., Rath, M., Zabidi, M.A., Pagani, M., Haberle, V., Kazmar, T., Catarino, R.R., et al. (2018). Resolving systematic errors in widely used enhancer activity assays in human cells. Nat. Methods *15*, 141–149. https://doi.org/10.1038/nmeth.4534.

25. Gschwind, A.R., Mualim, K.S., Karbalayghareh, A., Sheth, M.U., Dey, K.K., Jagoda, E., Nurtdinov, R.N., Xi, W., Tan, A.S., Jones, H., et al. (2023). An encyclopedia of enhancer-gene regulatory interactions in the human genome. Preprint at bioRxiv. https://doi.org/10.1101/2023.11.09.563812.

26. Batut, P.J., Bing, X.Y., Sisco, Z., Raimundo, J., Levo, M., and Levine, M.S. (2022). Genome organization controls transcriptional dynamics during development. Science *375*, 566–570. https://doi.org/10.1126/science.abi7178.

27. Chen, L.-F., Long, H.K., Park, M., Swigut, T., Boettiger, A.N., and Wysocka, J. (2023). Structural elements promote architectural stripe formation and facilitate ultra-long-range gene regulation at a human disease locus. Mol. Cell *83*, 1446–1461.e6. https://doi.org/10.1016/j.molcel.2023.03.009.

28. Loubiere, V., de Almeida, B.P., Pagani, M., and Stark, A. (2023). Developmental and housekeeping transcriptional programs display distinct modes of enhancer-enhancer cooperativity in Drosophila. Nat. Commun. *15*, 8584. https://doi.org/10.1038/s41467-024-52921-2.

29. Martinez-Ara, M., Comoglio, F., and van Steensel, B. (2023). Large-scale analysis of the integration of enhancer-enhancer signals by promoters. Elife *12*, RP91994. https://doi.org/10.7554/eLife.91994.

30. Lienert, F., Wirbelauer, C., Som, I., Dean, A., Mohn, F., and Schübeler, D. (2011). Identification of genetic elements that autonomously determine DNA methylation states. Nat. Genet. *43*, 1091–1097. https://doi.org/10.1038/ng.946.

31. Yang, P., Humphrey, S.J., Cinghu, S., Pathania, R., Oldfield, A.J., Kumar, D., Perera, D., Yang, J.Y.H., James, D.E., Mann, M., et al. (2019). Multi-omic profiling reveals dynamics of the phased progression of pluripotency. Cell Syst. *8*, 427–445.e10. https://doi.org/10.1016/j.cels.2019.03.012.

32. Spencley, A.L., Bar, S., Swigut, T., Flynn, R.A., Lee, C.H., Chen, L.-F., Bassik, M.C., and Wysocka, J. (2023). Co-transcriptional genome surveillance by HUSH is coupled to termination machinery. Mol. Cell *83*, 1623–1639.e8. https://doi.org/10.1016/j.molcel.2023.04.014.

33. Romeike, M., Spach, S., Huber, M., Feng, S., Vainorius, G., Elling, U., Versteeg, G.A., and Buecker, C. (2022). Transient upregulation of IRF1 during exit from naive pluripotency confers viral protection. EMBO Rep. *23*, e55375. https://doi.org/10.15252/embr.202255375.

34. Buecker, C., Srinivasan, R., Wu, Z., Calo, E., Acampora, D., Faial, T., Simeone, A., Tan, M., Swigut, T., and Wysocka, J. (2014). Reorganization of enhancer patterns in transition from naive to primed pluripotency. Cell Stem Cell *14*, 838–853. https://doi.org/10.1016/j.stem.2014.04.003.

35. Hansen, A.S., Hsieh, T.-H.S., Cattoglio, C., Pustova, I., Saldaña-Meyer, R., Reinberg, D., Darzacq, X., and Tjian, R. (2019). Distinct classes of chromatin loops revealed by deletion of an RNA-binding region in CTCF. Mol. Cell *76*, 395–411.e13. https://doi.org/10.1016/j.molcel.2019.07.039.

36. Rideout, W.M., Wakayama, T., Wutz, A., Eggan, K., Jackson-Grusby, L., Dausman, J., Yanagimachi, R., and Jaenisch, R. (2000). Generation of mice from wild-type and targeted ES cells by nuclear cloning. Nat. Genet. *24*, 109–110. https://doi.org/10.1038/72753.

37. Greenshpan, Y., Sharabi, O., Yegodayev, K.M., Novoplansky, O., Elkabets, M., Gazit, R., and Porgador, A. (2022). The contribution of the minimal promoter element to the activity of synthetic promoters mediating CAR expression in the tumor microenvironment. Int. J. Mol. Sci. *23*, 7431. https://doi.org/10.3390/ijms23137431.

38. Arcot, S.S., Flemington, E.K., and Deininger, P.L. (1989). The human thymidine kinase gene promoter: Deletion analysis and specific protein binding. J. Biol. Chem. *264*, 2343–2349. https://doi.org/10.1016/S0021-9258(18)94182-7.

39. Araki, K., Araki, M., and Yamamura, K. (1997). Targeted integration of DNA using mutant lox sites in embryonic stem cells. Nucleic Acids Res. *25*, 868–872. https://doi.org/10.1093/nar/25.4.868.

40. Branda, C.S., and Dymecki, S.M. (2004). Talking about a revolution: the impact of site-specific recombinases on genetic analyses in mice. Dev. Cell *6*, 7–28. https://doi.org/10.1016/S1534-5807(03)00399-X.

41. Gilpatrick, T., Lee, I., Graham, J.E., Raimondeau, E., Bowen, R., Heron, A., Downs, B., Sukumar, S., Sedlazeck, F.J., and Timp, W. (2020). Targeted nanopore sequencing with Cas9-guided adapter ligation. Nat. Biotechnol. *38*, 433–438. https://doi.org/10.1038/s41587-020-0407-5.

42. Porichis, F., Hart, M.G., Griesbeck, M., Everett, H.L., Hassan, M., Baxter, A.E., Lindqvist, M., Miller, S.M., Soghoian, D.Z., Kavanagh, D.G., et al. (2014). High throughput detection of miRNAs and gene-specific mRNA at the single-cell level by flow cytometry. Nat. Commun. *5*, 5641. https://doi.org/10.1038/ncomms6641.

43. Peng, T., Zhai, Y., Atlasi, Y., ter Huurne, M., Marks, H., Stunnenberg, H.G., and Megchelenbrink, W. (2020). STARR-seq identifies active, chromatin-masked, and dormant enhancers in pluripotent mouse embryonic stem cells. Genome Biol. *21*, 243. https://doi.org/10.1186/s13059-020-02156-3.

44. Agrawal, P., Blinka, S., Pulakanti, K., Reimer, M.H., Stelloh, C., Meyer, A.E., and Rao, S. (2021). Genome editing demonstrates that the −5 kb Nanog enhancer regulates Nanog expression by modulating RNAPII initiation and/or recruitment. J. Biol. Chem. *296*, 100189. https://doi.org/10.1074/jbc.RA120.015152.

45. Taylor, T., Sikorska, N., Shchuka, V.M., Chahar, S., Ji, C., Macpherson, N.N., Moorthy, S.D., de Kort, M.A.C., Mullany, S., Khader, N., et al. (2022). Transcriptional regulation and chromatin architecture maintenance are decoupled functions at the Sox2 locus. Genes Dev. *36*, 699–717. https://doi.org/10.1101/gad.349489.122.

46. Du, M., Stitzinger, S.H., Spille, J.-H., Cho, W.-K., Lee, C., Hijaz, M., Quintana, A., and Cissé, I.I. (2024). Direct observation of a condensate effect on super-enhancer controlled gene bursting. Cell *187*, 331–344.e17. https://doi.org/10.1016/j.cell.2023.12.005.

47. de Wit, E., Vos, E.S.M., Holwerda, S.J.B., Valdes-Quezada, C., Verstegen, M.J.A.M., Teunissen, H., Splinter, E., Wijchers, P.J., Krijger, P.H.L., and de Laat, W. (2015). CTCF binding polarity determines chromatin looping. Mol. Cell *60*, 676–684. https://doi.org/10.1016/j.molcel.2015.09.023.

48. Hsieh, T.-H.S., Cattoglio, C., Slobodyanyuk, E., Hansen, A.S., Rando, O.J., Tjian, R., and Darzacq, X. (2020). Resolving the 3D landscape of transcription-linked mammalian chromatin folding. Mol. Cell *78*, 539–553.e8. https://doi.org/10.1016/j.molcel.2020.03.002.

49. Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature *485*, 376–380. https://doi.org/10.1038/nature11082.

50. Rao, S.S.P., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., et al. (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell *159*, 1665–1680. https://doi.org/10.1016/j.cell.2014.11.021.

51. Chen, H., Levo, M., Barinov, L., Fujioka, M., Jaynes, J.B., and Gregor, T. (2018). Dynamic interplay between enhancer–promoter topology and gene activity. Nat. Genet. *50*, 1296–1303. https://doi.org/10.1038/s41588-018-0175-z.

52. Benabdallah, N.S., Williamson, I., Illingworth, R.S., Kane, L., Boyle, S., Sengupta, D., Grimes, G.R., Therizols, P., and Bickmore, W.A. (2019). Decreased enhancer-promoter proximity accompanying enhancer activation. Mol. Cell *76*, 473–484.e7. https://doi.org/10.1016/j.molcel.2019.07.038.

53. Alexander, J.M., Guan, J., Li, B., Maliskova, L., Song, M., Shen, Y., Huang, B., Lomvardas, S., and Weiner, O.D. (2019). Live-cell imaging reveals enhancer-dependent Sox2 transcription in the absence of enhancer proximity. eLife *8*, e41769. https://doi.org/10.7554/eLife.41769.

54. Li, J., Dong, A., Saydaminova, K., Chang, H., Wang, G., Ochiai, H., Yamamoto, T., and Pertsinidis, A. (2019). Single-molecule nanoscopy elucidates RNA polymerase II transcription at single genes in live cells. Cell *178*, 491–506.e28. https://doi.org/10.1016/j.cell.2019.05.029.

55. Li, J., Hsu, A., Hua, Y., Wang, G., Cheng, L., Ochiai, H., Yamamoto, T., and Pertsinidis, A. (2020). Single-gene imaging links genome topology, promoter-enhancer communication and transcription control. Nat. Struct. Mol. Biol. *27*, 1032–1040. https://doi.org/10.1038/s41594-020-0493-6.

56. Heist, T., Fukaya, T., and Levine, M. (2019). Large distances separate coregulated genes in living Drosophila embryos. Proc. Natl. Acad. Sci. USA *116*, 15062–15067. https://doi.org/10.1073/pnas.1908962116.

57. Xiao, J.Y., Hafner, A., and Boettiger, A.N. (2021). How subtle changes in 3D structure can create large changes in transcription. eLife *10*, e64320. https://doi.org/10.7554/eLife.64320.

58. Veitia, R.A. (2003). A sigmoidal transcriptional response: cooperativity, synergy and dosage effects. Biol. Rev. Camb. Philos. Soc. *78*, 149–170. https://doi.org/10.1017/s1464793102006036.

59. Sahu, B., Hartonen, T., Pihlajamaa, P., Wei, B., Dave, K., Zhu, F., Kaasinen, E., Lidschreiber, K., Lidschreiber, M., Daub, C.O., et al. (2021). Sequence determinants of human gene regulatory elements. Nature Genetics *54*, 283–294. https://doi.org/10.1101/2021.03.18.435942.

60. Nagy, A., Rossant, J., Nagy, R., Abramow-Newerly, W., and Roder, J.C. (1993). Derivation of completely cell culture-derived mice from early-passage embryonic stem cells. Proc. Natl. Acad. Sci. USA *90*, 8424–8428. https://doi.org/10.1073/pnas.90.18.8424.

61. Ewels, P.A., Peltzer, A., Fillinger, S., Patel, H., Alneberg, J., Wilm, A., Garcia, M.U., Di Tommaso, P., and Nahnsen, S. (2020). The nf-core framework for community-curated bioinformatics pipelines. Nat. Biotechnol. *38*, 276–278.

62. Robinson, J.T., Thorvaldsdottir, H., Turner, D., and Mesirov, J.P. (2023). igv.js: an embeddable JavaScript implementation of the Integrative Genomics Viewer (IGV). Bioinformatics *39*, btac830. https://doi.org/10.1093/bioinformatics/btac830.

Please cite this article in press as: Thomas et al., Enhancer cooperativity can compensate for loss of activity over large genomic distances, Molecular Cell (2024), https://doi.org/10.1016/j.molcel.2024.11.008

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
Stbl3 bacteria	propagated from the initial Invitrogen stock	N/A
Chemicals, peptides, and recombinant proteins		
NcoI-HF	NEB	Cat#R3193S
XhoI	NEB	Cat#R0146S
SphI-HF	NEB	Cat#R3182S
BamHI-HF	NEB	Cat#R3136S
NotI-HF	NEB	Cat#R3189S
BsrGI-HF	NEB	Cat#R3575S
Puromycin	InvivoGen	Cat#ant-pr-1
Ganciclovir	InvivoGen	Cat#sud-gcv
Neomycin	Sigma-Aldrich	Cat#G8168
Hygromycin B	Sigma-Aldrich	Cat#10843555001
Blasticidin	InvivoGen	Cat#ant-bl-1
Gibson Assembly mix	homemade	N/A
HyClone DMEM/F12 medium without HEPES	Cytiva (Formerly GE Healthcare Life Sciences)	Cat#SH30271.01
AlbuMAXTM II Lipid-Rich Bovine Serum Albumin	GIBCO	Cat#11021-029
serum-free B- 27 Supplement (50 x)	GIBCO	Cat#17504044
N2 supplement	homemade	N/A
MEM NEAA (100 x)	GIBCO	Cat#11140035
Penicillin-Streptomycin (5,000 U/mL)	GIBCO	Cat#15070063
Sodium Pyruvate (100 mM)	GIBCO	Cat#11360039
2-Mercaptoethanol	GIBCO	Cat#21985023
CHIR-99021	Selleckchem	Cat#S1263
PD0325901	Selleckchem	Cat#S1036
hLIF	VBCF Protein Technologies Facility	N/A
Poly-L-ornithine hydrobromide	Sigma-Aldrich	Cat#P4638
Laminin from Engelbreth-Holm-Swarm	Sigma-Aldrich	Cat#L2020
Trypsin-EDTA solution	Sigma-Aldrich	Cat#T3924
Fetal Bovine Serum	Sigma-Aldrich	Cat#F7524
Lipofectamine 2000 Transfection Reagent	Invitrogen	Cat#11668019
Critical commercial assays		
Invitrogen PrimeFlow RNA Assay Kit	Thermo Fisher Scientific	Cat#88-18005-204
Dual-Glo Luciferase Assay	Promega	Cat#E2920
Deposited data		
FACS data	This Paper	Mendeley Data: https://doi.org/10.17632/2j75sff6jw.2
ATAC Seq Data	This Paper	ENA, accession number PRJEB80727
Experimental models: Cell lines		
Mouse: v6.5 ES cells	Rideout et al.[36]	N/A
Mouse: v6.5 3xlox clone 1	This paper	N/A

(*Continued on next page*)

Please cite this article in press as: Thomas et al., Enhancer cooperativity can compensate for loss of activity over large genomic distances, Molecular Cell (2024), https://doi.org/10.1016/j.molcel.2024.11.008

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Mouse: v6.5 3xlox clone 2	This paper	N/A
Mouse: v6.5 3xlox clone 1 + *Map4k3 E1* (1.5 kb)	This paper	N/A
Mouse: v6.5 3xlox clone 1 + *Map4k3* E2 (1.5 kb)	This paper	N/A
Mouse: v6.5 3xlox clone 2 + *Map4k3* E2 (1.5 kb)	This paper	N/A
Mouse: v6.5 3xlox clone 1 + Rybp *E2* (1.5 kb)	This paper	N/A
Mouse: v6.5 3xlox clone 1 + eNanog (1.5 kb)	This paper	N/A
Mouse: v6.5 3xlox clone 2 + eNanog (1.5 kb)	This paper	N/A
Mouse: v6.5 3xlox clone 1 + SCR-rev (1.5 kb)	This paper	N/A
Mouse: v6.5 3xlox clone 2 + SCR-rev (1.5 kb)	This paper	N/A
Mouse: v6.5 3xlox clone 1 + *Fgf5*-PE (1.5 kb)	This paper	N/A
Mouse: v6.5 3xlox clone 2 + *Fgf5*-PE (1.5 kb)	This paper	N/A
Mouse: v6.5 3xlox clone 1 + *Map4k3 E1* (25 kb)	This paper	N/A
Mouse: v6.5 3xlox clone 2 + *Map4k3 E1* (25 kb)	This paper	N/A
Mouse: v6.5 3xlox clone 1 + *Map4k3 E2* (25 kb)	This paper	N/A
Mouse: v6.5 3xlox clone 2 + *Map4k3 E2* (25 kb)	This paper	N/A
Mouse: v6.5 3xlox clone 1 + Rybp *E2* (25 kb)	This paper	N/A
Mouse: v6.5 3xlox clone 2 + Rybp *E2* (25 kb)	This paper	N/A
Mouse: v6.5 3xlox clone 1 + eNanog (25 kb)	This paper	N/A
Mouse: v6.5 3xlox clone 2 + eNanog (25 kb)	This paper	N/A
Mouse: v6.5 3xlox clone 1 + SCR (25 kb)	This paper	N/A
Mouse: v6.5 3xlox clone 2 + SCR (25 kb)	This paper	N/A
Mouse: v6.5 3xlox clone 1 + SCR-rev (25 kb)	This paper	N/A
Mouse: v6.5 3xlox clone 2 + SCR-rev (25 kb)	This paper	N/A
Mouse: v6.5 3xlox clone 1 + *Map4k3 E1* (75 kb)	This paper	N/A
Mouse: v6.5 3xlox clone 2 + *Map4k3 E1* (75 kb)	This paper	N/A
Mouse: v6.5 3xlox clone 1 + *Map4k3 E2* (75 kb)	This paper	N/A
Mouse: v6.5 3xlox clone 2 + *Map4k3 E2* (75 kb)	This paper	N/A
Mouse: v6.5 3xlox clone 1 + Rybp *E2* (75 kb)	This paper	N/A
Mouse: v6.5 3xlox clone 2 + Rybp *E2* (75 kb)	This paper	N/A
Mouse: v6.5 3xlox clone 1 + eNanog (75 kb)	This paper	N/A
Mouse: v6.5 3xlox clone 2 + eNanog (75 kb)	This paper	N/A
Mouse: v6.5 3xlox clone 1 + SCR (75 kb)	This paper	N/A
Mouse: v6.5 3xlox clone 2 + SCR (75 kb)	This paper	N/A
Mouse: v6.5 3xlox clone 1 + SCR-rev (75 kb)	This paper	N/A
Mouse: v6.5 3xlox clone 2 + SCR-rev (75 kb)	This paper	N/A
Mouse: v6.5 3xlox clone 1 + Rybp *E1* (1.5 kb) clone 1 - clone 4	This paper	N/A
Mouse: v6.5 3xlox clone 1 + Rybp *E1* (25 kb) clone 1 - clone 4	This paper	N/A
Mouse: v6.5 3xlox clone 1 + *Sox2 E15* (1.5 kb) clone 1 - clone 4	This paper	N/A
Mouse: v6.5 3xlox clone 1 + *Sox2 E15* (25 kb) clone 1 - clone 4	This paper	N/A
Mouse: v6.5 3xlox clone 1 + *Sox2 E19* (1.5 kb) clone 1 - clone 4	This paper	N/A
Mouse: v6.5 3xlox clone 1 + *Sox2 E19* (25 kb) clone 1 - clone 4	This paper	N/A
Mouse: v6.5 3xlox clone 1 + *Sox2 E20* (1.5 kb) clone 1 - clone 2	This paper	N/A
Mouse: v6.5 3xlox clone 1 + *Sox2 E20* (25 kb) - clone 1 - clone 3	This paper	N/A
Mouse: v6.5 3xlox clone 1 + *Map4k3 E1* (25 kb) + *Map4k3 E2* (75 kb) clone 1 - clone 4	This paper	N/A
Mouse: v6.5 3xlox clone 1 + *Map4k3 E2* (25 kb) + eNanog (75 kb) clone 1 - clone 2	This paper	N/A
Mouse: v6.5 3xlox clone 1 + *Map4k3* E2 (25 kb) + *Map4k3 E1* (75 kb) clone 1 - clone 2	This paper	N/A

(*Continued on next page*)

Please cite this article in press as: Thomas et al., Enhancer cooperativity can compensate for loss of activity over large genomic distances, Molecular Cell (2024), https://doi.org/10.1016/j.molcel.2024.11.008

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Mouse: v6.5 3xlox clone 1 + *Map4k3 E2* (25 kb) + Rybp *E2* (75 kb) clone 1 - clone 2	This paper	N/A
Mouse: v6.5 3xlox clone 1 + eNanog (25 kb) + *Map4k3 E2* (75 kb) clone 1 - clone 2	This paper	N/A
Mouse: v6.5 3xlox clone 1 + *Map4k3 E1* (1.5 kb) + *Map4k3 E2* (25 kb) clone 1 - clone 2	This paper	N/A
Mouse: v6.5 3xlox clone 1 + *Map4k3 E2* (1.5 kb) + eNanog (75 kb) clone 1 - clone 2	This paper	N/A
Mouse: v6.5 3xlox clone 1 + *Map4k3 E1* (25 kb) + Rybp *E2* (75 kb) clone 1 - clone 6	This paper	N/A
Mouse: v6.5 3xlox clone 1 + eNanog (25 kb) + eNanog (75 kb) clone 1 - clone 6	This paper	N/A
Mouse: v6.5 3xlox clone 1 + eNanog (25 kb) + *Map4k3 E1* (75 kb) clone 1 - clone 5	This paper	N/A
Mouse: v6.5 3xlox clone 1 + Rybp *E2* (25 kb) + eNanog (75 kb) clone 1 - clone 6	This paper	N/A
Mouse: v6.5 3xlox clone 1 + Rybp *E2* (25 kb) + *Map4k3 E1* (75 kb) clone 1 - clone 6	This paper	N/A
Mouse: v6.5 3xlox clone 1 + *Map4k3 E1* (25 kb) + eNanog (75 kb) clone 1 - clone 6	This paper	N/A
Mouse: v6.5 3xlox clone 1 + Rybp *E2* (1.5 kb) + eNanog (25 kb) clone 1 - clone 6	This paper	N/A
Mouse: v6.5 3xlox clone 1 + eNanog (1.5 kb) + Rybp E2 (25 kb) clone 1 - clone 4	This paper	N/A
Mouse: v6.5 3xlox clone 1 + Rybp *E1* (25 kb) + eNanog (75 kb) clone 1 - clone 6	This paper	N/A
Mouse: v6.5 3xlox clone 1 + *Sox2 E15* (25 kb) + eNanog (75 kb) clone 1 - clone 6	This paper	N/A
Mouse: v6.5 3xlox clone 1 + *Sox2 E19* (25 kb) + eNanog (75 kb) clone 1 - clone 6	This paper	N/A
Mouse: v6.5 3xlox clone 1 + *Sox2 E20* (25 kb) + eNanog (75 kb) clone 1 - clone 2	This paper	N/A
Mouse: v6.5 3xlox clone 1 + *Map4k3 E2* (1.5 kb) + *Map4k3 E1* (75 kb) clone 1 - clone 2	This paper	N/A
Mouse: v6.5 3xlox clone 1 + *Map4k3 E2* (1.5 kb) + Rybp *E2* (75 kb) clone 1 - clone 2	This paper	N/A
Mouse: v6.5 3xlox clone 2 + *Map4k3 E2* (25 kb) + eNanog (75 kb) clone 1 - clone6	This paper	N/A
Mouse: v6.5 3xlox clone 2 + *Map4k3* E2 (25 kb) + Rybp E2 (75 kb) clone 1 - clone6	This paper	N/A
Mouse: v6.5 3xlox clone 2 + *Map4k3 E2* (25 kb) + *Map4k3 E1* (75 kb) clone 1 - clone6	This paper	N/A
Mouse: R1 mouse ESC	Nagy et al.[60]	N/A
Mouse: R1 Δ*Rybp E2* clone 1	This paper	N/A
Mouse: R1 Δ*Rybp* E2 clone 2	This paper	N/A
Mouse: R1 Δ*Map4k3 E1* clone 1	This paper	N/A
Mouse: R1 Δ*Map4k3 E1* clone 2	This paper	N/A
Mouse: R1 Δ*Map4k3 E2* clone 1	This paper	N/A
Mouse: R1 Δ*Map4k3 E2* clone 2	This paper	N/A
Recombinant DNA		
px330-U6-Chimeric_BB-CBh-hSpCas9	Addgene	Cat#42230
Fgf5-HAL-PE-pdTK-HAR	Thomas et al., [18]	N/A

(*Continued on next page*)

Please cite this article in press as: Thomas et al., Enhancer cooperativity can compensate for loss of activity over large genomic distances, Molecular Cell (2024), https://doi.org/10.1016/j.molcel.2024.11.008

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
BamHI - β-globin HAR - NotI	BioCat GmbH	N/A
targeting vector (pGemT-HAL-pdTK-spacer-TK-mCherry-pA-HAR)	This Paper	N/A
pGemT (circularised)	Promega	Cat#A362A
pGemT-FRT-lox66-FRT-blasticidin-deltaTK(7 additional versions of this plasmid with integration of Map4k3 E1, Map4k3 E2, Rybp E2, eNanog, SCR, SCR-rev and Fgf5-PE were also generated)	This Paper	N/A
pGemT-FRT3-loxm2/66-FRT3-puromycin-deltaTK (6 additional versions of this plasmid with integration of Map4k3 E1, Map4k3 E2, Rybp E2, eNanog, SCR and SCR-rev were also generated)	This Paper	N/A
pGemT-FRT5-lox2272/66-FRT5-hygromycin-deltaTK (6 additional versions of this plasmid with integration of Map4k3 E1, Map4k3 E2, Rybp E2, eNanog, SCR and SCR-rev were also generated)	This Paper	N/A
FlpO-expressing plasmid	Provided by the lab of Stefan Ameres	N/A
pGL3-SV40 promoter-luciferase-Map4k3 E1	This Paper	N/A
pGL3-SV40 promoter-luciferase-Map4k3 E2	This Paper	N/A
pGL3-SV40 promoter-luciferase-Rybp E2	This Paper	N/A
pGL3-SV40 promoter-luciferase-eNanog	This Paper	N/A
pGL3-SV40 promoter-luciferase-SCR	This Paper	N/A
pGL3-SV40 promoter-luciferase-SCR-rev	This Paper	N/A
Software and algorithms		
FlowJo (version 10.5.3)	BD Life Sciences - Biosciences	N/A
nf-core atacseq pipeline (version 2.1.2)	Ewels et al.[61]	N/A
Reform tool	https://github.com/gencorefacility/reform	N/A
Genrich (v0.6.1)	https://github.com/jsh58/Genrich	N/A
Other		
BD FACSAria III Cell Sorter	BD Life Sciences - Biosciences	N/A
Mus Musculus Castaneus DNA	Provided by the lab of Kikue Tachibana	N/A
HEK-293 DNA	Provided by the lab of Dea Slade	N/A

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

General culture conditions of mouse embryonic stem cells

For all work in this study we used the male v6.5 mouse embryonic stem cell line cultured as described in Thomas et al.[18]. For maintenance, cells were grown on 6/12-wells coated with first poly-L-ornithine hydrobromide (6 μg/ml in PBS, 1h at 37 °C) and then laminin (1.2 μg/ml in PBS, 1 hour at 37°C).

Cells were cultured in base medium DMEM/F12 without Hepes with 4 mg/mL Lipid-Rich Bovine Serum Albumin, 1x MACS NeuroBrew-21 with Vitamin A, 1x NEAA, 50 U/mL Penicillin-Streptomycin, 1 mM Sodium Pyruvate and 1x 2-Mercaptoethanol. The base medium was supplemented with 3.3 mM CHIR-99021, 0.8 mM PD0325901 and 10 ng/mL hLIF (provided by the VBCF Protein Technologies Facility, https://www.viennabiocenter.org/facilities/).

Please cite this article in press as: Thomas et al., Enhancer cooperativity can compensate for loss of activity over large genomic distances, Molecular Cell (2024), https://doi.org/10.1016/j.molcel.2024.11.008

For splitting, cells were treated with 1x Trypsin-EDTA solution at 37 °C until cells detached. Trypsinization was stopped with 2i+LIF medium with 10% FSC, cells were harvested by centrifugation at 300g for 3 min, resuspended in 2i+LIF and seeded in appropriate ratios.

Generation of reporter cell line

For KI of the reporter gene into the β-globin locus, we first generated a gRNA-expressing plasmid and a targeting vector. Therefore, a single gRNA was designed to target the locus roughly 1.5 kb downstream of the *Hbb-y* gene. As described before, the gRNA was cloned into Addgene plasmid #42230.[18] Due to a SNP in the underlying sequence, this gRNA recognises the C57BL/6 but not the 129/sv allele of v6.5 mouse ESCs, allowing us to introduce the reporter gene specifically into the β-globin-locus on the C57BL/6 allele.

We used Gibson assembly to generate the targeting vector to insert the reporter gene into the *β-globin*-locus unless noted otherwise. Homology arms were designed to disrupt the gRNA recognition sequence upon successful KI. All guide RNAs used in this study are listed in Table S1, primer sequences used for cloning in this study are listed in the Table S2, all single strand DNA oligos are listed in Table S4.

We first amplified a 77 bp long minimal TK promoter followed by the mCherry coding sequence by PCR and inserted the fragment into an NcoI-HF-digested (NEB) pGemT-vector (Promega). We amplified the left homology arm (637 bp) targeting the β-globin locus roughly 1.5 kb downstream of the *Hbb-y* gene from R1 ESC DNA by PCR. We inserted it upstream of the TK-promoter into the SphI-HF-digested (NEB) TK-mCherry plasmid, leaving the SphI-site intact. We used this SphI site to introduce a loxP-flanked puromycin-deltaTK cassette and a 1.5 kb spacer from inert human DNA between the left homology arm and the TK promoter. The loxP-flanked puromycin-deltaTK cassette was PCR-amplified from a previously generated targeting vector.[18] Using a forward primer that mapped to the loxP site upstream of the puromycin-deltaTK cassette but contained mutations, we turned the upstream loxP site into a lox71 site. The 1.5 kb spacer was amplified from HEK-293 genomic DNA, kindly provided by the lab of Dea Slade. We then inserted a 232 bp long BGH polyA sequence downstream of the mCherry coding sequence into the NotI-HF-digested (NEB) plasmid, leaving the NotI-site intact and introducing an additional BamHI-site. As we did not manage to PCR-amplify the right homology arm, we ordered the synthesised fragment (832 bp) integrated into a plasmid and flanked by a BamHI- (upstream) and a NotI-site (downstream) from BioCat GmbH. We obtained the right homology arm fragment by BamHI-HF- and NotI-HF-digestion (both NEB) and ligated it into the likewise BamHI-HF- and NotI-HF-digested HAL-pdTK-spacer-TK-mCherry-pA plasmid, downstream of the polyA site.

For KI of the reporter gene, v6.5 mouse ESCs[38] were cultured and transfected by lipofection with 400 ng of circular targeting vector, 400 ng of gRNA-containing plasmid and 4 μL of Lipofectamine 2000 Transfection Reagent (Invitrogen), as described before.[18] In brief, cells were transferred to a 10 cm dish the day after transfection and selection with Puromycin (2 μg/mL, InvivoGen) was started within 48 h of transfection. After a week of selection, single colonies were picked. Integration was validated by PCR with primers "KI validation 1 forward+reverse" (mapping upstream of the left homology arm in the genome and downstream of the left homology arm in the insert) and "KI validation 2 forward+reverse" (mapping upstream of the right homology arm in the insert and downstream of the right homology arm in the genome). All analysed clones were heterozygous, with the reporter gene being inserted on the C57BL/6 allele as judged by SNPs in Sanger-sequenced PCR products (Microsynth AG). Clones were expanded and subsequently transfected with a plasmid expressing Cre-recombinase to remove the puromycin-deltaTK cassette and leave a single lox71 site (in reverse orientation) behind. After one week of selection with Ganciclovir (500 ng/mL, InvivoGen), single clones were picked, and removal of the puromycin-deltaTK cassette was confirmed with primers "KI validation after Cre forward+reverse" (mapping up- and downstream of the excised cassette). Overlapping PCR products (primers "reporter validation 1-4 forward+reverse") spanning the entire insert, the homology arms and the genome up- and downstream of the insertion were submitted for Sanger sequencing to confirm the intactness of the β-globin-locus, the homology arms and the inserted sequence. We selected two clones with reporter gene integration for subsequent integration of the remaining LPs.

To introduce lox2272/71 and loxm2/71 sites roughly 25 and 75 kb upstream of the mCherry reporter gene, we designed gRNAs and cloned them into Addgene plasmid #42230. We also ordered single-stranded (ss) DNA oligos from Microsynth AG, containing the respective lox site (34 bp; in reverse orientation) and left and right homology arm. The homology arm at the 5' end of the respective oligo was 66 bp long, and the homology arm at the 3' end was 50 bp (150 bp total, including the lox site). Homology arms were designed to disrupt the gRNA recognition sequence upon successful KI.

We used two independent clones with reporter gene integration to sequentially knock-in the two lox sites. Therefore, we transfected 500 ng of gRNA-expressing plasmid, 1000 ng of ssDNA-oligo and 100 ng of plasmid expressing a fluorescent marker with 10 μl of Lipofectamine 2000 Transfection Reagent (Invitrogen). Transfected cells were processed as described before,[18] i.e., single fluorescent cells were sorted into 96-well plates. Successful insertion of the lox sites into the genome was initially confirmed by PCR with forward primers overlapping the integrated lox sites (and thus only binding to the modified alleles) and reverse primers downstream of the integration site ("lox KI 25/75 kb initial forward+reverse"). Selected clones were tested further with a more upstream forward primer binding to both alleles ("lox KI 25/75 kb validation forward"), and resulting PCR products were subcloned into a pGEM-T vector (Promega, pGEM-T Vector Systems) for subsequent Sanger sequencing of both alleles. The recognition sequence of the gRNA used for the lox site at 25 kb contains SNPs on the 129/sv allele, allowing us to integrate the lox2272/71 site on the C57BL/6 allele specifically. For integration at 75 kb, we did not design a gRNA overlapping a SNP. Instead, we chose heterozygous clones with the integration of the loxm2/71 site on the C57BL/6 allele, as judged by SNPs surrounding the integration site.

Please cite this article in press as: Thomas et al., Enhancer cooperativity can compensate for loss of activity over large genomic distances, Molecular Cell (2024), https://doi.org/10.1016/j.molcel.2024.11.008

METHOD DETAILS

Design of plasmids for enhancer integration

We generated three plasmids to integrate enhancers at 1.5, 25 and 75 kb distance upstream of the reporter gene by targeting the lox71, loxm2/71 and lox2272/71 sites. We first introduced a single FRT site into a pGemT-vector (Promega). Therefore, forward and reverse DNA oligonucleotides - containing the FRT sequence as well as the overhangs required for cloning - were ordered from Microsynth AG, annealed and cloned into a NcoI-HF- and SphI-HF- digested pGemT plasmid, as described before.[18] We generated three plasmids containing FRT, FRT3 and FRT5 sites, respectively. The original restriction sites were disrupted during this process, but an additional NcoI-site was introduced downstream of the FRT sites as part of the inserted oligo. We then ordered additional DNA oligos and introduced lox66 sites downstream of the FRT sites of the resulting NcoI-HF- and NotI-HF-digested plasmids (both NEB). For the plasmid containing an FRT site, we introduced a lox66 site; for the plasmid containing the FRT3 site, we introduced a loxm2/66 site; and for the plasmid with the FRT5 site, a lox2272/66 site.

We digested the resulting plasmids by NotI-HF (NEB) and inserted cassettes expressing fusions of antibiotic resistance genes with deltaTK downstream of the lox66 sites by Gibson assembly. We inserted blasticidin-deltaTK into the FRT-lox66-plasmid, puromycin-deltaTK into the FRT3-loxm2/66-plasmid and hygromycin-deltaTK into the FRT5-lox2272/66-plasmid. The forward primer used for amplification of the cassettes included an additional FRT site in the case of blasticidin-deltaTK, an additional FRT3 site in the case of hygromycin-deltaTK and an additional FRT5 site for puromycin-deltaTK. The NotI site between the lox and the newly inserted FRT site was restored and used to integrate enhancer sequences.

All in all, we generated the following three plasmids:

- pGemT-FRT-lox66-FRT-blasticidin-deltaTK
- pGemT-FRT3-loxm2/66-FRT3-puromycin-deltaTK and
- pGemT-FRT5-lox2272/66-FRT5-hygromycin-deltaTK.

These plasmids were digested by NotI-HF, and enhancer sequences amplified from *castaneus* mouse strain DNA (kindly provided by the lab of Kikuë Tachibana) were inserted by Gibson assembly. Primers for amplifying enhancer sequences were designed to encompass the central p300 peak (ChIP-seq from Buecker et al.[34]). The resulting plasmids were transfected together with a Cre-recombinase expressing plasmid into the reporter cell line for insertion of enhancers at the lox71 sites of the reporter locus.

Integration of enhancer sequences

To integrate enhancer sequences, 200,000 cells of the reporter cell line were plated and transfected by lipofection on the following day. Therefore, 250 ng of enhancer-containing plasmid, 1750 ng of Cre-recombinase expressing plasmid and 10 μL of Lipofectamine 2000 Transfection Reagent (Invitrogen) were used. Cells were selected with Blasticidin (10 μg/mL, InvivoGen), Hygromycin B (400 μg/mL, Sigma-Aldrich) or Puromycin (2 μg/mL, InvivoGen), depending on the selection cassette present in the enhancer plasmid. Single colonies were picked after a week of selection. Plasmid integration was validated by PCR with forward primers mapping to the integrated plasmid backbone (“integration forward”) and reverse primers mapping downstream of the respective lox site (“integration 1.5/75 reverse; for 25 kb: “lox KI 25 kb initial reverse” that was used before for validating KI of the lox site). Colonies with integration of enhancer plasmid were expanded and transfected with plasmid expressing FlpO-recombinase (a more active, codon-optimized version of Flp kindly provided by the lab of Stefan Ameres) to remove the plasmid backbone including the selection cassette (200,000 cells, 2 μg of FlpO-recombinase expressing plasmid, 10 μL of Lipofectamine 2000 Transfection Reagent (Invitrogen)). Cells were passaged, seeded at low density the day after transfection and selected with Ganciclovir (5 μg/mL, InvivoGen) for one week. Single colonies were picked. PCR confirmed the removal of the selection cassette with an enhancer-specific forward primer combined with a primer downstream of the respective lox71 site (“integration 1.5/75 reverse”; for 25 kb: “lox KI 25 kb initial reverse”). In addition, the PCR with primers “integration forward” and “integration 1.5/75 reverse” or “lox KI 25 kb initial reverse” was repeated. As the forward primer maps to the plasmid backbone, the absence of a band in that PCR confirms the complete removal of the plasmid backbone in the entire cell population.

Allele-specific primers up and downstream of the lox sites were designed. In the case of the lox71 site at 1.5 kb, a forward primer upstream of the lox71 site (“1.5 forward”, recognising both alleles) was combined with a reverse primer mapping to the spacer sequence that is only integrated on the C57BL/6 allele as part of the reporter gene and has been used before for confirming plasmid and enhancer integration (“integration 1.5 reverse”). For the other two lox sites, forward primers mapping just upstream of and partially overlapping the lox sites (“25/75 forward”) were combined with reverse primers that had been used before for confirming plasmid and enhancer integration and recognise both alleles (“integration 75 reverse”/“lox KI 25 kb initial reverse”). These primer combinations gave rise to PCR products spanning the insert and the surrounding genome. Thus, PCR products having the expected size confirmed the intactness of the inserted sequence and the surrounding genome. Selected PCR products were submitted for Sanger sequencing to further confirm the absence of mutations in the inserted enhancer sequences. The size of the SCR and SCR-rev prevented the amplification of a single PCR product spanning the entire insert. Instead, we performed two PCRs giving rise to partially overlapping PCR products that together span the entire insert and the surrounding genome (“1.5/25/75 forward+SCR-reverse” and “SCR-forward+integration 1.5/25/75 reverse” for SCR, “1.5/25/75 forward+SCR-forward” and

"SCR-reverse+integration 1.5/25/75 reverse" for SCR-rev). We introduced an initial set of enhancers (*Sox2 SCR*, *eNanog*, *Rybp E2*, *Map4k3 E1*, *Map4k3 E2*) into the two clones we derived from the initial knock-in of the reporter gene (see above). For each of the two initial reporter cell lines, we generated a single clone with the enhancer integration, treated with FlpO and selected a single clone after FlpO-treatment (resulting in a total of 2 clones for each enhancer integration after FlpO treatment, with each of them being derived from a different parental reporter cell line clone). Since the results in the two clones of the reporter cell line were highly reproducible, we integrated the remaining individual enhancers (*Rybp E1*, *Sox2 E15*, *Sox2 E19* and *Sox2 E20*) only into clone 1 of the initial reporter cell line. After Cre-mediated insertion of the enhancer-containing plasmid, two clones were independently expanded and transfected with FlpO. After selection, two clones per enhancer integration were expanded and verified, resulting in a total of four clones per single enhancer.

To integrate the dual enhancers, we selected 1-2 individual clonal cell lines before removing the plasmid backbone, i.e., before FlpO treatment. We integrated the second enhancer by transfecting the cells with Cre-recombinase and the second enhancer plasmid. We included both selections to ensure that both enhancers are integrated and selected single clones. We identified positive clones via PCR with the primers described above and removed both backbones by FlpO treatment afterwards. For all cell lines, extensive PCR validation ensured both enhancers were integrated. Multiple independently derived clones were selected for further analysis.

Cas9-targeted sequencing

To verify correct integrations and that the whole locus is not rearranged during the construction of the synthetic locus, we performed targeted nanopore sequencing with Cas9-guided adapter ligation for selected cell lines, both before and after enhancer integration. We designed two small pools of gRNA libraries targeting the modified reporter gene locus. The distance between each gRNA is 5 kb. Cas9 protein, tracRNA and crRNA were ordered from IDT. We cooperated with the Vienna BioCenter NGS core facility and followed the protocol for Cas9-targeted sequencing using the Cas9 sequencing Kit(SQK-CS9109) from Oxford Nanopore Technologies.

FACS analysis of mCherry expression

To assess reporter gene activation upon enhancer integration, mCherry levels were measured by FACS. Therefore, 200,000 cells were plated on a 6-well plate in 2i/LIF medium. For each experiment, the two parental reporter cell line clones without enhancer integration and an untransfected v6.5 control were included. After two days, cells were collected by adding trypsin-EDTA (Sigma-Aldrich). Trypsinisation was stopped after incubation at 37 °C for 7 minutes by adding 2i/LIF medium containing 10% serum. Cells were resuspended and centrifuged for 3 minutes at 300 g, the supernatant was removed, and the cell pellet was resuspended in 2i/LIF medium. mCherry-fluorescence was measured with a BD LSRFortessa Flow Cytometer (BD Life Sciences - Biosciences).

FACS sorting of mCherry-populations

Upon integration of the *Nanog* enhancer at 1.5 kb, we observed some clones with a fraction of mCherry-negative cells. We sorted mCherry-positive and -negative cells with a BD FACSAria III Cell Sorter (BD Life Sciences - Biosciences). Sorting gates were determined by measuring mCherry fluorescence of a clone without a fraction of mCherry-negative cells: All cells with mCherry signal lower than the lowest-expressing cells of that clone were sorted as negative, and all remaining cells were sorted as positive. Sorted cells were kept in culture, and mCherry levels were analysed with a BD LSRFortessa Flow Cytometer (BD Life Sciences - Biosciences) every few passages.

Luciferase assays

For luciferase assays, we used a pGL3 plasmid with the Firefly luciferase coding sequence under the control of the SV40 promoter followed by a poly-adenylation signal (Promega). The same enhancer sequences we integrated into the reporter locus were amplified by PCR from *castaneus* mouse strain DNA and inserted downstream of the poly-adenylation signal by Gibson assembly. The primer sequences used to generate these plasmids are indicated in Table S2. Luciferase assays were slightly adapted but otherwise performed as described previously.[18] 10,000 cells from the v6.5 cell line were plated on a 96-well plate and immediately transfected with 120 ng of enhancer-luciferase plasmid, 4 ng of Renilla control plasmid (Promega) and 0.62 μL of Lipofectamine 2000 Transfection Reagent (Invitrogen). The medium was removed 5-7 h after transfection, and fresh 2i/LIF medium was added. Firefly and Renilla luciferase activity were measured 48 h after transfection. Background-subtracted Firefly measurements were normalised to background-subtracted Renilla values. The resulting values were normalised to the control plasmid without enhancer integration. To detect statistically significant increases in luciferase activity compared to the no-enhancer control, we performed one-sided Welch one-sample t-tests on the resulting normalised values to assess whether they are significantly higher than 1 (the value of the no-enhancer control).

Enhancer KO and RT-qPCR

Enhancers at the *Map4k3* locus were deleted in male R1 WT ESCs[60] as described.[18] RNA was extracted from confluent 6-well plates, and resulting *Map4k3* expression levels were measured by RT-qPCR as described before.[18] All qPCR primers used in this study are listed in Table S3. Expression levels of each replicate were normalised first to Rpl13a and then to WT. To assess statistically significant decreases in expression upon enhancer deletion, we performed one-sided Welch one-sample t-tests on the resulting

Please cite this article in press as: Thomas et al., Enhancer cooperativity can compensate for loss of activity over large genomic distances, Molecular Cell (2024), https://doi.org/10.1016/j.molcel.2024.11.008

WT-normalized values to assess whether they are significantly lower than 1 (as all values are normalised to WT, a value of 1 corresponds to WT expression levels).

CTCF binding sites KO

Two CTCF binding sites are found between LP 25 kb and LP 75 kb. To delete the CTCF binding site, we designed one gRNA targeting the motif of CTCF peak 1, and two gRNAs surrounding the motif of CTCF peak 2. We first deleted the CTCF peak 1 in four cell lines (*eNanog* at 75 kb, *ME1* at 75 kb, *eNanog* at 75 kb with *ME2* at 25 kb and *ME1* at 75 kb with *ME2* at 25 kb). We choose two clones of CTCF peak 1 deletion from each cell line for further KO of CTCF peak 2. Several double KO clones of CTCF binding sites from each cell line are verified by PCR and Sanger sequencing.

RNA Flow-FISH

To validate the correlation between the levels of reporter gene fluorescence and the number of mRNAs, mCherry mRNA levels were quantified performing RNA Flow-FISH.

For this purpose, PrimeFlow RNA Assay Kit was purchased from Thermo Fisher Scientific (88–18005-204) along with Alexa Fluor 647 mCherry target probes. The assay was performed four times following the recommended protocol. Shortly, the cells were first fixed and then permeabilized using the PrimeFlow RNA Fixation/Permeabilization Buffers provided in the kit. Next, cells were incubated with mCherry target probe for the hybridization step, followed by signal amplification via incubation with corresponding pre-amplifiers, amplifiers and fluorescent label probes. Cells were then analyzed on a BD LSRFortessa Flow Cytometer (BD Life Sciences - Biosciences). The subsequent gating and data analysis was done in FlowJo software.

As negative controls, we also performed the entire protocol with samples containing cells without our synthetic locus, therefore lacking mCherry reporter gene; and samples lacking the mCherry target probes.

QUANTIFICATION AND STATISTICAL ANALYSIS

Quantification and statistical analysis of FACS data

In each FACS experiment, we always included the no enhancer control cell line to account for small changes in laser settings and where indicated the v6.5 WT control. In addition, we only compared experimental replicates that were analyzed in the same experiments.

The resulting FCS files were analysed with the FlowJo software (BD Life Sciences – Biosciences, version 10.5.3). We visualised cell populations as histograms normalised to the maximum cell count for each cell line to analyse the fluorescence distribution in a cell population. We calculated the mean of mCherry fluorescence with the FlowJo software to compare different cell lines and replicates. To account for background fluorescence, we subtracted the mean fluorescence of an untransfected v6.5 control from all values and used the resulting values for the subsequent analysis described below.

We calculated the average of mean fluorescence and plotted the resulting values for each measurement in bar graphs (see Figures). To assess statistically significant increases in mCherry fluorescence upon enhancer integration, we performed one-sided paired t-tests on cell lines with enhancer integration compared to their respective parental reporter cell line clone without enhancers.

To compare the distance dependency of different enhancers, we normalised the mean fluorescence of each enhancer construct as well as the "no enhancer" control to the clone carrying the enhancer at 1.5 kb. We performed one-sided paired t-tests on the normalised values to identify statistically significant increases compared to the respective "no enhancer" control.

To compare combinations of enhancers, we calculated the mean mCherry expression of all cell lines, either with one or the combination of two enhancers, and subtracted the mean mCherry expression of the ‘‘no enhancer’’ control from the mean expression of each cell line. We calculated the expected additive expression by adding the resulting (‘‘no enhancer’’ control subtracted) values. Normalized values from independently derived enhancer constructs were pooled and the mean expression was calculated for each replicate. The expected additive expression was calculated by adding the mean over all replicates. A one-sided paired t-test was performed to assess whether the dual enhancer cell line exceeds the expected additive expression.

Modelling of expected combined activities

The individual activity of enhancers in the 75 kb, 25 kb and 1.5 kb pads were scaled using the negative control conditions where the three pads contain inactive control sequences (-/-/-). Therefore, they correspond to fold-changes normalized to the basal activity of the core promoter expressed in log2. Hence, for a given combination C, $\log_2$ additive and multiplicative predicted values were computed using the following formulas:

$$C_{additive} = log\,2(2^{i_{75}} + 2^{i_{25}} + 2^{i_{1.5}} - 2)$$

$$C_{multiplicative} = i_{75} + i_{25} + i_{1.5}$$

Please cite this article in press as: Thomas et al., Enhancer cooperativity can compensate for loss of activity over large genomic distances, Molecular Cell (2024), https://doi.org/10.1016/j.molcel.2024.11.008

Where i_{75}, i_{25} *and* $i_{1.5}$ correspond to the $\log_2$ individual enhancer activities inserted 75 kb, 25 kb and 1.5 kb upstream of the core promoter, respectively. Finally, we fitted a multiplicative model with interaction term using $\log_2$ activity values and the *lm* function in R:

$$C_{activity} = \beta_0 + \beta_1 i_{75} + \beta_2 i_{25} + \beta_3 i_{1.5} + \beta_{12} i_{75} i_{25} + \beta_{13} i_{75} i_{1.5} + \beta_{23} i_{25} i_{1.5} + \beta_{123} i_{75} i_{25} i_{1.5} + \epsilon$$

Where β_0 is the intercept, and β_1, β_2 and β_3 coefficients represent the individual contributions of each enhancer's $\log_2$ activity to the overall activity contribution. β_{12}, β_{13} and β_{23} are the coefficients for the two-way interactions, and capture the combined effects when two enhancers are active simultaneously, beyond what would be expected from their individual effects alone. β_{123} is the coefficient of the three-way interaction, indicating how the simultaneous presence of enhancer activities at 75 kb, 25 kb, and 1.5 kb influences the core promoter activity. However, since no such combination was present in our dataset, this coefficient was automatically set to NA. Finally, ϵ is the error term.

The performance of each model was assessed using R-squared (R^2) computed with the following formula:

$$R^2 = \frac{SS_{regression}}{SS_{total}}$$

Where, $SS_{regression}$ and SS_{total} correspond to the Sum of Squares due to regression and the total Sum of Squares, respectively. For fitted models, R-squared values were adjusted (Adj. R^2) to account for the number of predictors.

RT-qPCR of endogenous genes in the synthetic locus

Cell lines with the SCR enhancer integrated at 1.5 kb, 25 kb, and 75 kb before and after FlpO treatment were selected to investigate the endogenous gene expression in the reporterlocus. We also included the v6.5 ES cell line and parental clone 1 cell line without any enhancers integrated as negative controls. RNA of each cell line was extracted from confluent 6-well plates, and was further reverse transcribed to cDNA. Two genes in the synthetic locus, *Olfr67* and *Hbb-*γ, were selected to analyse their expression level by RT-qPCR as described above. To assess statistically significant increases in expression upon enhancer integration, we performed one-sided Welch one-sample t-tests on the resulting v6.5-normalized values to assess whether they are significantly higher than 1 (as all values are normalised to v6.5, a value of 1 corresponds to v6.5 expression levels).

ATAC-seq analysis

For the ATAC-seq analysis performed by the Next Generation Sequencing facility at the Vienna BioCenter Core Facilities (VBCF) we seeded 100.000 cells and harvested 48h later, resuspending in Dulbecco's Phosphate Buffered Saline (PBS, Sigma-Aldrich).

We used cell lines with the SCR enhancer at 1.5 kb and 75 kb as well as *Map4k3 E2* at the same distances. V6.5 ES cell line and parental clone 2 without enhancer cell lines were used as controls. Over 100k cells with over 85% viability were required by the facility and the biological replicates were delivered on consecutive days. Samples were run on NovaSeq S4 PE150 XP.

We used the reform tool (https://github.com/gencorefacility/reform) to generate custom-made mm10 genome and annotation files for the modified cell lines with integration of the reporter gene, the landing pads and in some cases the different enhancers. Fastq files were analyzed with version 2.1.2 of the nf-core atacseq pipeline,[61] aligning to the custom genomes generated above. For the files aligned to the genome that only contains integration of the reporter gene and the landing pads, but not of the respective enhancers, we performed peak calling with Genrich (v0.6.1) and identified consensus peaks that overlap at least 50% between replicates with the samtools intersect tool. Peaks overlapping blacklisted regions (from http://mitra.stanford.edu/kundaje/akundaje/release/blacklists/mm10-mouse/mm10.blacklist.bed.gz) were removed with substractBed, and peaks as well as aligned reads were displayed in IGV[62] using the 'Group Autoscale' function.