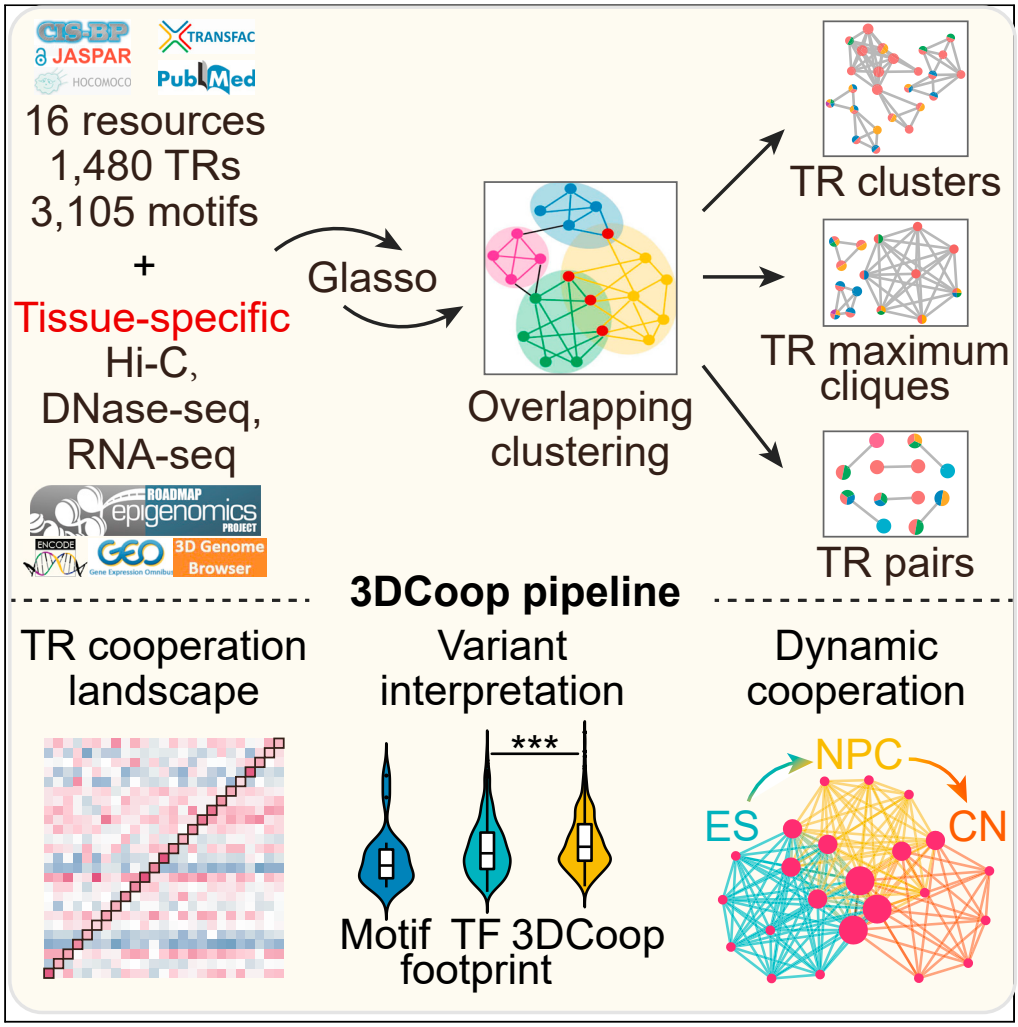
iScience

CellPress
OPEN ACCESS

Article

Interrogating cell type-specific cooperation of transcriptional regulators in 3D chromatin



Xianfu Yi, Zhanye Zheng, Hang Xu, ..., Hui Cheng, Lei Shi, Mulin Jun Li

yixfbio@gmail.com (X.Y.)
mulinli@connect.hku.hk (M.J.L.)

Highlights

Computational inference of transcriptional regulator (TR) cooperation in 3D chromatin

A landscape of 3D TR cooperation across 40 human tissue/cell types

TR cooperation can better interpret the disease-causal variants identified by GWAS

Cooperation of certain TRs shapes context-specific gene regulation in cell development

Yi et al., iScience *24*, 103468
December 17, 2021
https://doi.org/10.1016/j.isci.2021.103468

iScience

CellPress
OPEN ACCESS

Article

Interrogating cell type-specific cooperation of transcriptional regulators in 3D chromatin

Xianfu Yi,[1,2,*] Zhanye Zheng,[3] Hang Xu,[2] Yao Zhou,[2,3] Dandan Huang,[2,4] Jianhua Wang,[2,3] Xiangling Feng,[2,3] Ke Zhao,[2,3] Xutong Fan,[2,3] Shijie Zhang,[2,3] Xiaobao Dong,[2,5] Zhao Wang,[3] Yujun Shen,[3] Hui Cheng,[6] Lei Shi,[4] and Mulin Jun Li[2,3,7,8,*]

SUMMARY

Context-specific activities of transcription regulators (TRs) in the nucleus modulate spatiotemporal gene expression precisely. Using the largest ChIP-seq data and chromatin loops in the human K562 cell line, we initially interrogated TR cooperation in 3D chromatin via a graphical model and revealed many known and novel TRs manipulating context-specific pathways. To explore TR cooperation across broad tissue/cell types, we systematically leveraged large-scale open chromatin profiles, computational footprinting, and high-resolution chromatin interactions to investigate tissue/cell type-specific TR cooperation. We first delineated a landscape of TR cooperation across 40 human tissue/cell types. Network modularity analyses uncovered the commonality and specificity of TR cooperation in different conditions. We also demonstrated that TR cooperation information can better interpret the disease-causal variants identified by genome-wide association studies and recapitulate cell states during neural development. Our study characterizes shared and unique patterns of TR cooperation associated with the cell type specificity of gene regulation in 3D chromatin.

INTRODUCTION

A critical problem in functional genomics has long been the identification of spatiotemporal patterns of gene regulation in different biological conditions. Only tens of thousands of genes have been found in mammalian genomes, whereas over millions of *cis*-regulatory elements (CREs), such as promoters, enhancers, silencers, and insulators, regulate gene transcription precisely and control cell functions faithfully (Consortium, 2012; Consortium et al., 2020a). Such a great volume of discovered CREs, together with thousands of transcription regulators (TRs), which include transcription factors (TFs) and chromatin regulators, underpin the basal and tissue/cell type-specific transcription events in the nucleus. However, when and how these elements and TRs cooperatively regulate gene expression and cellular function remain largely unknown (Kazemian et al., 2013; Maston et al., 2006; Spitz and Furlong, 2012; Voss and Hager, 2014). Besides, the action of TRs usually depends on three-dimensional (3D) chromatin folding to modulate gene expression precisely and in turn shape context-specific genome organization (Kim and Shendure, 2019; Stadhouders et al., 2019). Therefore, accurate detection of cell type-specific TR cooperation in 3D chromatin would enhance the understanding of the complex gene regulation in cell differentiation and disease development (Petrovic et al., 2019; Tian et al., 2020).

Many genomic and epigenomic approaches for identifying *cis*-regulatory modules and the associated protein complexes in the one-dimensional (1D) genome space have been well documented (Gerstein et al., 2012; Guo and Gifford, 2017; Hardison and Taylor, 2012; Klemm et al., 2019). Recent chromosome conformation capture (3C)-derived techniques have led to a revolution in the detection of high-resolution genome-wide DNA interactions (Kempfer and Pombo, 2020; McCord et al., 2020; Robson et al., 2019; Xu et al., 2020). Several methods utilize TRs ChIP-seq (chromatin immunoprecipitation followed by sequencing) data, 3D chromatin conformation information, and protein-protein interaction (PPI) networks to infer TR cooperation in specific cell types, such as DBPnet, 3CPET, and HidPET (Djekidel et al., 2015; Wang et al., 2019; Zhang et al., 2016). However, they are highly reliant on the availability of ChIPped factors and loop-level chromatin interactions. Despite data limitations on the endogenous binding sites of sufficient TRs and high-resolution chromatin interactions for the majority of human or mouse tissue/cell types,

[1]School of Biomedical Engineering and Technology, Tianjin Medical University, Tianjin 300070, China

[2]Department of Bioinformatics, The Province and Ministry Co-sponsored Collaborative Innovation Center for Medical Epigenetics, Tianjin Medical University, Tianjin 300070, China

[3]Department of Pharmacology, Tianjin Key Laboratory of Inflammation Biology, School of Basic Medical Sciences, Tianjin Medical University, Tianjin 300070, China

[4]Department of Biochemistry and Molecular Biology, School of Basic Medical Sciences, Tianjin Medical University, Tianjin 300070, China

[5]Department of Genetics, School of Basic Medical Sciences, Tianjin Medical University, Tianjin 300070, China

[6]State Key Laboratory of Experimental Hematology, Chinese Academy of Medical Sciences, Tianjin 300070, China

[7]Department of Epidemiology and Biostatistics, Tianjin Key Laboratory of Molecular Cancer Epidemiology, National Clinical Research Center for Cancer, Tianjin Medical University Cancer Institute and Hospital, Tianjin Medical University, Tianjin 300070, China

[8]Lead contact

*Correspondence: yixfbio@gmail.com (X.Y.), mulinli@connect.hku.hk (M.J.L.)

https://doi.org/10.1016/j.isci.2021.103468

genomic DNase I footprinting enables the quantitative delineation of TR occupancy within native chromatin (Vierstra et al., 2020). Furthermore, chromatin loops can be accurately inferred from genome-wide contact maps of 3C-derived assays (Salameh et al., 2020; Schwessinger et al., 2020; Zhang et al., 2019). These technical advances allow computational predictions of global TR binding and short-range interaction among CREs across extensive tissue/cell types, which greatly facilitates the systematic investigation of context-specific TR cooperation in 3D chromatin.

In the present study, we initially present an analysis of 386 TR occupancy profiles in the K562 human myelogenous leukemia cell line, one of the ENCODE (Encyclopedia of DNA Elements) Tier 1 common cell types with the most ChIPped chromatin-associated proteins. By applying a graphical model and overlapping clustering on TR ChIP-seq profiles and high-resolution Hi-C interactions, we reveal interplay between many known and novel TRs sustaining basal and leukemia-specific gene regulation in 3D chromatin. To characterize context-specific TR cooperation across different biological conditions, we comprehensively integrate TR motif information, DNase-seq/ATAC-seq (DNase I-hypersensitive site sequencing/assay for transposase-accessible chromatin with high-throughput sequencing) data, and predict Hi-C chromatin loops for 40 human tissue/cell types. This marks the first time a tissue/cell type-shared and -specific 3D TR cooperation landscape has been constructed. Based on the network modularity analyses, disease-causal variant enrichment analyses, and dynamic TR cooperation identification during cell development, we demonstrate how these communications determine unique cellular functions and implicate disease pathogenesis.

RESULTS

Inferring 3D TR cooperation using ChIP-seq of 386 TRs and chromatin loops in K562 cells

A Gaussian graphical model (GGM) comprises items and lines between them (Epskamp et al., 2018). In the GGM, the line between two items captures partial correlation, which is controlled for all other items included (Bhushan et al., 2019). This feature makes a key advantage of GGM that it avoids spurious correlations. The nonparanormal graphical model, one derivative of GGM, is a semiparametric generalization for continuous variables and has emerged as an important tool for modeling dependency structure between items (Mulgrave and Ghosal, 2020; Xue and Zou, 2012; Zhang, 2019, 2020). These models can be incorporated to precisely infer the dependency structures of biomolecules (Liu et al., 2012; Yin and Li, 2011; Zhang et al., 2016). The graphical Lasso is a sparse penalized maximum likelihood estimator for the precision matrix of a multivariate Gaussian distribution when observations are limited (Friedman et al., 2008). We optimized the graphical Lasso algorithm on the TR network model applied in DBPnet (Zhang et al., 2016) to detect potential interplay among TRs in a high-dimensional chromatin environment (Figure S1A). To maximize the model's generalization capability and to minimize the impact of DNA-binding signals among the different TRs, measured either by ChIP-seq (including CUT&RUN and CUT&Tag technologies) or by computational footprinting, we used peak information instead of sequencing reads when constructing the TR-specific contact map together with Hi-C loops. In addition, we incorporated overlapping clustering to estimate network modules, which allows the premise that a single TR can be involved in multiple cooperation communities as *in vivo*. Based on these extended features and other improvements, we implemented a new pipeline, termed 3DCoop, to identify 3D TR cooperation by leveraging genome-wide TR-binding sites and high-resolution chromatin interactions (see STAR Methods for details). Here, we considered cooperation as something that is revealed by two TR molecules being in proximity to each other considering 3D context and tissue/cell type specificity, regardless of whether there is any biochemical/physical/sequential cooperation involved or not (Wang et al., 2019; Zhang et al., 2016).

We initially applied 3DCoop to the K562 human myelogenous leukemia cell line and investigated different levels of TR cooperation in 3D chromatin using Hi-C loops (Figure S1A). Briefly, we collected and uniformly processed 851 high-quality K562 ChIP-seq profiles of 386 TRs from ENCODE (Consortium et al., 2020a), Cistrome DB (Zheng et al., 2019), and ChIP-Atlas (Oki et al., 2018) (Figure S1B). These chromatin-associated TRs, with 18,510 peaks on average (Figure S2A), were classified into seven categories: TF, transcription cofactor, RNA-binding protein (RBP), chromatin remodeler, nuclear enzyme, polycomb group (PcG) protein, and other factors (Table S1). For 3D genome data, we used 10-kb chromatin loops called from K562 *in situ* Hi-C data using the Peakachu algorithm, yielding 16,629 significant interactions (Figure S2B). Peakachu has been used for accurately estimating chromatin loops even from low sequencing depth of a Hi-C library, which can facilitate the broader application of 3DCoop in widespread tissue/cell types (Salameh et al., 2020). Then, TR-specific contact maps were constructed by combining ChIP-seq peaks and Hi-C

significant interactions (Figure S2C). Based on these contact maps, the TR pair-wise correlation matrix was generated using the generalized Jaccard similarity. Finally, the precision matrix was estimated from the TR pair-wise correlation matrix by the copula nonparanormal graphical model (Liu et al., 2012) and then the TR cooperation was inferred based on the precision matrix by the ClusterONE overlapping clustering method (Nepusz et al., 2012) (see STAR Methods for details).

Benefiting from the available ChIA-PET (chromatin interaction analysis with paired-end tag sequencing) datasets of K562 TRs, we were able to evaluate the consistency between TR-specific contact maps constructed by 3DCoop and ChIA-PET interactions. Of 24,887 significant interactions detected by CTCF ChIA-PET, 17.33% of interactions were recovered by CTCF-specific contact maps (Figure S3A; $p < 0.0026$, permutation test, Figure S3B). Besides, RAD21/SMC2/ZNF143-specific contact maps recovered a comparable proportion (17.23%–17.27%) of interactions, which was nearly 2-fold of the recovery rate (7.38%–10.11%) by EP300/TBP/POLR2A-specific contact maps (Figure S3A). A similar trend was observed when using RAD21 ChIA-PET peaks (23.03%–23.33% versus 9.31%–13.10%, Figure S3C; $p < 0.0026$, permutation test, Figure S3D). Although the ChIA-PET datasets (low power in earlier experiment) are not golden standard to evaluate our TR-specific contact maps (de Wit and de Laat, 2012; Denker and de Laat, 2016), the comparisons between the 3DCoop TR-specific contact maps and ChIA-PET signals of functionally relevant proteins showed significant and reasonable results, which indicated the biological meaning of TR-specific contact maps and the effectiveness of the combination between TR binding events and chromatin loops for detecting TR cooperation.

Using TR-specific contact maps for each ChIPped factor, we investigated the activity of single-TR intervening chromatin interactions by combining the genome-wide effects of 3D binding proportion and 3D interaction proportion (Table S2, see STAR Methods for details). As expected, known cohesin subunits or architecture proteins, such as RAD21, SMC3, ZNF143, CTCF, and YY1, showed remarkable activity in 3D chromatin. In addition, MAZ, REST, ARID2, RCOR1, JUND, MYC, CBX3, TEAD4, BHLHE40, MAX, JUN, CTCFL, CC2D1A, IRF1, and ARID3A also displayed effects in the 3D genome, implying their potential function in regulating genome organization (Figure 1A). MAZ was recently reported as a novel factor with insulation property and contribution to genomic architectural organization (Ortabozkoyun-Kara et al., 2020). BHLHE40 has a new role in modulating CTCF loop formation (Hu et al., 2020). CTCFL has a new role in promoting regulatory chromatin interactions (Debruyne et al., 2019). These new findings largely support the effectiveness of 3DCoop procedures for identifying TR cooperation in 3D chromatin.

Assessment of predicted TR cooperation in 3D chromatin

Globally, 64 TR clusters, 222 TR maximum cliques, and 1,608 TR pairs were identified using the 3DCoop pipeline in K562 cells (Tables S3, S4, and S5). Interestingly, we observed potential interplay among many chromatin-associated RBPs such as HNRNPK-PCBP1 and PRPF4-U2AF1, which previous studies have documented (Xiao et al., 2019) and which direct PPI evidence supports (Figure 1B). By systematically collecting the predicted TR cooperation records from existing protein complexes detection methods, including all existing 3D methods (DBPnet [Zhang et al., 2016], 3CPET [Djekidel et al., 2015], and HidPET [Wang et al., 2019]) and two representative 1D methods (NMF [Giannopoulou and Elemento, 2013] and TICA [Perna et al., 2018]) (Table S6), we evaluated the recall of TR cooperation identified by 3DCoop in K562. Given TR pairs detected by the 3DCoop pipeline and collected TR pairs from other methods in K562, permutation analysis was performed to inspect the co-occurrence rate between 3DCoop and each of the selected methods. Taking the DBPnet method as an example, there were 29 TRs involved, wherein 23 TRs were shared by 3DCoop and DBPnet, incorporating 38 and 25 TR pairs for 3DCoop and DBPnet, respectively. Of 38 TR pairs, 10 (26.3%) were successfully detected by DBPnet, which was significantly higher than that detected by random sampling datasets ($p = 0.002$, permutation test). Overall, even in the absence of a gold standard, 3DCoop received significantly more shared TR pairs than expected by chance given the reported TR cooperation from most of the selected methods (Figure S4), suggesting its good performance in detecting the likely true TR cooperation.

3DCoop had several intrinsic properties from incorporating 3D interactome data. Besides, it can also be used to detect 1D TR cooperation without 3D information. The comparison between the 3D and 1D TR cooperation showed potential benefits by incorporating 3D interactome data. Briefly, 1D strategy resulted in 70 TR clusters, 252 TR maximum cliques, and 1,707 TR pairs in K562 (Figure S5A, Tables S7, S8, and S9). TR cooperation identified using 3D strategy largely overlapped with those detected by the 1D strategy, such

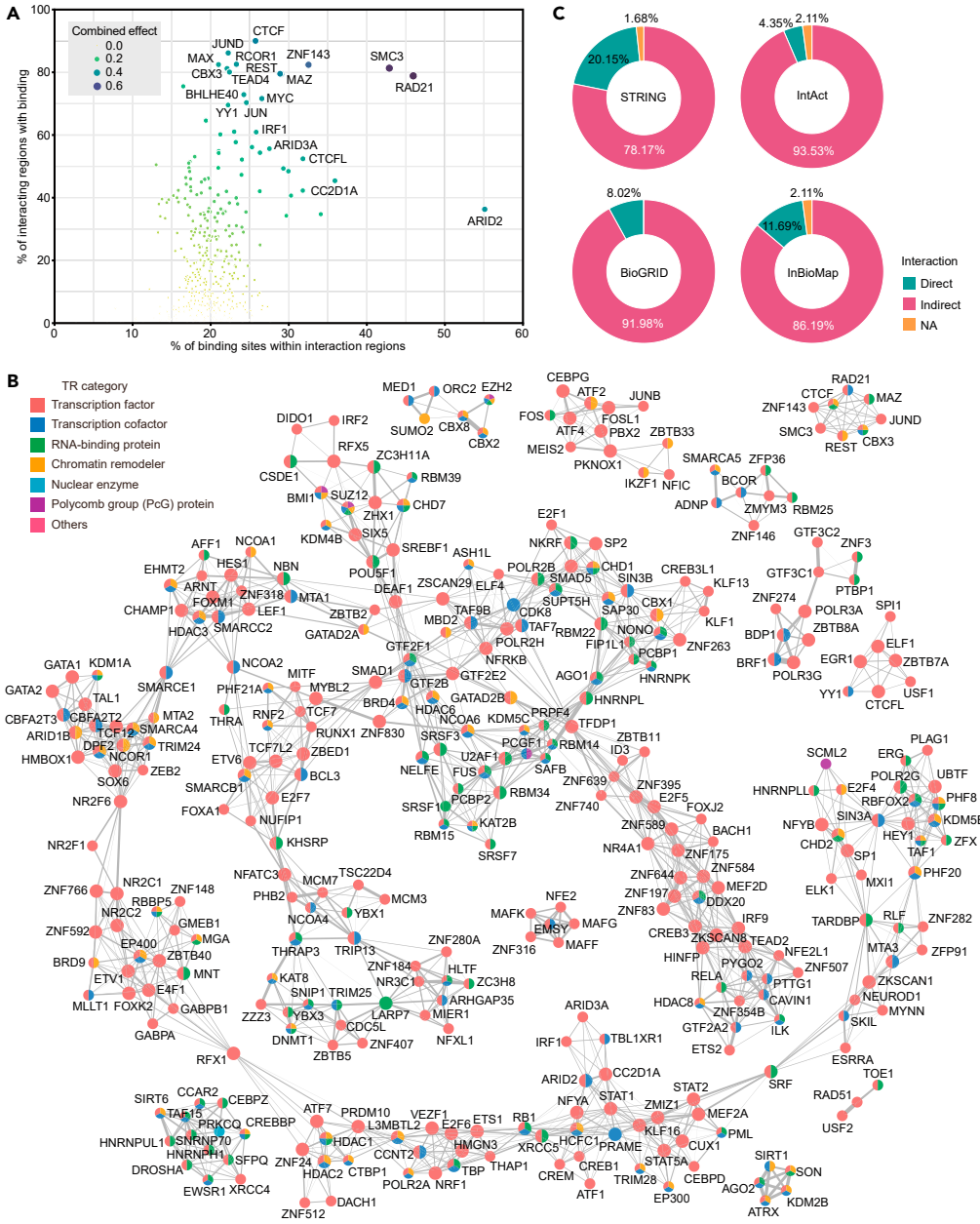


Figure 1. Evaluation of K562 TR cooperation identified by 3DCoop

(A) The activity of single-TR intervening chromatin interactions. The top20 TRs with the highest estimated effect were labeled.

(B) K562 TR cooperation network identified using 3DCoop pipeline. The categories for each TR were shown using pie plot. The size of TR node was weighted by its degree. Edges between TRs were weighted by their corresponding Glasso score.

(C) PPI evidence of detected TR cooperation from STRING, IntAct, BioGRID, and InBioMap databases.

See also Figures S1–S7.

as CTCF/cohesin complex CTCF-RAD21-SMC3, and hematopoietic factors GATA1-TAL1, which is consistent with previous findings (Zhang et al., 2016) (Figures 1B, S5A, and S5B). Compared with the 1D strategy, the 3D mode can detect many 3D-specific TR cooperation, which shows higher-order collaboration beyond linear DNA. For example, TR pairs of CTCF-MAZ and RAD21-MAZ only appeared in the results of 3D strategy, suggesting that MAZ might contribute to 3D chromatin regulation along with CTCF/cohesin complex. Consistently, a recent CRISPR screen identified that MAZ functions as a CTCF insulator co-factor and

RAD21 regulatory partner in Hox gene clusters (Ortabozkoyun-Kara et al., 2020). Besides, we found that, in general, 3D TR cooperation contains more indirect (1–2 intermediate proteins) TR-TR interactions than 1D TR cooperation (chi-squared test, Figure S5C), suggesting the 3D strategy is inclined to capture distal cooperation among TRs. Taken together, these evidences support that incorporation of 3D interactome data will benefit the identification of TR cooperation in high-dimensional chromatin environment.

To evaluate the performance of 3D TR cooperation detection in K562 cells, we first investigated the TR pairs with PPI evidence from different databases. Based on the PPI from the STRING database, 20.15% of TR pairs showed direct physical interaction, whereas 78.17% had indirect physical interaction, implying that they could be connected with other intermediate factors (Figure 1C). We obtained similar results using PPI data from the IntAct, BioGRID, and InBioMap databases, in which 4.35%–11.69% and 86.19%–93.53% of TR pairs had direct and indirect physical interactions, respectively (Figure 1C). These results suggest that other proteins or other collaborative mechanisms might mediate most of the detected TR cooperation. To build a control baseline for PPI evidence, we constructed all possible TR pairs using all TRs in the K562 cooperation network. We found that TR cooperation identified by 3DCoop received significantly more direct interactions in all four PPI databases ($p < 0.05$, chi-squared test, Figure S6A), suggesting the effectiveness of 3DCoop results. Here, TR pairs that had no physical interactions but could be connected by intermediate proteins were defined as indirect interactions. There are different degrees of indirect interactions, but the majority (>96%) indirect TR interactions were connected by one or two intermediate TRs (Figure S6B). Besides, we inspected the 3DCoop TR pairs using mass spectrometry results from *in situ* capture of chromatin interactions by biotinylated dCas9 (CAPTURE) at a well-characterized β-globin locus in K562 cells (Liu et al., 2017). For 42 involved TRs between 3DCoop and CAPTURE, 53.45% of the identified TR pairs were supported by the proteomic data of at least one captured locus. Moreover, nearly half of the identified TR pairs (44.83%) showed physical interaction in the PPI databases. The TR pairs supported by the proteomic data and PPI databases included GATA1-TAL1 and several SWI/SNF subunits (e.g., ARID1B, SMARCA4, SMARCC2) (Figure S7; Table S10). Taken together, these global and local evaluations demonstrate that 3DCoop has good detection power for capturing true TR cooperation in 3D chromatin.

TR cooperation underpins cell type-specific cellular function

Modularity analysis of the TR cooperation network can facilitate the interpretation of a cooperative relationship of multiple TRs for sustaining specific cellular functions. One obvious merit of the 3DCoop pipeline is that it allows overlapped TRs in different network modules. In addition to simple TR pairs, 3DCoop ends with two extra layers of TR cooperation, including overlapping TR clusters and TR maximum cliques. Therefore, a single TR could be attached to different clusters and maximum cliques, which is closer to the real situation of particular TR cooperation in cells. A clique that includes the largest possible number of vertices and provides a dense core is a maximum clique. It can be extended to produce plausible biological networks (Eblen et al., 2009). As local modules in the entire TR cooperation network, maximum cliques could be the best representatives for describing the biological function of TR cooperation. For each TR maximum clique identified by 3DCoop in K562 cells, the chromatin loop-associated genomic regions that encompass all TRs in the clique were extracted and analyzed with the GREAT function enrichment algorithm (McLean et al., 2010) (see STAR Methods for details). The KEGG pathways analysis can help us to interpret the cellular functions of identified TR cooperation from 3DCoop in K562 cells. By limiting a minimal 100 genomic regions, 149 of 222 (67%) TR maximum cliques were analyzed, yielding 186 significant KEGG pathways associated with clique-level TR cooperation. Among them, 74 TR maximum cliques were enriched in more than 10% significant pathways and 70 pathways received more than 10% TR maximum cliques (Figure S8, Table S11).

Based on hierarchical clustering, 70 pathways were classified into two major functional categories: leukemia/immune system-related pathways and non-immunological pathways (Figures 2A and S8). More than 40% of maximum cliques in the K562 cells were significantly enriched in leukemia-specific or immune-related pathways, such as acute myeloid leukemia (AML), chronic myeloid leukemia (CML), and B cell receptor signaling pathway, which indicates that a large number of cooperating TRs in K562 cells can sustain context-relevant biological functions. On the other hand, 74 maximum cliques could be roughly grouped into two clusters (Figures 2A and S8). TR cooperation in the first cluster, which was involved in most of the identified pathways, incorporated many architectural proteins, such as CTCF, cohesin, ZNF143, and YY1.

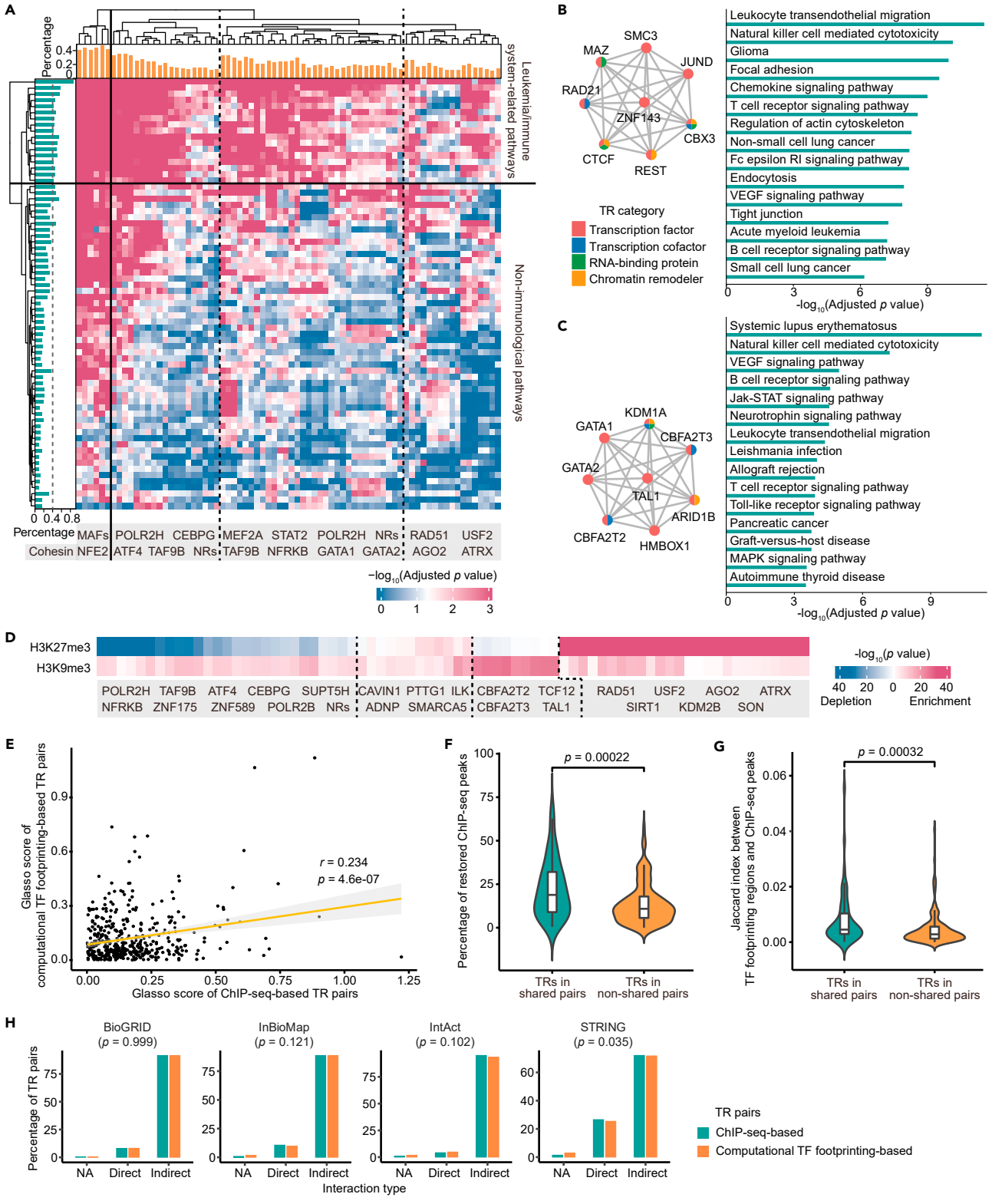

A
Percentage
0.4
0.2
0
Leukemia/immune system-related pathways
Non-immunological pathways
0 0.4 0.8
Percentage
MAFs POLR2H CEBPG MEF2A STAT2 POLR2H NRs RAD51 USF2
Cohesin NFE2 ATF4 TAF9B NRs TAF9B NFRKB GATA1 GATA2 AGO2 ATRX
−log10(Adjusted p value)
0 1 2 3
B
SMC3
MAZ
JUND
RAD21
ZNF143
CBX3
CTCF
REST
TR category
Transcription factor
Transcription cofactor
RNA-binding protein
Chromatin remodeler
Leukocyte transendothelial migration
Natural killer cell mediated cytotoxicity
Glioma
Focal adhesion
Chemokine signaling pathway
T cell receptor signaling pathway
Regulation of actin cytoskeleton
Non-small cell lung cancer
Fc epsilon RI signaling pathway
Endocytosis
VEGF signaling pathway
Tight junction
Acute myeloid leukemia
B cell receptor signaling pathway
Small cell lung cancer
0 3 6 9
-log10(Adjusted p value)
C
KDM1A
GATA1
CBFA2T3
GATA2
TAL1
ARID1B
CBFA2T2
HMBOX1
Systemic lupus erythematosus
Natural killer cell mediated cytotoxicity
VEGF signaling pathway
B cell receptor signaling pathway
Jak-STAT signaling pathway
Neurotrophin signaling pathway
Leukocyte transendothelial migration
Leishmania infection
Allograft rejection
T cell receptor signaling pathway
Toll-like receptor signaling pathway
Pancreatic cancer
Graft-versus-host disease
MAPK signaling pathway
Autoimmune thyroid disease
0 3 6 9
-log10(Adjusted p value)
D
H3K27me3
H3K9me3
POLR2H TAF9B ATF4 CEBPG SUPT5H CAVIN1 PTTG1 ILK CBFA2T2 TCF12 RAD51 USF2 AGO2 ATRX
NFRKB ZNF175 ZNF589 POLR2B NRs ADNP SMARCA5 CBFA2T3 TAL1 SIRT1 KDM2B SON
-log10(p value)
40 20 0 20 40
Depletion Enrichment
E
Glasso score of computational TF footprinting-based TR pairs
0.0 0.3 0.6 0.9
r = 0.234
p = 4.6e-07
0.00 0.25 0.50 0.75 1.00 1.25
Glasso score of ChIP-seq-based TR pairs
F
Percentage of restored ChIP-seq peaks
0 25 50 75 100
p = 0.00022
TRs in shared pairs
TRs in non-shared pairs
G
Jaccard index between TF footprinting regions and ChIP-seq peaks
0.00 0.02 0.04 0.06
p = 0.00032
TRs in shared pairs
TRs in non-shared pairs
H
BioGRID (p = 0.999)
InBioMap (p = 0.121)
IntAct (p = 0.102)
STRING (p = 0.035)
Percentage of TR pairs
NA Direct Indirect
Interaction type
TR pairs
ChIP-seq-based
Computational TF footprinting-based

Figure 2. Functional analysis of K562 TR maximum cliques from the cooperation network and consistency analysis between ChIP-seq-based results and computational TF footprinting-based results in K562

(A) Association matrix between TR maximum cliques and biological pathways. TR maximum cliques that obtained more than 10% significant pathways and KEGG pathways that were significantly enriched among more than 10% TR maximum cliques were used to plot the profile. Some representative TRs with multiple cooperation evidence were labeled. The percentage of significantly enriched pathways for each maximum clique and the percentage of maximum cliques with significant enrichment for each pathway were labeled on the top and left, respectively. Red for significant enrichment; blue for non-significant enrichment.

(B) The CTCF-cohesin-associated TR maximum clique and its top15 significantly enriched pathways. The categories for each TR were shown using pie plot. Pathways were ordered by adjusted *p* value for enrichment.

(C) The GATA1-TAL1-associated TR maximum clique and its top15 significantly enriched KEGG pathways.

(D) The association between TR maximum cliques and two repressive chromatin marks, H3K27me3 and H3K9me3. Colors in the heatmap represent the *p* value reported by GREAT for the colocalization, red for enrichment and blue for depletion.

(E) The correlation of Glasso scores between ChIP-seq-based TR pairs and computational TF footprinting-based TR pairs (Pearson correlation).

(F) TRs in shared pairs restore significantly more ChIP-seq peaks than TRs in non-shared pairs (two-tailed Mann-Whitney U test).

(G) The footprinting regions of TRs in shared pairs overlap significantly more ChIP-seq peaks than TRs in non-shared pairs (two-tailed Mann-Whitney U test).

(H) The distribution comparison of ChIP-seq-based TR cooperation and computational TF footprinting-based TR cooperation across PPI evidence from BioGRID, InBioMap, IntAct, and STRING databases (chi-squared test).

See also Figures S8–S10, and S13.

For example, the top enriched pathways of the maximum clique CBX3-CTCF-JUND-MAZ-RAD21-REST-SMC3-ZNF143 contained several hematopoietic cancer-related pathways (Figure 2B), suggesting that such TR cooperation may operate cell type-specific gene regulation in leukemia by modulating 3D chromatin organization (Hou et al., 2010). The remaining maximum cliques were grouped together and constituted the second main cluster. The cooperation of these TRs was inclined to separately control more specialized cellular functions in K562 cells. For example, the GATA1-GATA2-TAL1-associated and POLR2A-TBP-associated maximum cliques were highly scored in the network and connected to many pathways that determine immune and hematopoietic functions in particular (Figures 2C and S9). Together, the clique-pathway interrelated profile indicates that the TR maximum cliques detected by 3DCoop can recapitulate the context-dependent biological functions of transcriptional machinery in K562 cells.

To investigate the association between TR cooperation and chromatin states, we collected 13 epigenome profiles of K562 cells from the Roadmap Epigenomics Project (Roadmap Epigenomics et al., 2015). We tested the significance of colocalization between maximum clique-associated genomic regions and peak regions of specific epigenetic marks using GIGGLE enrichment analysis (Layer et al., 2018) (see STAR Methods for details). We found that the majority of TR cooperation significantly colocalized with active chromatin marks, such as DNase I hypersensitive site (DHS), H2A.Z, H3K27ac, and H3K4me1 (Figure S10, Table S12). The colocalization of active chromatin marks probably is not that surprising given that the maps are based on ChIPped factors and mostly function in open chromatin. However, the association between TR cooperation and repressive chromatin marks showed distinct patterns on different maximum cliques (Figure 2D). For example, several maximum cliques associated with DNA repair proteins (e.g., RAD51), chromatin remodelers (e.g., ATRX, SIRT1, and KDM2B), and RBPs (e.g., AGO2 and SON) were enriched in H3K27me3-marked regions (Figure S10). Interestingly, some maximum cliques associated with leukemia-specific factors (e.g., CBFA2T2, CBFA2T3, and TAL1) were enriched in H3K9me3-marked regions, suggesting that the cooperation of these TRs could also function as important complexes at repressive or bivalent chromatin. The colocalization analysis with chromatin marks revealed that the communications between TR cooperation and certain chromatin environments shape cell type-specific gene regulation.

Construction of the TR cooperation landscape across 40 human tissue/cell types

Previous attempts at detecting 3D TR cooperation were highly reliant on the availability of ChIPped factors and high-resolution chromatin interactions, which significantly limited the broader application to widespread tissue/cell types. Recent computational TF footprinting methods leveraged DNase-seq/ATAC-seq profiles and TF motifs to infer TF-binding sites and achieved high recovery rates of ChIP-seq peaks (Gusmao et al., 2016). To extend the identification of 3D TR cooperation in different conditions, we first systematically integrated 3,105 motifs of 1,480 human TRs by collecting and uniformly processing 16 existing TF motif resources, such as CIS-BP (Weirauch et al., 2014), JASPAR (Fornes et al., 2020), HOCOMOCO (Kulakovskiy et al., 2018), and others (Figure S11A, see STAR Methods for details). TFs and their co-factors received most of the motifs, yet RBPs, chromatin remodelers, nuclear enzymes, and PcG proteins were also included (Figures S11B and S11C), while 67.77% (1,003/1,480) TRs have more than one motif,

25.74% (381/1,480) contain two motifs and 42.03% (622/1,480) contain three motifs (Table S13). We then used PIQ computational TF footprinting based on these TR motifs, TR gene expression, and DNase-seq/ATAC-seq profiles on specific tissue/cell types to estimate genome-wide TR-binding events (Sherwood et al., 2014) (see STAR Methods for details). We also used the predicted 10-kb chromatin loops by Peakachu when tissue/cell type-specific Hi-C, open chromatin, and gene expression data were available (Figure S11D). Finally, the 3DCoop pipeline was used for identifying TR cooperation across different tissue/cell types by integrating TR motifs, open chromatin, gene expression, and chromatin loops (Figure S12, see STAR Methods for details).

We first applied this new strategy to the same 247 TRs using open chromatin, gene expression, and predicted chromatin loop data in K562 cells and compared the TR cooperation with the ChIP-seq-based results. Using computational TF footprinting, 77.5% of TR clusters and 85% of TR maximum cliques shared at least one TR pair with the ChIP-seq results (Figures S13A and S13B). For the TR pairs, we compared their Glasso scores between ChIP-seq-based results and computational TF footprinting-based results. The correlation coefficient (r = 0.234, Pearson correlation, Figure 2E) is significantly higher than those of random sampled ones (p < 1e-04, permutation test, Figure S13C). The comparison between ChIP-seq peaks and inferred TF footprinting regions showed that the TRs in shared pairs obtained more restored peaks and overlapped regions than the TRs in non-shared pairs (Figures 2F and 2G). Analysis of the results using PPI data also showed that most of the TR pairs predicted by computational TF footprinting could be validated and there was no significant difference compared with results of TR pairs predicted by ChIP-seq data (chi-squared test, Figures 2H, S13D, and S13E). These comparisons indicated the general concordance between computational TF footprinting-based results and ChIP-seq-based results.

Given the feasibility of the extended 3DCoop pipeline, we performed large-scale 3D TR cooperation analysis of 40 human tissue/cell types with condition-matched open chromatin, gene expression, and Hi-C data (Table S14). Among these tissue/cell types, 192–922 TRs (mean: 655, median: 717) were investigated according to their gene expression levels, and 18–86 clusters (mean, 48; median, 49), 193–55,626 maximum cliques (mean, 4,735; median, 3,246), and 2,732–12,188 TR pairs (mean, 8,287; median, 8,338) were detected (Figure S14, Table S15). Pairwise correlation and hierarchical clustering for TR pairs identified in more than 20% of the tissue/cell types briefly recapitulated the cell type origin (Figure S15A). For example, H1-derived pluripotent cells were mostly clustered in a separate group. Cells from the same lineage were also grouped together, such as A549, NHEK, HMEC, and HUVEC epithelial and endothelial-derived cells, or K562 and GM12878 blood-derived cells. HAP1 cells, which are near haploid and have one copy of almost every chromosome, showed the biggest difference from the others. However, tissue and cell type from the same organ generally diverged from each other, and the coherent pattern vanished when using tissue- or organ system-level clustering (Figures S15B and S15C), suggesting that TR cooperation could be fine-scale and cell type-specific.

Based on the shared pattern of TR cooperation in 3D chromatin across 40 human tissue/cell types, we classified all detected TR pairs into four types: (1) 0.12% (n = 148) tissue/cell type-shared TR pairs in more than 75% of tissue/cell types; (2) 4.69% (n = 5,828) tissue/cell type-relatively shared TR pairs in more than 25% but less than 75% of tissue/cell types; (3) 39.51% (n = 49,120) tissue/cell type-relatively specific TR pairs in at least two but less than 25% of tissue/cell types; (4) 55.68% (n = 69,216) tissue/cell type-specific TR pairs only in one tissue/cell type (Figure 3A, Table S16, see STAR Methods for details). Using gene expression quantification from all samples in the Genotype-Tissue Expression (GTEx) Project, we found that the expression correlation of the TR pairs showed a significant difference among the four TR cooperation categories (p < 2.22e-16, two-tailed Mann-Whitney U test, Figure 3B). As expected, the expression correlation of the TR pairs gradually decreased from the tissue/cell type-shared group to the tissue/cell type-specific group. Using the tissue-specific functional interactions from 145 tissues in HumanBase (Wong et al., 2018), we found that the number of tissues associated with the identified TR pairs showed a significant difference among the four TR cooperation categories (p < 0.05, two-tailed Mann-Whitney U test, Figure 3C). The tissue/cell type-specific TR pairs exist in least tissues (median = 28, mean = 36), whereas tissue/cell type-shared TR pairs exist in most tissues (median = 54, mean = 64). These results emphasized the distinct pattern of tissue/cell type commonality and specificity among the cooperating TRs.

Tissue/cell type commonality and specificity of TR cooperation

With the 148 tissue/cell type-shared TR pairs (Table S16), we constructed a tissue/cell type-shared TR cooperation network (Figure 3D). Some well-characterized TR cooperation clusters were identified in this

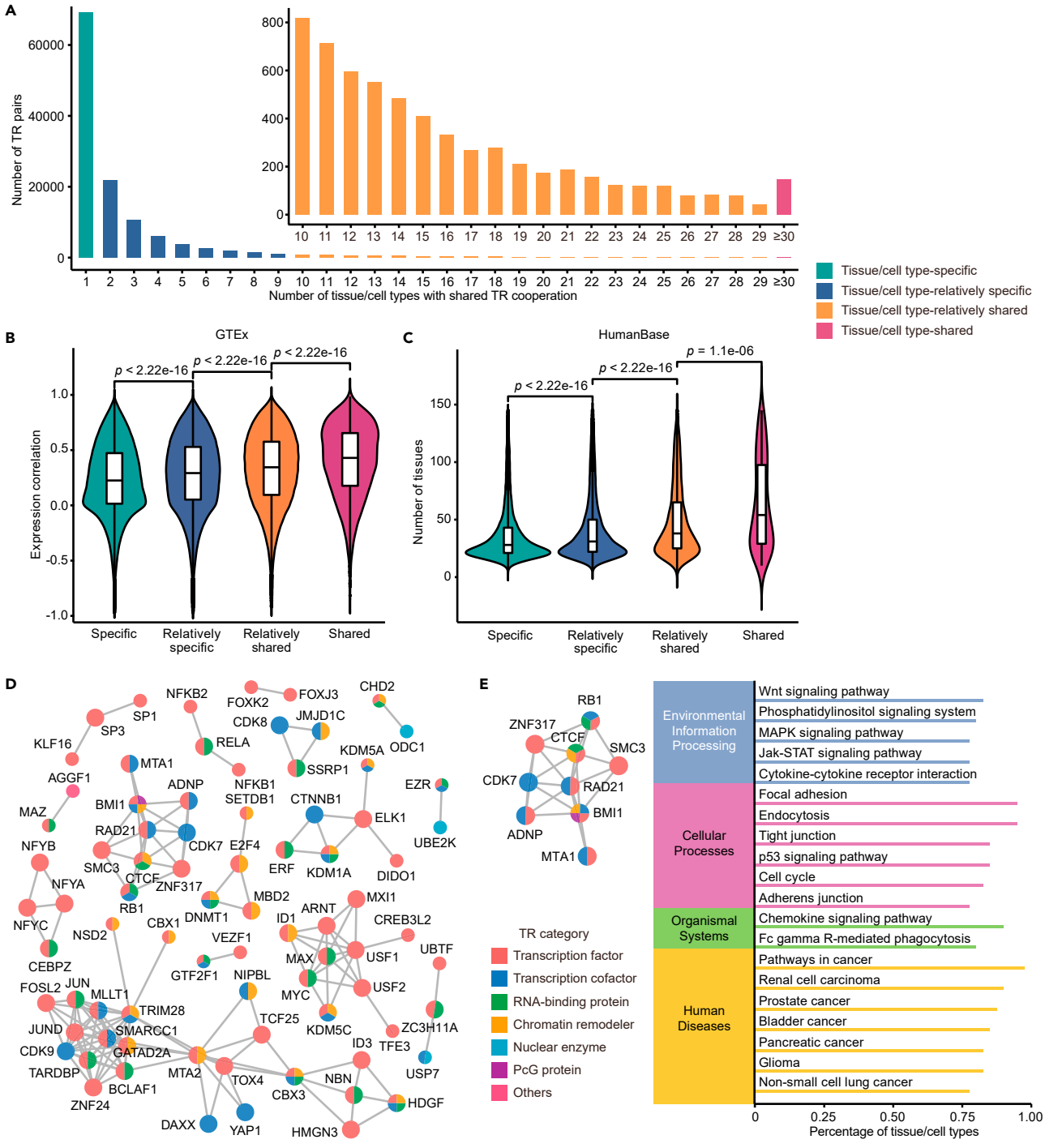


Figure 3. TR pairs classification and the shared TR cooperation network across 40 human tissue/cell types

(A) Summary information of TR pairs shared by different tissue/cell types. According to the shared pattern of TR cooperation in 3D chromatin across 40 human tissue/cell types, all detected TR pairs were classified into four categories, tissue/cell type-specific, tissue/cell type-relatively specific, tissue/cell type-relatively shared, and tissue/cell type shared.

(B) Comparison of expression correlation for TR pairs in four categories using GTEx samples. The two-tailed Mann-Whitney U test was used to test the significance.

(C) Comparison of tissue distribution for TR pairs in four categories using tissue-specific functional interactions across 145 tissues in HumanBase. The two-tailed Mann-Whitney U test was used to test the significance.

Figure 3. *Continued*

(D) Tissue/cell type-shared TR cooperation network. The categories for each TR were shown using pie plot. TR nodes with bigger size indicate that the degree of nodes was greater or equal to 2.

(E) CTCF-RAD21-SMC3 cluster and its associated KEGG pathways. The categories for each TR were shown using pie plot. KEGG pathways were ordered by the proportion of tissue/cell types having significant enrichment. KEGG pathways that were significantly enriched in more than 75% of tissue/cell types were shown. See also Figures S11, S12, and S14–S17.

network. For example, NF-kappa B is a pleiotropic TF present in almost all cell types and is composed of NFKB1 or NFKB2 bound to REL, RELA, or RELB. The heterodimeric RELA-NFKB1 complex appears to be the most abundant in gene regulation (Oeckinghaus and Ghosh, 2009), which is consistent with a NFKB1-RELA-NFKB2 cluster. CTCF-RAD21-SMC3 was associated with both active transcription machinery CDK7 and PcG/repressive factor BMI1 in the shared network. This cluster was enriched in many basic life activity-related pathways and cancer-related pathways (Figures 3E and S16A), implying that such TR cooperation might be important in sustaining the basal functions of cell survival. Interestingly, DNMT1-MBD2-SETDB1-E2F4 was identified as a cluster, and their potential genomic targets were related to many important signaling pathways such as the Notch, Wnt, and p53 signaling pathways (Figure S16B), which reveals a shared TR partner recruitment mechanism of DNMTs for DNA methylation across tissue/cell types (Hervouet et al., 2018). Similarly, other identified TR clusters, such as MAX-MYC-USF1 and CTNNB1-KDM1A-ELK1, have been documented and play important roles in various tissue/cell types. This tissue/cell type-shared TR cooperation network could provide a global reference for studying common transcriptional partners. Besides, based on the five immune cells (HAP1, GM12878, K562, KBM7, and THP-1), we built a blood-specific TR cooperation network (Figure S17A) and annotated the biological functions of clusters using GREAT. More than half of enriched pathways (54%, 47/87) connected to immune-related functions (see the mini-map in Figure S17B). When we zoomed in, we could find blood-related diseases (CML and AML) and many immune-related pathways, such as T cell receptor signaling pathway, B cell receptor signaling pathway, and Toll-like receptor signaling pathway (see the bold items in the main part in Figure S17B). The results also strengthened the association between the tissue/cell type-specific pathways and 3DCoop-detected TR cooperation in function-relevant tissue/cell types.

For the majority of TR cooperation detected in specific tissue/cell types (mean, 1,730; median, 1,647 for 40 tissue/cell types) (Figure S18A), we selected single tissue-specific TR cooperation from eight GTEx-matched tissues and evaluated their gene expression correlation. We found that TR pairs showed significantly higher expression correlation in matched tissue than in other tissues ($p < 2.22e\text{-}16$, two-tailed Mann-Whitney U test, Figure 4A), implying their tissue specificity at the transcription level. For 11 tissues shared by 3DCoop and HumanBase, we calculated the correlation between Glasso score from 3DCoop and posterior probability from HumanBase. TR pairs showed higher score correlation in matched tissue than in other tissues, even though it was not statistically significant ($p = 0.094$, one-tailed Mann-Whitney U test, Figure S18B). Tissue/cell type pair-wise comparisons showed that aorta-spleen ($n = 300$), H1-hESC-H1-NPC ($n = 230$), and K562-GM12878 ($n = 217$) had the most shared TR pairs (Figure S18C). Among the immune-related cell types, three myelogenous cell lines, i.e., K562, KBM7, and THP-1, were clustered together as compared with a lymphoblastoid cell line, i.e., GM12878. Consistent with that result, the TR cooperation pattern of HAP1 was relatively unique (Figure S18D). Moreover, the tissue/cell type-specific landscape of TR cooperation facilitated the exploration of their novel and context-dependent biological functions. For example, the basal expression of interferon (IFN)-induced genes (ISGs) is controlled through STAT2-IRF9 complexes, whose formation does not require IFN-I receptor signaling (Kessler et al., 1990). In the present study, we detected this tissue/cell type-relatively shared TR cooperation in 19 tissue/cell types. IFN can stimulate resting-state macrophages by switching from STAT2-IRF9 complexes to the complete ISGF3 complex containing STAT1, STAT2, and IRF9 (Au-Yeung et al., 2013; Platanitis et al., 2019). Consistently, STAT1-IRF9 and STAT1-STAT2 are two tissue/cell type-specific TR pairs that were detected in only THP-1 cells, a human monocytic cell line. The widespread presence of STAT2-IRF9 in multiple tissue/cell types and the unique role of STAT1-STAT2-IRF9 TR cooperation in stimulated monocytes demonstrated the tissue/cell type specificity of TR cooperation.

Studies on embryonic stem cells and other cell models have revealed that a small group of cell type-specific or lineage-specific TFs forms an interconnected autoregulatory loop to govern transcriptional programs in particular conditions (Boyer et al., 2005; Saint-Andre et al., 2016). Core transcription regulatory circuitry (CRC) constitutes several core TFs and their interconnected autoregulatory loop, both of which are critical for maintaining cell identity and cellular state. To investigate whether CRCs could support the tissue/cell

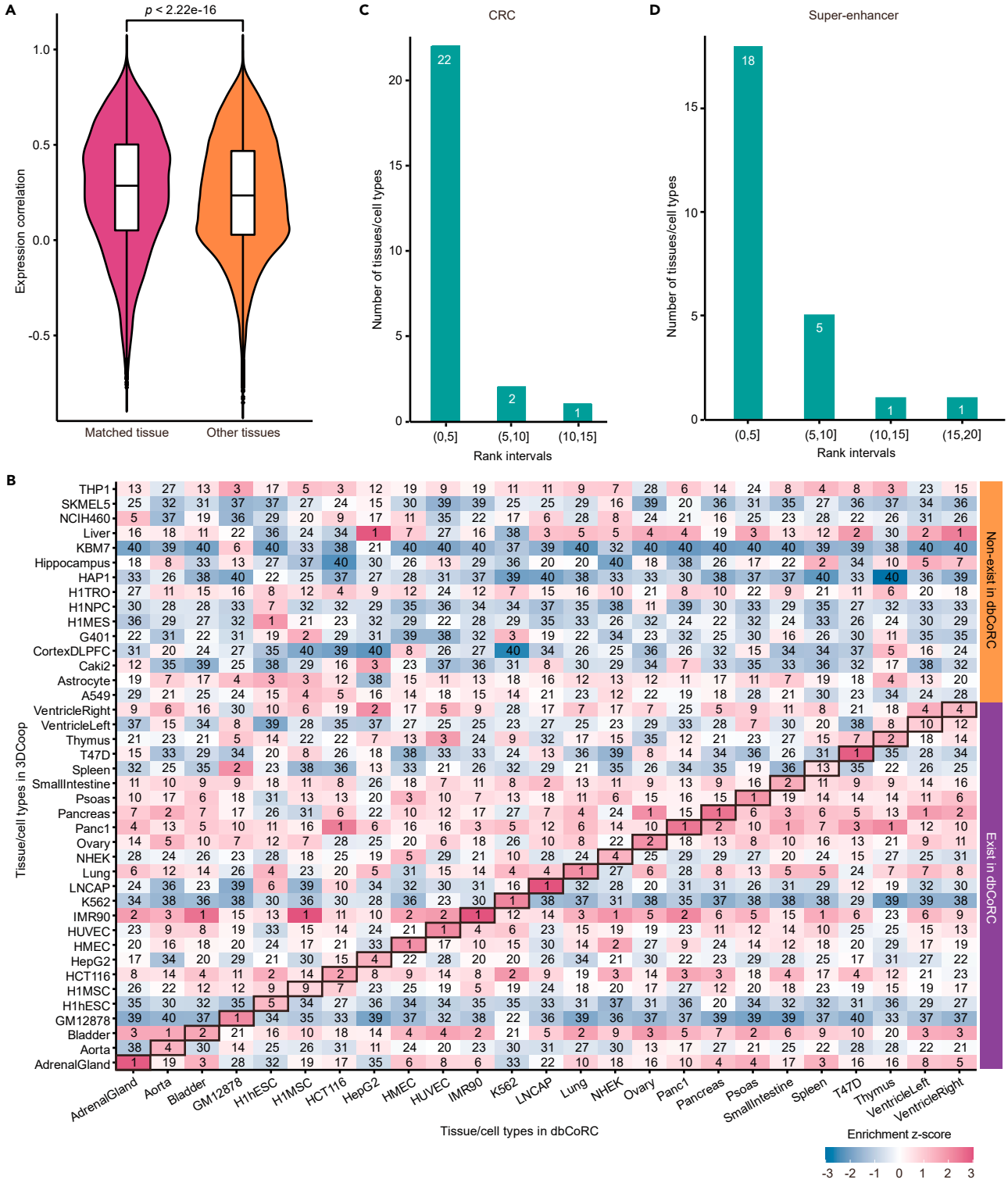


Figure 4. Tissue/cell type specificity of TR cooperation

(A) Comparison of expression correlation for tissue/cell type-specific TR pairs between matched tissues and other tissues. The two-tailed Mann-Whitney U test was used to test the significance.

Figure 4. *Continued*

(B) Evaluation of tissue/cell type specificity based on 3DCoop-detected TR pairs and CRC-derived TR pairs from dbCoRC. Colors in the heatmap represent the normalized enrichment z-score according to overlap ratio, red for enrichment. For each tissue/cell type in dbCoRC, tissue/cell types in 3DCoop were prioritized according to their overlap ratio, and the ranks were labeled.

(C) Distribution of prioritized ranks among 25 matchable tissue/cell types with CRC information in dbCoRC. The ranks were binned into three intervals, (0,5], (5,10], and (10,15].

(D) Distribution of prioritized ranks among 25 matchable tissue/cell types with super-enhancer regions in dbCoRC. The ranks were binned into four intervals, (0,5], (5,10], (10,15], and (15,20].

See also Figures S18 and S19.

type specificity of TR cooperation in tissue/cell type-matched conditions, for each tissue/cell type in 3DCoop, we compared the consistency between our detected TR pairs and CRC-derived TR pairs in dbCoRC (Huang et al., 2018b) and then prioritized tissue/cell types according to the overlap ratio (Figure 4B). Among 25 matched tissue/cell types with CRC information in dbCoRC, 22 (88%) ranked in the top5 and 24 (96%) ranked in the top10 (Figure 4C), and the majority of tissue/cell types (84%, 21/25) shows no significant difference between tissue/cell types not exist in dbCoRC and those unmatched in dbCoRC ($p > 0.05$, Mann-Whitney U test, Figure S19), suggesting good tissue/cell type agreement between TR cooperation and CRC. The comparison with tissue/cell type-specific TR pairs associated with super-enhancer regions from dbCoRC showed similar results (Figures 4D and S18E). Super enhancer and CRC-derived TR pairs in dbCoRC can help us to validate the cell type specificity of TR cooperation detected by 3DCoop in different biological angles. The consistency between TR pairs detected by 3DCoop and TR pairs extracted from dbCoRC among the different conditions greatly support the tissue/cell type specificity of our detected TR cooperation.

Information from tissue/cell type-specific TR cooperation facilitates the interpretation of disease-causal variants

The majority of disease risk loci identified by genome-wide association study (GWAS) is located in the non-coding genomic region; interpreting the biological mechanism underlying disease susceptibility continues to be a challenge. Conventionally, disease-causal regulatory variants are prioritized via statistical fine-mapping and functional annotations (Huang et al., 2018a, 2020), and the most plausible TFs associated with causal variants are predicted for functional follow-up. However, computational prediction of TFs altered by regulatory variants relies heavily on motif analysis but usually ignores the chromatin context and TF dependency at the variant locus. Although recent studies have systematically analyzed specific TFs binding to noncoding variants toward disease-associated risk loci using quantitative trait mapping or SNP-SELEX (single-nucleotide polymorphism evaluation by systematic evolution of ligands by exponential enrichment), they have been limited to a small fraction of TFs or disease-causal variants (Tehranchi et al., 2016; Yan et al., 2021). Here, we hypothesized that context-specific TFs with cooperation evidence are more likely linked to the true binding events at disease risk loci; therefore, the disease-causal variants should be enriched for the TR cooperation-associated loci identified in the present study rather than TR binding loci only estimated from motif scanning or computational TF footprinting.

To test this hypothesis, we first collected 7,747 candidate causal variants of 39 immune and non-immune diseases/traits derived from a GWAS fine-mapping study (Farh et al., 2015). By sampling matched control variants (Huang et al., 2021), we evaluated the genome-wide enrichment of the disease-causal variants on the putative binding sites of the top20 TRs (18 TRs were eventually used due to two TRs presenting cooperation evidence in <10 tissue/cell types, Table S14) with high activity intervening in chromatin interactions identified in this study across 40 human tissue/cell types. The putative TR-binding sites associated with disease-causal variants were predicted according to three strategies: motif scanning based on the TR position weight matrices (i.e., naïve strategy) and computational TF footprinting without (PIQ strategy) and with cooperation evidence (3DCoop) (see STAR Methods for details). As expected, the disease-causal variants were highly enriched in TR-binding sites when considering tissue/cell type context and cooperation evidence. Compared with the naïve and PIQ strategies, the 3DCoop pipeline yielded a larger proportion of significant TRs in more tissue/cell type-specific contexts (Figures 5A and 5B). Moreover, the p values of enrichment in the 3DCoop group were significantly smaller than those in the PIQ group ($p = 4.08e\text{-}12$, one-tailed paired t test, Figure 5C), suggesting that TR cooperation evidence could be used to prioritize candidate TFs associated with certain disease-causal regulatory variants.

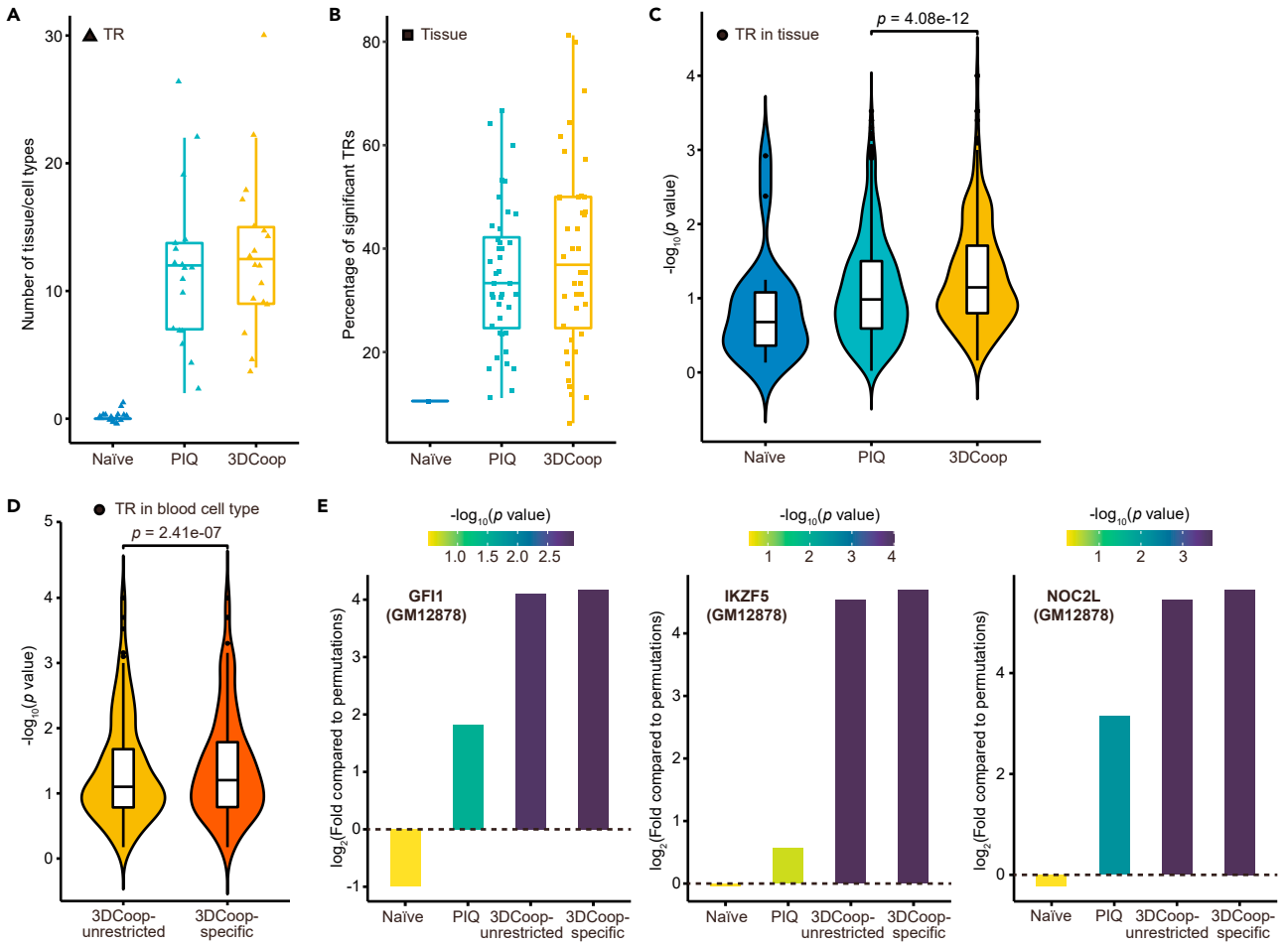


Figure 5. Interpreting disease-causal variants using tissue/cell type-specific TR cooperation

(A) Comparison of the number of tissue/cell types with significant disease-causal variants enrichment for each TR among naïve strategy, PIQ strategy, and 3DCoop.

(B) Comparison of the percentage of selected TRs with significant disease-causal variants enrichment for each tissue among naïve strategy, PIQ strategy, and 3DCoop. The naïve strategy has only one point because it cannot leverage context-dependent information.

(C) Comparison of logarithmic *p* value for disease-causal variants enrichment among naïve strategy, PIQ strategy, and 3DCoop. The one-tailed paired *t* test was used to test the significance between PIQ strategy and 3DCoop.

(D) Comparison of logarithmic *p* value for autoimmune disease-causal variants enrichment between 3DCoop-unrestricted group and 3DCoop-specific group. The one-tailed paired *t* test was used to test the significance.

(E) Comparison of significance for autoimmune disease-causal variants enrichment measured by permutation test on GM12878-specific TRs (GFI1, IKZF5, and NOC2L) among naïve strategy, PIQ strategy, 3DCoop-unrestricted group, and 3DCoop-specific group.

See also Figure S20.

As the majority of disease-causal regulatory variants show tissue/cell type specificity in phenotypically relevant contexts (Boix et al., 2021; Li et al., 2017a), we also investigated whether tissue/cell type-specific TR cooperation could facilitate the interpretation of TF binding specificity of disease-causal regulatory variants. By collecting 12,738 causal variants for 54 blood-related autoimmune diseases from CAUSALdb (Wang et al., 2020) and 102 blood trait/autoimmune disease-related TRs from the literature (Table S17), we tested the enrichment of autoimmune disease-causal variants on the putative binding sites of 102 TRs selected among five blood-derived cell types, namely, K562, GM12878, KBM7, HAP1, and THP-1 (see STAR Methods for details). Notably, the autoimmune disease-causal variants were more enriched in the binding sites with blood cell-specific cooperation evidence (3DCoop-specific) than the binding sites with unrestricted cooperation evidence (3DCoop-unrestricted) for each of the 102 selected TRs (p = 2.41e-07, one-tailed paired t test, Figure 5D). For example, GFI1, IKZF5, and NOC2L are highly expressed factors in GM12878 cells and are essential TRs in lymphopoiesis (Ma et al., 2014; Merkenschlager, 2010; van der Meer et al., 2010). We found that, compared with the randomly

sampled control variants, the autoimmune disease-causal variants exhibited similar binding enrichment patterns among these TRs, where the binding sites with blood cell-specific cooperation evidence received the best enrichment rather than those inferred from other strategies (Figure 5E). Taking the rs194747 associated with IBD (inflammatory bowel diseases) as an example, it affects GFI1 binding in GM12878, which involves the GFI1-CARM1 cooperation. This cooperation was supported by a previous investigation (Scheer and Zaph, 2017) and may play an important role in lymphoid cell differentiation and the development of autoimmune disease. Similar results were observed for several K562-specific TRs (Figure S20A). In addition, randomly selecting five tissue/cell types from other non-blood-derived tissue/cell types showed suboptimal results (Figure S20B). Together, these results indicate that context-dependent TR cooperation information can better interpret disease-causal variants identified by GWAS.

Characterization of lineage-specific TR cooperation during neural development

3D genome rewiring and the associated TR interplay facilitate the spatiotemporal control of gene expression during cell development. To demonstrate the ability of our 3DCoop for identifying changes in TR cooperation during sequential development stages, we constructed a TR interplay network and estimated potential TR cooperation in 3D chromatin across three cell states of mouse neural development (embryonic stem cell [ES], neural progenitor cell [NPC], cortical neuron [CN]) by systematically integrating ultra-high-resolution Hi-C interactions (Bonev et al., 2017), ATAC-seq, RNA-seq, and 1,636 motifs of 836 mouse TRs (Figure S21, Tables S18 and S19). Among 121 ES, 139 NPC, and 249 CN TRs filtered through their gene expression and available motifs, 68 TRs were shared among the three cell states (Figure 6A). 3DCoop identified 20/16/37 TR clusters, 30/20/68 maximum cliques, and 308/327/492 TR pairs for ES, NPC, and CN, respectively (Figure S22 and Tables S20–S28). Similar to the sharing pattern of the expressed TRs, ES and NPC had the most 72 shared TR pairs, more than the 46 TR pairs shared by NPC and CN, and the 28 TR pairs shared by ES and CN (Figure 6B). Network modularity analyses revealed shared TR cooperation networks during neural development, including a classical 3D TR cooperation cluster: Ctcf-Rad21-Smc3 (Figure 6C). We also constructed a TR cooperation network using TR pairs shared by any two cell states (Figure S23). Likewise, some TR cooperation in specific-state transition was clustered, such as Hcfc1-Yy1 in the ES-NPC stages. Hcfc1 loss-of-function mutations can disrupt neuronal and neural progenitor cells of the developing brain (Jolly et al., 2015). Hcfc1 collaborating with Yy1 (Michaud et al., 2013), a potential structuring factor (Weintraub et al., 2017), presents a candidate mechanism warranting further investigation during early neuronal development.

Some TR cooperation clusters were only identified in a particular cell state. Cooperation among the master regulators Nanog, Pou5f1, and Sox2 is essential for maintaining pluripotency (Chambers and Tomlinson, 2009), and this TR cooperation was only detected in ES. GREAT annotation analysis of the Nanog-Pou5f1-Sox2-associated genomic regions showed significant enrichment on circadian regulation of gene expression, circadian clock entrainment by photoperiod, and photoperiodism (Figure 6D). We also reveal a unique Klf-associated TR cooperation cluster, Chd1-Klf4-Klf5-Klf6-Mta3-Patz1-Pml-Sp3-Zfp281, in the 3D chromatin of ES, which provides evidence that Klf4 and its concrete partners are involved in organizing and regulating the pluripotency-associated 3D enhancer networks (Di Giammartino et al., 2019) (Figure 6E). In addition, we found that the cooperative members of Gabpa were dynamically changed during neural development. In CN, Gabpa was connected in particular to Elf2 and Elk1 instead of Etv proteins, suggesting a specific function of such TR cooperation in the developed neuron (Figure S22). A recent enhanced yeast one-hybrid assay has confirmed the Gabpa-Elf2-Elk1 physical interactions (Shrestha et al., 2019).

Notably, although Ctcf-Rad21-Smc3 cooperation was detected in all three cell types, their associated members varied dynamically during neuron differentiation (Figure 6F). GREAT annotation revealed a large difference in cellular functions driven by Ctcf-Rad21-Smc3-related TR cooperation in different stages (Figure 6G). Cell pluripotency-associated functions were generally enriched in ES. Biological functions such as anatomical structure regression and endothelial cell development were enriched in NPC. Meanwhile, more specialized functions, including cellular macromolecule localization, synaptic vesicle transport, and positive regulation of synapse maturation were enriched in CN. Together, these 3DCoop results are consistent with previous findings and imply that the cooperative partners for specific TRs may determine target gene recognition in a context-specific manner during cell development.

DISCUSSION

Combinatorial binding of chromatin-associated factors to CREs underpins dynamic gene regulation in the nucleus. Based on 386 ChIPped chromatin factors and high-resolution chromatin loops, we performed the

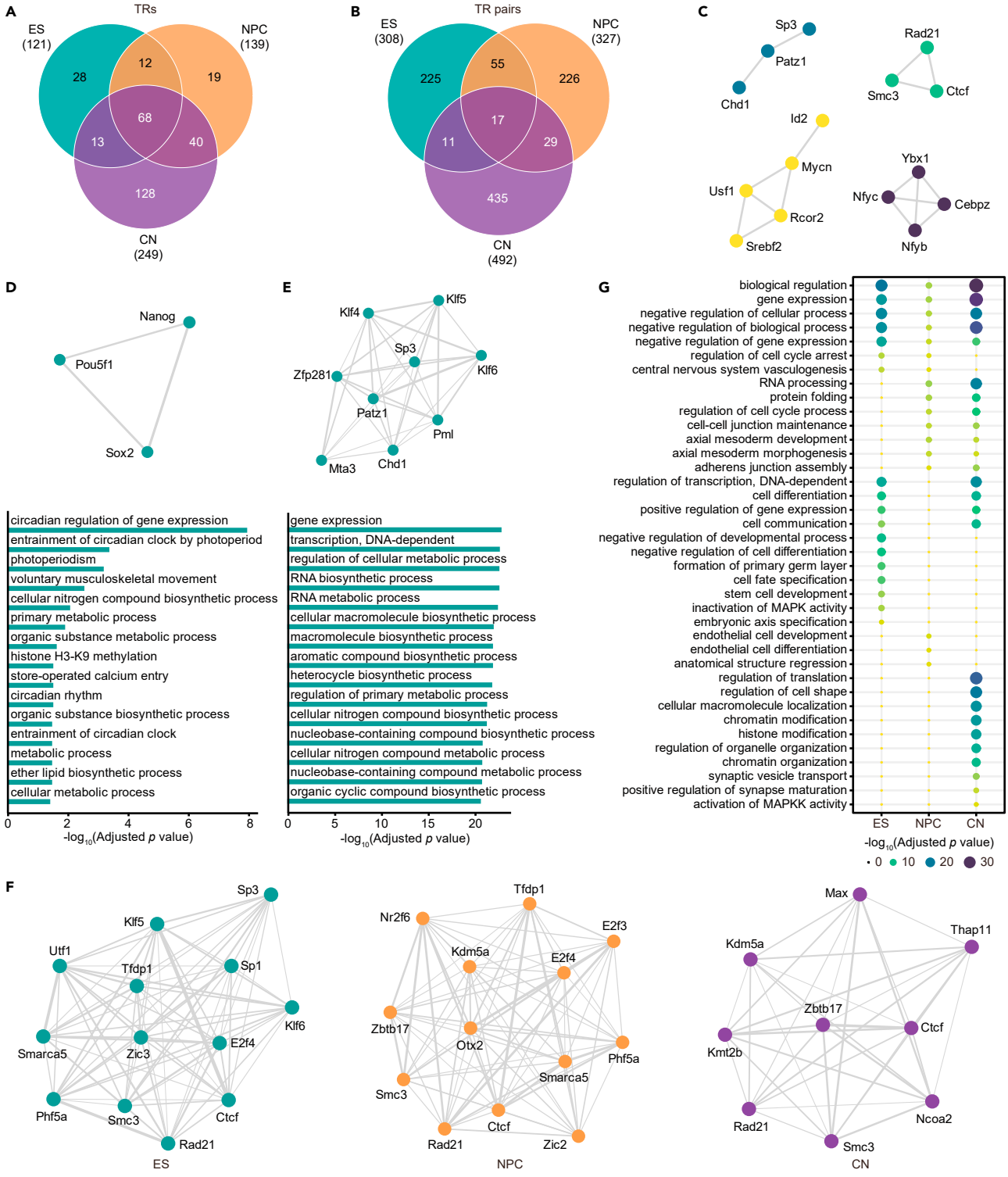


Figure 6. Lineage-specific TR cooperation during mouse neural development

(A) The Venn graph of qualified TRs among three cell states, ES, NPC, and CN.

(B) The Venn graph of TR pairs detected by 3DCoop among three cell states, ES, NPC, and CN.

(C) The TR cooperation network constructed from 17 TR pairs shared by all three cell states. TR clusters were colored for distinction.

(D) Nanog-Pou5f1-Sox2 cooperation cluster in ES and its top15 significantly enriched biological processes based on GREAT annotation analysis. Biological process terms from GO (Gene Ontology) were ordered by adjusted *p* value for enrichment.

Figure 6. *Continued*

(E) Chd1-Klf4-Klf5-Klf6-Mta3-Patz1-Pml-Sp3-Zfp281 cooperation cluster in ES and its top15 significantly enriched biological processes based on GREAT annotation analysis.

(F) The changes of associated TR members of Ctcf-Rad21-Smc3 cooperation during mouse neural development, from ES to NPC to CN. Edges between TRs were weighted by their corresponding Glasso score.

(G) Comparison of cellular functions driven by Ctcf-Rad21-Smc3-related TR cooperation in different stages. Cellular functions were characterized by significantly enriched biological process terms from GO based on GREAT annotation analysis.

See also Figures S21–S23.

largest TR cooperation analysis in K562 cells and demonstrated that TR cooperation largely contributes to leukemia-specific and immune-related function. Measurement of the activity of the TR intervening chromatin interactions showed that some TRs function as novel architectural proteins in regulating genome organization, such as MAZ (Ortabozkoyun-Kara et al., 2020), BHLHE40 (Hu et al., 2020), CTCFL (Debruyne et al., 2019), REST, and MAX. Consistent with recent findings (Van Nostrand et al., 2020; Xiao et al., 2019), we also observed several potential interplays among many chromatin-associated RBPs. So far, the categorization of TR cooperation and their chromatin arrangement in the 3D genome across different tissue/cell types remains challenging. To extend the genome-wide detection of TR cooperation in broad tissue/cell types, we leveraged computational TF footprinting to delineate a landscape of TR cooperation networks and to illustrate the previously uncharacterized patterns of TR cooperation across 40 human tissue/cell types. We also demonstrated how the tissue/cell type-specific TR cooperation information can aid the interpretation of disease-causal regulatory variants and dynamic TR cooperation during cell development.

Compared with the previous method DBPnet, the novelties of 3DCoop pipeline include using peaks than signals of ChIP-seq datasets, which makes it generalized, and incorporating overlapping clustering to estimate network modules, which allows the premise that a single TR can be involved in multiple cooperation communities as *in vivo*. Considering the difficulty to derive tissue/cell type-specific PPIs from experimental data or public resources such as STRING, we did not directly incorporate PPI information into the 3DCoop pipeline. Alternatively, tissue/cell type-specific epigenomics data accumulate intensively recently, including DNase-seq/ATAC-seq and ChIP-seq/CUT&Tag, which motivates us to generalize the 3DCoop on widespread conditions. In addition, although TR binding patterns (such as TR co-binding frequency and shared motif site frequency) were not used directly in 3DCoop pipeline, they were embodied implicitly. For example, the generalized Jaccard similarity based on the TR-specific contact maps reflects the TR co-binding in 3D chromatin. More than two-thirds TRs used in our generalized 3DCoop pipeline have two or three motifs. These motifs may encode co-binding patterns among multiple TRs.

Although we have delineated potential cooperation between many TRs in 3D chromatin across different tissue/cell types in the present study, the complete arrangement pattern of TR cooperation remains unknown (Morgunova and Taipale, 2017). Whether the cooperative bindings are aggregative, sequential, or competitive requires further investigation (Rao et al., 2021). Recent multi-way chromatin contacts identification methods have enabled capture of the simultaneous interactions among CREs, such as C-walks (Olivares-Chauvet et al., 2016), SPRITE (Quinodoz et al., 2018), and Trac-looping (Lai et al., 2018), which will greatly expand the searching space of potential TR communications. To incorporate more tissue/cell types, we used chromatin loops at 10-kb resolution identified by Peakachu to construct TR-specific contact maps. This compromise may introduce false-positive detection of TR cooperation, although good agreements among Hi-C and ChIA-PET data can be achieved. Micro-C and DNase Hi-C, utilizing micrococcal nuclease or DNase I to cut genomic DNA, can achieve mononucleosome resolution (Hsieh et al., 2015; Ma et al., 2015). Such fine-scale chromatin interaction capture techniques and high-throughput validation experiments in specific contexts (Stampfel et al., 2015) could facilitate more precise identification of TR cooperation in 3D chromatin. Moreover, large-scale 3D genome profiling of more tissue/cell types and development states will further extend the volume of TR cooperation identification (Jung et al., 2019; Zhang et al., 2020).

Only a small fraction of chromatin-associated factors has been ChIPped in diverse tissue/cell types (Consortium et al., 2020b; Lambert et al., 2018), yet high-density DNase I cleavage maps, together with abundant TR motifs, enable computational analysis of TF occupancy across the whole genome (Meuleman et al., 2020; Vierstra et al., 2020). The difference in TR cooperation between that inferred by TF footprinting and

ChIP-seq-based results is inevitable. It can be attributed to many intrinsic experimental features and biases of the applied methods. For example, ChIP-seq can capture the tethered binding of TRs that do not have DNA-binding domains or that indirectly interact with genomic DNA, while computational TF footprinting is highly reliant on cognate or composite recognition sequences that could be shared by different TRs (Jolma et al., 2015; Sung et al., 2014). In addition, computational footprinting analyses using DNase-seq or ATAC-seq contain distinct sequence biases, but existing footprinting methods selectively address these biases and exhibit varied performance on different TRs (He et al., 2014; Karabacak Calviello et al., 2019). The combinatory application of multiple tools, not only PIQ in the present study, and the introduction of reproducibility assessment could mitigate such biases (Gusmao et al., 2016; Karabacak Calviello et al., 2019). The recent Micro-Capture-C has also introduced single-base pair maps of ligation junctions for analyzing directional footprinting and protein-protein contacts at high resolution between viewpoints and interacting elements (Hua et al., 2021), which would significantly facilitate the identification of fine-scale TR cooperation in 3D chromatin.

The landscape of TR cooperation established by applying 3DCoop to 40 human tissue/cell types could be a useful resource in TR cooperation study and regulatory genomics. There are several potential biological applications based on the generated data. The identified tissue/cell type-shared TR cooperation network could provide a global reference for studying common transcriptional partners. Also, the TR cooperation detected by 3DCoop shows the tissue/cell type specificity, which can promote the regulatory genomics study in certain tissue/cell type as the data growing for more and more tissue/cell types. Tissue/cell type-specific TR cooperation can facilitate the interpretation of disease-causal variants identified by GWAS, especially for prioritizing candidate TFs associated with certain disease-causal regulatory variants. Moreover, the lineage-specific TR cooperation during tissue development could facilitate the relevant studies for gene regulation dynamics. We expect that the resource and 3DCoop pipeline, together with the new findings, will benefit mechanistic research on how the cooperation of certain TRs shapes cell type-specific gene regulation in development and disease progression.

Limitations of the study

Genome-wide systematic detection of TR cooperation across multiple tissue/cell types is experimental, expensive, and sometimes irreproducible. Because of the lack of ground truth in the field, a rigorous evaluation is currently infeasible given the available data, particularly for tissue/cell type-specific TR cooperation. Thus, experimental validation of some context-specific TR cooperation will improve the rigorousness of the predicted results and deserve further investigation.

STAR★METHODS

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

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- METHOD DETAILS
  - Collecting and processing data in K562
  - The 3DCoop pipeline
  - Comparison with ChIA-PET datasets
  - Measurement of TR activity in 3D chromatin
  - Comparison with PPI, methods, and experimental data
  - Functional annotation of TR maximum cliques
  - Colocalization analysis on epigenomic marks
  - Collecting and processing of human TR motifs
  - Extended 3DCoop using computational TF footprinting
  - Analysis of tissue/cell type specificity of TR cooperation
  - Evaluation of tissue/cell type-specific pattern with CRC
  - Interpretation of GWAS disease-causal variants with TR cooperation information
  - Analysis of TR cooperation in mouse neural development

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2021.103468.

ACKNOWLEDGMENTS

This work was supported by the following grants: the National Natural Science Foundation of China (32070675 to M.J.L., 31871327 to M.J.L., and 32000640 to Z.W.), and the Natural Science Foundation of Tianjin (19JCJQJC63600 to M.J.L., 19JCQNJC09000 to X.Y., and 18JCYBJC27300 to Y.S.). Part of the data used in the analyses described in this article was obtained from the ENCODE, the Roadmap Epigenomics Project, the 3D Genome Browser, and other resources. We appreciate all tool and resource providers.

AUTHOR CONTRIBUTIONS

X.Y. designed the experiments, conducted the computational analysis, and wrote the paper. M.J.L. conceived the study, designed the experiments, and wrote the paper. Z.Z., H.X., Y.Z., D.H., J.W., X.F., K.Z., and X.F. evaluated the computational tool and reviewed the manuscript. S.Z., X.D., Z.W., Y.S., H.C., and L.S. suggested the tool functions and reviewed the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: August 2, 2021
Revised: September 23, 2021
Accepted: November 12, 2021
Published: December 17, 2021

REFERENCES

Au-Yeung, N., Mandhana, R., and Horvath, C.M. (2013). Transcriptional regulation by STAT1 and STAT2 in the interferon JAK-STAT pathway. JAKSTAT *2*, e23931.

Bhushan, N., Mohnert, F., Sloot, D., Jans, L., Albers, C., and Steg, L. (2019). Using a Gaussian graphical model to explore relationships between items and variables in environmental psychology research. Front. Psychol. *10*, 1050.

Boix, C.A., James, B.T., Park, Y.P., Meuleman, W., and Kellis, M. (2021). Regulatory genomic circuitry of human disease loci by integrative epigenomics. Nature *590*, 300–307.

Bonev, B., Mendelson Cohen, N., Szabo, Q., Fritsch, L., Papadopoulos, G.L., Lubling, Y., Xu, X., Lv, X., Hugnot, J.P., Tanay, A., et al. (2017). Multiscale 3D genome rewiring during mouse neural development. Cell *171*, 557–572.e24.

Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G., et al. (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. Cell *122*, 947–956.

Chambers, I., and Tomlinson, S.R. (2009). The transcriptional foundation of pluripotency. Development *136*, 2311–2322.

Cheneby, J., Gheorghe, M., Artufel, M., Mathelier, A., and Ballester, B. (2018). ReMap 2018: an updated atlas of regulatory regions from an integrative analysis of DNA-binding ChIP-seq experiments. Nucl. Acids Res. *46*, D267–D275.

Consortium, E.P. (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57–74.

Consortium, E.P., Moore, J.E., Purcaro, M.J., Pratt, H.E., Epstein, C.B., Shoresh, N., Adrian, J., Kawli, T., Davis, C.A., Dobin, A., et al. (2020a). Expanded encyclopaedias of DNA elements in the human and mouse genomes. Nature *583*, 699–710.

Consortium, E.P., Snyder, M.P., Gingeras, T.R., Moore, J.E., Weng, Z., Gerstein, M.B., Ren, B., Hardison, R.C., Stamatoyannopoulos, J.A., Graveley, B.R., et al. (2020b). Perspectives on ENCODE. Nature *583*, 693–698.

de Wit, E., and de Laat, W. (2012). A decade of 3C technologies: insights into nuclear organization. Genes Dev. *26*, 11–24.

Debruyne, D.N., Dries, R., Sengupta, S., Seruggia, D., Gao, Y., Sharma, B., Huang, H., Moreau, L., McLane, M., Day, D.S., et al. (2019). BORIS promotes chromatin regulatory interactions in treatment-resistant cancer cells. Nature *572*, 676–680.

Denker, A., and de Laat, W. (2016). The second decade of 3C technologies: detailed insights into nuclear organization. Genes Dev. *30*, 1357–1382.

Di Giammartino, D.C., Kloetgen, A., Polyzos, A., Liu, Y., Kim, D., Murphy, D., Abuhashem, A., Cavaliere, P., Aronson, B., Shah, V., et al. (2019). KLF4 is involved in the organization and regulation of pluripotency-associated three-dimensional enhancer networks. Nat. Cell Biol. *21*, 1179–1190.

Djekidel, M.N., Liang, Z., Wang, Q., Hu, Z., Li, G., Chen, Y., and Zhang, M.Q. (2015). 3CPET: finding co-factor complexes from ChIA-PET data using a hierarchical Dirichlet process. Genome Biol. *16*, 288.

Eblen, J.D., Gerling, I.C., Saxton, A.M., Wu, J., Snoddy, J.R., and Langston, M.A. (2009). Graph algorithms for integrated biological analysis, with applications to type 1 diabetes data. In Clustering Challenges in Biological Networks, pp. 207–222.

Epskamp, S., Borsboom, D., and Fried, E.I. (2018). Estimating psychological networks and their accuracy: a tutorial paper. Behav. Res. Methods *50*, 195–212.

Farh, K.K., Marson, A., Zhu, J., Kleinewietfeld, M., Housley, W.J., Beik, S., Shoresh, N., Whitton, H., Ryan, R.J., Shishkin, A.A., et al. (2015). Genetic and epigenetic fine mapping of causal autoimmune disease variants. Nature *518*, 337–343.

Fornes, O., Castro-Mondragon, J.A., Khan, A., van der Lee, R., Zhang, X., Richmond, P.A., Modi, B.P., Correard, S., Gheorghe, M., Baranasic, D., et al. (2020). Jaspar 2020: update of the open-access database of transcription factor binding profiles. Nucl. Acids Res. *48*, D87–D92.

Friedman, J., Hastie, T., and Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. Biostatistics *9*, 432–441.

Gerstein, M.B., Kundaje, A., Hariharan, M., Landt, S.G., Yan, K.K., Cheng, C., Mu, X.J., Khurana, E., Rozowsky, J., Alexander, R., et al. (2012).

Architecture of the human regulatory network derived from ENCODE data. Nature *489*, 91–100.

Giannopoulou, E.G., and Elemento, O. (2013). Inferring chromatin-bound protein complexes from genome-wide binding assays. Genome Res. *23*, 1295–1306.

Grant, C.E., Bailey, T.L., and Noble, W.S. (2011). FIMO: scanning for occurrences of a given motif. Bioinformatics *27*, 1017–1018.

Guo, Y., and Gifford, D.K. (2017). Modular combinatorial binding among human trans-acting factors reveals direct and indirect factor binding. BMC Genomics *18*, 45.

Gusmao, E.G., Allhoff, M., Zenke, M., and Costa, I.G. (2016). Analysis of computational footprinting methods for DNase sequencing experiments. Nat. Methods *13*, 303–309.

Hardison, R.C., and Taylor, J. (2012). Genomic approaches towards finding cis-regulatory modules in animals. Nat. Rev. Genet. *13*, 469–483.

He, H.H., Meyer, C.A., Hu, S.S., Chen, M.W., Zang, C., Liu, Y., Rao, P.K., Fei, T., Xu, H., Long, H., et al. (2014). Refined DNase-seq protocol and data analysis reveals intrinsic bias in transcription factor footprint identification. Nat. Methods *11*, 73–78.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell. *38*, 576–589.

Hervouet, E., Peixoto, P., Delage-Mourroux, R., Boyer-Guittaut, M., and Cartron, P.F. (2018). Specific or not specific recruitment of DNMTs for DNA methylation, an epigenetic dilemma. Clin. Epigenetics *10*, 17.

Hou, C., Dale, R., and Dean, A. (2010). Cell type specificity of chromatin organization mediated by CTCF and cohesin. Proc. Natl. Acad. Sci. U S A *107*, 3651–3656.

Hsieh, T.H., Weiner, A., Lajoie, B., Dekker, J., Friedman, N., and Rando, O.J. (2015). Mapping nucleosome resolution chromosome folding in yeast by micro-C. Cell *162*, 108–119.

Hu, G., Dong, X., Gong, S., Song, Y., Hutchins, A.P., and Yao, H. (2020). Systematic screening of CTCF binding partners identifies that BHLHE40 regulates CTCF genome-wide distribution and long-range chromatin interactions. Nucl. Acids Res. *48*, 9606–9620.

Hua, P., Badat, M., Hanssen, L.L.P., Hentges, L.D., Crump, N., Downes, D.J., Jeziorska, D.M., Oudelaar, A.M., Schwessinger, R., Taylor, S., et al. (2021). Defining genome architecture at base-pair resolution. Nature *595*, 125–129.

Huang, D., Wang, Z., Zhou, Y., Liang, Q., Sham, P.C., Yao, H., and Li, M.J. (2021). vSampler: fast and annotation-based matched variant sampling tool. Bioinformatics *37*, 1915–1917.

Huang, D., Yi, X., Zhang, S., Zheng, Z., Wang, P., Xuan, C., Sham, P.C., Wang, J., and Li, M.J. (2018a). GWAS4D: multidimensional analysis of context-specific regulatory variant for human complex diseases and traits. Nucl. Acids Res. *46*, W114–W120.

Huang, M., Chen, Y., Yang, M., Guo, A., Xu, Y., Xu, L., and Koeffler, H.P. (2018b). dbCoRC: a database of core transcriptional regulatory circuitries modeled by H3K27ac ChIP-seq signals. Nucl. Acids Res. *46*, D71–D77.

Huang, D., Yi, X., Zhou, Y., Yao, H., Xu, H., Wang, J., Zhang, S., Nong, W., Wang, P., Shi, L., et al. (2020). Ultrafast and scalable variant annotation and prioritization with big functional genomics data. Genome Res. *30*, 1789–1801.

Hume, M.A., Barrera, L.A., Gisselbrecht, S.S., and Bulyk, M.L. (2015). UniPROBE, update 2015: new tools and content for the online database of protein-binding microarray data on protein-DNA interactions. Nucl. Acids Res. *43*, D117–D122.

Isakova, A., Groux, R., Imbeault, M., Rainer, P., Alpern, D., Dainese, R., Ambrosini, G., Trono, D., Bucher, P., and Deplancke, B. (2017). SMiLE-seq identifies binding motifs of single and dimeric transcription factors. Nat. Methods *14*, 316–322.

Jalili, V., Matteucci, M., Morelli, M.J., and Masseroli, M. (2017). MuSERA: multiple sample enriched region assessment. Brief Bioinform *18*, 367–381.

Jolly, L.A., Nguyen, L.S., Domingo, D., Sun, Y., Barry, S., Hancarova, M., Plevova, P., Vlckova, M., Havlovicova, M., Kalscheuer, V.M., et al. (2015). HCFC1 loss-of-function mutations disrupt neuronal and neural progenitor cells of the developing brain. Hum. Mol. Genet. *24*, 3335–3347.

Jolma, A., Kivioja, T., Toivonen, J., Cheng, L., Wei, G., Enge, M., Taipale, M., Vaquerizas, J.M., Yan, J., Sillanpaa, M.J., et al. (2010). Multiplexed massively parallel SELEX for characterization of human transcription factor binding specificities. Genome Res. *20*, 861–873.

Jolma, A., Yan, J., Whitington, T., Toivonen, J., Nitta, K.R., Rastas, P., Morgunova, E., Enge, M., Taipale, M., Wei, G., et al. (2013). DNA-binding specificities of human transcription factors. Cell *152*, 327–339.

Jolma, A., Yin, Y., Nitta, K.R., Dave, K., Popov, A., Taipale, M., Enge, M., Kivioja, T., Morgunova, E., and Taipale, J. (2015). DNA-dependent formation of transcription factor pairs alters their binding specificity. Nature *527*, 384–388.

Jung, I., Schmitt, A., Diao, Y., Lee, A.J., Liu, T., Yang, D., Tan, C., Eom, J., Chan, M., Chee, S., et al. (2019). A compendium of promoter-centered long-range chromatin interactions in the human genome. Nat. Genet. *51*, 1442–1449.

Karabacak Calviello, A., Hirsekorn, A., Wurmus, R., Yusuf, D., and Ohler, U. (2019). Reproducible inference of transcription factor footprints in ATAC-seq and DNase-seq datasets using protocol-specific bias modeling. Genome Biol. *20*, 42.

Kazemian, M., Pham, H., Wolfe, S.A., Brodsky, M.H., and Sinha, S. (2013). Widespread evidence of cooperative DNA binding by transcription factors in Drosophila development. Nucl. Acids Res. *41*, 8237–8252.

Kempfer, R., and Pombo, A. (2020). Methods for mapping 3D chromosome architecture. Nat. Rev. Genet. *21*, 207–226.

Kessler, D.S., Veals, S.A., Fu, X.Y., and Levy, D.E. (1990). Interferon-alpha regulates nuclear translocation and DNA-binding affinity of ISGF3, a multimeric transcriptional activator. Genes Dev. *4*, 1753–1765.

Kheradpour, P., and Kellis, M. (2014). Systematic discovery and characterization of regulatory motifs in ENCODE TF binding experiments. Nucl. Acids Res. *42*, 2976–2987.

Kim, S., and Shendure, J. (2019). Mechanisms of interplay between transcription factors and the 3D genome. Mol. Cell *76*, 306–319.

Klemm, S.L., Shipony, Z., and Greenleaf, W.J. (2019). Chromatin accessibility and the regulatory epigenome. Nat. Rev. Genet. *20*, 207–220.

Kulakovskiy, I.V., Vorontsov, I.E., Yevshin, I.S., Sharipov, R.N., Fedorova, A.D., Rumynskiy, E.I., Medvedeva, Y.A., Magana-Mora, A., Bajic, V.B., Papatsenko, D.A., et al. (2018). HOCOMOCO: towards a complete collection of transcription factor binding models for human and mouse via large-scale ChIP-Seq analysis. Nucl. Acids Res. *46*, D252–D259.

Lai, B., Tang, Q., Jin, W., Hu, G., Wangsa, D., Cui, K., Stanton, B.Z., Ren, G., Ding, Y., Zhao, M., et al. (2018). Trac-looping measures genome structure and chromatin accessibility. Nat. Methods *15*, 741–747.

Lambert, S.A., Jolma, A., Campitelli, L.F., Das, P.K., Yin, Y., Albu, M., Chen, X., Taipale, J., Hughes, T.R., and Weirauch, M.T. (2018). The human transcription factors. Cell *172*, 650–665.

Layer, R.M., Pedersen, B.S., DiSera, T., Marth, G.T., Gertz, J., and Quinlan, A.R. (2018). GIGGLE: a search engine for large-scale integrated genome analysis. Nat. Methods *15*, 123–126.

Li, M.J., Li, M., Liu, Z., Yan, B., Pan, Z., Huang, D., Liang, Q., Ying, D., Xu, F., Yao, H., et al. (2017a). cepip: context-dependent epigenomic weighting for prioritization of regulatory variants and disease-associated genes. Genome Biol. *18*, 52.

Li, T., Wernersson, R., Hansen, R.B., Horn, H., Mercer, J., Slodkowicz, G., Workman, C.T., Rigina, O., Rapacki, K., Staerfeldt, H.H., et al. (2017b). A scored human protein-protein interaction network to catalyze genomic interpretation. Nat. Methods *14*, 61–64.

Liu, H., Han, F., Yuan, M., Lafferty, J., and Wasserman, L. (2012). High-dimensional semiparametric Gaussian copula graphical models. Ann. Stat. *40*, 2293–2326.

Liu, X., Zhang, Y., Chen, Y., Li, M., Zhou, F., Li, K., Cao, H., Ni, M., Liu, Y., Gu, Z., et al. (2017). In situ capture of chromatin interactions by biotinylated dCas9. Cell *170*, 1028–1043.e1019.

Ma, C.A., Pusso, A., Wu, L., Zhao, Y., Hoffmann, V., Notarangelo, L.D., Fowlkes, B.J., and Jain, A. (2014). Novel INHAT repressor (NIR) is required for early lymphocyte development. Proc. Natl. Acad. Sci. U S A. *111*, 13930–13935.

Ma, W., Ay, F., Lee, C., Gulsoy, G., Deng, X., Cook, S., Hesson, J., Cavanaugh, C., Ware, C.B., Krumm, A., et al. (2015). Fine-scale chromatin interaction maps reveal the cis-regulatory landscape of human lincRNA genes. Nat. Methods *12*, 71–78.

Maston, G.A., Evans, S.K., and Green, M.R. (2006). Transcriptional regulatory elements in the human genome. Annu. Rev. Genomics Hum. Genet. *7*, 29–59.

Matys, V., Kel-Margoulis, O.V., Fricke, E., Liebich, I., Land, S., Barre-Dirrie, A., Reuter, I., Chekmenev, D., Krull, M., Hornischer, K., et al. (2006). TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes. Nucl. Acids Res. *34*, D108–D110.

McCord, R.P., Kaplan, N., and Giorgetti, L. (2020). Chromosome conformation capture and beyond: toward an integrative view of chromosome structure and function. Mol. Cell *77*, 688–708.

McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., and Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. Nat. Biotechnol. *28*, 495–501.

Merkenschlager, M. (2010). Ikaros in immune receptor signaling, lymphocyte differentiation, and function. FEBS Lett. *584*, 4910–4914.

Meuleman, W., Muratov, A., Rynes, E., Halow, J., Lee, K., Bates, D., Diegel, M., Dunn, D., Neri, F., Teodosiadis, A., et al. (2020). Index and biological spectrum of human DNase I hypersensitive sites. Nature *584*, 244–251.

Michaud, J., Praz, V., James Faresse, N., Jnbaptiste, C.K., Tyagi, S., Schutz, F., and Herr, W. (2013). HCFC1 is a common component of active human CpG-island promoters and coincides with ZNF143, THAP11, YY1, and GABP transcription factor occupancy. Genome Res. *23*, 907–916.

Morgunova, E., and Taipale, J. (2017). Structural perspective of cooperative transcription factor binding. Curr. Opin. Struct. Biol. *47*, 1–8.

Mulgrave, J.J., and Ghosal, S. (2020). Bayesian inference in nonparanormal graphical models. Bayesian Anal. *15*, 449–475.

Nepusz, T., Yu, H., and Paccanaro, A. (2012). Detecting overlapping protein complexes in protein-protein interaction networks. Nat. Methods *9*, 471–472.

Oeckinghaus, A., and Ghosh, S. (2009). The NF-kappaB family of transcription factors and its regulation. Cold Spring Harb Perspect. Biol. *1*, a000034.

Oki, S., Ohta, T., Shioi, G., Hatanaka, H., Ogasawara, O., Okuda, Y., Kawaji, H., Nakaki, R., Sese, J., and Meno, C. (2018). ChIP-Atlas: a data-mining suite powered by full integration of public ChIP-seq data. EMBO Rep. *19*, e46255.

Olivares-Chauvet, P., Mukamel, Z., Lifshitz, A., Schwartzman, O., Elkayam, N.O., Lubling, Y., Deikus, G., Sebra, R.P., and Tanay, A. (2016). Capturing pairwise and multi-way chromosomal conformations using chromosomal walks. Nature *540*, 296–300.

Orchard, S., Ammari, M., Aranda, B., Breuza, L., Briganti, L., Broackes-Carter, F., Campbell, N.H., Chavali, G., Chen, C., del-Toro, N., et al. (2014). The MIntAct project–IntAct as a common curation platform for 11 molecular interaction databases. Nucl. Acids Res. *42*, D358–D363.

Ortabozkoyun-Kara, H., Huang, P.-Y., Cho, H., Narendra, V., Leroy, G., Skok, J.A., Tsirigos, A., Mazzoni, E.O., and Reinberg, D. (2020). A CRISPR screen identifies Myc-associated zinc finger protein (MAZ) as an insulator functioning at CTCF boundaries in Hox clusters. https://doi.org/10.1101/2020.08.25.267237.

Oughtred, R., Stark, C., Breitkreutz, B.J., Rust, J., Boucher, L., Chang, C., Kolas, N., O'Donnell, L., Leung, G., McAdam, R., et al. (2019). The BioGRID interaction database: 2019 update. Nucl. Acids Res. *47*, D529–D541.

Pachkov, M., Balwierz, P.J., Arnold, P., Ozonov, E., and van Nimwegen, E. (2013). SwissRegulon, a database of genome-wide annotations of regulatory sites: recent updates. Nucl. Acids Res. *41*, D214–D220.

Perna, S., Pinoli, P., Ceri, S., and Wong, L. (2018). TICA: transcriptional interaction and coregulation analyzer. Genom. Proteomics Bioinform. *16*, 342–353.

Petrovic, J., Zhou, Y., Fasolino, M., Goldman, N., Schwartz, G.W., Mumbach, M.R., Nguyen, S.C., Rome, K.S., Sela, Y., Zapataro, Z., et al. (2019). Oncogenic notch promotes long-range regulatory interactions within hyperconnected 3D cliques. Mol. Cell *73*, 1174–1190.e12.

Platanitis, E., Demiroz, D., Schneller, A., Fischer, K., Capelle, C., Hartl, M., Gossenreiter, T., Muller, M., Novatchkova, M., and Decker, T. (2019). A molecular switch from STAT2-IRF9 to ISGF3 underlies interferon-induced gene transcription. Nat. Commun. *10*, 2921.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842.

Quinodoz, S.A., Ollikainen, N., Tabak, B., Palla, A., Schmidt, J.M., Detmar, E., Lai, M.M., Shishkin, A.A., Bhat, P., Takei, Y., et al. (2018). Higher-order inter-chromosomal hubs shape 3D genome organization in the nucleus. Cell *174*, 744–757.e24.

Rao, S., Ahmad, K., and Ramachandran, S. (2021). Cooperative binding between distant transcription factors is a hallmark of active enhancers. Mol. Cell *81*, 1651–1665.e4.

Rao, S.S., 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.

Roadmap Epigenomics, C., Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., Kheradpour, P., Zhang, Z., Wang, J., et al. (2015). Integrative analysis of 111 reference human epigenomes. Nature *518*, 317–330.

Robson, M.I., Ringel, A.R., and Mundlos, S. (2019). Regulatory landscaping: how enhancer-promoter communication is sculpted in 3D. Mol. Cell *74*, 1110–1122.

Saint-Andre, V., Federation, A.J., Lin, C.Y., Abraham, B.J., Reddy, J., Lee, T.I., Bradner, J.E., and Young, R.A. (2016). Models of human core transcriptional regulatory circuitries. Genome Res. *26*, 385–396.

Salameh, T.J., Wang, X., Song, F., Zhang, B., Wright, S.M., Khunsriraksakul, C., Ruan, Y., and Yue, F. (2020). A supervised learning framework for chromatin loop detection in genome-wide contact maps. Nat. Commun. *11*, 3428.

Scheer, S., and Zaph, C. (2017). The lysine methyltransferase G9a in immune cell differentiation and function. Front. Immunol. *8*, 429.

Schwessinger, R., Gosden, M., Downes, D., Brown, R.C., Oudelaar, A.M., Telenius, J., Teh, Y.W., Lunter, G., and Hughes, J.R. (2020). DeepC: predicting 3D genome folding using megabase-scale transfer learning. Nat. Methods *17*, 1118–1124.

Servant, N., Varoquaux, N., Lajoie, B.R., Viara, E., Chen, C.J., Vert, J.P., Heard, E., Dekker, J., and Barillot, E. (2015). HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. Genome Biol. *16*, 259.

Sherwood, R.I., Hashimoto, T., O'Donnell, C.W., Lewis, S., Barkal, A.A., van Hoff, J.P., Karun, V., Jaakkola, T., and Gifford, D.K. (2014). Discovery of directional and nondirectional pioneer transcription factors by modeling DNase profile magnitude and shape. Nat. Biotechnol. *32*, 171–178.

Shrestha, S., Sewell, J.A., Santoso, C.S., Forchielli, E., Carrasco Pro, S., Martinez, M., and Fuxman Bass, J.I. (2019). Discovering human transcription factor physical interactions with genetic variants, novel DNA motifs, and repetitive elements using enhanced yeast one-hybrid assays. Genome Res. *29*, 1533–1544.

Spitz, F., and Furlong, E.E. (2012). Transcription factors: from enhancer binding to developmental control. Nat. Rev. Genet. *13*, 613–626.

Stadhouders, R., Filion, G.J., and Graf, T. (2019). Transcription factors and 3D genome conformation in cell-fate decisions. Nature *569*, 345–354.

Stampfel, G., Kazmar, T., Frank, O., Wienerroither, S., Reiter, F., and Stark, A. (2015). Transcriptional regulators form diverse groups with context-dependent regulatory functions. Nature *528*, 147–151.

Sung, M.H., Guertin, M.J., Baek, S., and Hager, G.L. (2014). DNase footprint signatures are dictated by factor dynamics and DNA sequence. Mol. Cell *56*, 275–285.

Szklarczyk, D., Gable, A.L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N.T., Morris, J.H., Bork, P., et al. (2019). STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucl. Acids Res. *47*, D607–D613.

Tang, Z., Luo, O.J., Li, X., Zheng, M., Zhu, J.J., Szalaj, P., Trzaskoma, P., Magalska, A., Wlodarczyk, J., Ruszczycki, B., et al. (2015). CTCF-mediated human 3D genome architecture reveals chromatin topology for transcription. Cell *163*, 1611–1627.

Tehranchi, A.K., Myrthil, M., Martin, T., Hie, B.L., Golan, D., and Fraser, H.B. (2016). Pooled ChIP-Seq links variation in transcription factor binding to complex disease risk. Cell *165*, 730–741.

Tian, D., Zhang, R., Zhang, Y., Zhu, X., and Ma, J. (2020). MOCHI enables discovery of heterogeneous interactome modules in 3D nucleome. Genome Res. *30*, 227–238.

van der Meer, L.T., Jansen, J.H., and van der Reijden, B.A. (2010). Gfi1 and Gfi1b: key regulators of hematopoiesis. Leukemia *24*, 1834–1843.

Van Nostrand, E.L., Freese, P., Pratt, G.A., Wang, X., Wei, X., Xiao, R., Blue, S.M., Chen, J.Y., Cody, N.A.L., Dominguez, D., et al. (2020). A large-scale binding and functional map of human RNA-binding proteins. Nature *583*, 711–719.

Vierstra, J., Lazar, J., Sandstrom, R., Halow, J., Lee, K., Bates, D., Diegel, M., Dunn, D., Neri, F., Haugen, E., et al. (2020). Global reference mapping of human transcription factor footprints. Nature *583*, 729–736.

Vorontsov, I.E., Kulakovskiy, I.V., and Makeev, V.J. (2013). Jaccard index based similarity measure to compare transcription factor binding site models. Algorithms Mol. Biol. *8*, 23.

Voss, T.C., and Hager, G.L. (2014). Dynamic regulation of transcriptional states by chromatin and transcription factors. Nat. Rev. Genet. *15*, 69–81.

Wang, J., Huang, D., Zhou, Y., Yao, H., Liu, H., Zhai, S., Wu, C., Zheng, Z., Zhao, K., Wang, Z., et al. (2020). CAUSALdb: a database for disease/trait causal variants identified using summary statistics of genome-wide association studies. Nucl. Acids Res. *48*, D807–D816.

Wang, R., Wang, Y., Zhang, X., Zhang, Y., Du, X., Fang, Y., and Li, G. (2019). Hierarchical cooperation of transcription factors from integration analysis of DNA sequences, ChIP-Seq and ChIA-PET data. BMC Genomics *20*, 296.

Wang, Y., Song, F., Zhang, B., Zhang, L., Xu, J., Kuang, D., Li, D., Choudhary, M.N.K., Li, Y., Hu, M., et al. (2018). The 3D genome browser: a web-based browser for visualizing 3D genome organization and long-range chromatin interactions. Genome Biol. *19*, 151.

Weintraub, A.S., Li, C.H., Zamudio, A.V., Sigova, A.A., Hannett, N.M., Day, D.S., Abraham, B.J., Cohen, M.A., Nabet, B., Buckley, D.L., et al. (2017). YY1 is a structural regulator of enhancer-promoter loops. Cell *171*, 1573–1588.e1528.

Weirauch, M.T., Yang, A., Albu, M., Cote, A.G., Montenegro-Montero, A., Drewe, P., Najafabadi, H.S., Lambert, S.A., Mann, I., Cook, K., et al. (2014). Determination and inference of eukaryotic transcription factor sequence specificity. Cell *158*, 1431–1443.

Wong, A.K., Krishnan, A., and Troyanskaya, O.G. (2018). Giant 2.0: genome-scale integrated analysis of gene networks in tissues. Nucl. Acids Res. *46*, W65–W70.

Xiao, R., Chen, J.Y., Liang, Z., Luo, D., Chen, G., Lu, Z.J., Chen, Y., Zhou, B., Li, H., Du, X., et al. (2019). Pervasive chromatin-RNA binding protein interactions enable RNA-based regulation of transcription. Cell *178*, 107–121.e118.

Xie, Z., Hu, S., Blackshaw, S., Zhu, H., and Qian, J. (2010). hPDI: a database of experimental human protein-DNA interactions. Bioinformatics *26*, 287–289.

Xu, H., Zhang, S., Yi, X., Plewczynski, D., and Li, M.J. (2020). Exploring 3D chromatin contacts in gene regulation: the evolution of approaches for the identification of functional enhancer-promoter interaction. Comput. Struct. Biotechnol. J. *18*, 558–570.

Xue, L.Z., and Zou, H. (2012). Regularized rank-based estimation of high-dimensional nonparanormal graphical models. Ann. Stat. *40*, 2541–2571.

Yan, J., Qiu, Y., Ribeiro Dos Santos, A.M., Yin, Y., Li, Y.E., Vinckier, N., Nariai, N., Benaglio, P., Raman, A., Li, X., et al. (2021). Systematic analysis of binding of transcription factors to noncoding variants. Nature *591*, 147–151.

Yevshin, I., Sharipov, R., Kolmykov, S., Kondrakhin, Y., and Kolpakov, F. (2019). GTRD: a database on gene transcription regulation-2019 update. Nucl. Acids Res. *47*, D100–D105.

Yin, J., and Li, H. (2011). A sparse conditional Gaussian graphical model for analysis of genetical genomics data. Ann. Appl. Stat. *5*, 2630–2650.

Zhang, C., Xu, Z., Yang, S., Sun, G., Jia, L., Zheng, Z., Gu, Q., Tao, W., Cheng, T., Li, C., et al. (2020). tagHi-C reveals 3D chromatin architecture dynamics during mouse hematopoiesis. Cell Rep. *32*, 108206.

Zhang, K., Li, N., Ainsworth, R.I., and Wang, W. (2016). Systematic identification of protein combinations mediating chromatin looping. Nat. Commun. *7*, 12249.

Yin, Y., Morgunova, E., Jolma, A., Kaasinen, E., Sahu, B., Khund-Sayeed, S., Das, P.K., Kivioja, T., Dave, K., Zhong, F., et al. (2017). Impact of cytosine methylation on DNA binding specificities of human transcription factors. Science *356*, eaaj2239.

Zhang, Q. (2019). Direct estimation of differential networks under high-dimensional nonparanormal graphical models. Can. J. Stat. *48*, 187–203.

Zhang, Q. (2020). Testing differential gene networks under nonparanormal graphical models with false discovery rate control. Genes (Basel) *11*, 167.

Zhang, S., Chasman, D., Knaack, S., and Roy, S. (2019). In silico prediction of high-resolution Hi-C interaction matrices. Nat. Commun. *10*, 5449.

Zhao, Y., and Stormo, G.D. (2011). Quantitative analysis demonstrates most transcription factors require only simple models of specificity. Nat. Biotechnol. *29*, 480–483.

Zheng, R., Wan, C., Mei, S., Qin, Q., Wu, Q., Sun, H., Chen, C.H., Brown, M., Zhang, X., Meyer, C.A., et al. (2019). Cistrome data browser: expanded datasets and new tools for gene regulatory analysis. Nucl. Acids Res. *47*, D729–D735.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
K562 ChIP-seq	ENCODE; Cistrome DB; ChIP-Atlas	https://www.encodeproject.org/; http://cistrome.org/db/; https://chip-atlas.org/
K562 *in situ* Hi-C	Rao et al., 2014	GEO: GSE63525
K562 CTCF ChIA-PET	ENCODE	ENCODE: ENCSR000CAC
K562 RAD21 ChIA-PET	ENCODE	ENCODE: ENCSR000FDB
PPI data	STRING; IntAct; BioGRID; InBioMap	https://string-db.org/; https://www.ebi.ac.uk/intact/; https://thebiogrid.org/; http://www.lagelab.org/resources
K562 epigenomic marks ChIP-seq	Roadmap Epigenomics project	https://roadmapproject.org/
Human TRs and motifs	Cis-BP 1.02; JASPAR 2018; HOCOMOCO v11; etc.	http://cisbp.ccbr.utoronto.ca/; http://jaspar.genereg.net/; https://hocomoco11.autosome.ru/; see STAR Methods for details
Human tissue/cell types Hi-C	3D Genome Browser	http://3dgenome.fsm.northwestern.edu/; see Table S14 for details
Human tissue/cell types DNase-seq/ATAC-seq	ENCODE; Roadmap; GEO	See Table S14 for details
Human tissue/cell types RNA-seq	ENCODE; Roadmap; GEO	See Table S14 for details
Tissue-specific functional interactions	HumanBase	https://humanbase.flatironinstitute.org/
CRC data	dbCoRC	http://dbcorc.cam-su.org/
GWAS disease-causal variants	PICS GWAS fine-mapping study; CAUSALdb	https://pubs.broadinstitute.org/pubs/finemapping/; http://mulinlab.org/causaldb
Mouse TRs and motifs	Cis-BP 1.02; JASPAR 2018; HOCOMOCO v11; etc.	http://cisbp.ccbr.utoronto.ca/; http://jaspar.genereg.net/; https://hocomoco11.autosome.ru/; see STAR Methods for details
Mouse Hi-C	3D Genome Browser	http://3dgenome.fsm.northwestern.edu/; see Table S19 for details
Mouse ATAC-seq	GEO	See Table S19 for details
Mouse RNA-seq	GEO	See Table S19 for details
Software and algorithms		
MuSERA (v2.3)	Jalili et al., 2017	http://www.bioinformatics.deib.polimi.it/genomic_computing/MuSERA/
HiC-Pro (v2.10.0)	Servant et al., 2015	https://nservant.github.io/HiC-Pro/
Peakachu (v1.1.2)	Salameh et al., 2020	https://github.com/tariks/peakachu
bedtools (v2.27.1)	Quinlan and Hall, 2010	https://bedtools.readthedocs.io/en/latest/
R (v3.5.1)	R Core Team, 2018	https://www.r-project.org/
huge (v1.2.7)	R package	https://cran.r-project.org/web/packages/huge/index.html
ClusterONE (v1.0)	Nepusz et al., 2012	https://paccanarolab.org/cluster-one/
igraph (v1.2.4)	R package	https://igraph.org/r/
GREAT (v3.0.0)	McLean et al., 2010	http://great.stanford.edu/public/html/
GIGGLE (v0.6.3)	Layer et al., 2018	https://github.com/ryanlayer/giggle
MACRO-APE (v3.0.2)	Vorontsov et al., 2013	https://opera.autosome.ru/macroape/

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
fpc (v2.2-3)	R package	https://cran.r-project.org/web/packages/fpc/index.html
PIQ (v1.3)	Sherwood et al., 2014	https://bitbucket.org/thashim/piq-single/src/master/
vSampler (v1.2)	Huang et al., 2020	http://mulinlab.org/vsampler/
FIMO from MEME suite (v5.3.0)	Grant et al., 2011	https://meme-suite.org/meme/doc/fimo.html
tidyverse (v1.2.1)	R package	https://www.tidyverse.org/
ggnetwork (v0.5.8)	R package	https://cran.rstudio.com/web/packages/ggnetwork/index.html
3DCoop	This paper	https://github.com/mulinlab/3Dcoop

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Mulin Jun Li (mulinli@connect.hku.hk).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- This paper analyses existing, publicly available data. These accession numbers for the datasets are listed in the key resources table.
- All original code has been deposited at GitHub (https://github.com/mulinlab/3Dcoop) and is publicly available as of the date of publication. It has been listed in the key resources table.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

METHOD DETAILS

Collecting and processing data in K562

The ChIP-seq datasets in K562 were systematically collected from three databases, ENCODE (Consortium et al., 2020a), Cistrome DB (Zheng et al., 2019), and ChIP-Atlas (Oki et al., 2018). Only datasets with no severe auditing problems (that is, without orange and red audit signs) in ENCODE, no poor quality (that is, with six green QC signs in Cistrome DB website) in Cistrome DB, and peaks with *q* value less than 1e-05 in ChIP-Atlas, were collected. Based on the high-quality ChIP-seq datasets, different profiles for the same TR were collapsed by MuSERA (v2.3) (Jalili et al., 2017). The categories of human TRs were assigned according to their classification in ENCODE and literature mining: transcription factor (TF), transcription cofactor, RNA-binding protein (RBP), chromatin remodeler, nuclear enzyme, polycomb group (PcG) protein, and other factors. The K562 *in situ* Hi-C dataset was downloaded from GSE63525 (Rao et al., 2014), and processed by HiC-Pro (v2.10.0) (Servant et al., 2015) in 10-kb resolution on the human GRCh37/hg19 genome. The significant interactions from *in situ* Hi-C dataset were identified using Peakachu (v1.1.2) (Salameh et al., 2020). ChIA-PET chromatin interactions were downloaded from ENCODE (ENCSR000CAC for CTCF and ENCSR000FDB for RAD21) (Consortium et al., 2020a; Tang et al., 2015).

The 3DCoop pipeline

We optimized the graphical Lasso algorithm based on the TR network model applied in DBPnet (Zhang et al., 2016) to detect potential interplay among TRs in a high-dimensional chromatin environment (Figure S1A). To maximize the generalization capability of the model and minimize the impact of DNA-binding signal among different TRs which measured either by ChIP-seq (including CUT&RUN and CUT&Tag technologies) or by computational footprinting, we used narrow peak information instead of raw sequencing reads when constructing TR-specific contact map together with Hi-C loops. Firstly, the TR peaks were

mapped to 10-kb significant Hi-C interactions using bedtools (v2.27.1) (Quinlan and Hall, 2010). The significant interaction was assigned as the TR-specific contact map when a peak could be mapped to either end of this interaction. Based on the TR-specific contact maps, we calculated the generalized Jaccard similarity to construct the TR pair-wise correlation matrix by considering the interaction intensity of each TR-associated contact. Then the graphical Lasso algorithm (Glasso) was adopted and the precision matrix was estimated from the TR pair-wise correlation matrix to reduce the false positive rate of potential TRs dependency. The copula nonparanormal graphical model was used with the huge package (v1.2.7) in R (Liu et al., 2012). To estimate network modules, we incorporated overlapping clustering method to compute communities and maximum cliques using ClusterONE (v1.0) (Nepusz et al., 2012), which allows that single TR can involve in multiple cooperation communities. Finally, the igraph R package (v1.2.4) was used to analyse and extract TR clusters, TR maximum cliques, and TR pairs. Based on these extended features and other improvements, we implemented a new pipeline, termed 3DCoop, to identify 3D TR cooperation by leveraging genome-wide TR binding sites and high-resolution chromatin interactions. Except for the 3D mode, 3DCoop pipeline could also be used in 1D mode without 3D chromatin loops. The genome was binned into non-overlapping 10-kb windows with makewindows command from bedtools (v2.27.1), and the peaks of ChIP-seq were mapped to genome bins to get TR-specific contacts. Then the following steps were same to 3D mode.

Comparison with ChIA-PET datasets

The ChIA-PET data on limited TRs could help us briefly interrogate the derived contact maps for several classical architecture proteins such as CTCF and cohesion. The pairtopair command from bedtools (v2.27.1) with parameter "-type both" was used to overlap the contact maps and ChIA-PET data to calculate the recovering rate. To make the results more interpretable, we used the permutation to build a null distribution (baseline) of TR-specific contact maps. In K562, we permuted 14,969 contacts which equals the contacts for CTCF from aggregated contact maps of all investigated TRs. With 386 (the total number of investigated TRs in K562) permutations, we overlapped each of permuted contact maps with the CTCF ChIA-PET loops and calculated the recovering rate. Similar permutation was also done for cohesin complex component RAD21.

Measurement of TR activity in 3D chromatin

The activity of single TR intervening chromatin interactions was measured by combining its 3D binding proportion and 3D interaction proportion. The 3D binding proportion was measured by the proportion of TR binding sites within interaction regions over total TR binding sites. The 3D interaction proportion was measured by the proportion of TR-specific interactions over total Hi-C significant interactions. For comparison of the effect among TRs, these two proportions were converted to odds ratio by dividing their average proportion followed by 0-1 range normalization, respectively. Then, the normalized odds ratios were combined by product to get the final estimated effect in 3D chromatin, in which the bigger score indicates the stronger activity of TR intervening chromatin interactions.

Comparison with PPI, methods, and experimental data

We collected the human physical PPI data from four databases, STRING (Szklarczyk et al., 2019), IntAct (Orchard et al., 2014), BioGRID (Oughtred et al., 2019), and InBioMap (Li et al., 2017b). TR pairs that had the physical interactions were defined as direct interaction. TR pairs that had no physical interactions but could be connected by intermediate proteins were defined as indirect interactions. Other TR pairs were defined as not available (NA) interactions. We also collected K562 TRs cooperation data for all existing 3D-based methods (3CPET (Djekidel et al., 2015), DBPnet (Zhang et al., 2016), and HidPET (Wang et al., 2019)) and two representative 1D methods (NMF (Giannopoulou and Elemento, 2013) and TICA (Perna et al., 2018)). The comparison was performed in three steps. Firstly, for each existing method, the TRs shared with 3DCoop were extracted. Secondly, the TR pairs related to shared TRs were used and compared with existing methods to retrieve the overlapping proportion. Finally, permutation analysis was used to detect the statistical significance of overlapping by randomly selecting TR pairs from shared TRs when the number of sampled TR pairs was equal to the number of TR pairs detected by 3DCoop, and calculating the overlaps with TR pairs detected by existing methods. The *p* value was calculated as the number of random sampling datasets with overlapping proportion greater than or equal to real overlapping proportion in 10,000 permutations. Only the physical interactions from PPI databases were considered in the comparison of TR pairs with mass spectrometry-based TR pairs at the β-globin locus via *in situ* capture of chromatin interactions by biotinylated dCas9 (CAPTURE) in K562 (Liu et al., 2017).

Functional annotation of TR maximum cliques

We extracted the 10-kb genomic regions shared by all TRs in the maximum clique from TR-specific contact maps. The top5,000 regions with the highest interaction intensity were kept for functional annotation. Only the TR maximum cliques with more than 100 unique genomic regions were analysed. Functional annotation and enrichment for associated genomic regions of each maximum clique were performed using GREAT (v3.0.0) (McLean et al., 2010) with default parameters, and the whole genome was used as the background. The KEEG pathways were extracted from the GREAT results and the pathways which had the Benjamini-Hochberg adjusted *p* values less than 0.05 were considered as significantly enriched terms. The maximum cliques and pathways were hierarchically clustered with complete-linkage clustering method based on the adjusted *p* values of enrichment. Only TR maximum cliques enriched in more than 10% significant pathways and KEGG pathways received more than 10% TR maximum cliques were analysed and visualized in detail.

Colocalization analysis on epigenomic marks

DNase I hypersensitive site (DHS), H2A.Z, H3K4me1, H3K4me2, H3K4me3, H3K9ac, H3K9me1, H3K9me3, H3K27ac, H3K27me3, H3K36me3, H3K79me2, and H4K20me1 ChIP-seq peaks were obtained from the Roadmap Epigenomics project (Roadmap Epigenomics et al., 2015). The unique genomic regions for each TR maximum clique, extracted by the aforementioned method, were used to investigate potential co-localization with 13 epigenomic marks using GIGGLE (Layer et al., 2018). GIGGLE identified and ranked the significance of shared genomic loci between TR maximum clique related genomic regions and epigenome marks. The GIGGLE Fisher's two-tailed test and odds ratio with directional information were used to interpret the enrichment or depletion of epigenomic mark on TR maximum clique.

Collecting and processing of human TR motifs

We systematically collected human TRs and motifs from 16 existing TF motif resources, including Cis-BP 1.02 (Weirauch et al., 2014), JASPAR 2018 (Fornes et al., 2020), HOCOMOCO v11 (Kulakovskiy et al., 2018), TRANSFAC (Matys et al., 2006), ENCODE-motifs (Kheradpour and Kellis, 2014), HOMER motifs (Heinz et al., 2010), SwissRegulon (Pachkov et al., 2013), hPDI (Xie et al., 2010), UniPROBE (Hume et al., 2015), ReMap2 (Cheneby et al., 2018), GTRD (Yevshin et al., 2019), SELEX (Jolma et al., 2010), and four literature (Isakova et al., 2017; Jolma et al., 2013, 2015; Yin et al., 2017). To reduce the motif redundancy from different resources, we used MACRO-APE (v3.0.2) (Vorontsov et al., 2013) to select the best up to three distinct motifs by measuring the similarity of available motifs for each TR. The optimal cluster number was selected based on Calinski-Harabasz index using the pamk function from the fpc package (v2.2-3). When there was more than one motif in a cluster, the best motif was kept according to motif width ($6 \leq W \leq 18$) and the higher information content (but not outlier).

Extended 3DCoop using computational TF footprinting

For human tissue/cell types with available Hi-C data and Peakachu-based chromatin loops in the 3D Genome Browser (Wang et al., 2018), we searched tissue/cell type-matched DNase-seq and RNA-seq data from the GEO database and published resources (Consortium et al., 2020a; Roadmap Epigenomics et al., 2015), resulting in 40 tissue/cell types with necessary profiles. The DNase-seq dataset for the THP-1 cell line was not found, and its ATAC-seq dataset was used instead. The BAM file of open chromatin assay for each dataset was used and converted to genome GRCh37/hg19. For RNA-seq, the TPM (Transcripts Per Million) was calculated and used to represent the gene expression level. When these datasets were prepared, the genome-wide TR binding events were estimated based on TR motifs, TR gene expression, and DNase-seq/ATAC-seq profiles. Firstly, the TRs were filtered by their gene expression level in certain tissue/cell type, and only TRs with sufficient expression (average TPM ≥ 10) were kept. Secondly, TF footprinting sites were identified based on the DNase-seq/ATAC-seq profiles and uniformly integrated TR motifs by using PIQ (v1.3) with default parameters (Sherwood et al., 2014). The purity value of 0.7, which means that 70% of instances of motif matches could be true binding sites, were used to filter the PIQ results. For TRs with more than one motif, the match instances of all motifs were merged by bedtools (v2.27.1) to represent the genome-wide footprinting. After that, the following steps were the same as the 3DCoop pipeline for K562.

Analysis of tissue/cell type specificity of TR cooperation

We examined the tissue/cell type specificity of selected TR pairs based on the proportion of tissue/cell types containing them. TR pairs were classified into four categories according to their shared patterns.

Specifically, TR pairs that were shared in no less than 75% ($n \geq 30$) tissue/cell types were classified as tissue/cell type-shared TR pairs. TR pairs that were shared by no less than 25% but less than 75% ($10 \leq n < 30$) tissue/cell types were classified as tissue/cell type-relatively shared TR pairs. TR pairs that were shared by at least two but less than 25% tissue/cell types ($2 \leq n < 10$) were classified as tissue/cell type-relatively specific TR pairs. The remaining TR pairs which were detected only in one tissue/cell type ($n = 1$) were classified as tissue/cell type-specific TR pairs. For tissue/cell type-shared TR pairs which were shared in no less than 75% ($n \geq 30$) tissue/cell types, we constructed a shared TR cooperation network and annotated the function of certain TR cluster according to the GREAT annotation procedure described before for the investigation of TR maximum cliques in K562. To evaluate the classification of TR cooperation, we investigated the expression correlation of TR pairs in different TR cooperation categories using GTEx RNA-seq samples. Besides, the tissue-specific functional interactions from 145 tissues were downloaded from HumanBase to evaluate the classification and the tissue/cell type specificity of TR cooperation. Based on the five immune cells (HAP1, GM12878, K562, KBM7, and THP-1), the blood-specific TR cooperation network was built by extracting the TR pairs existing in at least four immune cells. The clusters which contain at least three TRs were extracted and then annotated using GREAT. The KEGG pathways which were significant in at least 50% immune cells were gathered and analyzed in detail. The immune-related pathways were extracted from KEGG pathways categories 5.1 (Immune system) and 6.6 (Immune disease).

Evaluation of tissue/cell type-specific pattern with CRC

dbCoRC contains core transcription regulatory circuitry (CRC) models from 188 human samples through computational analysis of H3K27ac ChIP-seq data (Huang et al., 2018b). We matched 40 human tissue/cell types with dbCoRC tissue/cell types, and resulted in 25 tissue/cell types. Then, we extracted TRs from CRCs and associated super-enhancer regions. For certain CRC- or super-enhancer regions-derived TR in each of 25 tissue/cell types, we counted the number of TR pairs related to this TR and calculated its proportion to total TR pairs detected by 3DCoop in each of 40 tissue/cell types. Then we averaged the proportion of all CRC- or super-enhancer regions-derived TRs to get the overlap ratio in each of 40 tissue/cell types. Finally, the 40 tissue/cell types were prioritized and ranked according to the overlap ratio in each of the 25 tissue/cell types. For intuitive comparison and visualization, the overlap ratio was normalized to z-score to represent the relevance of CRC- or super-enhancer regions-derived TRs in each of 25 tissue/cell types.

Interpretation of GWAS disease-causal variants with TR cooperation information

7,747 candidate causal variants of 39 immune and non-immune diseases/traits were collected from PICS GWAS fine-mapping study (Farh et al., 2015), and the control variants were sampled with matched allele frequency using vSampler (Huang et al., 2021). We used three computational strategies to estimate the associated TR binding in focused variant locus: motif scanning based on the TR position weight matrices using FIMO from the MEME suite (v5.3.0) (Grant et al., 2011) (naïve strategy), PIQ computational TF footprinting using both TR motif and open chromatin profile of specific tissue/cell type (Sherwood et al., 2014) (PIQ strategy), and 3Dcoop inferred TR binding event with cooperation evidence in context-specific 3D chromatin (3DCoop). Using 10,000 permutations, we tested the enrichment of the GWAS disease-causal variants on the putative binding sites derived from above three strategies for several TRs showing high activity intervening chromatin interactions. To investigate the tissue/cell type specificity of TR cooperation for explaining disease-causal variant in phenotypically relevant contexts, we further incorporated 12,738 causal variants for 54 blood-related autoimmune diseases from CAUSALdb (Wang et al., 2020) and 102 blood traits/autoimmune diseases-related TRs from literatures. Five blood-derived cell types (including K562, GM12878, KBM7, HAP1, and THP-1) and associated TR cooperation information were used to inspect the enrichment of autoimmune disease-causal variants on the putative binding sites of the 102 selected TRs, between the binding sites with blood cell-specific cooperation evidence (3DCoop-specific) and the binding sites with unrestricted cooperation evidence (3DCoop-unrestricted). The enrichment statistics for selected TRs were compared using one-tailed paired *t* test among different strategies.

Analysis of TR cooperation in mouse neural development

For mouse embryonic stem cell (ES), neural progenitor cell (NPC), and cortical neuron (CN), Peakachu-based Hi-C chromatin loops were collected from the 3D Genome Browser on mouse GRCm38/mm10 genome (Wang et al., 2018). Mouse TR motifs were collected from 10 resources, including Cis-BP 1.02 (Weirauch et al., 2014), JASPAR 2018 (Fornes et al., 2020), HOCOMOCO v11 (Kulakovskiy et al., 2018), SwissRegulon (Pachkov et al., 2013), UniPROBE (Hume et al., 2015), GTRD (Yevshin et al., 2019), HOMER motifs (Heinz et al., 2010), MEME motifs (Zhao and Stormo, 2011) and two literature (Isakova et al., 2017;

Jolma et al., 2013), and were integrated using same procedure as compiling human TR motifs. The cell state-matched ATAC-seq/RNA-seq datasets were downloaded from the GEO database. Similar to the 3DCoop pipeline used for 40 human tissue/cell types, TR clusters, TR maximum cliques, and TR pairs were estimated accordingly. Other methods for analysing TR cooperation in mouse were the same as the aforementioned descriptions.