



**HAL**  
open science

# TADs and Their Borders: Free Movement or Building a Wall?

Li-Hsin Chang, Sourav Ghosh, Daan Noordermeer

► **To cite this version:**

Li-Hsin Chang, Sourav Ghosh, Daan Noordermeer. TADs and Their Borders: Free Movement or Building a Wall?. *Journal of Molecular Biology*, 2020, 432 (3), pp.643-652. 10.1016/j.jmb.2019.11.025 . hal-02995869

**HAL Id: hal-02995869**

**<https://hal.science/hal-02995869>**

Submitted on 19 Nov 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13

# **TADs and their borders: free movement or building a wall?**

Li-Hsin Chang<sup>1</sup>, Sourav Ghosh<sup>1</sup>, Daan Noordermeer<sup>1†</sup>

Affiliation:

<sup>1</sup> Institute for Integrative Biology of the Cell (I2BC), CEA, CNRS, University Paris-sud,  
University Paris-Saclay, Gif-sur-Yvette, France

<sup>†</sup> Correspondence: daan.noordermeer@i2bc.paris-saclay.fr

Key words: TADs; Topologically Associating Domains; CTCF; loop extrusion; Hi-C

14 **Abstract**

15 The tridimensional (3D) organization of mammalian genomes combines structures from  
16 different length scales. Within this organization, Topologically Associating Domains (TADs)  
17 are visible in Hi-C heat maps at the sub-megabase scale. The integrity of TADs is important  
18 for correct gene expression, but in a context-dependent and variable manner. The correct  
19 structure and function of TADs require the binding of the CTCF protein at both borders,  
20 which appears to block an active and dynamic mechanism of ‘Cohesin-mediated loop  
21 extrusion’. As a result, mammalian TADs appear as so-called ‘loop domains’ in Hi-C data,  
22 which are the focus of this review. Moreover, we present a reanalysis of TADs from three  
23 “golden-standard” mammalian Hi-C data sets. Despite the prominent presence of TADs in  
24 Hi-C heat maps from all studies, we find consistently that regions within these domains are  
25 only moderately insulated from their surroundings. Moreover, single-cell Hi-C and super-  
26 resolution microscopy have revealed that the structure of TADs and the position of their  
27 borders can vary from cell-to-cell. The function of TADs as units of gene regulation may thus  
28 require additional aspects, potentially incorporating the mechanism of loop extrusion as well.  
29 Recent developments in single-cell and multi-contact genomics and super-resolution  
30 microscopy assays will be instrumental to link TAD formation and structure to their function  
31 in transcriptional regulation.

## 32 **TADs as functional units of gene regulation**

33 A major biophysical challenge for cells is to compact their large genome—up to 2 meters for  
34 a mammalian cell—into a nucleus that is only a few microns in diameter, all while ensuring  
35 both proper genome function and the programmable flexibility required for cellular  
36 differentiation [1]. The first order of DNA compaction, the wrapping of 147 bp of DNA  
37 around a histone octamer core to form the nucleosome ('10 nm fiber'), is well agreed upon  
38 [2-4]. Subsequent higher-order chromatin organization is formed by more diverse  
39 combinations of DNA structures, ranging from chromatin loops and contact domains at the  
40 smallest scales to chromosome territories that occupy large volumes within interphase nuclei  
41 [5, 6].

42 The development of Chromosome Conformation Capture (3C) and genomics-based  
43 derivatives have greatly advanced our capacity to explore structural aspects of 3D genome  
44 architecture [7, 8]. Particularly Hi-C has proven to be instrumental in our understanding of  
45 3D genome organization, by providing information on all possible genome-wide interactions  
46 within a single experiment [9]. When analyzed at sufficient resolution (typically below 100  
47 kb), Hi-C experiments reveal the existence of distinct structural domains [10]. These domains,  
48 referred to as Topologically Associating Domains (TADs), are characterized by increased  
49 intra-domain interactions and insulation from their surroundings, and thus appear as  
50 'pyramids' in Hi-C heat maps [10, 11] (Fig. 1A). Whereas TADs in mammalian cells are  
51 generally identified with median sizes just below one megabase, in other organisms they may  
52 manifest themselves at considerably different sizes [10-16]. Moreover, studies with  
53 increasingly higher resolution have suggested the presence of additional finer-scale levels of  
54 sub-domains ('sub-TADs'), nested within mammalian TADs [17-20].

55 Initial correlative studies revealed enrichment within TADs of co-regulated gene  
56 clusters, enhancer-promoter pairs and replication domains [10, 11, 21, 22]. Moreover, a large  
57 fraction of the borders that surround mammalian TADs bind the CTCF protein (CCCTC-  
58 binding factor), the only somatic insulator protein in somatic cells [10, 18] (Fig. 1A). Until  
59 now, CTCF binding at large subsets of TAD borders has only been reported in vertebrates  
60 (e.g. [10, 15]). Other TAD borders can be linked to (at least) two other biological phenomena.  
61 First, the presence of highly expressed genes can introduce a strong insulation between  
62 surrounding regions in the Hi-C heat map of a diverse range of organisms [10, 23-25].  
63 Second, a sharp separation may be observed between regions bound by histone modifications  
64 associated with opposing transcriptional and regulatory output [12, 18, 26]. This latter

65 separation, in mammals usually at the smaller sub-domain levels, appears not to directly  
66 involve the borders though, but rather to be caused by the self-aggregation of regions that  
67 carry the same epigenetic marks due to phase separation, thereby forming so-called ‘contact  
68 domains’ or ‘compartmental domains’ ([5, 18, 27-29], see also [30] in this JMB issue for  
69 further discussion). Unless mentioned otherwise, in the remainder of this review we focus on  
70 TADs where CTCF is bound at the borders.

71 Why CTCF is present at a large fraction of mammalian TAD borders, and if it is linked  
72 to transcriptional regulation, has been addressed in various pathologies and developmental  
73 contexts (see Table 1 for summary). The CTCF insulator protein was previously described as  
74 an ‘enhancer blocker’ that precludes transcriptional activation when bound in-between an  
75 enhancer-promoter pair [31]. Lupiáñez and colleagues showed that a group of congenital  
76 hand malformations could be traced to deletions, inversions and duplications around the  
77 *EPHA4* gene that had in common that they created chromosomal configurations where  
78 CTCF-bound sites between the enhancers of the *EPHA4* gene and neighboring inactive genes  
79 were removed [32]. In turn, ectopic enhancer-promoter interactions could be formed that  
80 activated these neighboring genes, which in turn caused the various hand malformations [32].  
81 Similarly, in Cooks syndrome, a duplication that spanned parts of two TADs and CTCF-  
82 bound border in between, resulted in the formation of a new TAD (‘neo-TAD’) where the  
83 genes encountered a reorganized regulatory landscape that imposed aberrant activation [33].  
84 In further analogy, mutations or epi-mutations of CTCF sites at TAD borders resulted in  
85 ectopic enhancer-promoter interactions and inappropriate gene activation in cancer cells as  
86 well [34, 35]. Combined, these studies suggest that the enhancer blocking function of CTCF  
87 is achieved by positioning genes and non-relevant enhancers in separate TADs.

88 More recent studies have provided nuance to the impact of perturbed CTCF binding at  
89 TAD borders. At the *Shh* locus, the limb-specific ZRS enhancer is located near the TAD  
90 border, where it is surrounded by multiple CTCF sites. Deletion of CTCF sites, either one or  
91 multiple, reduced *Shh* promoter-ZRS enhancer contacts but only had moderate influence on  
92 *Shh* gene activity. Removal of multiple CTCF sites increased interactions with neighboring  
93 sequences, which coincided with a mildly increased activity of a gene that is contained within  
94 this region [36, 37]. Similarly, at the *Sox9* locus, a fusion with the neighboring TAD required  
95 the deletion of 8 CTCF sites at and near the border. Yet, upon this fusion, the effect on gene  
96 activity was relatively mild, with *Sox9* activity moderately reduced and a twofold  
97 upregulation of the lowly expressed neighboring *Kcnj2* gene [38]. Combined, these studies  
98 show that TAD borders can be composed of more complex ensembles of CTCF-bound sites.

99 Moreover, the effect of TAD-border perturbations on gene expression may be relatively  
100 moderate. Intriguingly, acute depletion of CTCF in embryonic stem cells resulted in similar  
101 moderate changes to transcriptional output. After two days of CTCF removal, insulation  
102 between TADs was strongly reduced, yet the transcriptional deregulation of only nine  
103 hundred genes could be assigned to TAD deregulation [39]. In a specific cell type, the  
104 contribution of individual TADs on gene expression may thus be relatively small. This effect  
105 may be related to the cell-type specificity of enhancers, which reduces the chance of  
106 inappropriate enhancer-promoter contacts within a given cell type [40, 41].

107 In conclusion, genes and their associated enhancers are clustered in TADs, which act as  
108 functional but variable and context-dependent units of gene regulation. Demarcation and  
109 insulation of TADs are primarily dependent on the correct location of their borders, which are  
110 strongly enriched for the binding of the CTCF insulator protein (see also Table 1).

### 111 **Cohesin-dependent loop extrusion and TAD formation**

112 TADs can be identified from a Hi-C heat map, where they appear as pyramid-shaped  
113 domains with increased intra-domain interactions (Fig. 1A). Initial TAD-calling algorithms  
114 used a ‘Directionality Index’ to identify bin-by-bin transitions in the asymmetry of contacts,  
115 thus indicating the switch from one insulated TAD to the next [10]. More recently,  
116 algorithms to determine ‘Insulation Scores’ have been developed, which determine localized  
117 minima in interaction frequency by looking at a more regional scale, thus being more robust  
118 to individual outlying bins ([42, 43]; see [30] for further considerations).

119 How punctuated CTCF binding at TAD borders can result in increased intra-domain  
120 contacts, remains not fully established. An important mechanistic suggestion comes from the  
121 Hi-C data itself, where many TADs with CTCF at their borders display strongly enriched  
122 signal on the top of the pyramid [18]. This punctuated signal indicates that the borders of  
123 these domains loop together, which is particularly frequently observed when CTCF is bound  
124 at both sides. These TADs, with both increased intra-domain interactions and looping of the  
125 borders, are therefore also referred to as ‘loop domains’ [18, 44] (Fig. 1B). This feature  
126 distinguishes these domains from conventional chromatin loops, which are often associated  
127 with enhancer-promoter contacts, and ‘contact / compartmental domains’, which are  
128 associated with the aggregation of domains carrying identical histone modifications [5, 12, 18,  
129 26, 45] (Fig. 1B).

130 CTCF binds non-symmetric DNA motifs in the genome, where it often co-localizes  
131 with the ring shaped Cohesin complex [46-48]. Intriguingly, up to 90% of CTCF motifs that

132 are at the base of TAD-spanning loops are in a convergent orientation [18, 49, 50]. The  
133 importance of the correct orientation of CTCF motifs has been experimentally confirmed by  
134 artificial inversions, which result in reduced TAD insulation [49, 51].

135 How CTCF motif orientation guides TAD formation may be explained through the co-  
136 localization with the Cohesin complex. Based on this link, a ‘Cohesin-dependent loop  
137 extrusion’ model has been developed, which reproduces Hi-C data *in silico* with high degrees  
138 of similarity to experimentally obtained data [52-55]. In this model, the Cohesin complex  
139 randomly associates with the chromatin fiber, where it starts the progressive formation of a  
140 DNA loop through bi-directional extrusion (Fig. 1C). Loop extrusion is blocked, on one side  
141 only, when a CTCF molecule is encountered that is bound in the correct orientation. When  
142 Cohesin has encountered CTCF on both borders of the TAD, the loop is stabilized, as  
143 reflected by a visible loop in the Hi-C map (Fig. 1C). The visible enrichment of intra-domain  
144 interactions indicates that this process is simultaneously accompanied by more transient DNA  
145 contacts within the extruded loop, suggesting it adopts a more crumpled configuration.  
146 Together, these two patterns of 3D interactions are reflected as a loop domain in the Hi-C  
147 heat map (Fig. 1B).

148 Importantly, to reconstruct TADs *in silico*, the loop extrusion in the model needs to be  
149 a dynamic process. This implies that the Cohesin complex does not only continuously  
150 associate with the DNA at random positions in the genome, but also dissociates from the  
151 DNA after a certain amount of time [54]. In practice, this dissociation will mostly occur after  
152 Cohesin has co-occupied a CTCF-bound site for a certain period of time. Various  
153 experimental studies have dissected how the combined action of the Cohesin complex and  
154 CTCF influences TAD structure, thereby confirming key aspects of the Cohesin-dependent  
155 loop extrusion model. Rapid elimination of the Cohesin component Rad21/Scc1, or of its  
156 loading factor Nipbl, results in elimination of all TADs that appear as loop domains. In  
157 contrast, contact domains appear mostly untouched or even increased in structure [44, 56-58].  
158 These studies therefore uncouple the process of loop domain formation from that of contact  
159 domains. Loss of TADs upon mutations in the ATPase domains of Smc3, another Cohesin  
160 complex member, further confirms that loop extrusion is energy consuming, and thus an  
161 active mechanism [59]. In the same study, a subset of TADs was identified where one or both  
162 borders showed contiguous enriched contacts within the entire domain, visible as  
163 “architectural stripes”. These stripes are mostly found at TAD borders that combine strong  
164 insulation and multiple instances of CTCF binding with a nearby Cohesin loading area [59].  
165 At such borders, the Cohesin complex will be blocked instantly on one side. According to

166 loop extrusion model, such unidirectional loop formation is indeed predicted to manifest  
167 itself as horizontal and vertical stripes in the Hi-C matrix. Similar to Cohesin removal, the  
168 rapid removal of CTCF from TAD borders reduced the insulated nature of loop domains,  
169 whereas contact domains remained intact as well [39]. Interestingly though, the elimination  
170 of Cohesin and CTCF differentially affected genome compaction: whereas overall  
171 compaction was not affected by CTCF elimination, the loss of Cohesin association caused a  
172 global decompaction of the genome [39, 57]. This supports the model that Cohesin can create  
173 loops in the absence of CTCF, but that it needs to be blocked at specific sites in the genome  
174 to create noticeable loop domains.

175 An intriguing addition to the model of loop formation has been identified upon the  
176 deletion of Wapl, the protein that is required for the release of Cohesin from the DNA [58,  
177 60]. As expected from reinforced residence of the Cohesin complex at TAD borders, intra-  
178 TAD interactions become stronger. Unexpectedly though, TADs increased in size as well,  
179 which was visible as multiple TADs fusing into a single loop domain [60]. Prolonged  
180 blocking of Cohesin at a TAD borders may therefore result in a ‘read through’ of loop  
181 extrusion towards the next site occupied by CTCF in the correct orientation. This  
182 phenomenon may be due to a ‘traffic jam’ effect, whereby multiple TAD borders cluster  
183 together [61]. In turn, this may be caused by the dynamics of CTCF binding to the DNA,  
184 whose dissociation is considerably more dynamic than that of the Cohesin complex (~1-2  
185 minutes vs. ~20 minutes residence time) [62]. Cohesin-dependent loop extrusion may  
186 therefore not only depend on dynamic association and dissociation of the Cohesin complex,  
187 but may also involve a dynamic CTCF-binding component.

188 Combined, the results from *in silico* modeling and experimental studies provide  
189 accumulating evidence that TADs, as visible as loop domains in Hi-C heat maps, are shaped  
190 by a dynamic and active process that consists of a constant and energy-consuming bi-  
191 directional extrusion of DNA loops by the Cohesin complex. Within this process, CTCF  
192 functions to temporarily fix extruded chromatin loops at defined sites in the genome, which  
193 in turn creates specific insulated TADs.

## 194 **TADs as structural units of genome organization: consistency or** 195 **heterogeneity?**

196 How TADs are structured and how they are involved in context-depending transcriptional  
197 regulation is increasingly understood. In contrast, important questions remain on how stable  
198 TADs are over time, if they are consistent between different cells, and what is the nature of



199 their internal structure. Most insights on TAD organization have been obtained from Hi-C  
200 studies on large cell populations, whose averaged interaction data preclude the interpretation  
201 of truly dynamic aspects.

202         Nonetheless, a more detailed appraisal of population-based Hi-C heat maps does  
203 provide valuable insights on how the chromatin is organized within and between TADs.  
204 Whereas pairing of CTCF-bound borders in a loop domain is represented by the strong signal  
205 at the top of the triangle in the Hi-C heat map, the internal organization of a TAD is described  
206 by the signal within the triangle itself. Moreover, the degree of insulation between  
207 neighboring TADs is measured by the depletion of signal between two TADs (Fig. 1A;  
208 absence of signal in the ‘inversed’ pyramid between TADs). Similarly, the strength of  
209 borders between TADs influences insulation as well. Strong borders, visible as sharp valleys  
210 in the Hi-C heat map, are indicative of a separation between highly discrete TADs. In  
211 contrast, the presence of a more gradual valley between two TADs indicates a border that is  
212 more transitional and dynamic in nature (Fig. 2A). Inspection of high-resolution Hi-C data  
213 reveals that many TADs are separated by zones of transition that can span a considerable  
214 distance (Fig. 2B, see below for additional characterization). TAD-calling algorithms will  
215 call discrete borders within such zones of transition, yet the actual insulation at these zones is  
216 more gradual (Fig. 2B; dashed versus solid line in Hi-C heat map). In turn, according to the  
217 population-based Hi-C data, sequences near the borders do not strictly localize within one  
218 TAD (Fig. 2B; virtual 4C tracks). Although the presence of zones of transition may be due to  
219 artifacts in the Hi-C assay, similar interaction patterns are observed in the crosslink- and  
220 ligation-free DamC assay as well [63]. A more biological explanation for the appearance of  
221 these zones of transition may be that TAD borders can be dynamic, linked both to the  
222 discontinuous nature of CTCF binding and the observed clustering of CTCF binding sites at  
223 the majority of TAD borders [62, 64, 65]. In line with the model of Cohesin-dependent loop  
224 extrusion, the actual position of a TAD border may therefore not be fixed but vary within  
225 these defined zones of transition over time and between cells.

226         More recently developed single-cell Hi-C methods have started to shed light into the  
227 structural organization within and between TADs [56, 66-68]. Compared to population-based  
228 Hi-C, chromatin interactions within individual cells tend to be more stochastic and to often  
229 span multiple TADs. Yet, combination of single-cell Hi-C data from multiple individual cells  
230 tends to faithfully reproduce TAD structure as observed in population-based Hi-C data [67,  
231 68]. TADs in population-based Hi-C data may thus represent the statistically most prominent

232 average 3D configuration, which is built up from a variety of structures that is present within  
233 the cell population.

234 To get a more detailed and quantitative insight into the structural organization of TADs,  
235 we reanalyzed currently highest resolution mouse and human Hi-C data sets, both of which  
236 are commonly used as reference contact maps [13, 18]. Whereas TADs appear as (relatively)  
237 discrete pyramids in the Hi-C heat map of both data sets, a comparison of interaction  
238 frequencies between equally distant loci that are located either within or between TADs  
239 reveals only a maximum twofold degree of insulation (Fig. 2C). A similar moderate  
240 difference was previously reported for lower resolution mouse Hi-C data as well [10, 69]. A  
241 comparison of TAD structure and borders between the both mouse embryonic stem cell data  
242 sets confirms their similarities (Fig. S1A-D). The calling of these structural aspects of TAD  
243 organization is therefore highly robust, despite considerable experimental and analytical  
244 differences. Interestingly, the maximum difference between intra- and inter-TAD interaction  
245 frequencies is only observed at a considerable distance from the borders between the TADs  
246 (Fig. 2C; dashed lines at 290 and 785 kb). This distance for maximum insulation exceeds the  
247 span of many TADs and appears not reduced at smaller TADs (Fig. 2C and Fig S1E). The  
248 zones of transition that separate TADs therefore encompass both the borders and (at least  
249 partially) the domains themselves. In turn, these results suggest that the insulating capacity of  
250 TADs is relatively weak, yet this capacity scales with the size of the TAD.

251 Importantly though, a low level of insulation is not incompatible with the model of loop  
252 extrusion, as the extrusion process does not prevent the resulting loops in different domains  
253 from crossing borders and intermingling with each other. In contrast, it may suggest that the  
254 creation of regulatory units within TADs is not determined by the insulated nature of these  
255 domains alone. Rather, it may be that the ongoing loop extrusion that structures TADs is  
256 involved in guiding enhancer-promoter contacts as well. Indirect support for this hypothesis  
257 may be found at architectural stripes, where active (super-)enhancer elements that are in close  
258 proximity to stripe anchors have considerably elevated contacts with other regulatory  
259 elements [59]. Prevalent promoter-enhancer contacts between chromosomes in dipteran  
260 insects (including transvection) and an artificial mammalian system using an ectopic super-  
261 enhancer and its endogenous target gene have ruled out though that the formation of  
262 regulatory contacts depend on loop extrusion alone [70, 71].

263 Recent developments in super-resolution microscopy, particularly the combination of  
264 STORM (STochastic Optical Reconstruction Microscopy) with Oligopaint labeling [72],  
265 constitute an alternative approach to characterize TAD structure and insulation in individual

266 cells. Multiple studies in *Drosophila* cells were able to confirm the presence of TADs, albeit  
267 with a certain degree of heterogeneity and intermingling [73-75]. The small size of these  
268 domains and their overlap with specific histone modifications suggest they are more  
269 associated with contact domains, rather than the process of Cohesin-dependent loop extrusion.  
270 In mammals, larger spatially segregated nano-compartments that correspond to TADs could  
271 also be visualized in individual cells [76]. Similar to Hi-C data, an enrichment of intra-TAD  
272 interactions could be detected, yet with relatively little specificity. Moreover, frequent  
273 contacts between TADs were detected as well, but with a reduced consistency between cells.  
274 Similar patterns of moderately enriched, 2- to 3-fold, intra-TAD interactions were reported in  
275 a study that mapped large number of pair-wise interactions at a genome-wide scale [77].

276 Interestingly, the depletion of Cohesin in these single-cell microscopy studies did not  
277 result in a loss of TAD-like structures in individual cells, which contradicts published  
278 population-based Hi-C data [44, 57, 76]. In contrast, the preferential positioning of  
279 boundaries at CTCF bound sites was lost, which is more in line with single-cell Hi-C studies  
280 [67, 68]. The overall 3D organization of TADs may therefore be shaped by additional nested  
281 structures, including smaller loop domains that are linked to epigenetic state.

## 282 **Conclusions and Perspectives**

283 In this review, we have discussed the current evidence for the existence of TADs, how they  
284 are structured and how they are involved in transcriptional regulation. Whereas their presence  
285 can be unambiguously detected in population-based Hi-C heat maps, both quantitative  
286 interpretation of Hi-C data and the comparison of single-cell Hi-C and microscopy data  
287 suggest a certain degree of heterogeneity in TAD structure. As such, these data imply that  
288 TADs are dynamic structures, both regarding their internal structure and the position of their  
289 borders, with a relatively low degree of insulation from their surroundings. How this lack of  
290 structural consistency can be reconciled with their function as regulatory units in the genome  
291 remains to be determined. Here, we hypothesize that the setting of transcriptional programs  
292 may incorporate the process of TAD formation through loop extrusion as well.

293 To unambiguously link TAD formation and structure to transcriptional regulation, we  
294 will need to comprehensively address these aspects at highest precision in individual cells.  
295 Particularly instrumental will be the recent developments in super-resolution microscopy,  
296 including the combined visualization of DNA conformation and RNA [78, 79], and the  
297 development of genomics-based approaches that can identify and quantify ensembles of  
298 chromatin contacts. Recently, several technologies have emerged that either allow genome-

299 wide non-targeted characterization of chromatin hubs (C-walks [80], GAM (Genome  
300 Architecture Mapping) [81], SPRITE [82], and ChIA-Drop [83]) or the quantitative  
301 measurement of three-way interactions at defined sites (Multi-Contact 4C [61] and Tri-C  
302 [84]). The use of these technologies, and potential further developments, together with  
303 advanced modeling approaches and cellular models with modified aspects of TAD structure,  
304 will provide the tools to comprehensively unravel the dynamic nature of TADs.

## 305 **Methods**

### 306 *Reanalysis of published Hi-C data*

307 Unprocessed sequencing data from published Hi-C studies was obtained from the GEO data  
308 repository: human GM12878 B-lymphoblastoid cells: GSE63525 [18] and mouse embryonic  
309 stem cells: GSE96107 [13]. Data was processed at 5 kb resolution using the HiC-Pro tool,  
310 using standard settings and ICE-normalization of the resulting matrices [85]. TADs and their  
311 borders were called using the TADtool from the 5 kb Hi-C matrices based on insulation score  
312 [43]. Lower resolution Hi-C data [10], including normalized matrices and positions of TADs,  
313 were downloaded from <http://chromosome.sdsc.edu/mouse/hi-c/download.html>. For the  
314 comparison of TAD borders in mouse embryonic stem cell Hi-C data (Fig. S1C), data from  
315 [10] was converted to mm10 using the LiftOver tool (<http://genome.ucsc.edu>).  
316 Autosomal read pairs spanning 5 kb – 2 Mb were filtered for presence within TADs,  
317 followed by sorting for being within the same TAD or in different TADs. Median interaction  
318 signal for all read pairs was calculated based on category and distance.  
319

## 320 **Acknowledgments**

321 We thank the members of the Noordermeer lab for useful discussion. L.C. and S.G. are  
322 supported by funding from the Agence Nationale de Recherche to D.N. (grants ANR-14-  
323 ACHN-0009-01, ANR-16-TERC-0027-01 and ANR-17-CE12-0001-02).  
324

## 325 **Declarations of interest**

326 None  
327

## 328 **Author contributions**

329 S.G. reanalyzed published Hi-C data. L.-H.C. and D.N. wrote the manuscript with input from  
330 S.G.

331 **References**

- 332 [1] Mozziconacci J, Merle M, Lesne A. The 3D genome encodes gene regulation during  
333 development. *J Mol Biol. This issue (2019)*.
- 334 [2] Khorasanizadeh S. The nucleosome: from genomic organization to genomic  
335 regulation. *Cell*. 116 (2004) 259-272.
- 336 [3] Schalch T, Duda S, Sargent DF, Richmond TJ. X-ray structure of a tetranucleosome  
337 and its implications for the chromatin fibre. *Nature*. 436 (2005) 138-141.
- 338 [4] Fussner E, Strauss M, Djuric U, Li R, Ahmed K, Hart M, et al. Open and closed  
339 domains in the mouse genome are configured as 10-nm chromatin fibres. *EMBO Rep*. 13  
340 (2012) 992-996.
- 341 [5] Rowley MJ, Corces VG. Organizational principles of 3D genome architecture. *Nat*  
342 *Rev Genet*. 19 (2018) 789-800.
- 343 [6] Stadhouders R, Fillion GJ, Graf T. Transcription factors and 3D genome conformation  
344 in cell-fate decisions. *Nature*. 569 (2019) 345-354.
- 345 [7] Dekker J, Rippe K, Dekker M, Kleckner N. Capturing chromosome conformation.  
346 *Science*. 295 (2002) 1306-1311.
- 347 [8] Davies JO, Oudelaar AM, Higgs DR, Hughes JR. How best to identify chromosomal  
348 interactions: a comparison of approaches. *Nat Methods*. 14 (2017) 125-134.
- 349 [9] Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragozy T, Telling A,  
350 et al. Comprehensive mapping of long-range interactions reveals folding principles of the  
351 human genome. *Science*. 326 (2009) 289-293.
- 352 [10] Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, et al. Topological domains in  
353 mammalian genomes identified by analysis of chromatin interactions. *Nature*. 485 (2012)  
354 376-380.
- 355 [11] Nora EP, Lajoie BR, Schulz EG, Giorgetti L, Okamoto I, Servant N, et al. Spatial  
356 partitioning of the regulatory landscape of the X-inactivation centre. *Nature*. 485 (2012) 381-  
357 385.
- 358 [12] Sexton T, Yaffe E, Kenigsberg E, Bantignies F, Leblanc B, Hoichman M, et al.  
359 Three-dimensional folding and functional organization principles of the *Drosophila* genome.  
360 *Cell*. 148 (2012) 458-472.
- 361 [13] Bonev B, Mendelson Cohen N, Szabo Q, Fritsch L, Papadopoulos GL, Lubling Y, et  
362 al. Multiscale 3D Genome Rewiring during Mouse Neural Development. *Cell*. 171 (2017)  
363 557-572 e524.

364 [14] Liu C, Cheng YJ, Wang JW, Weigel D. Prominent topologically associated domains  
365 differentiate global chromatin packing in rice from Arabidopsis. *Nat Plants*. 3 (2017) 742-748.

366 [15] Kaaij LJT, van der Weide RH, Ketting RF, de Wit E. Systemic Loss and Gain of  
367 Chromatin Architecture throughout Zebrafish Development. *Cell Rep*. 24 (2018) 1-10 e14.

368 [16] Fishman V, Battulin N, Nuriddinov M, Maslova A, Zlotina A, Strunov A, et al. 3D  
369 organization of chicken genome demonstrates evolutionary conservation of topologically  
370 associated domains and highlights unique architecture of erythrocytes' chromatin. *Nucleic  
371 Acids Res*. 47 (2019) 648-665.

372 [17] Phillips-Cremins JE, Sauria ME, Sanyal A, Gerasimova TI, Lajoie BR, Bell JS, et al.  
373 Architectural protein subclasses shape 3D organization of genomes during lineage  
374 commitment. *Cell*. 153 (2013) 1281-1295.

375 [18] Rao SSP, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, et al.  
376 A 3D map of the human genome at kilobase resolution reveals principles of chromatin  
377 looping. *Cell*. 159 (2014) 1665-1680.

378 [19] Hsieh T-HS, Slobodyanyuk E, Hansen AS, Cattoglio C, Rando OJ, Tjian R, et al.  
379 Resolving the 3D landscape of transcription-linked mammalian chromatin folding. *bioRxiv*.  
380 (2019) 638775.

381 [20] Krietenstein N, Abraham S, Venev SV, Abdennur N, Gibcus JH, Hsieh T-HS, et al.  
382 Ultrastructural details of mammalian chromosome architecture. *bioRxiv*. (2019) 639922.

383 [21] Sanyal A, Lajoie BR, Jain G, Dekker J. The long-range interaction landscape of gene  
384 promoters. *Nature*. 489 (2012) 109-113.

385 [22] Shen Y, Yue F, McCleary DF, Ye Z, Edsall L, Kuan S, et al. A map of the cis-  
386 regulatory sequences in the mouse genome. *Nature*. 488 (2012) 116-120.

387 [23] Hou C, Li L, Qin ZS, Corces VG. Gene density, transcription, and insulators  
388 contribute to the partition of the *Drosophila* genome into physical domains. *Mol Cell*. 48  
389 (2012) 471-484.

390 [24] Le TB, Imakaev MV, Mirny LA, Laub MT. High-resolution mapping of the spatial  
391 organization of a bacterial chromosome. *Science*. 342 (2013) 731-734.

392 [25] Eser U, Chandler-Brown D, Ay F, Straight AF, Duan Z, Noble WS, et al. Form and  
393 function of topologically associating genomic domains in budding yeast. *Proc Natl Acad Sci  
394 U S A*. 114 (2017) E3061-E3070.

395 [26] Noordermeer D, Leleu M, Splinter E, Rougemont J, De Laat W, Duboule D. The  
396 dynamic architecture of Hox gene clusters. *Science*. 334 (2011) 222-225.

397 [27] Larson AG, Elnatan D, Keenen MM, Trnka MJ, Johnston JB, Burlingame AL, et al.  
398 Liquid droplet formation by HP1alpha suggests a role for phase separation in  
399 heterochromatin. *Nature*. 547 (2017) 236-240.

400 [28] Strom AR, Emelyanov AV, Mir M, Fyodorov DV, Darzacq X, Karpen GH. Phase  
401 separation drives heterochromatin domain formation. *Nature*. 547 (2017) 241-245.

402 [29] Tatavosian R, Kent S, Brown K, Yao T, Duc HN, Huynh TN, et al. Nuclear  
403 condensates of the Polycomb protein chromobox 2 (CBX2) assemble through phase  
404 separation. *J Biol Chem*. 294 (2019) 1451-1463.

405 [30] de Wit E. TADs as the caller calls them. *J Mol Biol*. This issue (2019).

406 [31] Merkenschlager M, Nora EP. CTCF and Cohesin in Genome Folding and  
407 Transcriptional Gene Regulation. *Annu Rev Genomics Hum Genet*. 17 (2016) 17-43.

408 [32] Lupianez DG, Kraft K, Heinrich V, Krawitz P, Brancati F, Klopocki E, et al.  
409 Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer  
410 interactions. *Cell*. 161 (2015) 1012-1025.

411 [33] Franke M, Ibrahim DM, Andrey G, Schwarzer W, Heinrich V, Schopflin R, et al.  
412 Formation of new chromatin domains determines pathogenicity of genomic duplications.  
413 *Nature*. 538 (2016) 265-269.

414 [34] Flavahan WA, Drier Y, Liao BB, Gillespie SM, Venteicher AS, Stemmer-  
415 Rachamimov AO, et al. Insulator dysfunction and oncogene activation in IDH mutant  
416 gliomas. *Nature*. 529 (2016) 110-114.

417 [35] Hnisz D, Weintraub AS, Day DS, Valton AL, Bak RO, Li CH, et al. Activation of  
418 proto-oncogenes by disruption of chromosome neighborhoods. *Science*. 351 (2016) 1454-  
419 1458.

420 [36] Paliou C, Guckelberger P, Schopflin R, Heinrich V, Esposito A, Chiariello AM, et al.  
421 Preformed chromatin topology assists transcriptional robustness of Shh during limb  
422 development. *Proc Natl Acad Sci U S A*. (2019).

423 [37] Williamson I, Kane L, Devenney PS, Anderson E, Kilanowski F, Hill RE, et al.  
424 Developmentally regulated Shh expression is robust to TAD perturbations. *bioRxiv*. (2019)  
425 609941.

426 [38] Despang A, Schopflin R, Franke M, Ali S, Jerkovic I, Paliou C, et al. Functional  
427 dissection of the Sox9-Kcnj2 locus identifies nonessential and instructive roles of TAD  
428 architecture. *Nat Genet*. 51 (2019) 1263-1271.



429 [39] Nora EP, Goloborodko A, Valton AL, Gibcus JH, Uebersohn A, Abdennur N, et al.  
430 Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from  
431 Genomic Compartmentalization. *Cell*. 169 (2017) 930-944 e922.

432 [40] Heintzman ND, Hon GC, Hawkins RD, Kheradpour P, Stark A, Harp LF, et al.  
433 Histone modifications at human enhancers reflect global cell-type-specific gene expression.  
434 *Nature*. 459 (2009) 108-112.

435 [41] Visel A, Blow MJ, Li Z, Zhang T, Akiyama JA, Holt A, et al. ChIP-seq accurately  
436 predicts tissue-specific activity of enhancers. *Nature*. 457 (2009) 854-858.

437 [42] Crane E, Bian Q, McCord RP, Lajoie BR, Wheeler BS, Ralston EJ, et al. Condensin-  
438 driven remodelling of X chromosome topology during dosage compensation. *Nature*. 523  
439 (2015) 240-244.

440 [43] Kruse K, Hug CB, Hernandez-Rodriguez B, Vaquerizas JM. TADtool: visual  
441 parameter identification for TAD-calling algorithms. *Bioinformatics*. 32 (2016) 3190-3192.

442 [44] Rao SSP, Huang SC, Glenn St Hilaire B, Engreitz JM, Perez EM, Kieffer-Kwon KR,  
443 et al. Cohesin Loss Eliminates All Loop Domains. *Cell*. 171 (2017) 305-320 e324.

444 [45] Tolhuis B, Palstra RJ, Splinter E, Grosveld F, de Laat W. Looping and interaction  
445 between hypersensitive sites in the active beta-globin locus. *Mol Cell*. 10 (2002) 1453-1465.

446 [46] Kim TH, Abdullaev ZK, Smith AD, Ching KA, Loukinov DI, Green RD, et al.  
447 Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. *Cell*.  
448 128 (2007) 1231-1245.

449 [47] Parelho V, Hadjur S, Spivakov M, Leleu M, Sauer S, Gregson HC, et al. Cohesins  
450 functionally associate with CTCF on mammalian chromosome arms. *Cell*. 132 (2008) 422-  
451 433.

452 [48] Wendt KS, Yoshida K, Itoh T, Bando M, Koch B, Schirghuber E, et al. Cohesin  
453 mediates transcriptional insulation by CCCTC-binding factor. *Nature*. 451 (2008) 796-801.

454 [49] de Wit E, Vos ES, Holwerda SJ, Valdes-Quezada C, Verstegen MJ, Teunissen H, et al.  
455 CTCF Binding Polarity Determines Chromatin Looping. *Mol Cell*. 60 (2015) 676-684.

456 [50] Tang Z, Luo OJ, Li X, Zheng M, Zhu JJ, Szalaj P, et al. CTCF-Mediated Human 3D  
457 Genome Architecture Reveals Chromatin Topology for Transcription. *Cell*. 163 (2015) 1611-  
458 1627.

459 [51] Guo Y, Xu Q, Canzio D, Shou J, Li J, Gorkin DU, et al. CRISPR Inversion of CTCF  
460 Sites Alters Genome Topology and Enhancer/Promoter Function. *Cell*. 162 (2015) 900-910.

461 [52] Alipour E, Marko JF. Self-organization of domain structures by DNA-loop-extruding  
462 enzymes. *Nucleic Acids Res*. 40 (2012) 11202-11212.

463 [53] Sanborn AL, Rao SS, Huang SC, Durand NC, Huntley MH, Jewett AI, et al.  
464 Chromatin extrusion explains key features of loop and domain formation in wild-type and  
465 engineered genomes. *Proc Natl Acad Sci U S A*. 112 (2015) E6456-6465.

466 [54] Fudenberg G, Imakaev M, Lu C, Goloborodko A, Abdennur N, Mirny LA. Formation  
467 of Chromosomal Domains by Loop Extrusion. *Cell Rep*. 15 (2016) 2038-2049.

468 [55] Nuebler J, Fudenberg G, Imakaev M, Abdennur N, Mirny LA. Chromatin  
469 organization by an interplay of loop extrusion and compartmental segregation. *Proc Natl*  
470 *Acad Sci U S A*. 115 (2018) E6697-E6706.

471 [56] Gassler J, Brandao HB, Imakaev M, Flyamer IM, Ladstatter S, Bickmore WA, et al.  
472 A mechanism of cohesin-dependent loop extrusion organizes zygotic genome architecture.  
473 *EMBO J*. 36 (2017) 3600-3618.

474 [57] Schwarzer W, Abdennur N, Goloborodko A, Pekowska A, Fudenberg G, Loe-Mie Y,  
475 et al. Two independent modes of chromatin organization revealed by cohesin removal.  
476 *Nature*. 551 (2017) 51-56.

477 [58] Wutz G, Varnai C, Nagasaka K, Cisneros DA, Stocsits RR, Tang W, et al.  
478 Topologically associating domains and chromatin loops depend on cohesin and are regulated  
479 by CTCF, WAPL, and PDS5 proteins. *EMBO J*. 36 (2017) 3573-3599.

480 [59] Vian L, Pekowska A, Rao SSP, Kieffer-Kwon KR, Jung S, Baranello L, et al. The  
481 Energetics and Physiological Impact of Cohesin Extrusion. *Cell*. 173 (2018) 1165-1178  
482 e1120.

483 [60] Haarhuis JHI, van der Weide RH, Blomen VA, Yanez-Cuna JO, Amendola M, van  
484 Ruiten MS, et al. The Cohesin Release Factor WAPL Restricts Chromatin Loop Extension.  
485 *Cell*. 169 (2017) 693-707 e614.

486 [61] Allahyar A, Vermeulen C, Bouwman BAM, Krijger PHL, Verstegen M, Geeven G, et  
487 al. Enhancer hubs and loop collisions identified from single-allele topologies. *Nat Genet*. 50  
488 (2018) 1151-1160.

489 [62] Hansen AS, Pustova I, Cattoglio C, Tjian R, Darzacq X. CTCF and cohesin regulate  
490 chromatin loop stability with distinct dynamics. *Elife*. 6 (2017).

491 [63] Redolfi J, Zhan Y, Valdes-Quezada C, Kryzhanovska M, Guerreiro I, Iesmantavicius  
492 V, et al. DamC reveals principles of chromatin folding in vivo without crosslinking and  
493 ligation. *Nat Struct Mol Biol*. 26 (2019) 471-480.

494 [64] Madani Tonekaboni SA, Mazrooei P, Kofia V, Haibe-Kains B, Lupien M. CREAM:  
495 Clustering of genomic REgions Analysis Method. *bioRxiv*. (2018) 222562.

496 [65] Kentepozidou E, Aitken SJ, Feig C, Stefflova K, Ibarra-Soria X, Odom DT, et al.  
497 Clustered CTCF binding is an evolutionary mechanism to maintain topologically associating  
498 domains. *bioRxiv*. (2019) 668855.

499 [66] Nagano T, Lubling Y, Stevens TJ, Schoenfelder S, Yaffe E, Dean W, et al. Single-cell  
500 Hi-C reveals cell-to-cell variability in chromosome structure. *Nature*. 502 (2013) 59-64.

501 [67] Flyamer IM, Gassler J, Imakaev M, Brandao HB, Ulianov SV, Abdennur N, et al.  
502 Single-nucleus Hi-C reveals unique chromatin reorganization at oocyte-to-zygote transition.  
503 *Nature*. 544 (2017) 110-114.

504 [68] Stevens TJ, Lando D, Basu S, Atkinson LP, Cao Y, Lee SF, et al. 3D structures of  
505 individual mammalian genomes studied by single-cell Hi-C. *Nature*. 544 (2017) 59-64.

506 [69] Dekker J, Mirny L. The 3D Genome as Moderator of Chromosomal Communication.  
507 *Cell*. 164 (2016) 1110-1121.

508 [70] Pirrotta V. Transvection and chromosomal trans-interaction effects. *Biochim Biophys*  
509 *Acta*. 1424 (1999) M1-8.

510 [71] Noordermeer D, de Wit E, Klous P, van de Werken H, Simonis M, Lopez-Jones M, et  
511 al. Variegated gene expression caused by cell-specific long-range DNA interactions. *Nat Cell*  
512 *Biol*. 13 (2011) 944-951.

513 [72] Beliveau BJ, Boettiger AN, Avendano MS, Jungmann R, McCole RB, Joyce EF, et al.  
514 Single-molecule super-resolution imaging of chromosomes and in situ haplotype  
515 visualization using Oligopaint FISH probes. *Nat Commun*. 6 (2015) 7147.

516 [73] Boettiger AN, Bintu B, Moffitt JR, Wang S, Beliveau BJ, Fudenberg G, et al. Super-  
517 resolution imaging reveals distinct chromatin folding for different epigenetic states. *Nature*.  
518 529 (2016) 418-422.

519 [74] Cattoni DI, Cardozo Gizzi AM, Georgieva M, Di Stefano M, Valeri A, Chamousset D,  
520 et al. Single-cell absolute contact probability detection reveals chromosomes are organized  
521 by multiple low-frequency yet specific interactions. *Nat Commun*. 8 (2017) 1753.

522 [75] Szabo Q, Jost D, Chang JM, Cattoni DI, Papadopoulos GL, Bonev B, et al. TADs are  
523 3D structural units of higher-order chromosome organization in *Drosophila*. *Sci Adv*. 4 (2018)  
524 eaar8082.

525 [76] Bintu B, Mateo LJ, Su JH, Sinnott-Armstrong NA, Parker M, Kinrot S, et al. Super-  
526 resolution chromatin tracing reveals domains and cooperative interactions in single cells.  
527 *Science*. 362 (2018).

528 [77] Finn EH, Pegoraro G, Brandao HB, Valton AL, Oomen ME, Dekker J, et al.  
529 Extensive Heterogeneity and Intrinsic Variation in Spatial Genome Organization. *Cell*. 176  
530 (2019) 1502-1515 e1510.

531 [78] Cardozo Gizzi AM, Cattoni DI, Fiche J-B, Espinola SM, Gurgo J, Messina O, et al.  
532 Microscopy-Based Chromosome Conformation Capture Enables Simultaneous Visualization  
533 of Genome Organization and Transcription in Intact Organisms. *Molecular Cell*. 74 (2019)  
534 212-222.e215.

535 [79] Mateo LJ, Murphy SE, Hafner A, Cinquini IS, Walker CA, Boettiger AN. Visualizing  
536 DNA folding and RNA in embryos at single-cell resolution. *Nature*. 568 (2019) 49-54.

537 [80] Olivares-Chauvet P, Mukamel Z, Lifshitz A, Schwartzman O, Elkayam NO, Lubling  
538 Y, et al. Capturing pairwise and multi-way chromosomal conformations using chromosomal  
539 walks. *Nature*. 540 (2016) 296-300.

540 [81] Beagrie RA, Scialdone A, Schueler M, Kraemer DC, Chotalia M, Xie SQ, et al.  
541 Complex multi-enhancer contacts captured by genome architecture mapping. *Nature*. 543  
542 (2017) 519-524.

543 [82] Quinodoz SA, Ollikainen N, Tabak B, Palla A, Schmidt JM, Detmar E, et al. Higher-  
544 Order Inter-chromosomal Hubs Shape 3D Genome Organization in the Nucleus. *Cell*. 174  
545 (2018) 744-757 e724.

546 [83] Zheng M, Tian SZ, Capurso D, Kim M, Maurya R, Lee B, et al. Multiplex chromatin  
547 interactions with single-molecule precision. *Nature*. 566 (2019) 558-562.

548 [84] Oudelaar AM, Davies JOJ, Hanssen LLP, Telenius JM, Schwessinger R, Liu Y, et al.  
549 Single-allele chromatin interactions identify regulatory hubs in dynamic compartmentalized  
550 domains. *Nat Genet*. 50 (2018) 1744-1751.

551 [85] Servant N, Varoquaux N, Lajoie BR, Viara E, Chen CJ, Vert JP, et al. HiC-Pro: an  
552 optimized and flexible pipeline for Hi-C data processing. *Genome Biol*. 16 (2015) 259.

553 [86] Andrey G, Montavon T, Mascrez B, Gonzalez F, Noordermeer D, Leleu M, et al. A  
554 switch between topological domains underlies HoxD genes collinearity in mouse limbs.  
555 *Science*. 340 (2013) 1195.

556 [87] Sun JH, Zhou L, Emerson DJ, Phyo SA, Titus KR, Gong W, et al. Disease-Associated  
557 Short Tandem Repeats Co-localize with Chromatin Domain Boundaries. *Cell*. 175 (2018)  
558 224-238 e215.

559

## 560 **Figure and table legends**

### 561 ***Figure 1. TADs and loop extrusion.***

- 562 **A.** Schematic of two TADs in a Hi-C heat map, with the arrow indicating the border  
563 between the TADs. Simulated CTCF binding is depicted below, with convergently  
564 oriented CTCF peaks at the borders of both TADs (orientation of CTCF sites shown as  
565 blue and red triangles).
- 566 **B.** Appearance of different types of chromatin structures in a Hi-C heat map. Left: cartoon  
567 of 3D chromatin structure; right: corresponding simplified Hi-C heat maps.
- 568 **C.** Model for Cohesin-dependent loop extrusion. After random association of the Cohesin  
569 complex with the chromosome (**1, 2**), a loop will start to bi-directionally extrude (**2, 3**).  
570 Extrusion on either side will continue until a correctly oriented CTCF-bound site is  
571 encountered (left blue triangle; **4**). A complete block is only achieved when both sides  
572 have encountered convergently orientated CTCFs (blue and red triangles; **5**).

### 573 ***Figure 2. The insulated nature of TADs and their borders.***

- 574 **A.** Schematic of strong and weak TAD borders. Weak TAD borders are separated by a  
575 ‘zone of transition’ that appears as a gradual valley between two TADs.
- 576 **B.** Reanalyzed Hi-C data showing a nearly 100 kb ‘zone of transition’ between two TADs  
577 in mouse embryonic stem cells [13]. Top: Hi-C heat map with the algorithmically called  
578 TAD border indicated with dotted lines, and the more gradual transition between the  
579 TADs with a solid line. Middle: virtual 4C signal for two bins at either side of the zone  
580 of transition [12]. Bottom: published CTCF ChIP-seq signal in mouse embryonic stem  
581 cells [22].
- 582 **C.** Genome-wide assessment of TAD insulation in human GM12878 B-lymphoblastoid  
583 cells [18] (red) and mouse embryonic stem cells [13] (blue) using reanalyzed Hi-C data  
584 at 5 kb resolution. Left: median genome-wide interaction counts for intra-TAD (solid  
585 lines) and inter-TAD (dashed lines) read-pairs at distances from 5 kb to 2 Mb. Right:  
586 enrichment of intra-TAD signal at different distance scales. Vertical dotted lines indicate  
587 the distance between read pairs with maximum signal enrichment (human B-  
588 lymphoblastoid cells: 290 kb – mouse embryonic stem cells: 785 kb).

### 589 ***Table 1. Key studies showing the influence of TAD borders on gene expression***

590

591 **Supplemental figure legend**

592 **Figure S1. Comparison of reanalyzed Hi-C data in mouse embryonic stem cells**

593 **A.** Example of published and reanalyzed Hi-C data from two different studies in mouse  
594 embryonic stem cells. Top: 40 kb resolution Hi-C data from Dixon *et al.* [10]; bottom: 10  
595 kb resolution Hi-C data from Bonev *et al.* [13]. Despite biological, experimental and  
596 analytical differences (left), an overall similar framework of TADs is called. Differences  
597 between both data sets are indicated above (see also panel **B.**). 10 Mb region was  
598 selected for the presence of multiple types of differences between data sets.

599 **B.** Differences in TAD calling between two different studies in mouse embryonic stem cells.  
600 Except for a minor fraction of TADs (category “Disagreement spanning multiple TADs”,  
601 making up 2% of TADs in Dixon *et al.*), differences can be traced back to TADs or  
602 individual borders that are called in only one of two data sets (see panel **A.** for examples  
603 of categories).

604 **C.** Differences in TAD border calling between two different studies in mouse embryonic  
605 stem cells. Two-thirds of TAD borders that can be unambiguously linked between data  
606 sets map within 80 kb, despite the initial difference in resolution (40 kb *vs.* 5 kb  
607 resolution) and mapping to different versions of the mouse genome (mm10 *vs.* mm9) for  
608 both data sets.

609 **D.** Genome-wide assessment of TAD insulation between two different studies in mouse  
610 embryonic stem cells. In both data sets, a comparably moderate degree of insulation and  
611 a maximum insulation at a considerable distance from the borders is observed. Left:  
612 median genome-wide interaction counts for intra-TAD (solid lines) and inter-TAD  
613 (dashed lines) read-pairs. Right: enrichment of intra-TAD signal at different distance  
614 scales. Notice that data from Dixon *et al.* is visualized at a 40 kb resolution.

615 **E.** Genome-wide assessment of differences in insulation between TADs in mouse  
616 embryonic stem with spans below or above median size (median = 615 kb) [13]. Smaller  
617 TADs attain higher intra-TAD signal enrichment, yet this maximum occurs at a larger  
618 distance from the borders as compared to large TADs. Left: median genome-wide  
619 interaction counts. Right: enrichment of intra-TAD signal at different distance scales.  
620 TADs with sizes below median are limited to their maximum span, with signal becoming  
621 highly variable in the last bins due to scarcity of the signal (dotted line).