

Additional file 2: Supplementary Figures S1–S12

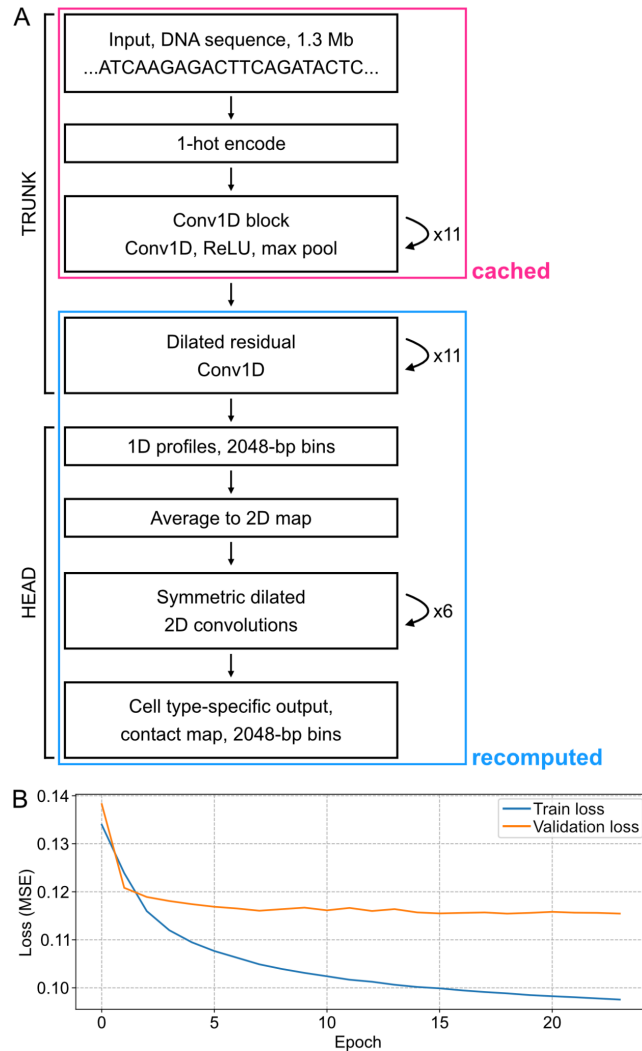


Figure S1. Akita architecture and example fine-tuning training curves.

A) Schematic of the Akita architecture, adapted from Fudenberg et al. [34]. A one-hot-encoded DNA sequence is passed through an initial 1D convolutional block followed by a series of dilated residual blocks, producing 1D profiles binned at 2048 bp resolution — these layers constitute the trunk. The head then averages the 1D profiles into a 2D representation and applies symmetric dilated 2D convolutions to yield cell-type-specific contact maps. The pink box marks the output of the convolutional tower (a Conv1D–ReLU–max-pool block repeated 11 times), which is cached; the turquoise box marks the layers responsible for long-distance communication and transformation of the 1D representation to a 2D map, which are recomputed during Akita^{SF} sequence optimization.

B) Shown are training dynamics from framework-transfer fine-tuning (TensorFlow to PyTorch) for one Akita model fine-tuned to mESC data [36]. Training loss (blue) and validation loss (orange) both exhibit a large improvement during the first epoch, followed by rapid stabilization and plateauing over subsequent epochs.

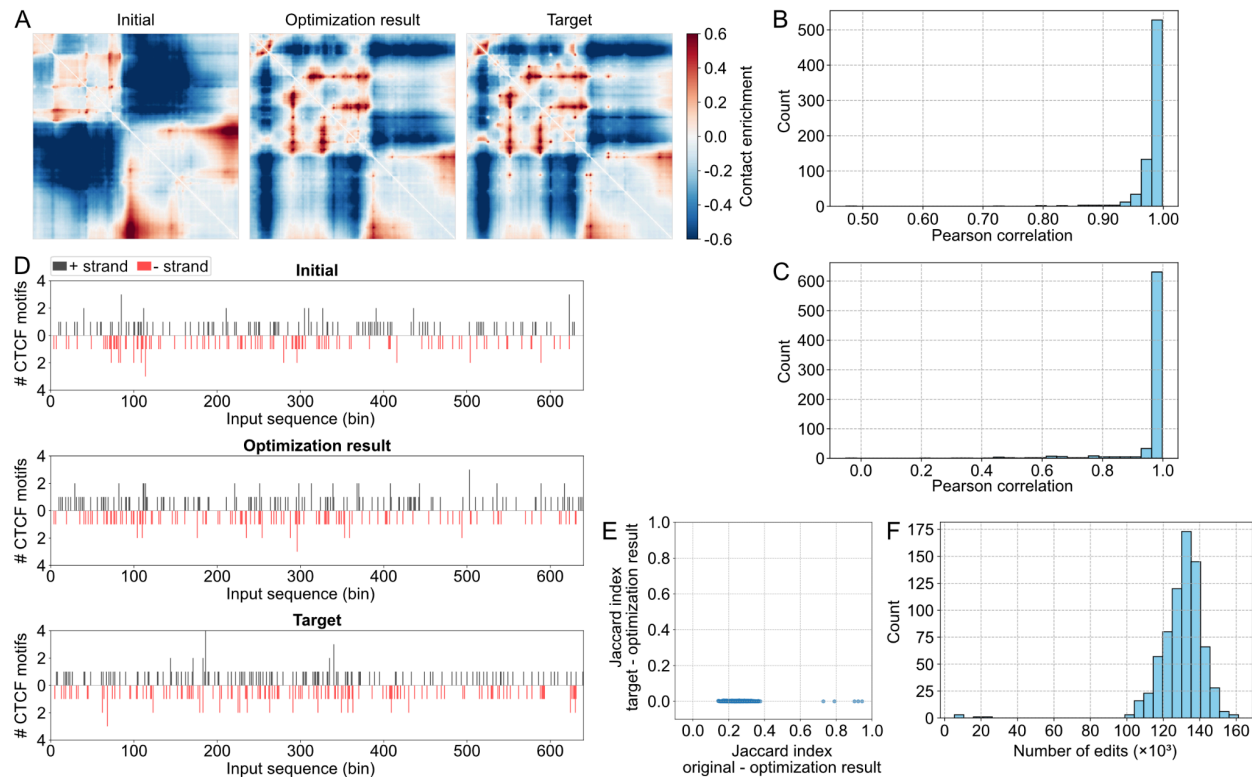


Figure S2. Akita_{PT} combined with the Ledidi optimizer can “rewire” genome folding.

A) Three contact maps are shown: the original genome folding prediction, the optimized genome folding prediction, and the target genome folding prediction.

B) Histogram of Pearson's correlation between the optimized and target genome folding predictions for 725 genome folding “rewiring” optimizations.

C) As in B, but optimizations were initialized from shuffled sequences (k-mer-preserving shuffles with featureless predicted maps) rather than genomic windows. The large majority of optimized maps were highly similar to their targets (Pearson's $r > 0.9$), comparable to genomic initialization.

D) CTCF motif landscapes for the initial (top), optimized (middle), and the sequence corresponding to the target maps (bottom). The x-axis spans 640 bins (~1.3 Mb). Positive black bars indicate forward-strand CTCF motifs, and negative red bars indicate reverse-strand motifs.

E) Overlap fraction of CTCF motifs between the initial and optimized sequences, and between the optimized and target sequences. Motifs were considered overlapping only if they matched in both genomic location and strand. The optimized sequences remain more similar to the original sequences than to those underlying the target folding patterns.

F) Histogram showing the number of accepted sequence edits required to achieve map-to-map transformations for 725 genome folding “rewiring” optimizations. Edits were allowed across the entire input sequence (~1.3 Mb).

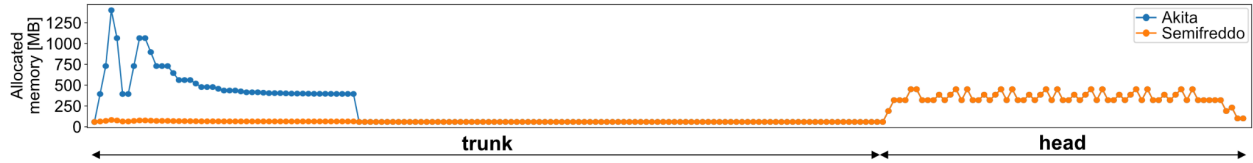


Figure S3. Akita^{SF} vs. full Akita memory usage.

Allocated GPU memory per layer for the full Akita model (blue) and the Akita^{SF} model (orange), averaged over 2,400 predictions with central CTCF insertions. The x-axis indicates layer indices in forward order. Akita^{SF} bypasses the memory-intensive convolutional tower, avoiding the large memory peak.

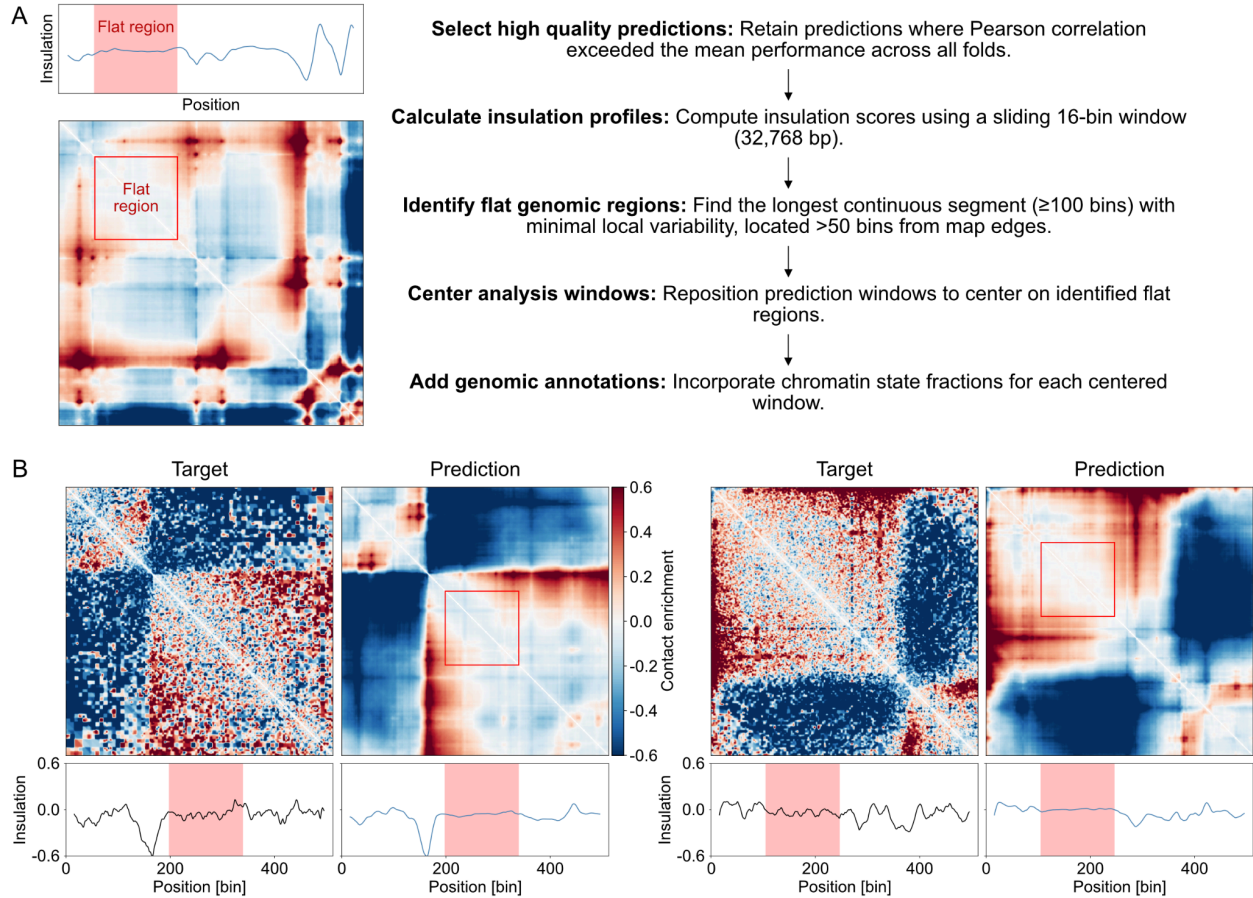


Figure S4. Selection of optimization input sequences.

A) Selection of optimization input sequences. We began with a single cross-validation fold (1/8 of the training data). For each prediction window in this fold, we computed Pearson's correlation between the model prediction and the target. Only windows with Pearson's r above the fold-specific average were retained, ensuring that optimization starts from regions the model predicts well. Next, we identified long continuous segments (≥ 100 bins, ~ 200 kb) in which the insulation profile showed minimal local variability and which were located away from map edges (>50 bins from either edge). Prediction windows were then centered on these flat regions. For each selected window, chromatin states were computed as in [6].

B) Two example target-prediction pairs with detected flat regions highlighted as red rectangles on the predicted maps. Insulation profiles aligned below each map also show the corresponding flat regions (red), indicating where the selection criteria were met.

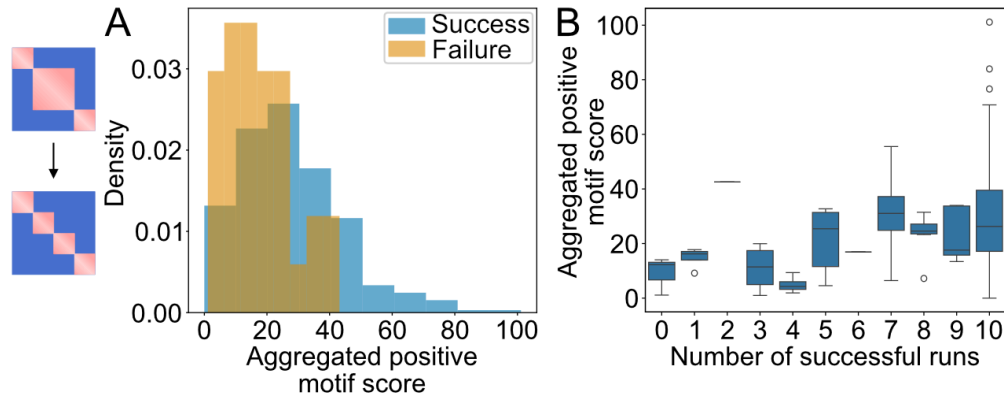
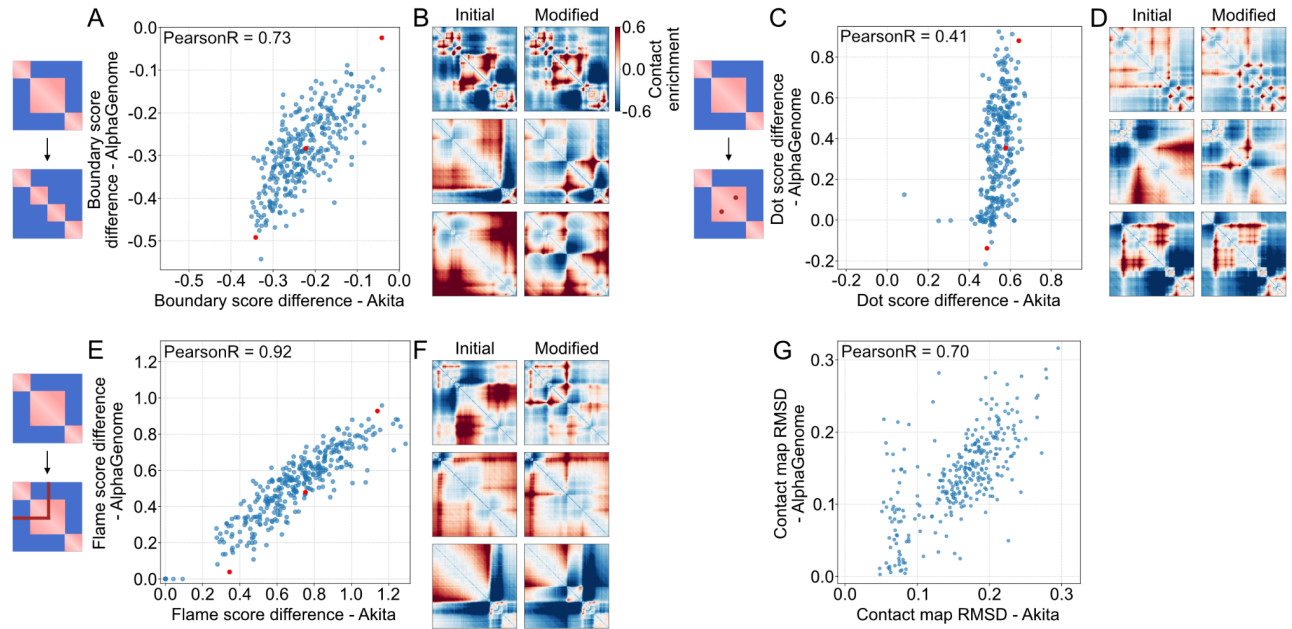


Figure S5. Boundary optimization success depends on aggregated positive CTCF motif scores.

A) To understand why strong boundary optimizations failed for some sequences, we computed an aggregated positive motif score, defined as the sum of all positive CTCF motif scores across the entire edited sequence prior to optimization. Comparison of successfully optimized sequences (green) with failed sequences (red) revealed that failed optimizations tend to exhibit lower aggregated positive motif scores. This metric was more predictive of optimization success than other pre-existing sequence features, including: (i) the number of positive motif instances, (ii) the sum of all motif scores (positive and negative), and (iii) the maximum motif score.

B) Each of the 174 sequences was subjected to strong boundary optimization across 10 independent runs. Boxplots show the distribution of aggregated positive motif scores as a function of the number of successful optimizations (out of 10 independent runs). Sequences with more frequent optimization success tend to have higher aggregated positive motif scores. Note that only one sequence displayed either 2 or 6 successful runs.



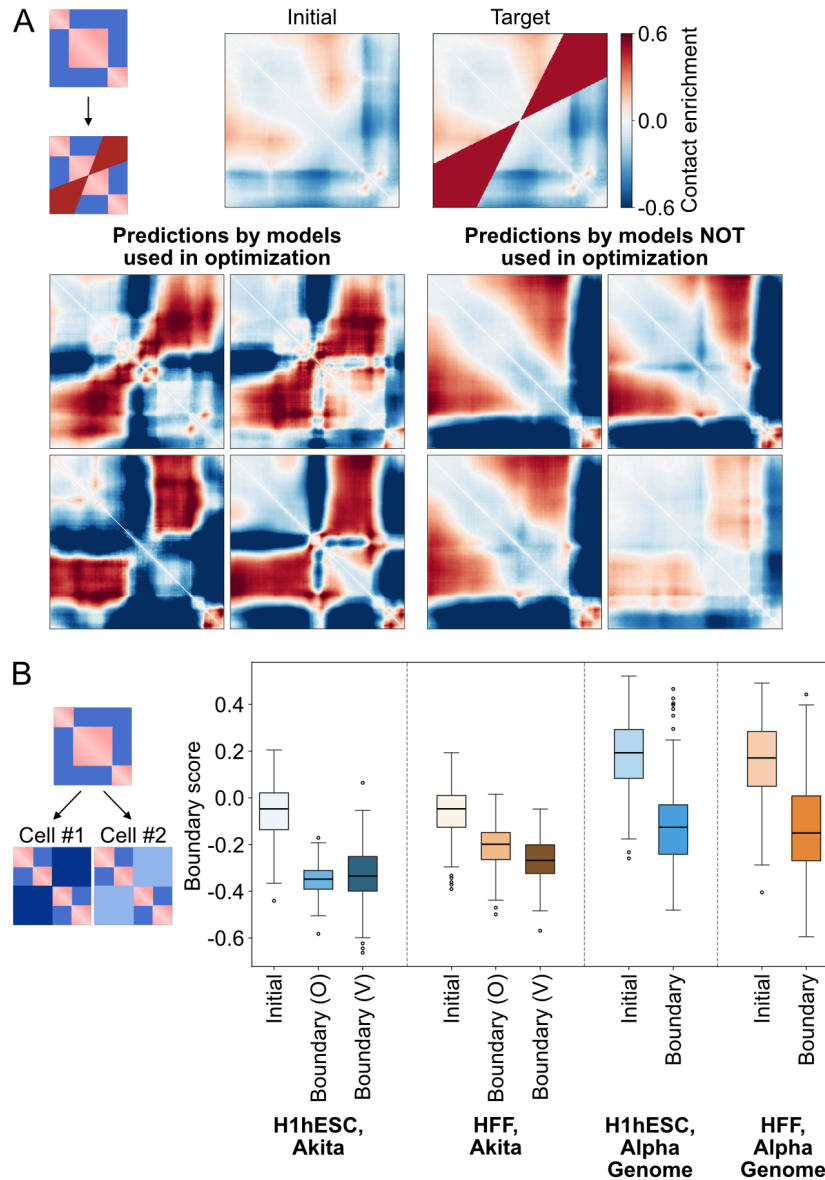


Figure S7. Failed validation of fountain designs and cell-type-specific boundaries.

A) Unsuccessful fountain designs using models trained on B-cell data [48], where fountain structures were characterized without data aggregation. *Top*: contact map before optimization showing a flat genomic region in the center, and the target map with a fountain mask applied. *Bottom*: predictions from four models used in the multi-model fountain optimization show visible fountain-like structures, while predictions from four models not used in optimization (serving as validation models) do not exhibit fountain-like structures, precluding visual validation of the designs. We note that a more stringent validation with AlphaGenome was also unsuccessful for fountains designed with this protocol.

B) Unsuccessful cell-type-specific boundary design. The optimization objective was to create a sequence resulting in a strong boundary in H1hESC (blue) and a weak boundary in HFF (orange). Boxplots show the distributions of boundary scores: (i) H1hESC boundary scores before optimization, after optimization as predicted by two Akita models used in optimization,

and after optimization as predicted by two Akita models not used in optimization (validation models); (ii) HFF boundary scores before optimization, after optimization as predicted by two Akita models used in optimization, and after optimization as predicted by two validation models; (iii) H1hESC boundary scores before and after optimization as predicted by AlphaGenome; (iv) HFF boundary scores before and after optimization as predicted by AlphaGenome. Both the Akita optimization and validation models show that the designed boundaries are stronger in H1hESC than in HFF. However, the distributions of boundary scores calculated from AlphaGenome predictions overlap substantially, failing to validate the cell-type-specific design.

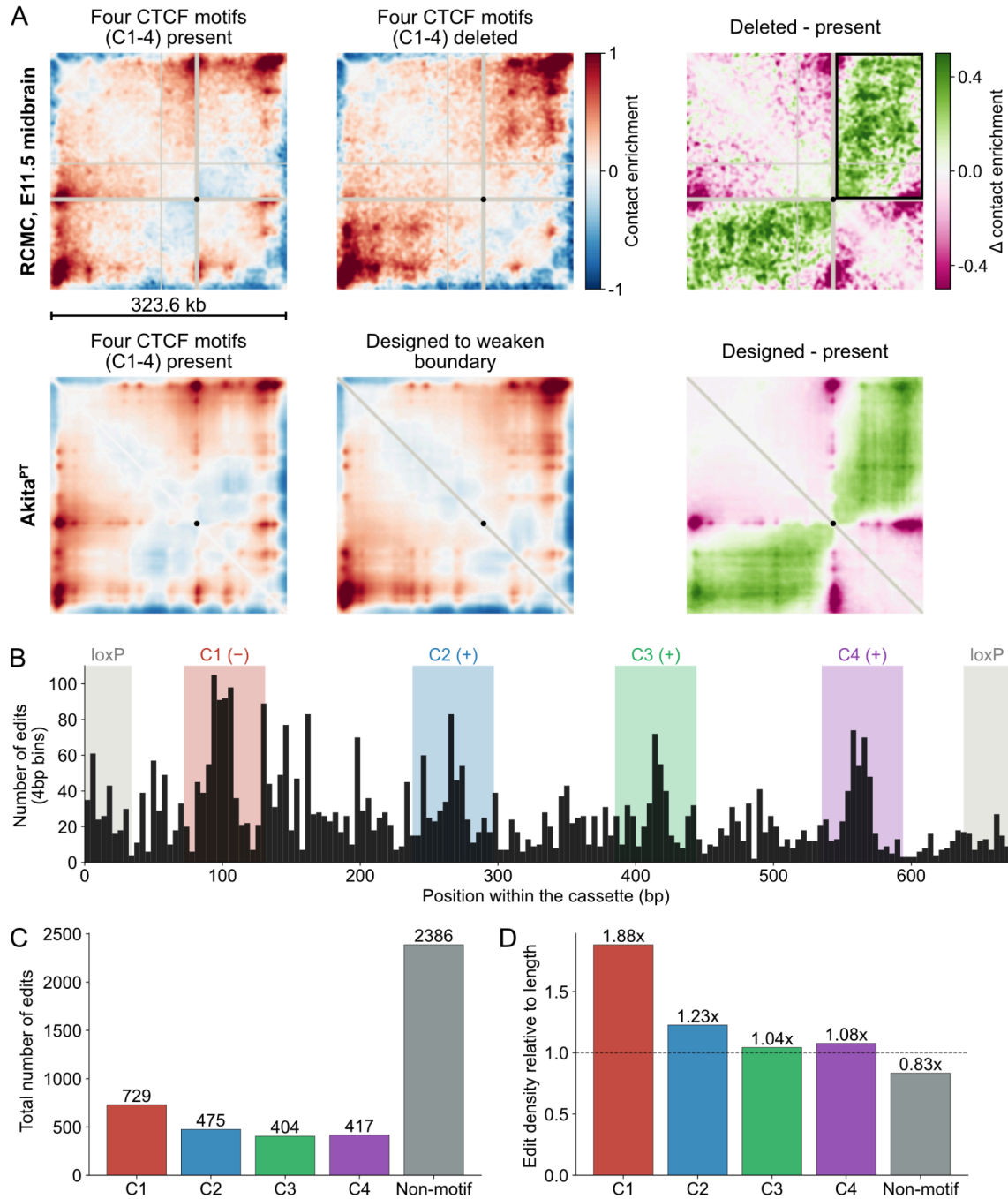


Figure S8. AkitaSF recovers the causal CTCF site at an experimentally validated boundary.

Chakraborty et al. [43] replaced the 23.9 kb CTCF cluster separating the *Ano1* enhancers from *Fgf3/Fgf4/Fgf15* with a 672 bp cassette carrying only the four CTCF core motifs (C1–C4), flanked by loxP sites. This allele is viable and retains boundary function; Cre-mediated excision removes all four motifs and abolishes the boundary. We use the cassette-containing allele as the reference sequence throughout.

A) Observed/expected contact maps across the TAD spanning the boundary (chr7:144,574,464–144,900,096, 2,048 bp bins); the black dot on the diagonal marks the

position of the cassette. *Top*: region-capture Micro-C from Chakraborty et al., with the cassette present (left), after Cre excision (center), and their difference (right). The grey line marks the cassette itself, which is absent from the excised allele and therefore has no comparable data. The black rectangle overlaid on the difference map marks the region over which the loss was computed during design. *Bottom*: Akita_{PT} predictions for the cassette-containing sequence (left), for a sequence designed to remove the boundary (center), and their difference (right).

B) Distribution of accepted edits across the cassette, pooled over 100 independent optimizations. The optimization target was the predicted cassette map with the measured excision difference from (A) added across the boundary; edits were permitted anywhere in the 672 bp cassette. All 100 runs converged, closing 79–83% of the target difference with 35–53 substitutions each. loxP sites are shaded grey and the four CTCF cores (with 20bp flanks) are color-coded.

C) Total edits within each CTCF core (± 20 bp) and in the non-motif remainder of the cassette.

D) Edit density normalized to region length, so that 1.0 is the density expected if edits were placed at random throughout the cassette. All four core motifs are edited above background levels with the most frequent edits at C1. This design outcome is consistent with the experimental finding that deleting C1 alone, but not C2-C4, recapitulates the phenotype of the full cluster deletion.

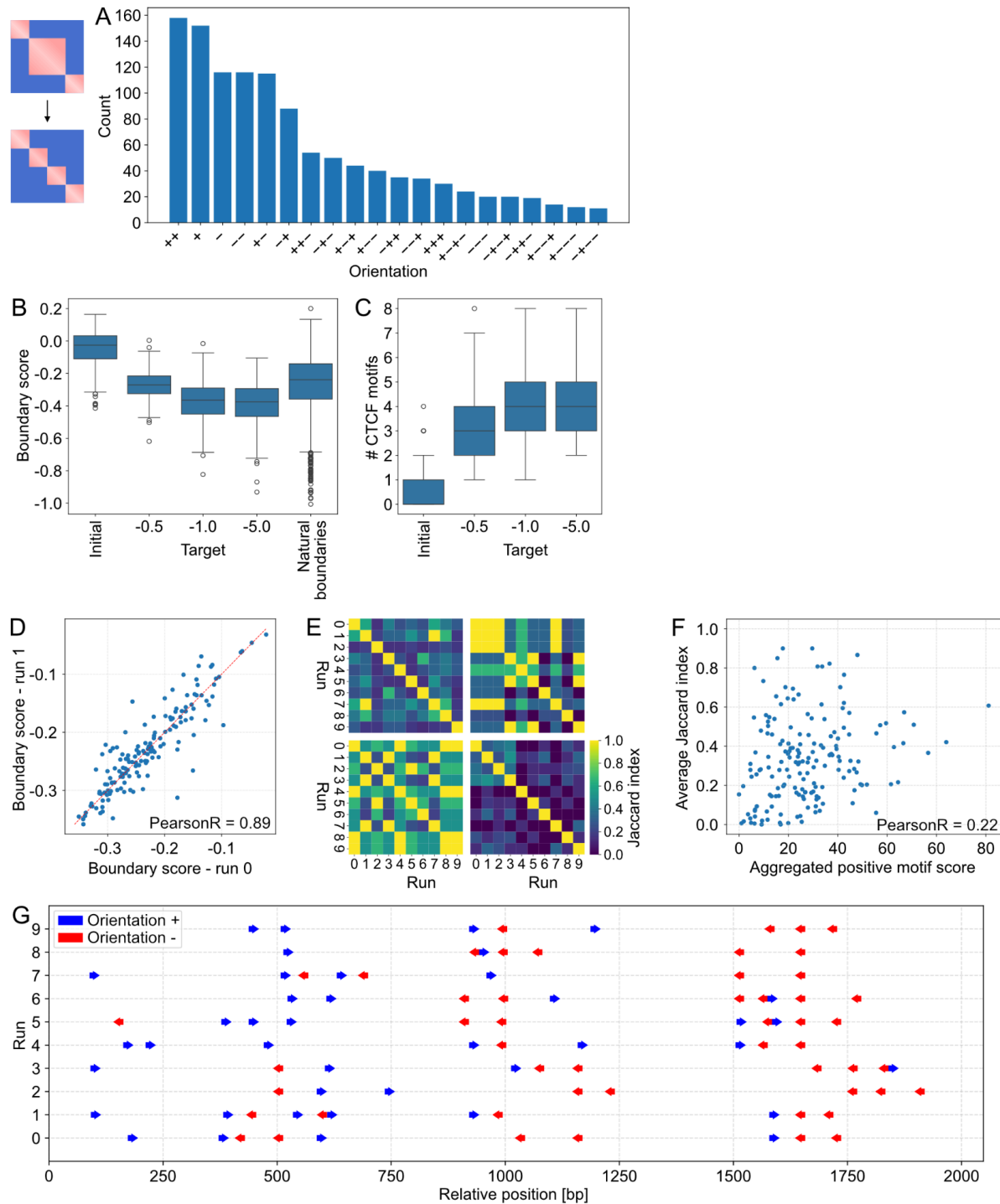


Figure S9. Sequence features of designed boundaries.

A) Histogram of CTCF motifs orientations across all boundary optimization experiments (targets -0.2, -0.3, -0.4, -0.5). Only orientations with counts exceeding 10 are shown for clarity.

B) Boundary scores for natural boundaries (estimated from >4,000 mouse boundaries), initial genome-folding predictions, and predictions for sequences optimized toward strong (-0.5) and extreme (-1, -5) boundary strengths. Boundary scores after optimization plateau at extreme target values, indicating that such strong boundaries are not achievable within our framework.

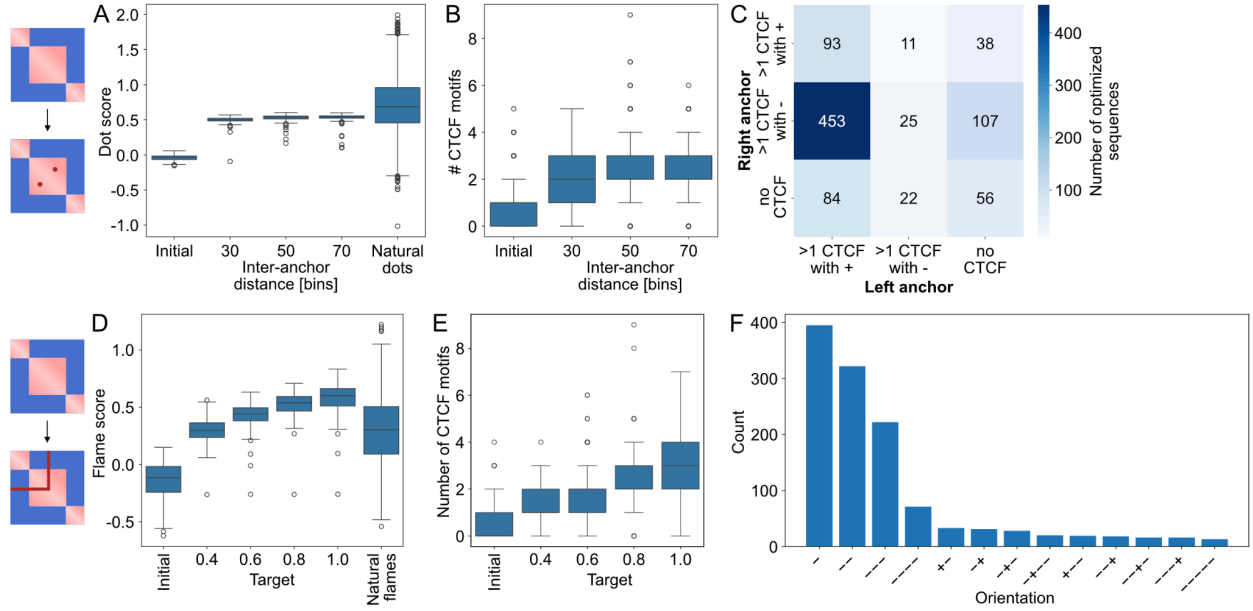
C) Number of detected CTCF motifs for initial sequences and sequences optimized toward strong (-0.5) and extreme (-1, -5) boundary strengths. The number of CTCF motifs does not scale proportionally with increasingly strong target boundaries.

D) Reproducibility of boundary designs. Boundary score differences were calculated for two independent runs of the same optimization (same targets, same genomic regions). The results are highly consistent (Pearson's $r > 0.8$).

E) Four examples of regions for which all 10 independent optimization runs were successful. For each pair of runs, we calculated the Jaccard index between the sets of detected CTCF motifs. Some regions show high consistency across runs, while others yield more diverse solutions.

F) For each optimized sequence, we computed the average Jaccard index of CTCF motif sets across all runs versus the aggregated positive motif score prior to optimization. The scatterplot shows that sequences with higher aggregated positive motif scores exhibit higher average Jaccard indices (Pearson's $r = 0.22$), indicating lower diversity of optimized CTCF configurations. This suggests that optimization tends to reuse pre-existing motif-like features when such features are more abundant in the initial sequence.

G) CTCF motif positions detected in ten independent optimizations as in Figure 3D for a strong (target = -0.5) boundary optimizations but performed with a much lower $\lambda = 0.01$. These runs produce more CTCF motifs and exhibit greater diversity in CTCF motifs across independent optimizations.



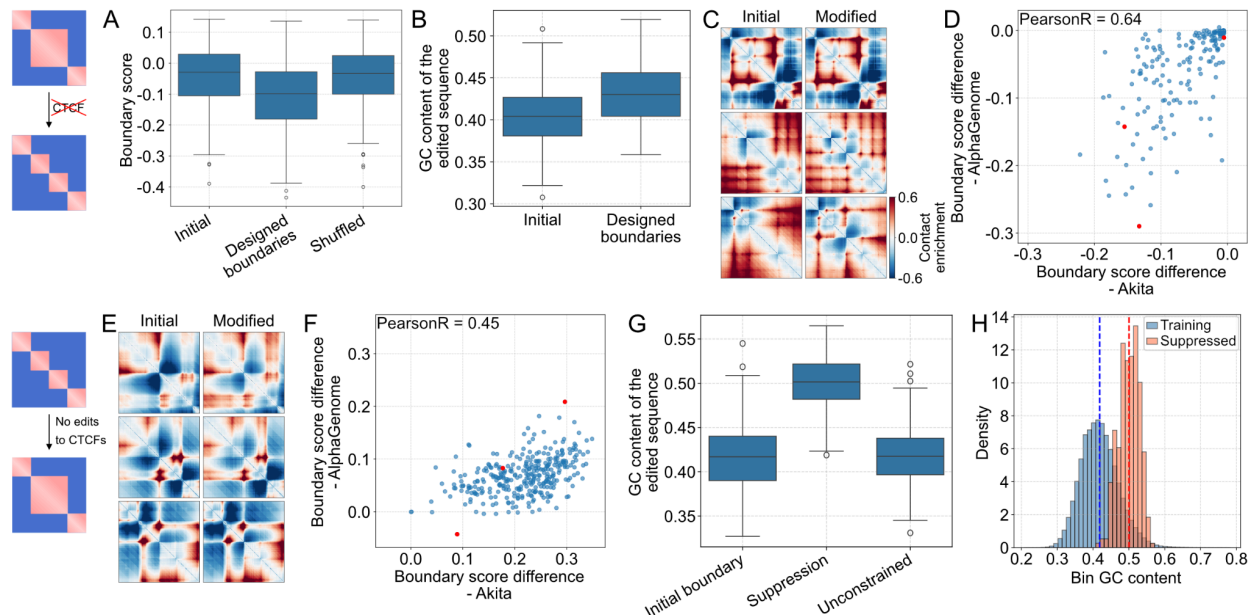


Figure S11. Boundary design without CTCF motifs and boundary suppression without modifying pre-existing CTCFs.

A) Boundary scores for initial sequence predictions and for sequences edited to form boundaries in the absence of CTCF motifs, and for dinucleotide-preserving shuffled controls.

B) GC content within the edited subsequences for the initial sequences and after weak boundary design without introducing CTCF motifs.

C) Three examples of AlphaGenome predictions illustrating original sequences and their edited counterparts, which gain weak boundaries in the middle of the region despite lacking CTCF motifs. Examples are ordered by boundary score from weakest to strongest insulation (top to bottom, corresponding to red dots in (D)).

D) Quantitative validation of weak boundaries created without CTCF motifs. Boundary score differences (optimized – initial) computed using Akita and AlphaGenome predictions show concordance (Pearson's $r > 0.6$).

E) Representative AlphaGenome predicted contact maps for three sequences before and after locked-motif boundary suppression. Examples are ordered from strongest to weakest boundary suppression (top to bottom). The color scale is the same as in panel C.

F) Scatterplot of boundary score differences (optimized – initial) computed for Akita and AlphaGenome predictions. The cross-model correlation is moderate (Pearson's $r > 0.4$).

G) GC content within the edited region for initial sequences and after locked-motif boundary suppression. GC content increases significantly following suppression but not in unrestricted control.

H) Distribution of per-bin GC content across Akita training data (blue) and within the edited bins of boundary-suppressed sequences (orange), showing that optimized sequences remain within the training distribution.

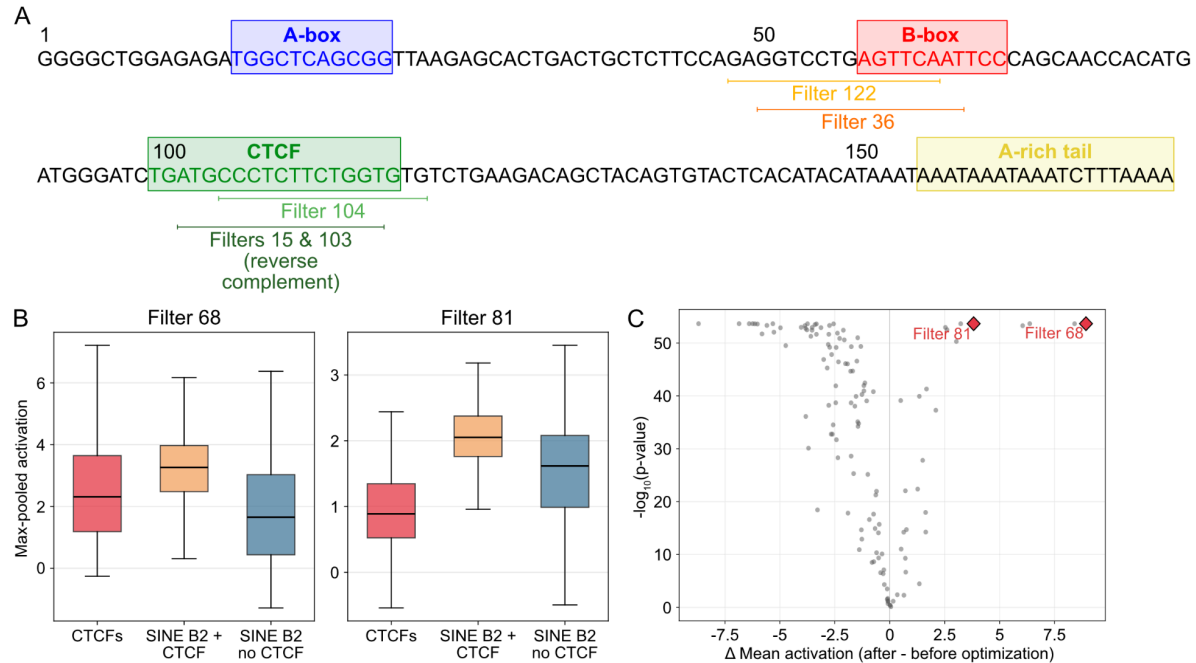


Figure S12. Akita recognizes SINE B2 elements and can use their features to suppress boundaries without deleting existing CTCF sites.

A) Schematic of a B2_Mm2 SINE B2 element with annotated structural features: A-box, B-box, reverse-orientation CTCF motif, and A-rich tail. Filters strongly activated by SINE B2+CTCF elements were identified, and their learned sequence preferences were compared against the B2_Mm2 consensus sequence to assess correspondence (**Methods**). Identified matches include filters recognizing the B-box RNA Polymerase III internal promoter (filters 36 and 122) and the SINE B2-specific CTCF motif variant (filters 15, 103, and 104).

B) Activation of fourth convolutional layer filters (receptive field ~135 bp, sufficient to capture a SINE B2 element) across three sequence groups: isolated strong CTCF motifs, SINE B2 elements containing CTCF motifs (SINE B2+CTCF), and SINE B2 elements lacking CTCF motifs (SINE B2 no CTCF). Filters 68 and 81 show the highest specificity for the SINE B2+CTCF group.

C) Filters most specifically activated by SINE B2+CTCF elements relative to both SINE B2 elements lacking CTCFs and isolated CTCF motifs are also among the most strongly activated in boundary-suppressed sequences compared to their pre-optimization counterparts (filters 68 and 81, ranked 1st and 5th by activation change).