Trans-regulation of heterochromatin underlies genetic variation in 3D genome contacts

Haley J Fortin[1,2], Anna Z Struba[1], Christopher L Baker[1,2*]

1 The Jackson Laboratory, Bar Harbor, ME 04609, USA

2 Graduate School of Biomedical Sciences, Tufts University, Boston, MA 02111, USA

* Corresponding author. Tel: +1 802 288 6365; E-mail: christopher.baker@jax.org

Authors' contributions

CLB and HJF conceptualized the study and contributed to methodology. CLB provided resources. HJF and AZS investigated the study. HJF and CLB performed data curation and formal analysis. HJF and CLB performed visualization. HJF and CLB wrote the manuscript. HJF and CLB revised and edited the manuscript. CLB performed supervision. CLB performed funding acquisition.

Competing interests

The authors declare that they have no competing interests.

Keywords

3D genome organization, embryonic stem cells, epigenetics, genetics, genomics

Abstract

Background

Genetic variation drives phenotypic diversity and disease susceptibility. *Trans*-acting genetic variation coordinates genome-wide chromatin changes, yet the molecular mechanisms underlying this regulation remain unclear. Here, we use the power of mouse genetics to investigate how genetic variation at *trans*-acting loci regulates 3D chromatin interactions.

Results

Using HiChIP to map H3K27ac-associated regulatory elements in C57BL/6J and DBA/2J embryonic stem cells (ESCs), we identified 4,962 strain-differential interactions, 71% of which overlapped chromatin accessibility quantitative trait loci (QTL), establishing chromatin interaction variation is predominantly heritable. These differential interactions showed coordinated changes in chromatin state and gene expression, with stronger interactions associated with increased accessibility and transcription. Notably, loci regulated in *trans* exhibited a unique chromatin signature where weaker interactions were enriched for H3K9me3-marked heterochromatin. Analysis of F1 hybrids revealed dominant repressive effects, consistent with heterochromatin-mediated *trans*-regulation. To causally test this mechanism, we generated reciprocal congenic mouse strains carrying a Chr13 *trans*-QTL region. Integrated multiomic profiling of congenic ESC lines demonstrated that this single locus coordinates changes in H3K9me3, H3K27ac, chromatin accessibility, and 3D contact frequency at hundreds of distal targets. Consistent with our expectations, 73-83% of differential interactions at Chr13 *trans*-QTL targets changed in the predicted direction, demonstrating that heterochromatin-mediated *trans*-regulation coordinately regulates hundreds of regulatory loci.

Conclusions

This work establishes heterochromatin formation as a mechanism by which genetic variation at *trans*-acting loci coordinates changes across chromatin accessibility, histone modifications, and 3D genome organization, providing a framework for understanding how early developmental chromatin states could generate phenotypic variation while preserving essential developmental programs.

Background

In humans, genetic variation has a profound impact on normal health and disease. The majority of loci associated with variation in complex traits are concentrated in *cis* regulatory elements (CREs), such as promoters, enhancers, and insulators, rather than coding regions [1–3]. Of interest, many of these regulatory elements are active during early development [1,4], suggesting potential developmental origins of adult-onset diseases. The activity of regulatory elements is modulated through DNA and histone modifications, chromatin accessibility, and three-dimensional (3D) genome architecture [5–8]. Given that 99% of the human genome is non-coding, understanding how genetic variation influences CREs is critical for deciphering genome function [9,10], normal development, and, when these processes goes awry, the etiology of developmental disorders and disease.

Sequence variation can alter gene expression locally in *cis*, i.e., on the same chromosome (Chr), or distally, on another chromosome, in *trans*. *Cis*-regulatory variants affect nearby genes through changes in local CREs, whereas *trans*-regulatory effects are mediated by diffusible factors, often transcription factors or chromatin regulators, that can act on many CREs throughout the genome [11]. There is a growing interest in addressing how *trans*-regulation impacts complex traits and disease [12–18]. While effect sizes, tissue-specificity, and multiple testing burden similarly challenge both human and mouse mapping studies [15,16], the reduced power to detect *trans* quantitative trail loci (QTL) in human cohorts is largely driven by allele frequency. Rare variants often exhibit larger effects but require very large sample sizes for detection, whereas common variants, that typically exhibit weaker effects, can be mapped with smaller cohorts [19–21]. In contrast, mapping populations such as the mouse Diversity Outbred (DO) and BXD recombinant inbred panel effectively make rare variants common, enabling more powerful mapping of *trans*-QTL [22–25]. Using the BXD recombinant inbred panel, whose parental strains C57BL/6J (B6) and DBA/2J (D2) differ at over 5,000,000 single nucleotide variants (SNVs) [26], we previously identified both *cis* and *trans* quantitative trait loci (QTL) regulating chromatin accessibility, histone modifications, and gene expression in mouse embryonic stem cells (mESCs) and male germ cells [27,28]. In ESCs, we mapped several *trans*-QTL hotspots on Chrs 4, 5, 7, 12, and 13 that regulate coordinated changes in both chromatin accessibility and gene expression, impacting gene networks important for pluripotency and cell fate determination [28]. Interestingly, genes associated with development and disease have also been independently mapped to closely overlapping intervals within the Chr13 region [27,29–33]. This

69 Chr13 region contains a cluster of KRAB zinc-finger proteins (KZFPs), known transcriptional repressors that act
70 in *trans* to establish heterochromatin. Despite success in identifying *trans*-QTL in both mice and humans
71 [27,28,34–36], their molecular mechanisms remain poorly understood.

72 Chromosome conformation capture (3C) technologies [37–42] have revealed regulatory roles for 3D
73 genome organization in gene regulation [43–48], from regulatory loops connecting CREs to their target genes
74 [49–52] to topologically associated domains [53–55]. Local genetic variation alters 3D genome structure in *cis*,
75 helping determine how SNVs are associated with disease risk [56] and how structural variants (SVs) impact
76 development of disorders ranging from rheumatoid arthritis and sex reversal to limb malformations [57–59].
77 Systems genetic approaches have also identified how genetic variation among human lymphoblastoid cell lines
78 influences diverse features of chromatin state and 3D genome architecture [60]. While these studies have
79 established how *cis*-acting genetic variants disrupt 3D chromatin organization, the mechanisms by which *trans*-
80 acting genetic variation influences 3D genome architecture remain largely unexplored. Even when *trans*-QTL
81 are detected, the molecular mechanisms by which single genetic loci coordinate chromatin state changes at
82 hundreds of distal genomic regions are largely unknown. Understanding these mechanisms is essential for
83 interpreting how genetic variation influences complex traits and disease susceptibility, and the contribution of
84 variation to early developmental regulation.

85 Here, we investigated molecular mechanisms underlying *trans*-regulation of 3D genome organization.
86 Using HiChIP to map H3K27ac-associated chromatin interactions in B6 and D2 mouse embryonic stem cells,
87 we identified that contact anchors of differential interactions substantially overlapped QTL associated with
88 chromatin accessibility (caQTL) and gene expression (eQTL). Integrated epigenomic profiling revealed that
89 heterochromatin formation distinguished *trans* from *cis* regulatory mechanisms. To causally test this mechanism,
90 we performed reciprocal genetic perturbations by exchanging a *trans*-QTL region between otherwise isogenic
91 B6 and D2 genetic backgrounds, enabling independent validation of predicted chromatin changes in both
92 directions. Multiomic profiling of these congenic mESCs validated that this *trans*-acting locus coordinates
93 chromatin state at hundreds of distal regulatory loci with high concordance. Together, these findings establish
94 *trans*-regulation through heterochromatin formation as a key mechanism that coordinates changes in the
95 regulatory landscape.

Results

Genetic variation alters chromatin interactions and coordinates gene expression

To investigate how genetic variation impacts 3D genome architecture at CREs, we performed HiChIP [39] on mESC lines derived from B6 and D2 embryos. HiChIP was performed using the histone modification H3K27ac to enrich for 3D interactions at active enhancers. In total, this approach identified a combined set of 90,098 high-confidence interactions across both genetic backgrounds, defined as those spanning greater than 5 kb, a false discovery rate (FDR) < 0.01, and supported by at least four paired-end tag (PET) counts in at least two samples. Independently derived replicate lines from the same strain showed high concordance, while between-strain comparisons revealed reduced concordance, indicating that genetic background, rather than line-specific variation, drives chromatin organization differences (Fig. 1a and Additional file 1: Fig. S1a). These interactions predominantly connected enhancers to promoters (e-p, 84%), or enhancers to other enhancers (e-e, 14%). To understand the functional impact of 3D contacts, we integrated RNA-seq data previously collected from the same ESCs [28]. We observed that gene expression at connected promoters increases with increasing number of chromatin interactions (All pairwise comparisons Tukey p < 2e-16; Fig. 1b), consistent with previous findings [61,62]. Furthermore, genes annotated for “stem cell population maintenance” (MGI GO:0019827) are enriched among the upper quartile of connected genes (Fisher’s Exact Test p < 0.032), consistent with the observation that highly connected genes represent cell type–specific function [63,64]. Among the high-confidence interactions, 4,962 (5.5%) were significantly different between B6 and D2 ESCs (Benjamini-Hochberg corrected p < 0.2; Fig. 1c). To validate HiChIP results, we used 3C-qPCR [65] on three differential interactions (DIs) and confirmed the directional trends of strain-specific interaction patterns (Additional file 1: Fig. S2-4).

Differential chromatin interactions between B6 and D2 ESCs correspond with coordinated changes in chromatin accessibility and gene expression. The *Zfp951* locus provides an example of coordinated regulation. Specifically, B6 shows higher PET counts at both e-e (Interaction 1) and e-p (Interaction 2) interactions compared to D2 (Fig. 1d, e). Increased 3D contact in B6 is consistent with higher epigenetic marks associated with increased CRE activity (e.g., chromatin accessibility and H3K4me3 enrichment) in the loop anchor region (Fig. 1D yellow highlight) and higher *Zfp951* expression in B6 (Fig. 1f). This pattern suggests a shared regulatory mechanism coordinating changes in CRE activity, 3D contacts, and gene expression. Genome-wide, DIs showed coordinated strain-specific changes in both chromatin accessibility and interaction strength (Fig. 1g). The strain-

124 specific bias seen for *Zfp951* expression generally holds for all DIs (Fig. 1h); chromatin interactions that are
125 stronger in either B6 or D2 coincide with higher expression of anchor-connected genes. In contrast, genes
126 connected to non-DIs, such as the housekeeping gene *Gapdh*, showed minimal expression bias ($\log_2$(B6/D2) =
127 0.05416) (Fig. 1h, Additional file 1: Fig. S1b, c). Together, these data demonstrate that genetic variation
128 coordinately modulates CRE activity, 3D interaction strength, and expression of interacting genes, but leaves
129 open the question of the mechanism of regulation.

130 **Genetic variation regulates chromatin interactions through both *cis*- and *trans*-acting mechanisms**

131 To test whether coordinated changes in chromatin state and gene expression reflect shared genetic
132 control, we determined if DI anchors are enriched for genetically-regulated CREs by examining overlap with
133 previously identified caQTL targets [28]. For instance, both gene expression of *Zfp951* and accessibility of an
134 upstream CRE (Chr5:104,727,903 bp) were mapped to the same Chr5 locus (Fig. 2a, b, c), with B6 haplotypes
135 conferring higher activity. In total, five interaction anchors contacting *Zfp951* overlap eight *cis*-caQTL all showing
136 higher B6 accessibility [28], demonstrating that individual anchors can harbor multiple caQTL. This pattern
137 extends genome-wide; 71.7% of DI anchors are enriched for caQTL targets at one (42.6%) or both (29.1%)
138 anchors (QTL at one and both anchors $p < 0.001$, Fig. 2d; Additional file 1: Fig. S1d, e), indicating that contact
139 frequency at anchors for most DIs are under genetic regulation. Further, these data suggest that genetic
140 regulation coordinates both chromatin accessibility and 3D genome structure.

141 Given QTL enrichment at DI anchors, we tested whether parental interaction strength in B6 and D2
142 correlates with QTL effect sizes measured across the BXD panel [28]. Interaction strength showed moderate but
143 highly significant positive correlation with chromatin accessibility at *cis* (Pearson r = 0.41, p = 1.4e-105) and *trans*
144 (r = 0.45, p = 7.1e-100) caQTL targets (Fig. 2e, f). We further explored the link between 3D contacts and gene
145 regulation (Fig. 2g, h) and identified a similar correlation at *cis* (r = 0.34, p = 2.5e-30) and *trans* (r = 0.15, p =
146 4.1e-3) eQTL targets. Beyond correlation magnitude, directional concordance provides stronger evidence for
147 coordinated regulation; loci where the B6 allele shows higher interaction strength might be expected to exhibit
148 B6-biased molecular phenotypes. Indeed, at the categorical level, we found significant allele concordance
149 between interaction strength and chromatin accessibility or gene expression at both *cis*- and *trans*-QTL targets
150 (Fig. 2i, Pearson's Chi-squared test all $p < 6.7e-6$). The high allelic concordance between 3D architecture,

151 chromatin accessibility, and gene expression at both *cis*- and *trans*-QTL suggests shared regulatory mechanisms
152 coordinate these molecular layers. Moreover, the presence of *trans*-acting effects indicates that diffusible
153 chromatin regulators mediate long-range genetic control of chromatin organization.

154 **Heterochromatin formation regulates *trans*-acting chromatin interactions**

155 While both *cis* and *trans*-QTL showed coordinated effects on chromatin accessibility and 3D interactions,
156 the molecular mechanism distinguishing these regulatory modes remained unclear. We hypothesized that
157 distinct mechanisms underlie *trans* and *cis* regulation. To test this, we examined enrichment of DNA-binding
158 factors at interaction anchors that overlap *trans*-regulated loci (i.e., *trans*-targets). Supporting our previous
159 observations [28], we identified TRIM28 (KAP1) as the most highly enriched factor at *trans*-targets (Fig. 3a).
160 Given TRIM28's role in establishing heterochromatin through H3K9me3 deposition [66,67], we hypothesized that
161 *trans*-targets would show enrichment for repressive chromatin marks. Specifically, at loci with DIs, we predicted
162 H3K9me3 enrichment in the strain with weaker contacts and H3K27ac enrichment in the strain with stronger
163 contacts. To test this, we performed H3K9me3 and H3K27ac ChIP-seq in independently derived mESC lines
164 from three B6 and three D2 mice (Additional file 1: Fig. S5). At a putative CRE downstream from *Sox13* (Fig. 3b,
165 c, d), D2 shows weaker chromatin interactions than B6 and increased H3K9me3 compared to B6 at the anchor
166 overlapping a *trans*-target (Fig. 3b, yellow highlight). Conversely, the same *trans*-target showed increased
167 H3K27ac in B6 compared to D2. To broadly capture this trend, we quantified the level of both H3K27ac and
168 H3K9me3 at both *cis* and *trans*-targets and correlated strain-specific changes in histone modifications with
169 changes in chromatin accessibility (Fig. 3e, f, g, h). When considering H3K27ac, we identified a strong positive
170 correlation at *cis*- (Pearson $r = 0.83$, $p < 2.2e-16$) and *trans*- ($r = 0.91$, $p < 2.2e-16$) caQTL targets, indicating that
171 both signify active chromatin. The key distinction in chromatin state is with H3K9me3. While *cis*-targets showed
172 little correlation between H3K9me3 and accessibility (Fig. 3g), *trans*-targets showed a strong negative correlation
173 ($r = -0.66$, $p = 5.8e-143$; Fig. 3h), indicating that heterochromatin formation specifically characterizes *trans*-
174 regulation. Critically, *trans*-targets were significantly more likely than *cis*-targets to show discordant chromatin
175 states – defined as opposing changes in accessibility and H3K9me3 enrichment ($p < 2.2e-16$, odds ratio = 20.49,
176 logistic regression; Fig. 3i), establishing heterochromatin formation as a defining feature of *trans*-regulation.
177 Together, these data support a molecular mechanism in which genetic variation in *trans*-acting factors drives

178 TRIM28-mediated heterochromatin formation via H3K9me3 deposition, restricting chromatin accessibility and
179 weakening 3D contacts at distal loci.

180 **Chr13 *trans*-QTL coordinates chromatin accessibility and gene expression in an allele-specific manner**

181 To determine the effect of *trans*-regulation on coordinated chromatin regulation and gene expression, we
182 focused on a compound *trans*-QTL on Chr 13 that regulated the expression of 59 genes (eGenes) and 2,033
183 putative CREs (*trans*-targets) with LOD $\geq$ 5 [28]. Spatial coordination of *trans*-targets and eGenes was assessed
184 for proximity enrichment using a modified Region Associated Differentially expressed gene approach [68], testing
185 whether chromatin accessibility peaks cluster near transcriptionally regulated genes. Chr13 *trans*-targets
186 showed higher enrichment near Chr13 eGenes compared to all other *trans*-QTL eGenes (Fig. 4a), indicating
187 distal *trans*-regulation drives coordinated changes in chromatin accessibility and gene expression at co-localized
188 target loci. Critically, this spatial coordination was directionally concordant with QTL effects; chromatin peaks
189 with higher accessibility associated with the B6 allele at the QTL are specifically enriched near eGenes with
190 higher expression when the genotype at the QTL is B6 (Fig. 4b, left), and the reciprocal pattern held for D2 (Fig.
191 4b, right). These analyses demonstrate that independently measured chromatin accessibility and gene
192 expression are spatially coordinated in an allele-specific manner, supporting *cis*-propagation of *trans*-QTL
193 through localized chromatin remodeling.

194 To determine the nature of the QTL effect on chromatin accessibility, we performed ATAC-seq on three
195 independently derived mESCs from (B6xD2)F1 hybrids. Focusing on high-confidence caQTL (LOD > 8), *trans*-
196 caQTL showed dominant repression, where F1 hybrid mESCs resembled the parent with lower accessibility
197 rather than intermediate (additive) values. This is exemplified by the Chr13 *trans*-target at Chr14:122776397,
198 where the F1 showed the same (dominant) repressed accessibility as the D2 parent (Fig. 4c, left). In contrast, a
199 *cis*-QTL target at Chr5:104727903 showed additive inheritance with F1 chromatin accessibility intermediate
200 between the parent lines (Fig. 4c, right). This distinction between *trans* and *cis* effects extended genome-wide;
201 *trans*-caQTL showed dominant repression (median r = -0.95), while *cis*-caQTL showed near-additive inheritance
202 (median r = -0.04; Wilcoxon $p < 2.2e-16$, effect size r = 0.61; Fig. 4d). Chr13 *trans*-caQTL targets specifically
203 showed strong dominant repression (median = -1.04), with the D2 allele at the Chr13 QTL associated with
204 reduced accessibility at most targets (84%, N = 1,709) (Fig. 4e).

205 The dominant repressive inheritance pattern supports a model in which the Chr13 QTL encodes one or
206 more diffusible repressors that establish heterochromatin at *trans*-targets. To causally test whether this locus
207 coordinates heterochromatin formation and 3D architecture, we generated reciprocal congenic mouse strains
208 that exchange the Chr13 *trans*-QTL. The first strain, B6.D2-(rs30079041-rs51756099)/Bakr (hereafter referred
209 to as B6-13D2), was bred to contain the D2 Chr13 congenic region (Chr13:52,765,995-71,064,928) on an
210 otherwise B6 background. Conversely, D2.B6-(rs29224292-rs51756099)/Bakr (hereafter referred to as D2-
211 13B6) was bred to carry the B6 Chr13 congenic region (Chr13:47,066,720-71,064,928) on an otherwise D2
212 background (Fig. 4f). Together, these congenic lines provide ideal resources to investigate the impact of the
213 Chr13 *trans*-QTL on heterochromatin formation.

214 **Congenic strains confirm Chr13 *trans*-QTL drives heterochromatin formation at *trans*-targets**

215 To test whether the Chr13 QTL drives *trans*-regulation of heterochromatin, we tested predictions based
216 on QTL effects of BXDs on reciprocal mESCs from congenic strains. If Chr13 is causal, swapping the QTL
217 region should alter chromatin state according to the genotype introduced. Specifically, B6-13D2 should
218 resemble D2 at *trans*-targets, while D2-13B6 should resemble B6. We tested these predictions using
219 H3K9me3 and H3K27ac ChIP-seq, and ATAC-seq in B6, D2, B6-13D2, and D2-13B6 mESCs (Additional file 1:
220 Fig. S6).

221 If repression is mediated by H3K9me3-marked heterochromatin, we predicted *trans*-targets should show
222 increased H3K9me3 in B6-13D2 (with D2 Chr13) compared to B6. Strikingly, 77% of Chr13 *trans*-caQTL targets
223 showed the predicted H3K9me3 increase in B6-13D2 (Fig. 5a, g, left; exact binomial test $p < 2.2e-16$).
224 Reciprocally, 89% showed decreased H3K9me3 in D2-13B6 (with B6 Chr13) compared to D2 (Fig. 5b, left, g; p
225 $< 2.2e-16$). Of the remaining targets (324 loci predicted to be repressed when the allele at the QTL is B6), 69%
226 showed the inverse pattern where H3K9me3 increased in D2-13B6 and decreased in B6-13D2 (Fig. 5b, c, right,
227 g). Consistent with H3K9me3-mediated repression, active chromatin marks showed the reciprocal pattern as
228 predicted. Targets gaining H3K9me3 in B6-13D2 showed reduced chromatin accessibility (77% of predicted
229 decreases validated; Fig. 5c, g) and reduced H3K27ac (94% validated; Fig. 5e, g). Similarly, targets losing
230 H3K9me3 in D2-13B6 gained accessibility and H3K27ac (Fig. 5d, f, g). Further, *trans*-targets that changed in the
231 predicted direction showed significantly larger effect sizes than those that did not (Fig. 5h), indicating non-

232 validated targets likely reflect measurement noise rather than biological exceptions. These reciprocal epigenomic
233 changes, validated through two independent genetic perturbations across three chromatin features, directly
234 demonstrate that Chr13 causally regulates hundreds of distal *trans*-targets through heterochromatin formation.

235 **Heterochromatin formation alters 3D contacts at *trans*-QTL target sites**

236 To validate that Chr13-driven heterochromatin formation alters 3D interactions at *trans*-targets, we
237 repeated H3K27ac HiChIP in B6 and D2, while including B6-13D2 and D2-13B6 mESCs (two technical replicates
238 per strain; Additional file 1: Fig. S7). We hypothesized that introducing the D2 Chr13 region in B6-13D2 would
239 reduce chromatin interaction at *trans*-targets (compared to B6), by increasing heterochromatin. Further,
240 introducing the B6 Chr13 region into D2-13B6 would strengthen interactions (compared to D2), by reducing
241 heterochromatin. Consistent with our hypothesis, at the *Litaf* locus we observed increased interactions in D2-
242 13B6 compared to D2 (Fig. 6a, b, Interactions 1 and 2 $p < 0.027$). Reciprocally, at the same locus we observe
243 weaker interactions in B6-13D2 compared to B6 (Fig 6a, b; Interactions 1 and 2, $p < 1.51e-3$). This is particularly
244 noteworthy because *Litaf* is a Chr13 eQTL target, highlighting a potential functional link between chromatin
245 architecture and gene regulation.

246 To test if heterochromatin formation underlies these changes in chromatin interactions, we examined
247 H3K9me3 enrichment at the interaction anchors. Critically, H3K9me3 showed the reciprocal pattern to 3D
248 contacts. At *Litaf*, D2-13B6 exhibited significantly lower H3K9me3 than D2 at the Chr13 *trans*-target; and B6-
249 13D2 showed increased H3K9me3 compared to B6 (Fig. 6c; XD $p = 4.81e-8$, BQ $p = 2.98e-15$). These examples
250 demonstrate that the Chr13 QTL drives heterochromatin formation at interaction anchors, coincident with altered
251 3D interaction strength. Further, in agreement with a heterochromatin-mediated mechanism, H3K27ac showed
252 the opposite patterns, with D2-13B6 exhibiting higher levels at the same anchor compared to D2 at *Litaf* (Fig.
253 6d; $p < 9.24-103$), while B6-13D2 showed lower H3K27ac than B6 (Fig 6d; BQ $p = 4.54e-119$). Consistent DI
254 patterns and opposing H3K9me3/H3K27ac profiles are observed at additional *trans*-regulated loci, including
255 *Zfp516* and *Fam241a* (Additional file 1: Fig. S8-9). Together these data confirm that Chr13 regulates
256 heterochromatin formation and 3D organization at these *trans*-targets.

257 To assess whether these examples reflect genome-wide patterns, we tested whether genetic differences
258 identified between parental strains (B6 and D2) predict changes in congenic strains. Remarkably, 73-83% of

259 differential interactions changed in the predicted direction (Fig. 7a, b, c; Exact binomial test: all $p < 1.5e-8$)
260 Furthermore, interactions that change contact frequency as predicted generally showed larger effect sizes than
261 those that did not (Fig. 7d), indicating non-validated targets may reflect measurement noise. These validation
262 rates, achieved across two independent genetic perturbations and testing reciprocal predictions, demonstrate
263 that Chr13 causally regulates chromatin state at hundreds of distal loci. The consistent validation across
264 H3K9me3, H3K27ac, chromatin accessibility, and 3D contact frequency establishes *trans*-regulation of
265 heterochromatin as a mechanism for coordinated regulation of the epigenome (Fig. 7e).

266 **Discussion**

267 Here, we establish heterochromatin formation as the molecular mechanism distinguishing *trans* from *cis*
268 regulatory effects on 3D genome organization in mESCs. *Trans*-regulated regions of the genome show a 20-fold
269 enrichment for discordant H3K9me3 and chromatin accessibility, providing a unique epigenetic signature
270 compared to *cis*-QTL. This mechanism is supported genetically by the dominant repressive pattern observed in
271 F1 hybrids. Critically, reciprocal genetic perturbations demonstrated that a single locus causally coordinates
272 chromatin state at hundreds of distal regulatory elements with high concordance across molecular assays.
273 Validation rates ranged from 73-83% for 3D contact frequency to 89-94% for H3K9me3 and H3K27ac,
274 demonstrating that the coordinated mechanism is robust across the epigenome. Together, these findings explain
275 how *trans*-acting genetic variation achieves coordinated regulation across molecular layers through targeted
276 heterochromatin deposition.

277 Consistent with previous observations [60,62,69,70], we find that 3D genome interactions are positively
278 correlated with gene expression and active histone modifications, with contact frequency at individual loci
279 associated with increased transcription. Prior work in human induced pluripotent stem cells (iPSCs)
280 demonstrated that genetic variation affecting contact propensity results in coordinated changes in H3K27ac
281 [71,72]; however, these studies did not distinguish *cis* from *trans* regulation. Our findings extend beyond
282 correlations by establishing causal relationships through reciprocal genetic perturbations and identifying
283 heterochromatin formation as a defining feature of *trans*-regulation. Notably, the 20-25% of cases where
284 increased contacts did not correspond to increased expression suggest additional regulatory mechanisms
285 beyond direct connection between contact frequency and expression and warrant future investigation.

286 The functional genomic, genetic, and positional evidence continue to point to variation in genes that
287 encode KRAB Zinc-Finger Proteins (KZFPs) as probable molecular basis for *trans*-regulation. Firstly, KZFPs
288 provide DNA-binding specificity to recruit TRIM28, and chromatin modifying enzymes, to form heterochromatin
289 [66,73–77]. Consistent with this, TRIM28 is enriched at *trans*-QTL targets and a cluster of KZFPs is found in the
290 Chr13 QTL [78]. Second, the unique association of *trans*-QTL with H3K9me3 and their dominant repressive
291 effect in F1 hybrids indicates a diffusible silencing factor [27,28]. Third, KZFPs are highly variable within species,
292 indicating a source of rapid genetic variation [78,79]. The Chr13 region represents a compound QTL containing
293 structurally variant KZFP genes, each encoding distinct DNA-binding specificities; this genetic architecture

294 explains how a single locus coordinates changes at hundreds of *trans*-targets. Fourth, KZFPs have been
295 demonstrated to impact transcriptional repression through heterochromatin spreading from their DNA binding
296 sites [80], supporting the regional *cis*-acting effects we see on enhancer-promoter contacts. Lastly, the epigenetic
297 engineering approach of CRISPR interference (CRISPRi) [81] leverages a nuclease-dead Cas9 fused to a KRAB
298 domain to induce heterochromatin and silence target genomic loci, directly demonstrating *trans*-regulation of
299 heterochromatin. Building on this, recent work used CRISPRi to show that targeted H3K9me3 deposition led to
300 a reduction in 3D contacts between enhancers and promoters [69], demonstrating direct effect of
301 heterochromatin on 3D contacts in an inducible system.

302 We propose that *trans*-regulation of heterochromatin formation represents an evolutionarily
303 advantageous mechanism for generating regulatory variation without disrupting core transcription factor (TF)
304 networks. Classical models of *trans*-regulation invoke changes in TF expression or sequence, which would
305 coordinately affect all targets genome-wide. However, developmental TFs and their core gene regulatory
306 networks are under strong purifying selection due to their essential roles in lineage specification; and large
307 variation in these factors likely results in severe phenotypic consequences. In contrast, KZFP DNA-binding
308 domains evolve rapidly through structural variation [82,83], and specifically at Chromosome 13 in mice [78],
309 potentially enabling population-level variation in which genomic regions are differentially targeted for
310 heterochromatin formation based on the repertoire of zinc-finger domains present in an individual. Recent cross-
311 species comparisons between human and macaque cells demonstrate that *trans*-regulatory divergence is
312 specifically enriched at transposable elements bearing KZFP footprints [84], providing evolutionary support that
313 heterochromatin variation drives regulatory divergence between populations while preserving core
314 developmental networks. Further, heterochromatin acts as a barrier to pioneer transcription factor binding during
315 cellular reprogramming [85,86], and depletion of a specific KZFP reduced this barrier and enhanced
316 reprograming efficiency [87]. The coordinated regulation we observe here reveals a powerful and
317 underappreciated layer of genome control where a single locus affects heterochromatin deposition and therefore
318 chromatin accessibility and 3D contacts at hundreds of regulatory elements. By allowing population-level
319 variation in heterochromatin targeting rather than disrupting essential TF function, this mechanism may provide
320 standing regulatory variation necessary for evolutionary adaptation while maintaining developmental robustness.

The developmental implications of heterochromatin-mediated *trans*-regulation established in early pluripotency may extend to adult phenotypes. Heterochromatin can act as a barrier to lineage specification [85]; and variation in early heterochromatin formation may be propagated, creating a persistent epigenetic state that influences lineage commitment and cellular function. Supporting functional consequences, the Chr13 region investigated here has been associated with multiple developmental phenotypes in mice including craniofacial development, limb malformation, lupus susceptibility, meiotic recombination frequency, and imprint stability [27,29–33]. The mechanism we establish in mouse ESCs appears conserved in human biology. Inter-individual variation in heterochromatin landscapes across diverse human iPSC lines correlates with differentiation propensity for early neural fates versus other lineages [88], consistent with the model that heterochromatin variation creates tunable barriers affecting lineage-specific transcription factor access. This suggests population-level variation in heterochromatin may contribute to phenotypic diversity in humans through the same mechanism. Further, KZFPs show dynamic expression beyond pluripotency, specifically upregulated in adult immune cells and neurons [83]. Consistent with functional relevance, genomic regions containing KZFPs in humans contain signatures of recent selection and are associated with adult disease [89]. Together, these observations suggest that early variation in heterochromatin establishment may act to constrain developmental outcomes and contribute to phenotypic variation in complex traits.

Our study provides strong genetic and epigenetic evidence that heterochromatin formation mediates *trans*-regulation of 3D genome organization and chromatin state, but several limitations remain. First, all experiments were performed in mESCs, and whether similar mechanisms operate in differentiated tissues, other species, or in vivo is unknown. Second, the Chr13 congenic interval spans a large region and contains multiple KZFPs, making it unclear whether *trans*-regulation reflects the combined effects of this region or a single causal gene. Fine-mapping, and targeted perturbations, such as overexpression or CRISPR-based manipulation, will be needed to identify the causal gene(s). Finally, while reciprocal congenic experiments validate chromatin changes, functional consequences for pluripotency, lineage specification, and organismal phenotypes remain to be tested; and extension to human systems will be essential to establish relevance to complex traits and disease.

Conclusions

We establish heterochromatin formation as a defining chromatin signature distinguishing *trans* from *cis* regulatory mechanisms in mouse embryonic stem cells. Through reciprocal genetic perturbations, we

349 demonstrate that genetic variation at a single *trans*-acting locus regulates chromatin accessibility, histone
350 modifications, and 3D interactions at hundreds of distal regulatory elements. This work provides a framework for
351 understanding how *trans*-regulation operates mechanistically through targeted heterochromatin formation to
352 reshape regulatory landscapes and suggests that genetic variation in early developmental chromatin states may
353 contribute to phenotypic diversity in complex traits and disease susceptibility.

354 **Methods**

355 **Mice**

356 All mice were obtained from The Jackson Laboratory (Bar Harbor, ME) including C57BL/6J (JR# 000664) and
357 DBA/2J (JR# 000671) parental inbred strains. B6 and D2 parents were crossed to the F3 generation prior to
358 marker-based genotyping and subsequent backcrossing for a minimum of ten generations for form the B6.D2-
359 (rs30079041-rs51756099)/Bakr (JR# 37111) and D2.B6-(rs29224292-rs51756099)/Bakr (JR# 37112) congenic
360 lines. Congenic genotypes were confirmed using MiniMUGA array [90] (Neogen) for which the background
361 analysis report is available for a representative mouse (B6-13D2: Additional file 2; D2-13B6: Additional file 3).
362 All animal experiments were approved by the Animal Care and Use Committee of The Jackson Laboratory
363 (Summary #04008). Congenic lines are available upon request.

364 **Derivation and maintenance of mouse embryonic stem cells**

365 Derivation of mouse embryonic stem cells for B6-13D2 and D2-13B6 strains were performed, after
366 backcrossing for a minimum of 10 generations, using protocols outlined in Czechanski et al. [91]. For all
367 cultures, mouse ESCs were seeded with irradiated mouse embryonic fibroblasts (MEFs) from B6 animals and
368 cultured in DMEM high glucose base medium supplemented with 15% fetal bovine serum (FBS, Lonza, cat.
369 no. 14-501F lot no. 0000217266), 1X Pen/Strep, 2 mM GlutaMAX, 1 mM sodium pyruvate, 0.1 mM MEM-
370 NEAA, 0.1 mM 2-mercaptoethanol, 103 IU LIF, 1 μM PD0325901, and 3 μM CHIR99021. Mouse ESCs were
371 expanded for 2–3 days to reach ~70% confluency for molecular assays. For functional genomic assays, mouse
372 ESCs were enzymatically disassociated into single cell suspensions using trypsin and MEFs were depleted by
373 incubation on gelatine coated plates for 20 minutes to allow to settle. All cell lines tested negative for
374 mycoplasma and are available upon request. Genotyping was performed to determine zygosity of congenic
375 region and sex of cell lines. Primers for KASP genotyping used SNVs within the congenic region to distinguish
376 between B6 and D2 backgrounds. All primers used for genotyping are listed in Additional file 4, Table S1.

377 **ATAC-seq sample preparation and sequencing**

378 ATAC-seq experiments were performed, using two or three independent cultures from the same cell line per
379 strain, following the FAST-ATAC method as described [92], using 100,000 cells. DNA quality and quantity was
380 determined using High Sensitivity D5000 ScreenTape (Agilent Technologies). Libraries were sequenced to a
381 read depth of about 40 million reads per sample on an Illumina NovaSeq X Plus with a 150 bp read length.

382 **Chromatin Immunoprecipitation (ChIP) sample preparation and sequencing**

383 ChIP-seq experiments performed with B6 and D2 used three independently derived cell lines from each strain.
384 Validation ChIP-seq experiments performed with B6, D2, B6-13D2, and D2-13B6 used two or three
385 independent cultures from the same cell line for each strain. For each ESC ChIP-seq library, 3 million cells
386 were harvested and fixed as previously described [27]. Cell lysis, chromatin fragmentation, and immuno-
387 precipitation were performed as described [93], with minor modifications. For H3K27ac ChIP, sodium butyrate
388 was added to hypotonic lysis buffer, MNase buffer, RIPA buffer (post-overnight antibody conjugation), RIPA
389 wash buffer, and TE to inhibit histone deacetylase activity. Immunoprecipitation was performed using
390 antibodies against H3K27ac (12 µL, abcam, ab4729, lot# GR312658-1; Cell Signaling Technology, 8173, lot#
391 9) and H3K9me3 (6 µL, Active Motif, 39161, lot# 09919003). ChIP-seq libraries were constructed using the
392 KAPA HyperPrep Kit (Roche Sequencing and Life Science). Quality and concentration of libraries were
393 assessed using the High Sensitivity D5000 ScreenTape (Agilent Technologies) and KAPA Library
394 Quantification Kit (Roche Sequencing and Life Sciences). Libraries were sequenced to a read depth of about
395 40 million reads per sample (75 bp) on an Illumina NextSeq or NovaSeq X Plus.

396 **HiChIP sample preparation and sequencing**

397 *Restriction enzyme based HiChIP*

398 HiChIP experiments were performed as described [39], using two samples from independently derived cell
399 lines per strain. We used 10 million cells, DpnII (NEB, R05443), and the following 3C modifications [94] to
400 improve digestion, incorporation, ligation, and removal of dangling ends. Specifically, 1) biotin incorporation
401 was performed at 23°C for 4 hours with interval shaking (10 sec at 900 rpm, 5 min off). 2) Ligation was
402 performed at 16°C for 4 hours with interval shaking (1 min at 500 rpm, 4 min off). 3) Samples were sheared
403 using a Diagenode Biorupter with the following parameters: tube size = 1.5 mL, time on = 30 sec, time off = 30
404 sec, cycle # = 8. 4) A total of 3 µL of H3K27ac antibody (Abcam #ab4729) was added to the sample and
405 incubated at 4°C overnight with rotation. 5) After ChIP washes, samples were resuspended in 20 µL removal of
406 biotin-dATP master mix: 5 µL 10X NEB Buffer 2.1, 0.125 µL of dATP and dGTP at 10 mM each, 5 µL T4 DNA
407 polymerase (NEB, M0203), and 39.8 µL distilled water. Samples were incubated at 20°C for 4 hours with
408 interval shaking (1 min at 400 rpm, 4 min off). Lastly, size selection of libraries was performed using a two-
409 sided selection with AMPureXP beads. Briefly, 22.5 µL of beads were added to the sample to capture

410 fragments less than 700 bp. Supernatant was transferred to a new tube and 13.5 µL of beads were added to
411 the supernatant sample to capture fragments greater than 300 bp. Libraries were eluted in 50 µL 10 mM Tris
412 pH 8.0. Libraries were quantified using qPCR against Illumina primers and then paired end sequenced with
413 read length of 75 bp to a sequencing depth of 200 million reads per sample over two Illumina NextSeq runs.
414 *Dovetail HiChIP MNase Kit*
415 The Dovetail HiChIP MNase Kit was used for HiChIP for validation experiments repeating B6, D2, and
416 including B6-13D2 and D2-13B6. HiChIP was performed using two samples from independent cultures from
417 the same cell line per strain. The User Guide [95] (version 2.0) was used to process samples of 10 million
418 cells. H3K27ac (Cell Signaling Technology, 8173, lot #9) was used for immunoprecipitation. Libraries were
419 paired end sequenced with a read length of 150 bp to a sequencing depth of 300 million reads per sample
420 (over two flow cell lanes) on an Illumina NovaSeq X Plus.
421 **ATAC-seq and ChIP-seq data processing**
422 Illumina adaptors were trimmed from ChIP reads using Trimmomatic [96] (version 0.39) and then aligned to
423 mouse reference genome (mm10), modified to incorporate known D2 variants reported in Mouse Genomes
424 Project Database, REL-1505 [26,97] using hisat2 [98] (version 2.2.1). Duplicate reads were removed using
425 Picard Tools [99] (version 2.0.1), and peaks were called for each sample using MACS [100] (version 1.4.2).
426 While independent peak calling was performed on H3K27ac and H3K9me3 to determine quality and
427 reproducibility, to compare ATAC- and ChIP-seq signal at QTL targets the comprehensive set of open
428 chromatin locations (i.e., peakome) previously developed [28] was used to generate read counts across all
429 samples using bedtools [101] (version 2.9.2) similar to ATAC-seq. Read matrices were TMM-normalized and
430 $\log_2$-transformed downstream differential analysis of histone modification using edgeR [102] (version 3.34.1).
431 For visualization on genome browser tracks with karyoploteR [103] (version 1.20.3) we used bwa [104]
432 (version 0.7.5; filter mapping quality < 30, unmapped, secondary, QC fail, duplicates, and supplementary
433 reads) alignment to generate bigwig files.
434 **HiChIP data processing and analysis**
435 *Restriction enzyme-based HiChIP*
436 HiChIP reads were processed using HiC-Pro [105] (v2.11.4) with the following parameters:
437 N_CPU = 4, SORT_RAM = 768M

438 REFERENCE_GENOME = mm10

439 LIGATION_SITE = GATCGATC

440 MIN_FRAG_SIZE = 30

441 MAX_FRAG_SIZE =100000

442 MIN_INSERT_SIZE = 30

443 MAX_INSERT_SIZE = 600.

444 Long-range interactions were called with hichipper [106] (v.0.7.8b0) using H3K27ac ChIP-seq peaks generated
445 in this paper. PET counts were imported into RStudio (version 4.1) using diffloop [107] (v1.22.0). Before
446 filtering data, read matrices were TMM-normalized. Anchors of interactions were not merged and interactions
447 less than 5,000 bp were filtered out. P-values were adjusted using an FDR cutoff of 0.01. Interactions were
448 filtered for at least four read counts in at least two samples before identifying differential interactions using
449 DESeq2 [108] within the diffloop [107] package. Significant differential interactions were cutoff at FDR $\leq$ 0.2.
450 Differential interactions (DIs) showed higher average PET counts than non-differential interactions (Wilcoxon
451 rank sum test $p < 2.2e-16$; Additional file 1: Fig. S10), consistent with increased statistical power to detect
452 differences at well-covered loci.

453 *Dovetail HiChIP MNase Kit*

454 HiChIP reads were processed using Dovetail's HiChIP Loop Calling [109] (readthedocs Revision 8f2e0b16).
455 Reads were aligned to mm10 using bwa [104] (v0.7.18). pairtools [110] (version 1.1.2) was used to record and
456 sort valid ligation events, remove PCR duplicates, and generate .pairs and bam files. Samtools [111] (version
457 1.19.2) was used to create and index the final bam file. Libraries were verified to have no-dup read pairs $\geq$
458 50%, about 20% or greater of the total mapped non-dup pairs to be in *cis* and > 1 kb in length, and total reads
459 in 1 kb around center of ChIP peaks > 2%. Filtered pairs were converted to Hi-C Pro valid pairs format. Loops
460 were called using FitHiChIP [112] (version 11.0) with the following parameters in the bias correction/coverage
461 bias configuration file:

462 Interaction type = peak to peak

463 Bin size = 2500

464 Lower distance threshold = 10000

465 Upper distance threshold = 2000000

466 Bias types = coverages bias regression

467 Merge filtering was employed (option 1)

468 q-value = 0.01

469 Chromosome sizes correspond to mm10.

470 Peak-to-peak interactions were imported into RStudio (version 4.1). diffloop [107] (version 1.22.0) was used to
471 adjust p-values using an FDR cutoff of 0.01. Interactions were filtered for at least four read counts in at least
472 two samples. PET counts were normalized, and edgeR [102] (version 3.34.1) was used to identify differential
473 interactions (FDR < 0.2).

474 **Spatial Enrichment Analysis**

475 Region Associated Differentially expressed gene (RAD) framework [68] was adapted to test whether chromatin
476 accessibility changes and gene expression changes regulated by *trans*-QTL spatially co-localize. The genomic
477 background was defined as genes expressed in ESCs consistent with cutoff used for QTL mapping [28], rather
478 than all annotated genes genome-wide. Strand-specific signed distances between each gene's TSS to nearby
479 chromatin accessibility peaks were calculated, with negative values indicating peak positions upstream of the
480 TSS in the direction of transcription and positive values indicating downstream positions. Distances were
481 binned into 100 kb intervals spanning -500 kb to +500 kb relative to the TSS. Binomial tests were used to
482 determine significance treating each gene-peak connection within 500 kb as an independent trial (binom.test in
483 R, alternative = 'greater'), with the expected probability defined by the genome-wide proportion of focal genes
484 (p = M_focal / N_total), and Benjamini-Hochberg correction was applied for multiple testing within each
485 analysis.

486 **3C-qPCR**

487 3C-qPCR experiments were performed as previously described [65] with the following modifications. Briefly, 1)
488 10 million cells were fixed with fresh formaldehyde at a final concentration of 1% for 10 minutes with rotation
489 before quenching the reaction with 10X glycine for 5 minutes with rotation. Cells were pelleted, washed with
490 PBS, and flash frozen; 2) 800 U restriction enzyme (NEB, #R0101) and 1 mM ATP were used for digestion; 3)
491 4,000 U ligase (NEB, #M0202) was used for an overnight incubation at 16°C; 4) Qiagen MaXtract HD tubes
492 were used for 3C sample purification. BACs overlapping loci of interest (Additional file 4, Table S2) were used
493 for PCR control templates. Primers used for 3C-qPCR experiments are listed in Additional file 4.

494 **Declarations**

495 **Ethics approval and consent to participate**

496 All animal procedures were approved by the Animal Care and Use Committee of The Jackson Laboratory
497 under protocol number 04008.

498 **Availability of data and materials**

499 Original data discussed in this study in NCBI's Gene Expression Omnibus and are accessible through GEO
500 Series accession number GSE312918 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE312918).
501 Along with raw sequencing data, processed data tables in the accession include raw read counts for ATAC-
502 seq, H3K27ac ChIP-seq, H3K9me3 ChIP-seq, and HiChIP. Additionally, there are bigwig files for each ATAC,
503 ChIP, and HiChIP sample. H3K4me3 ChIP-seq [27], ATAC-seq, and RNA-seq [28] for B6 and D2 were
504 collected previously and are available through GEO accessions (H3K4me3: GSE113192,
505 https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE113192; ATAC-seq and RNA-seq: GSE164935,
506 https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE164935).

507 **Funding**

508 This work was supported by the National Institute of General Medical Sciences (NIGMS) grant R35GM133724.

509 **Acknowledgements**

510 We would like to thank all the members of the Baker laboratory for comments and discussion. Additionally, we
511 would like to thank Steve Munger for comments on the manuscript and Natalie Powers and Petko Petkov for
512 their early breeding of the congenic lines. We recognize contributions from The Jackson Laboratory Genome
513 Technologies services. The Jackson Laboratory scientific services are supported in part through the National
514 Cancer Institute's Cancer Core Grant P30CA034196.

References

1. Maurano MT, Humbert R, Rynes E, Thurman RE, Haugen E, Wang H, et al. Systematic Localization of Common Disease-Associated Variation in Regulatory DNA. Science. American Association for the Advancement of Science; 2012;337:1190–5. https://doi.org/10.1126/science.1222794

2. Tak YG, Farnham PJ. Making sense of GWAS: using epigenomics and genome engineering to understand the functional relevance of SNPs in non-coding regions of the human genome. Epigenetics & Chromatin. 2015;8:57. https://doi.org/10.1186/s13072-015-0050-4

3. Ward LD, Kellis M. Evidence of abundant purifying selection in humans for recently-acquired regulatory functions. Science. 2012;337:1675–8. https://doi.org/10.1126/science.1225057

4. Guan Y, Liu H, Ma Z, Li S-Y, Park J, Sheng X, et al. Dnmt3a and Dnmt3b-Decommissioned Fetal Enhancers are Linked to Kidney Disease. J Am Soc Nephrol. 2020;31:765–82. https://doi.org/10.1681/ASN.2019080797

5. Ma Y, Buttitta L. Chromatin organization changes during the establishment and maintenance of the postmitotic state | Epigenetics & Chromatin | Full Text. Epigenetics & Chromatin [Internet]. 2017 [cited 2021 June 22];10. https://doi.org/10.1186/s13072-017-0159-8

6. Fagnocchi L, Cherubini A, Hatsuda H, Fasciani A, Mazzoleni S, Poli V, et al. A Myc-driven self-reinforcing regulatory network maintains mouse embryonic stem cell identity. Nat Commun. 2016;7:11903. https://doi.org/10.1038/ncomms11903

7. Atlasi Y, Stunnenberg HG. The interplay of epigenetic marks during stem cell differentiation and development | Nature Reviews Genetics. Nature Reviews Genetics. 2017;18:643–58. https://doi.org/10.1038/nrg.2017.57

8. Dixon JR, Jung I, Selvaraj S, Shen Y, Antosiewicz-Bourget JE, Lee AY, et al. Chromatin architecture reorganization during stem cell differentiation. Nature. Nature Publishing Group; 2015;518:331–6. https://doi.org/10.1038/nature14222

9. Javierre BM, Burren OS, Wilder SP, Kreuzhuber R, Hill SM, Sewitz S, et al. Lineage-Specific Genome Architecture Links Enhancers and Non-coding Disease Variants to Target Gene Promoters. Cell. Elsevier; 2016;167:1369-1384.e19. https://doi.org/10.1016/j.cell.2016.09.037

10. Zhang F, Lupski JR. Non-coding genetic variants in human disease. Hum Mol Genet. 2015;24:R102–10. https://doi.org/10.1093/hmg/ddv259

11. Rockman MV, Kruglyak L. Genetics of global gene expression. Nat Rev Genet. Nature Publishing Group; 2006;7:862–72. https://doi.org/10.1038/nrg1964

12. Liu X, Mefford JA, Dahl A, He Y, Subramaniam M, Battle A, et al. GBAT: a gene-based association test for robust detection of trans-gene regulation. Genome Biology. 2020;21:211. https://doi.org/10.1186/s13059-020-02120-1

13. Mostafavi H, Spence JP, Naqvi S, Pritchard JK. Systematic differences in discovery of genetic effects on gene expression and complex traits. Nat Genet. Nature Publishing Group; 2023;55:1866–75. https://doi.org/10.1038/s41588-023-01529-1

14. Tyler AL, Mahoney JM, Keller MP, Baker CN, Gaca M, Srivastava A, et al. Transcripts with high distal heritability mediate genetic effects on complex metabolic traits. Nat Commun. Nature Publishing Group; 2025;16:5507. https://doi.org/10.1038/s41467-025-61228-9

554 15. Westra H-J, Peters MJ, Esko T, Yaghootkar H, Schurmann C, Kettunen J, et al. Systematic identification of
555 trans eQTLs as putative drivers of known disease associations. Nat Genet. Nature Publishing Group;
556 2013;45:1238–43. https://doi.org/10.1038/ng.2756

557 16. Wheeler HE, Ploch S, Barbeira AN, Bonazzola R, Andaleon A, Fotuhi Siahpirani A, et al. Imputed gene
558 associations identify replicable trans-acting genes enriched in transcription pathways and complex traits. Genet
559 Epidemiol. 2019;43:596–608. https://doi.org/10.1002/gepi.22205

560 17. Yao DW, O'Connor LJ, Price AL, Gusev A. Quantifying genetic effects on disease mediated by assayed
561 gene expression levels. Nat Genet. Nature Publishing Group; 2020;52:626–33. https://doi.org/10.1038/s41588-
562 020-0625-2

563 18. Albert FW, Bloom JS, Siegel J, Day L, Kruglyak L. Genetics of trans-regulatory variation in gene
564 expression. Wittkopp PJ, editor. eLife. eLife Sciences Publications, Ltd; 2018;7:e35471.
565 https://doi.org/10.7554/eLife.35471

566 19. Visscher PM, Wray NR, Zhang Q, Sklar P, McCarthy MI, Brown MA, et al. 10 Years of GWAS Discovery:
567 Biology, Function, and Translation. Am J Hum Genet. 2017;101:5–22.
568 https://doi.org/10.1016/j.ajhg.2017.06.005

569 20. Cirulli ET, Goldstein DB. Uncovering the roles of rare variants in common disease through whole-genome
570 sequencing. Nat Rev Genet. Nature Publishing Group; 2010;11:415–25. https://doi.org/10.1038/nrg2779

571 21. Zuk O, Schaffner SF, Samocha K, Do R, Hechter E, Kathiresan S, et al. Searching for missing heritability:
572 Designing rare variant association studies. Proceedings of the National Academy of Sciences. Proceedings of
573 the National Academy of Sciences; 2014;111:E455–64. https://doi.org/10.1073/pnas.1322563111

574 22. Churchill GA, Gatti DM, Munger SC, Svenson KL. The Diversity Outbred Mouse Population. Mamm
575 Genome. 2012;23:713–8. https://doi.org/10.1007/s00335-012-9414-2

576 23. Svenson KL, Gatti DM, Valdar W, Welsh CE, Cheng R, Chesler EJ, et al. High-Resolution Genetic
577 Mapping Using the Mouse Diversity Outbred Population. Genetics. 2012;190:437–47.
578 https://doi.org/10.1534/genetics.111.132597

579 24. Williams RW, Gu J, Qi S, Lu L. The genetic structure of recombinant inbred mice: High-resolution
580 consensus maps for complex trait analysis. Genome Biol. 2001;2:preprint0007.1. https://doi.org/10.1186/gb-
581 2001-2-8-preprint0007

582 25. Taylor BA, Wnek C, Kotlus BS, Roemer N, MacTaggart T, Phillips SJ. Genotyping new BXD recombinant
583 inbred mouse strains and comparison of BXD and consensus maps. Mammalian Genome. 1999;10:335–48.
584 https://doi.org/10.1007/s003359900998

585 26. Keane TM, Goodstadt L, Danecek P, White MA, Wong K, Yalcin B, et al. Mouse genomic variation and its
586 effect on phenotypes and gene regulation. Nature. 2011;477:289–94. https://doi.org/10.1038/nature10413

587 27. Baker CL, Walker M, Arat S, Ananda G, Petkova P, Powers NR, et al. Tissue-Specific Trans Regulation of
588 the Mouse Epigenome. Genetics. Genetics; 2019;211:831–45. https://doi.org/10.1534/genetics.118.301697

589 28. Byers C, Spruce C, Fortin HJ, Hartig EI, Czechanski A, Munger SC, et al. Genetic control of the
590 pluripotency epigenome determines differentiation bias in mouse embryonic stem cells. The EMBO Journal.
591 John Wiley & Sons, Ltd; 2021;40:e109445. https://doi.org/10.15252/embj.2021109445

592 29. Johnson KR, Lane PW, Ward-Bailey P, Davisson MT. Mapping the Mouse Dactylaplasia Mutation,Dac,and
593 a Gene That Controls Its Expression,mdac. Genomics. 1995;29:457–64.
594 https://doi.org/10.1006/geno.1995.9981

595 30. Kano H, Kurahashi H, Toda T. Genetically regulated epigenetic transcriptional activation of retrotransposon
596 insertion confers mouse dactylaplasia phenotype. PNAS. National Academy of Sciences; 2007;104:19034–9.
597 https://doi.org/10.1073/pnas.0705483104

598 31. Plamondon JA, Harris MJ, Mager DL, Gagnier L, Juriloff DM. The clf2 gene has an epigenetic role in the
599 multifactorial etiology of cleft lip and palate in the A/WySn mouse strain. Birth Defects Research Part A: Clinical
600 and Molecular Teratology. 2011;91:716–27. https://doi.org/10.1002/bdra.20788

601 32. Swanzey E, McNamara TF, Apostolou E, Tahiliani M, Stadtfeld M. A Susceptibility Locus on Chromosome
602 13 Profoundly Impacts the Stability of Genomic Imprinting in Mouse Pluripotent Stem Cells. Cell Reports.
603 Elsevier; 2020;30:3597-3604.e3. https://doi.org/10.1016/j.celrep.2020.02.073

604 33. Treger RS, Pope SD, Kong Y, Tokuyama M, Taura M, Iwasaki A. The Lupus Susceptibility Locus Sgp3
605 Encodes the Suppressor of Endogenous Retrovirus Expression SNERV. Immunity. 2019;50:334-347.e9.
606 https://doi.org/10.1016/j.immuni.2018.12.022

607 34. Skelly DA, Czechanski A, Byers C, Aydin S, Spruce C, Olivier C, et al. Mapping the Effects of Genetic
608 Variation on Chromatin State and Gene Expression Reveals Loci That Control Ground State Pluripotency. Cell
609 Stem Cell. Elsevier; 2020;27:459-469.e8. https://doi.org/10.1016/j.stem.2020.07.005

610 35. Cheung VG, Spielman RS. Genetics of human gene expression: mapping DNA variants that influence
611 gene expression. Nat Rev Genet. 2009;10:595–604. https://doi.org/10.1038/nrg2630

612 36. Mattioli K, Oliveros W, Gerhardinger C, Andergassen D, Maass PG, Rinn JL, et al. Cis and trans effects
613 differentially contribute to the evolution of promoters and enhancers. Genome Biology. 2020;21:210.
614 https://doi.org/10.1186/s13059-020-02110-3

615 37. Fullwood MJ, Wei C-L, Liu ET, Ruan Y. Next-generation DNA sequencing of paired-end tags (PET) for
616 transcriptome and genome analyses. Genome Res. 2009;19:521–32. https://doi.org/10.1101/gr.074906.107

617 38. Mifsud B, Tavares-Cadete F, Young AN, Sugar R, Schoenfelder S, Ferreira L, et al. Mapping long-range
618 promoter contacts in human cells with high-resolution capture Hi-C. Nat Genet. Nature Publishing Group;
619 2015;47:598–606. https://doi.org/10.1038/ng.3286

620 39. Mumbach MR, Rubin AJ, Flynn RA, Dai C, Khavari PA, Greenleaf WJ, et al. HiChIP: efficient and sensitive
621 analysis of protein-directed genome architecture. Nat Methods. 2016;13:919–22.
622 https://doi.org/10.1038/nmeth.3999

623 40. Fang R, Yu M, Li G, Chee S, Liu T, Schmitt AD, et al. Mapping of long-range chromatin interactions by
624 proximity ligation-assisted ChIP-seq. Cell Res. Nature Publishing Group; 2016;26:1345–8.
625 https://doi.org/10.1038/cr.2016.137

626 41. Krijger PHL, Geeven G, Bianchi V, Hilvering CRE, de Laat W. 4C-seq from beginning to end: A detailed
627 protocol for sample preparation and data analysis. Methods. 2020;170:17–32.
628 https://doi.org/10.1016/j.ymeth.2019.07.014

629 42. van Berkum NL, Lieberman-Aiden E, Williams L, Imakaev M, Gnirke A, Mirny LA, et al. Hi-C: A Method to
630 Study the Three-dimensional Architecture of Genomes. J Vis Exp. 2010;1869. https://doi.org/10.3791/1869

631 43. Beagrie RA, Scialdone A, Schueler M, Kraemer DCA, Chotalia M, Xie SQ, et al. Complex multi-enhancer
632 contacts captured by genome architecture mapping. Nature. Nature Publishing Group; 2017;543:519–24.
633 https://doi.org/10.1038/nature21411

634 44. Dekker J, Mirny L. The 3D Genome as Moderator of Chromosomal Communication. Cell. 2016;164:1110–
635 21. https://doi.org/10.1016/j.cell.2016.02.007

636 45. Bortle KV, Corces VG. Nuclear Organization and Genome Function. Annu Rev Cell Dev Biol. 2012;28:163–
637 87. https://doi.org/10.1146/annurev-cellbio-101011-155824

638 46. Bonev B, Cavalli G. Organization and function of the 3D genome. Nat Rev Genet. 2016;17:661–78.
639 https://doi.org/10.1038/nrg.2016.112

640 47. Gorkin DU, Leung D, Ren B. The 3D Genome in Transcriptional Regulation and Pluripotency. Cell Stem
641 Cell. 2014;14:762–75. https://doi.org/10.1016/j.stem.2014.05.017

642 48. Hug CB, Vaquerizas JM. The Birth of the 3D Genome during Early Embryonic Development. Trends in
643 Genetics. Elsevier; 2018;34:903–14. https://doi.org/10.1016/j.tig.2018.09.002

644 49. Carter D, Chakalova L, Osborne CS, Dai Y-F, Fraser P. Long-range chromatin regulatory interactions in
645 vivo. Nature Genetics. 2002;32:623–6. https://doi.org/10.1038/ng1051

646 50. Li G, Ruan X, Auerbach RK, Sandhu KS, Zheng M, Wang P, et al. Extensive Promoter-Centered
647 Chromatin Interactions Provide a Topological Basis for Transcription Regulation. Cell. 2012;148:84–98.
648 https://doi.org/10.1016/j.cell.2011.12.014

649 51. Sanyal A, Lajoie BR, Jain G, Dekker J. The long-range interaction landscape of gene promoters. Nature.
650 Nature Publishing Group; 2012;489:109–13. https://doi.org/10.1038/nature11279

651 52. Tolhuis B, Palstra R-J, Splinter E, Grosveld F, de Laat W. Looping and Interaction between Hypersensitive
652 Sites in the Active β-globin Locus. Molecular Cell. 2002;10:1453–65. https://doi.org/10.1016/S1097-
653 2765(02)00781-5

654 53. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, et al. Topological domains in mammalian genomes
655 identified by analysis of chromatin interactions. Nature. Nature Publishing Group; 2012;485:376–80.
656 https://doi.org/10.1038/nature11082

657 54. Nora EP, Lajoie BR, Schulz EG, Giorgetti L, Okamoto I, Servant N, et al. Spatial partitioning of the
658 regulatory landscape of the X-inactivation centre. Nature. 2012;485:381–5.
659 https://doi.org/10.1038/nature11049

660 55. Smith EM, Lajoie BR, Jain G, Dekker J. Invariant TAD Boundaries Constrain Cell-Type-Specific Looping
661 Interactions between Promoters and Distal Elements around the *CFTR* Locus. The American Journal of
662 Human Genetics. 2016;98:185–201. https://doi.org/10.1016/j.ajhg.2015.12.002

663 56. Martin P, McGovern A, Orozco G, Duffus K, Yarwood A, Schoenfelder S, et al. Capture Hi-C reveals novel
664 candidate genes and complex long-range interactions with related autoimmune risk loci. Nat Commun. Nature
665 Publishing Group; 2015;6:10069. https://doi.org/10.1038/ncomms10069

666 57. Franke M, Ibrahim DM, Andrey G, Schwarzer W, Heinrich V, Schöpflin R, et al. Formation of new
667 chromatin domains determines pathogenicity of genomic duplications. Nature. 2016;538:265–9.
668 https://doi.org/10.1038/nature19800

669 58. Lupiáñez DG, Kraft K, Heinrich V, Krawitz P, Brancati F, Klopocki E, et al. Disruptions of Topological
670 Chromatin Domains Cause Pathogenic Rewiring of Gene-Enhancer Interactions. Cell. 2015;161:1012–25.
671 https://doi.org/10.1016/j.cell.2015.04.004

672 59. Cova G, Glaser J, Schöpflin R, Prada-Medina CA, Ali S, Franke M, et al. Combinatorial effects on gene
673 expression at the Lbx1/Fgf8 locus resolve split-hand/foot malformation type 3. Nat Commun. Nature Publishing
674 Group; 2023;14:1475. https://doi.org/10.1038/s41467-023-37057-z

675 60. Gorkin DU, Qiu Y, Hu M, Fletez-Brant K, Liu T, Schmitt AD, et al. Common DNA sequence variation
676 influences 3-dimensional conformation of the human genome. Genome Biology. 2019;20:255.
677 https://doi.org/10.1186/s13059-019-1855-4

678 61. Peng Y, Xiong D, Zhao L, Ouyang W, Wang S, Sun J, et al. Chromatin interaction maps reveal genetic
679 regulation for quantitative traits in maize. Nat Commun. 2019;10:2632. https://doi.org/10.1038/s41467-019-
680 10602-5

681 62. Zhao L, Wang S, Cao Z, Ouyang W, Zhang Q, Xie L, et al. Chromatin loops associated with active genes
682 and heterochromatin shape rice genome architecture for transcriptional regulation. Nat Commun.
683 2019;10:3640. https://doi.org/10.1038/s41467-019-11535-9

684 63. Javierre BM, Burren OS, Wilder SP, Kreuzhuber R, Hill SM, Sewitz S, et al. Lineage-Specific Genome
685 Architecture Links Enhancers and Non-coding Disease Variants to Target Gene Promoters. Cell.
686 2016;167:1369-1384.e19. https://doi.org/10.1016/j.cell.2016.09.037

687 64. Schoenfelder S, Javierre B-M, Furlan-Magaril M, Wingett SW, Fraser P. Promoter Capture Hi-C: High-
688 resolution, Genome-wide Profiling of Promoter Interactions. JoVE (Journal of Visualized Experiments).
689 2018;e57320. https://doi.org/10.3791/57320

690 65. Hagège H, Klous P, Braem C, Splinter E, Dekker J, Cathala G, et al. Quantitative analysis of chromosome
691 conformation capture assays (3C-qPCR). Nat Protoc. 2007;2:1722–33. https://doi.org/10.1038/nprot.2007.243

692 66. Friedman JR, Fredericks WJ, Jensen DE, Speicher DW, Huang XP, Neilson EG, et al. KAP-1, a novel
693 corepressor for the highly conserved KRAB repression domain. Genes Dev. 1996;10:2067–78.
694 https://doi.org/10.1101/gad.10.16.2067

695 67. Moosmann P, Georgiev O, Thiesen H-J, Hagmann M, Schaffner W. Silencing of RNA Polymerases II and
696 III-Dependent Transcription by the KRAB Protein Domain of KOX1, a Krüppel-Type Zinc Finger Factor.
697 Biological Chemistry. De Gruyter; 1997;378:669–78. https://doi.org/10.1515/bchm.1997.378.7.669

698 68. Guo Y, Xue Z, Yuan R, Li JJ, Pastor WA, Liu W. RAD: a web application to identify region associated
699 differentially expressed genes. Bioinformatics. 2021;37:2741–3. https://doi.org/10.1093/bioinformatics/btab075

700 69. Ray-Jones H, Sung CK, Chan LT, Haglund A, Artemov P, Della Rosa M, et al. Genetic coupling of
701 enhancer activity and connectivity in gene expression control. Nat Commun. Nature Publishing Group;
702 2025;16:970. https://doi.org/10.1038/s41467-025-55900-3

703 70. Shi C, Zhao D, Butler J, Frantzeskos A, Rossi S, Ding J, et al. Multi-omics analysis in primary T cells
704 elucidates mechanisms behind disease-associated genetic loci. Genome Biology. 2025;26:26.
705 https://doi.org/10.1186/s13059-025-03492-y

706 71. Greenwald WW, Li H, Benaglio P, Jakubosky D, Matsui H, Schmitt A, et al. Subtle changes in chromatin
707 loop contact propensity are associated with differential gene regulation and expression. Nat Commun.
708 2019;10:1054. https://doi.org/10.1038/s41467-019-08940-5

709 72. Arthur TD, Nguyen JP, Henson BA, D'Antonio-Chronowska A, Jaureguy J, Silva N, et al. Multiomic QTL
710 mapping reveals phenotypic complexity of GWAS loci and prioritizes putative causal variants. Cell Genom.
711 2025;5:100775. https://doi.org/10.1016/j.xgen.2025.100775

712 73. Thiesen HJ. Multiple genes encoding zinc finger domains are expressed in human T cells. New Biol.
713 1990;2:363–74.

714 74. Bellefroid EJ, Poncelet DA, Lecocq PJ, Revelant O, Martial JA. The evolutionarily conserved Krüppel-
715 associated box domain defines a subfamily of eukaryotic multifingered proteins. Proceedings of the National

716 Academy of Sciences. Proceedings of the National Academy of Sciences; 1991;88:3608–12.
717 https://doi.org/10.1073/pnas.88.9.3608

718 75. Schultz DC, Ayyanathan K, Negorev D, Maul GG, Rauscher FJ. SETDB1: a novel KAP-1-associated
719 histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic
720 genes by KRAB zinc-finger proteins. Genes Dev. 2002;16:919–32. https://doi.org/10.1101/gad.973302

721 76. Lupo A, Cesaro E, Montano G, Zurlo D, Izzo P, Costanzo P. KRAB-Zinc Finger Proteins: A Repressor
722 Family Displaying Multiple Biological Functions. Curr Genomics. 2013;14:268–78.
723 https://doi.org/10.2174/1389202911314999000 2

724 77. Sripathy SP, Stevens J, Schultz DC. The KAP1 Corepressor Functions To Coordinate the Assembly of De
725 Novo HP1-Demarcated Microenvironments of Heterochromatin Required for KRAB Zinc Finger Protein-
726 Mediated Transcriptional Repression. Mol Cell Biol. 2006;26:8623–38. https://doi.org/10.1128/MCB.00487-06

727 78. Francis BA, Gozashti L, Costello K, Kasahara T, Harringmeyer OS, Lilue J, et al. Complete genome
728 assemblies of two mouse subspecies reveal structural diversity of telomeres and centromeres. Nat Genet.
729 Nature Publishing Group; 2025;57:2852–62. https://doi.org/10.1038/s41588-025-02367-z

730 79. Bruno M, Maisha S, Mitra A, Costello K, Watkins-Chow D, Logsdon GA, et al. Young KRAB-zinc finger
731 gene clusters are highly dynamic incubators of ERV-driven genetic heterogeneity in mice [Internet]. bioRxiv;
732 2025 [cited 2025 July 3]. p. 2025.02.26.640358. https://doi.org/10.1101/2025.02.26.640358

733 80. Groner AC, Meylan S, Ciuffi A, Zangger N, Ambrosini G, Dénervaud N, et al. KRAB–Zinc Finger Proteins
734 and KAP1 Can Mediate Long-Range Transcriptional Repression through Heterochromatin Spreading. PLOS
735 Genetics. Public Library of Science; 2010;6:e1000869. https://doi.org/10.1371/journal.pgen.1000869

736 81. Larson MH, Gilbert LA, Wang X, Lim WA, Weissman JS, Qi LS. CRISPR interference (CRISPRi) for
737 sequence-specific control of gene expression. Nat Protoc. 2013;8:2180–96.
738 https://doi.org/10.1038/nprot.2013.132

739 82. Bruno M, Maisha S, Mitra A, Costello K, Watkins-Chow D, Logsdon GA, et al. Young KRAB-zinc finger
740 gene clusters are highly dynamic incubators of ERV-driven genetic heterogeneity in mice. bioRxiv.
741 2025;2025.02.26.640358. https://doi.org/10.1101/2025.02.26.640358

742 83. Imbeault M, Helleboid P-Y, Trono D. KRAB zinc-finger proteins contribute to the evolution of gene
743 regulatory networks. Nature. Nature Publishing Group; 2017;543:550–4. https://doi.org/10.1038/nature21683

744 84. Hansen TJ, Fong SL, Day JK, Capra JA, Hodges E. Human gene regulatory evolution is driven by the
745 divergence of regulatory element function in both cis and trans. Cell Genom. 2024;4:100536.
746 https://doi.org/10.1016/j.xgen.2024.100536

747 85. Becker JS, Nicetto D, Zaret KS. H3K9me3-Dependent Heterochromatin: Barrier to Cell Fate Changes.
748 Trends Genet. 2016;32:29–41. https://doi.org/10.1016/j.tig.2015.11.001

749 86. Soufi A, Donahue G, Zaret KS. Facilitators and Impediments of the Pluripotency Reprogramming Factors'
750 Initial Engagement with the Genome. Cell. 2012;151:994–1004. https://doi.org/10.1016/j.cell.2012.09.045

751 87. Kaemena DF, Yoshihara M, Beniazza M, Ashmore J, Zhao S, Bertenstam M, et al. B1 SINE-binding
752 ZFP266 impedes mouse iPSC generation through suppression of chromatin opening mediated by
753 reprogramming factors. Nat Commun. 2023;14:488. https://doi.org/10.1038/s41467-023-36097-9

754 88. Kim S-K, Seo S, Stein-O'Brien G, Jaishankar A, Ogawa K, Micali N, et al. Individual variation in the
755 emergence of anterior-to-posterior neural fates from human pluripotent stem cells. Stem Cell Reports.
756 2024;19:1336–50. https://doi.org/10.1016/j.stemcr.2024.07.004

757 89. Perdomo-Sabogal Á, Nowick K. Genetic Variation in Human Gene Regulatory Factors Uncovers
758 Regulatory Roles in Local Adaptation and Disease. Genome Biol Evol. 2019;11:2178–93.
759 https://doi.org/10.1093/gbe/evz131

760 90. Sigmon JS, Blanchard MW, Baric RS, Bell TA, Brennan J, Brockmann GA, et al. Content and Performance
761 of the MiniMUGA Genotyping Array: A New Tool To Improve Rigor and Reproducibility in Mouse Research.
762 Genetics. 2020;216:905–30. https://doi.org/10.1534/genetics.120.303596

763 91. Czechanski A, Byers C, Greenstein I, Schrode N, Donahue LR, Hadjantonakis A-K, et al. Derivation and
764 characterization of mouse embryonic stem cells (mESCs) from permissive and non-permissive strains. Nat
765 Protoc. 2014;9:559–74. https://doi.org/10.1038/nprot.2014.030

766 92. Corces MR, Buenrostro JD, Wu B, Greenside PG, Chan SM, Koenig JL, et al. Lineage-specific and single-
767 cell chromatin accessibility charts human hematopoiesis and leukemia evolution. Nat Genet. Nature Publishing
768 Group; 2016;48:1193–203. https://doi.org/10.1038/ng.3646

769 93. Billings T, Parvanov ED, Baker CL, Walker M, Paigen K, Petkov PM. DNA binding specificities of the long
770 zinc-finger recombination protein PRDM9. Genome Biology. 2013;14:R35. https://doi.org/10.1186/gb-2013-14-
771 4-r35

772 94. Belaghzal H, Dekker J, Gibcus JH. Hi-C 2.0: An optimized Hi-C procedure for high-resolution genome-wide
773 mapping of chromosome conformation. Methods. 2017;123:56–65. https://doi.org/10.1016/j.ymeth.2017.04.004

774 95. Dovetail-HiChip-MNase-Kit-Version-2.0.pdf [Internet]. [cited 2025 July 7]. https://cantatabio.com/wp-
775 content/uploads/2023/12/Dovetail-HiChip-MNase-Kit-Version-2.0.pdf. Accessed 7 July 2025

776 96. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics.
777 2014;30:2114–20. https://doi.org/10.1093/bioinformatics/btu170

778 97. Yalcin B, Wong K, Agam A, Goodson M, Keane TM, Gan X, et al. Sequence based characterization of
779 structural variation in the mouse genome. Nature. 2011;477:326–9. https://doi.org/10.1038/nature10432

780 98. Pertea M, Kim D, Pertea G, Leek JT, Salzberg SL. Transcript-level expression analysis of RNA-seq
781 experiments with HISAT, StringTie, and Ballgown. Nat Protoc. 2016;11:1650–67.
782 https://doi.org/10.1038/nprot.2016.095

783 99. Picard Tools - By Broad Institute [Internet]. [cited 2025 July 9]. https://broadinstitute.github.io/picard/.
784 Accessed 9 July 2025

785 100. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, et al. Model-based Analysis of ChIP-
786 Seq (MACS). Genome Biol. 2008;9:R137. https://doi.org/10.1186/gb-2008-9-9-r137

787 101. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics.
788 2010;26:841–2. https://doi.org/10.1093/bioinformatics/btq033

789 102. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression
790 analysis of digital gene expression data. Bioinformatics. 2010;26:139–40.
791 https://doi.org/10.1093/bioinformatics/btp616

792 103. Gel B, Serra E. karyoploteR: an R/Bioconductor package to plot customizable genomes displaying
793 arbitrary data. Bioinformatics. 2017;33:3088–90. https://doi.org/10.1093/bioinformatics/btx346

794 104. Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics.
795 2009;25:1754–60. https://doi.org/10.1093/bioinformatics/btp324

796 105. Servant N, Varoquaux N, Lajoie BR, Viara E, Chen C-J, Vert J-P, et al. HiC-Pro: an optimized and flexible
797 pipeline for Hi-C data processing. Genome Biology. 2015;16:259. https://doi.org/10.1186/s13059-015-0831-x

798 106. Lareau CA, Aryee MJ. hichipper: a preprocessing pipeline for calling DNA loops from HiChIP data. Nat
799 Methods. 2018;15:155–6. https://doi.org/10.1038/nmeth.4583

800 107. Lareau CA, Aryee MJ. diffloop: a computational framework for identifying and analyzing differential DNA
801 loops from sequencing data. Bioinformatics. 2018;34:672–4. https://doi.org/10.1093/bioinformatics/btx623

802 108. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with
803 DESeq2. Genome Biol [Internet]. Springer Science and Business Media LLC; 2014 [cited 2025 July 7];15.
804 https://doi.org/10.1186/s13059-014-0550-8

805 109. HiChIP 0.1 documentation [Internet]. [cited 2025 July 7].
806 https://hichip.readthedocs.io/en/latest/before_you_begin.html. Accessed 7 July 2025

807 110. Open2C, Abdennur N, Fudenberg G, Flyamer IM, Galitsyna AA, Goloborodko A, et al. Pairtools: From
808 sequencing data to chromosome contacts. Ay F, editor. PLoS Comput Biol. Public Library of Science (PLoS);
809 2024;20:e1012164. https://doi.org/10.1371/journal.pcbi.1012164

810 111. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format
811 and SAMtools. Bioinformatics. 2009;25:2078–9. https://doi.org/10.1093/bioinformatics/btp352

812 112. Bhattacharyya S, Chandra V, Vijayanand P, Ay F. Identification of significant chromatin contacts from
813 HiChIP data by FitHiChIP. Nat Commun. Nature Publishing Group; 2019;10:4221.
814 https://doi.org/10.1038/s41467-019-11950-y

Figures

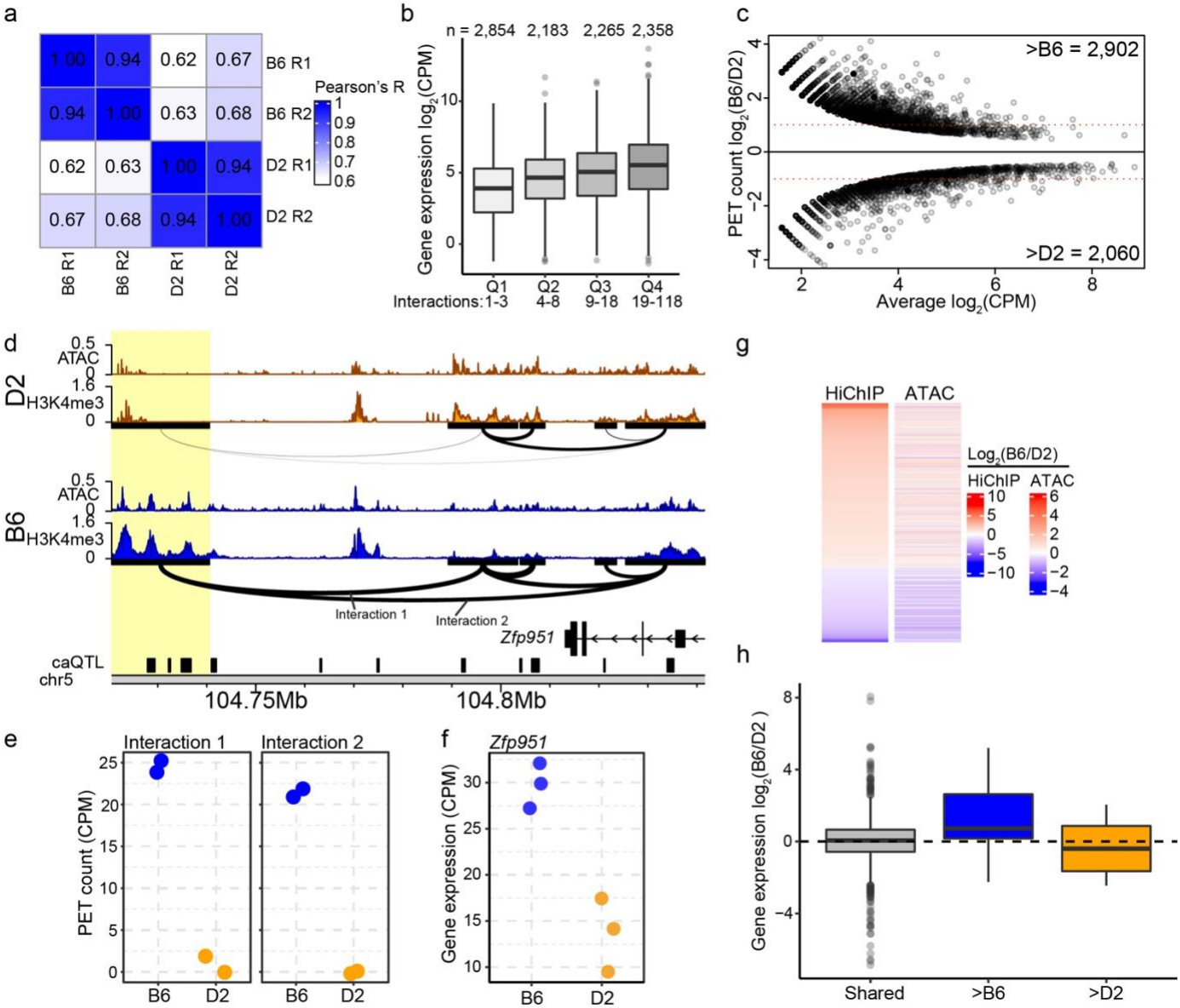

Fig. 1. Genetic variation coordinates chromatin interactions, accessibility, and gene expression. a Pearson correlation of H3K27ac HiChIP from biological replicates (R1 and R2) of mouse embryonic stem cells (ESCs) derived from B6 and D2 strains. **b** Box and whisker plots indicating mean gene expression values from RNA-seq for B6 and D2 mESCs (for all figure boxes: first quartile, median, and third quartile, whiskers: minimum and maximum, outliers are indicated as dots). Genes were stratified into quartiles (Q1-4) based on the number of contacts per promoter (interactions below graph, gene number above). One-way ANOVA: $F_{(3, 9656)} = 232$, $p < 2e\text{-}16$; Tukey HSD post-hoc test: all pairwise comparisons between quartiles were significant. **c** MA-plot for differential interactions B6 and D2 ESCs (FDR < 0.2). **d** Genome browser profiles of H3K4me3 ChIP-seq and ATAC-seq at the *Zfp951* locus for B6 (blue) and D2 (orange). DIs are drawn below H3K4me3 tracks. The thickness of the arc connecting two anchors represent contact frequency. Black boxes represent chromatin accessibility regions mapped as quantitative trait loci (caQTL) [28]. Yellow highlight – contact anchor overlapping multiple QTL targets. **e**. HiChIP PET counts for B6 and D2 from interactions 1 (p = 7.44e-9) and 2 (p = 4.48e-10) at *Zfp951* locus. **f** Expression of *Zfp951* from B6 and D2 ESCs (p = 2.81e-5). **g** Heatmap indicating fold-change (B6/D2) in HiChIP PET counts for DIs (left) and corresponding fold-change for ATAC-seq signal at peaks that overlap contact anchors (right). **h** Box and whisker plots of fold-change (B6/D2) in RNA-seq signal at genes which overlap contact anchors for interactions without significant differences (shared, n = 943), and either greater in B6 (n = 17) or D2 (n = 6).

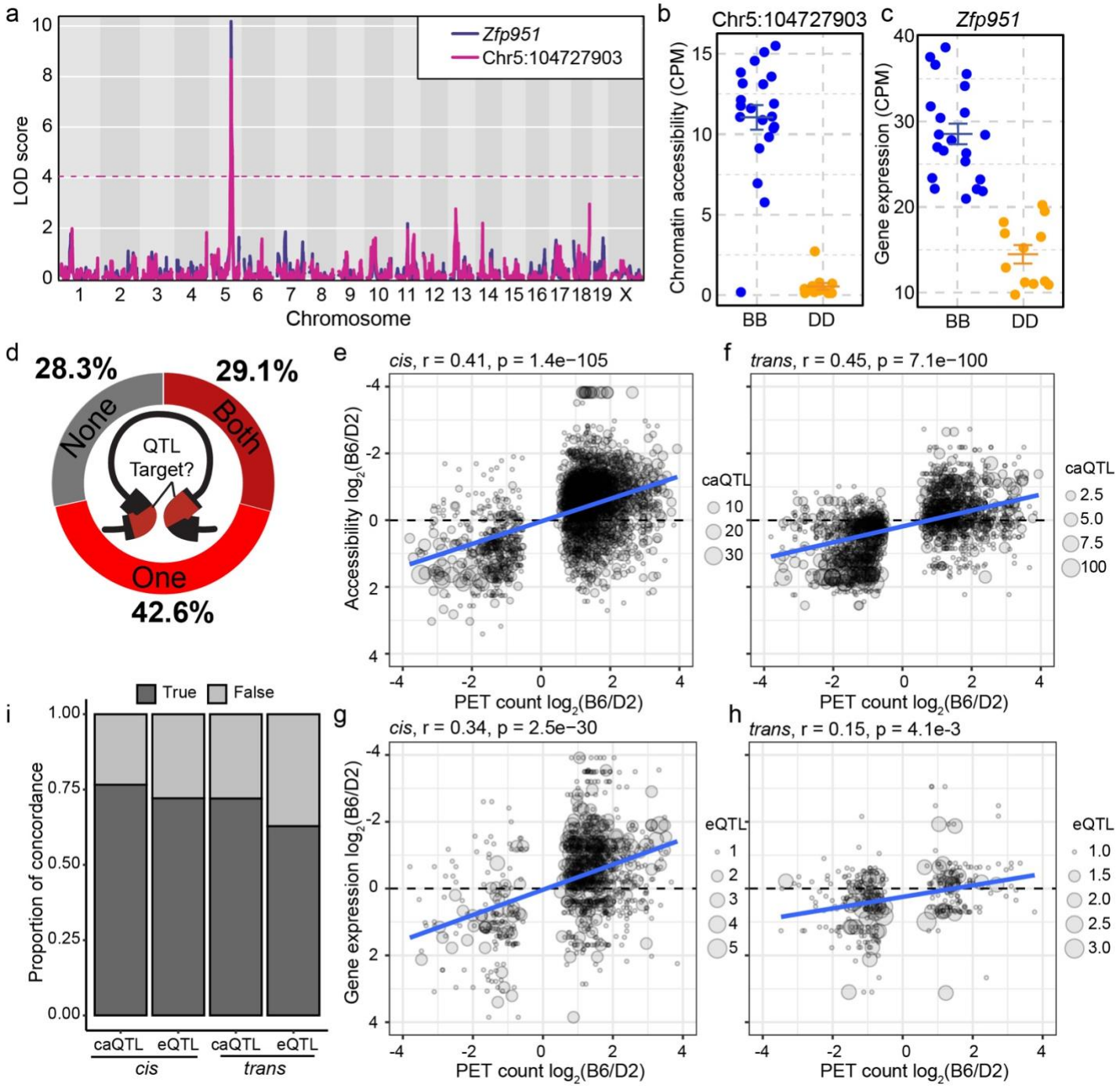

Fig. 2. Differential interactions overlap QTL targets and show allelic concordance across regulatory layers. **a** LOD score plot for *Zfp951* and a regulatory element QTL (Fig 1d, leftmost caQTL target in highlighted anchor). Both chromatin accessibility and expression of *Zfp951* are regulated locally on Chr 5. **b** Chromatin accessibility at *Zfp951* regulatory element (Fig 2a pink line) in 33 BXD ESC lines based on genotype at the QTL locus (line represents mean ± standard error). **c** Gene expression of *Zfp951* in 33 BXD ESC lines. **d** Donut plot indicating proportions of DIs for which contact anchors overlap caQTL targets at one, both, or no anchors. Correlation between PET count and chromatin accessibility (**e**, **f**) or gene expression (**g**, **h**), separated into *cis*- (left) and *trans*-associations (right). **i** Proportion of allele concordance between interaction strength and chromatin accessibility (caQTL) or gene expression (eQTL). *cis* caQTL: χ^2 = 529.5, df = 2, p < 2.2e-16; *trans* caQTL: χ^2 = 440.2, df = 2 p < 2.2e-16; *cis* eQTL: χ^2 = 148.2, df = 1 p < 2.2e-16; *trans* eQTL χ^2 = 20.3, df = 1 = 6.7e-6.

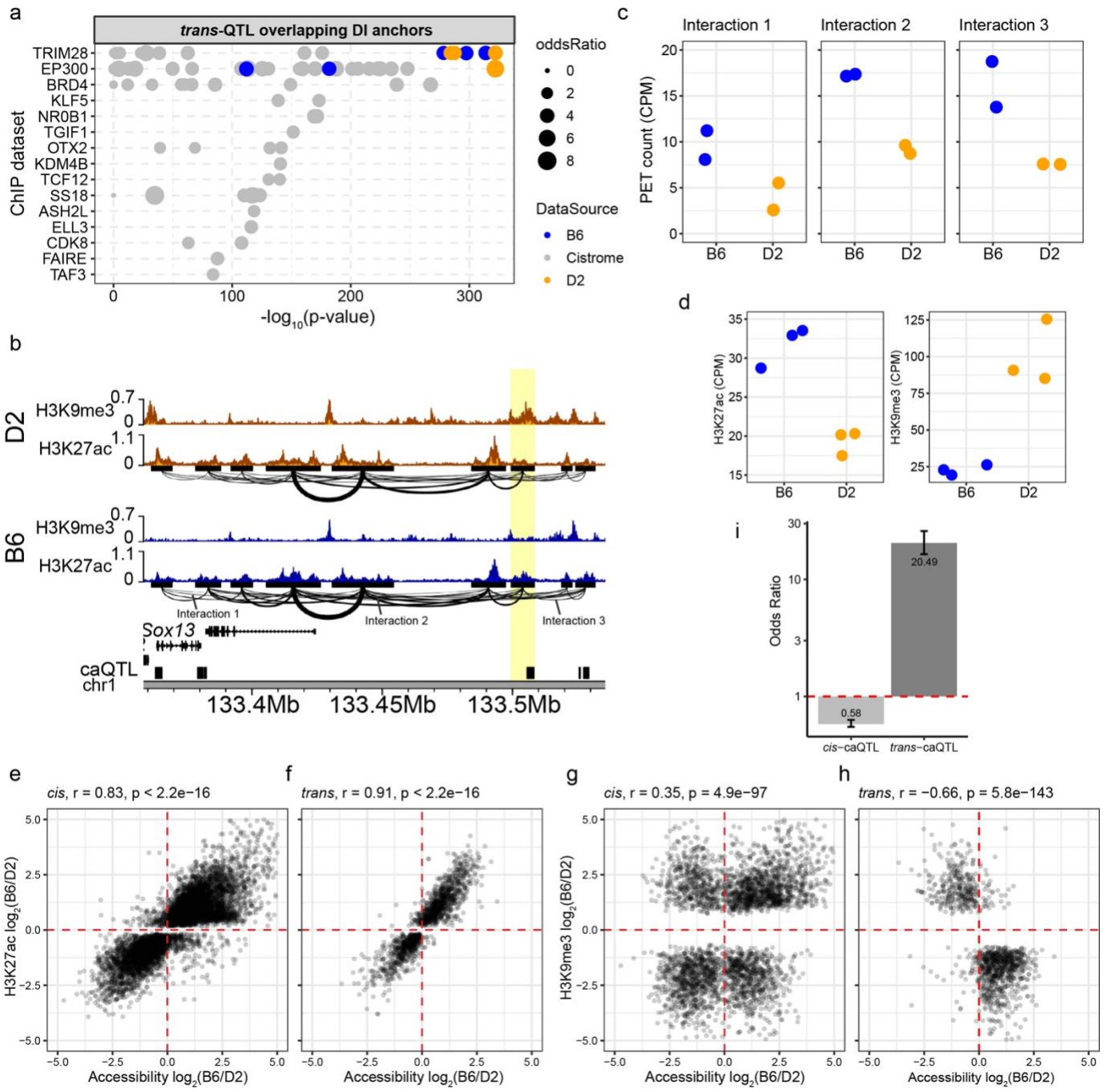

Fig. 3. Heterochromatin formation distinguishes *trans* from *cis* regulatory mechanisms. a Enrichment of the top 15 DNA-binding proteins at *trans*-QTL targets overlapping DI anchors. Each dot represents a ChIP-seq experiment either from publicly available database (grey - cistrome), or our previously published B6 (blue) and D2 (orange) datasets. Binding sites were evaluated for overlap between *trans*-targets and DI contact anchors (Fisher's exact test, top 15 factors with q-value cutoff < 0.01). **b** Genome profiles for H3K27ac and H3K9me3 ChIP enrichment from D2 (orange) and B6 (blue) at *Sox13* locus. HiChIP interactions indicated below H3K27ac tracks. Highlight indicates a contact anchor that overlaps a *trans*-target (black box). **c** HiChIP PET counts for three significant DIs indicated in panel *b* (Interaction 1 p = 9.53e-3, Interaction 2 p = 1.05e-2, Interaction 3 p = 3.04e-3). **d** Difference in H3K27ac (left) and H3K9me3 (right) at caQTL target highlighted in panel *b* between B6 and D2 (H3K27ac p = 2.24e-3, H3K9me3 p = 3.58e-3). Correlation between fold change (B6/D2) in H3K27ac enrichment and chromatin accessibility at *cis*- and *trans*-QTL targets (**e, f**). Correlation between fold change in H3K9me3 enrichment and chromatin accessibility at *cis*- and *trans*-QTL targets (**g, h**). **i** Odds ratio showing that *trans*-targets are 20.49 times more likely to show discordant chromatin states resulting in high H3K9me3 and low ATAC.

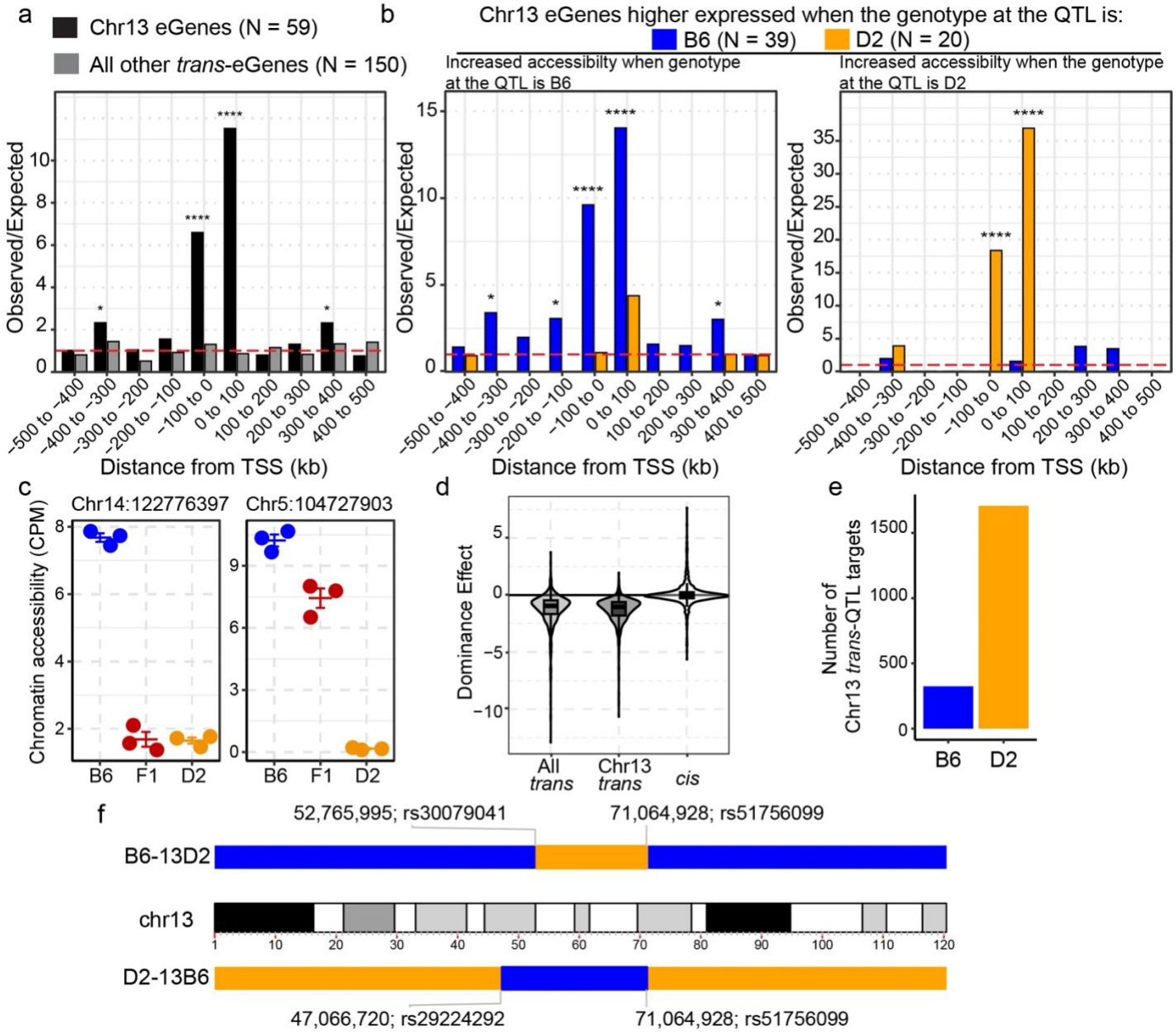

Fig. 4. Chr13 *trans*-QTL coordinates spatial regulation of chromatin accessibility and gene expression through dominant repression. **a** Observed-over-expected ratio for the spatial proximity of Chr13 *trans*-targets to transcription start sites (TSS) of either Chr13 eGenes (black) or all other *trans*-eGenes (grey). Red dashed line is null expectation (binomial test, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$). **b** Similar to *a* except stratifying genes that either higher expressed when Chr13 QTL haplotype is B6 (blue) or D2 (orange). Spatial proximity was tested for peaks with higher chromatin accessibility when Chr13 is B6 (left) or D2 (right). **c** Scatterplot of chromatin accessibility based on B6, D2, or F1-hybrid genotype of ESCs at two caQTL. Left – representative Chr13 *trans*-target indicating dominant repressive effect on chromatin accessibility. Right - *cis*-QTL target representative of additive effect (line – mean, bars – standard error). **d** Distribution of dominance effect values in all *trans*-targets (N = 836), Chr13 *trans*-targets (N = 575), and all *cis*-caQTL (N = 7049). Negative values indicate repressive effect; values near zero indicate additive effect. **e** Classification of Chr13 *trans*-targets by repressive allele. To establish directional predictions targets were stratified by which QTL genotype reduces chromatin accessibility, consistent with dominant repression. **f** Schematic of chromosome 13 showing the base-pair position and identifying SNP for congenic regions in B6-13D2 (D2 region on B6 background) and D2-13B6 (B6 region on D2 background) strains.

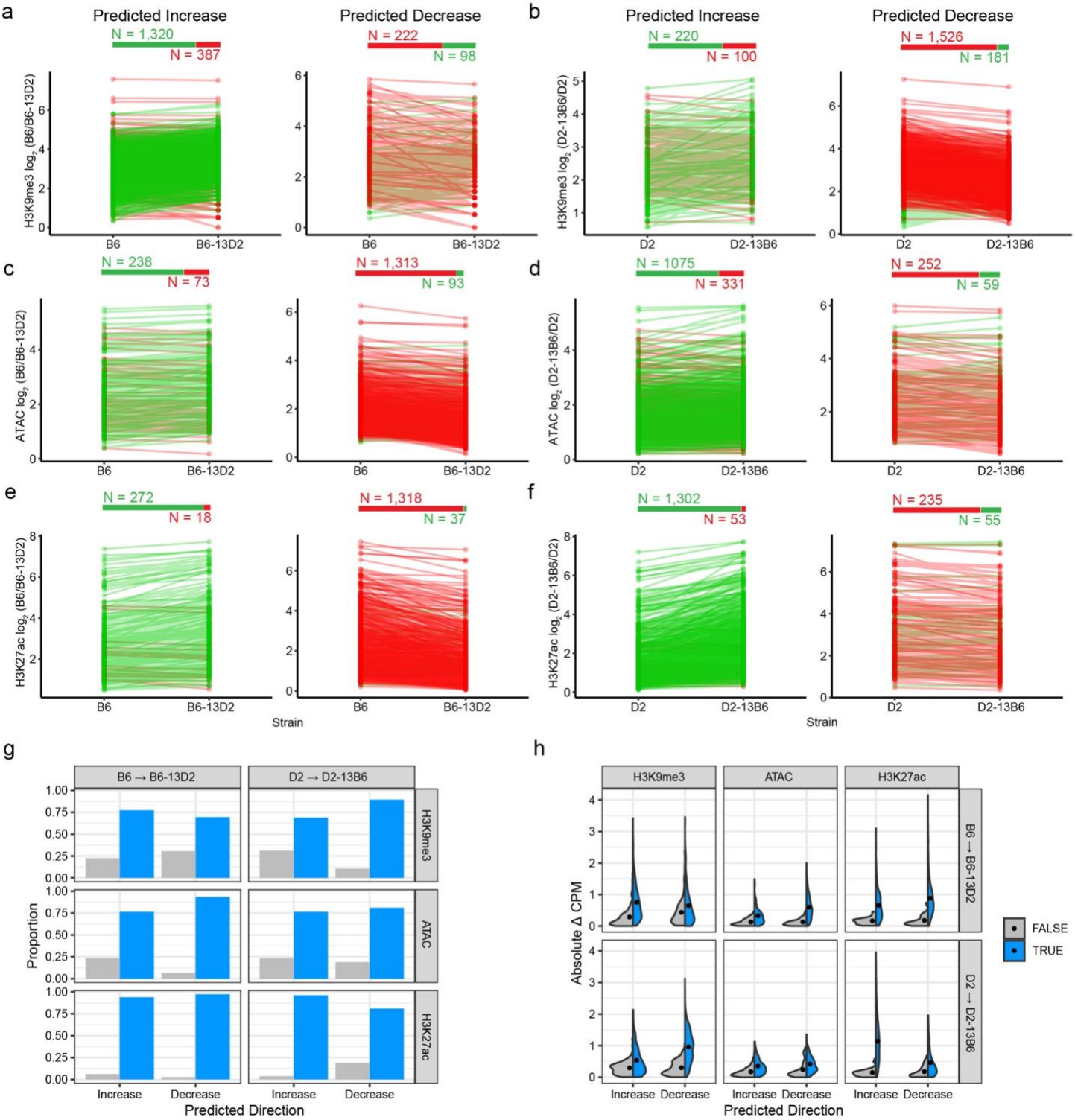

Fig. 5. Chr13 *trans*-QTL regulates heterochromatin formation at hundreds of distal loci. **a-f** Histone modification or chromatin accessibility at Chr13 *trans*-targets in parental (B6 and D2) and reciprocal congenic (B6-13D2 and D2-13B6) strains for H3K9me3 ChIP-seq (**a, b**), ATAC-seq (**c, d**), and H3K27ac ChIP-seq (**e, f**). Each line represents a single *trans*-target, connecting measurements between the parental strain (left) and its corresponding congenic (right). B6→B6-13D2 panels (**a, c, e**) test the effect of introducing the D2 Chr13 region into the B6 background, while D2→D2-13B6 panels (**b, d, f**) test the reciprocal introduction of the B6 Chr13 region into the D2 background. Within each panel, loci are separated by predicted direction of change based on QTL effect direction from BXD mapping: "Predicted Increase" (left) indicates targets expected to gain signal in the congenic strain, while "Predicted Decrease" (right) indicates targets expected to lose signal. Green lines indicate increasing values, while red lines indicate decreasing values. Numbers above each plot show total

883 *trans*-targets tested that increase (green) or decrease (red). Statistical significance assessed by exact binomial
884 test. H3K9me3 enrichment: **a** increase $p < 2.2e-16$, decrease $p = 3.3e-12$; **b** increase $p = 1.7e-11$, decrease p
885 $< 2.2e-16$ ATAC-seq: **c** increase $p < 2.2e-16$, decrease $p < 2.2e-16$; **d** increase $p < 2.2e-16$, decrease $p <$
886 $2.2e-16$. H3K27ac: **e** increase $p < 2.2e-16$, decrease $p < 2.2e-16$; **f** increase $p < 2.2e-16$, decrease $p < 2.2e-$
887 16. **g** Observed proportion of Chr13 *trans*-targets for which chromatin state changes in the direction predicted
888 (i.e., True) based on QTL effect and dominant repression. **h** Violin plot showing distribution of absolute
889 changes in counts per million (CPM) at Chr13 *trans*-QTL targets. *Trans*-targets that change as predicted (blue)
890 or do not change as predicted (grey). Dots indicated median value.

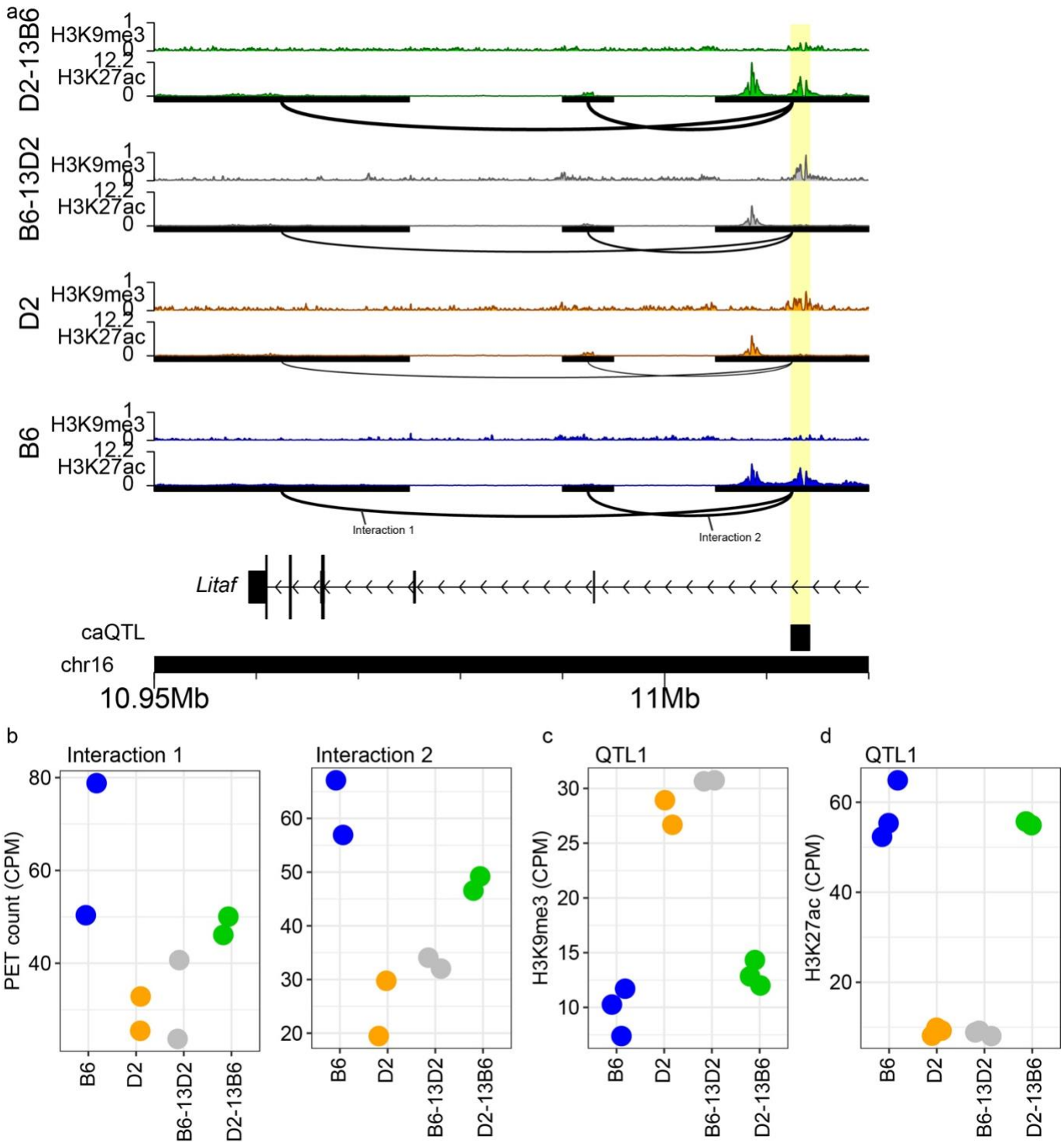

Fig. 6. Chr13 coordinates heterochromatin and 3D architecture at *Litaf* locus. **a** Genome profiles for H3K27ac and H3K9me3 ChIP enrichment from B6 (blue), D2 (orange), B6-13D2 (grey) and D2-13B6 (green) at the *Litaf* locus. HiChIP interactions indicated below H3K27ac tracks with the thickness of the arc connecting two anchors representing contact frequency. Highlight indicates a *trans*-target (black box) overlapping a contact anchor and variable histone modifications. **b** HiChIP PET counts for all four strains for two interactions indicated in panel a (Interaction 1: BD p = 4.67e-4, BB13D p = 3.14e-4, DD13B p = 2.68e-2; Interaction 2: BD p = 9.94e-5, BB13D p = 1.51e-3, DD13B p = 3.15e-3). **c** H3K9me3 enrichment at Chr13 *trans*-target highlighted panel a (BD p = 8.25e-12, BB13D p = 2.98e-15, DD13B p = 4.81e-8). **d** H3K27ac enrichment at Chr13 *trans* caQTL target highlighted panel a (BD p = 6.62e-108, BB13D p = 4.54e-119, DD13B p = 9.24e-103). BD is the comparison between B6 and D2; BB13D is the comparison between B6 and B6-13D2; DD13B is the comparison between D2 and D2-13B6.

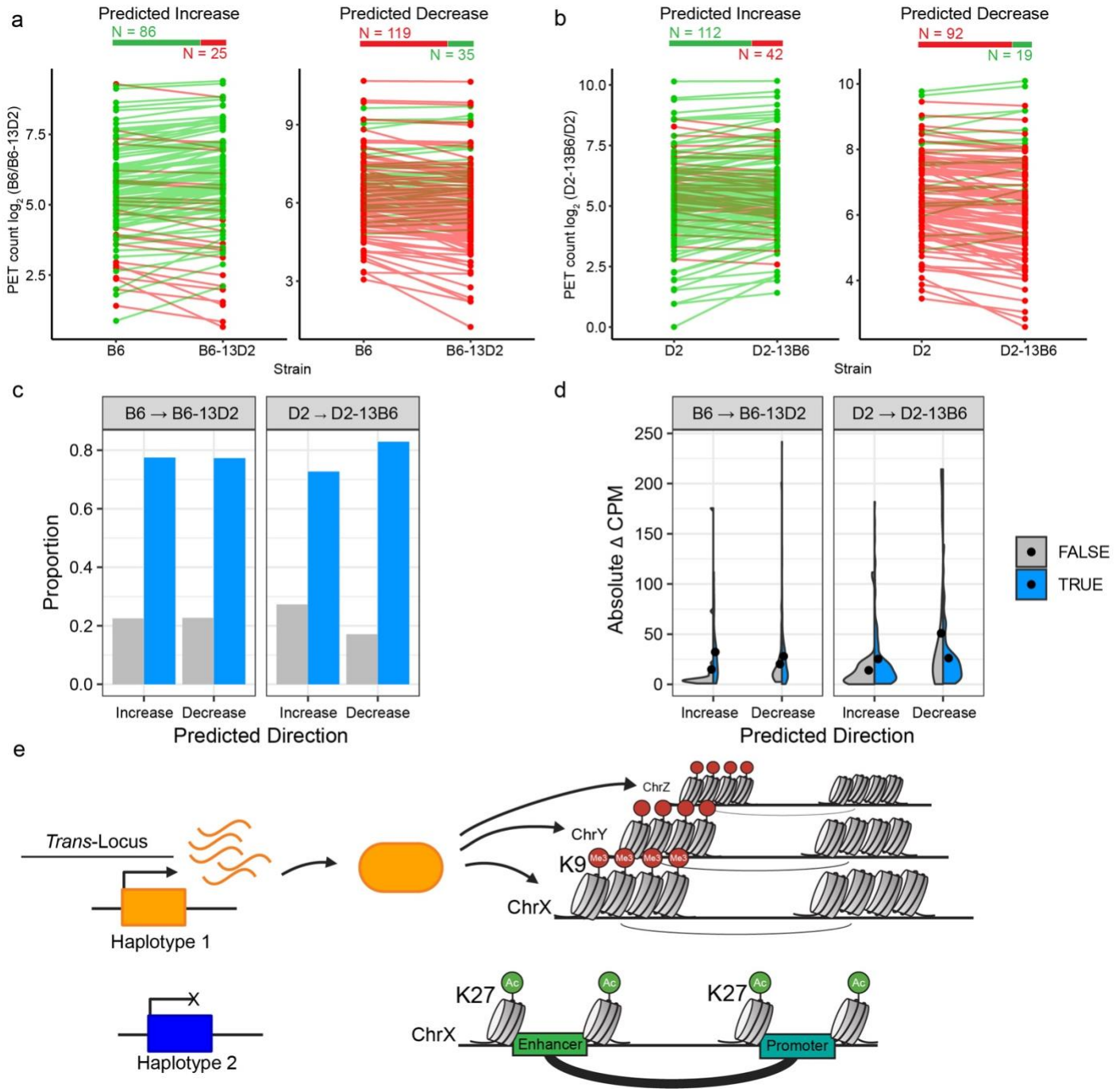

Fig. 7. Heterochromatin coordinates genome-wide changes in 3D architecture. **a**, **b** Validation of predicted changes in 3D contact frequency at Chr13 *trans*-targets using H3K27ac HiChIP. Experimental design and visualization as in Fig. 5a-f, with each line representing a single interaction overlapping a Chr13 *trans*-target. B6→B6-13D2 (**a**) and D2→D2-13B6 (**b**). Loci were split by predicted direction of change based on BXD QTL mapping. Green lines indicate increasing PET counts; redlines indicate decreasing PET counts. Numbers above show total interactions tested that increase (green) or decrease (red). Exact binomial test: **a** increase p = 5.03e-9, decrease p = 6.59e-12; **b** increase p = 1.53e-8, decrease p = 1.12e-12. **c** Observed proportion of interactions that changed as predicted (blue). **d** Distribution of absolute changes in PET counts at Chr13 *trans*-targets, showing validated predictions (blue) and non-validated targets (grey). Dots indicate median values. **e** Model illustrating *trans*-regulation through heterochromatin formation. In this example, the orange haplotype *encodes* diffusible repressor(s) that establish H3K9me3-marked heterochromatin at distal *trans*-targets (ChrX, Y, Z), reducing chromatin accessibility and weakening 3D contacts. Whereas the blue haplotype lacks active repressor(s), permitting open chromatin and stronger enhancer-promoter interactions and gene expression at the same distal locus.