

Additional file 1 for:

**Reconstructing sequence-grammar trajectories enables interpretable
and tunable *cis*-regulatory element design**

Mingqian Ma^{1,9}, Wanjuan Bu^{2,9}, Guoqing Liu^{3,9}, Yuxuan Liu^{4,5,9}, Sizhen Liu², Zhen Zhao^{4,5},
Shijie Yao², Qingru Hua^{2,6}, Yujie Zhang², Cuiting Zhong^{2,7}, Haitao Huang^{7,8}, Pan Deng⁷,
Peiran Jin⁷, Qijin Yin⁷, Chuan Cao⁷, Haiguang Liu⁷, Mo Xu^{4,5,*}, Yuan He^{7,*}, Tao Qin^{7,*}, Zeyu
Chen^{2,*}

1. Machine Learning Department, School of Computer Science, Carnegie Mellon University,
Pittsburgh, PA, USA

2. Department of Cell Biology, MOE Engineering Research Center of Regenerative Medicine,
School of Basic Medical Sciences, Peking University Health Science Center, Beijing, China

3. Microsoft Research AI for Science, Cambridge, UK

4. National Institute of Biological Sciences, Beijing, China

5. Tsinghua Institute of Multidisciplinary Biomedical Research, Tsinghua University, Beijing,
China

6. Department of Biochemistry and Biophysics, University of Rochester, Rochester, NY, USA

7. Zhongguancun Academy, Beijing, China

8. College of Computer Science and Electronic Engineering, Hunan University, Changsha,
China

9. These authors contributed equally to this work.

* Corresponding authors: xumo@nibs.ac.cn (M.X.), heyuan@bza.edu.cn (Y.H.),
gintao@bza.edu.cn (T.Q.) or zeyuchen@hsc.pku.edu.cn (Z.C.)

Supplementary Methods

PPO Optimization and Reward Processing

Quality-Based Reward Shaping. To accelerate convergence, we applied multiplicative reward shaping based on four reward thresholds. Sequences with $R_{base} \geq 7$ received a 1.5× multiplier; sequences with $5 \leq R_{base} < 7$ received a 1.3× multiplier; sequences with $2 \leq R_{base} < 5$ received a 1.1× multiplier; sequences with $R_{base} < 2$ were unscaled. Final rewards were clipped to $[-2.0, 10.0]$.

Temporal Credit Assignment. Because the Malinois reward model evaluates complete 200-bp sequences, the terminal reward was distributed uniformly across the final 5 generated nucleotide positions, providing a denser learning signal than full-sequence terminal-only assignment.

Constrained Sequence Generation. To ensure all generated sequences contained only valid DNA nucleotides, we masked all non- $\{A,C,G,T\}$ vocabulary logits to negative infinity before nucleus sampling (top-p = 0.9, temperature = 1.0).

PPO Implementation. We optimized the SFT-initialized policy using Proximal Policy Optimization with the clipped surrogate objective:

$$L_{ppo}(\theta) = E_t [\min R(\theta) \hat{A}_t, \text{clip}(R_t(\theta), 1 - \epsilon, 1 + \epsilon) \hat{A}_t]$$

where $r_t(\theta) = \frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{old}}(a_t|s_t)}$ is the probability ratio, $\hat{A}_t$ is the advantage estimate computed using

Generalized Advantage Estimation (GAE), and $\epsilon = 0.2$ is the clipping parameter.

The full training objective combined policy, value, entropy, and KL divergence terms:

$$L(\theta) = L_{ppo}(\theta) + c_v L_{value}(\theta) - c_e H_\pi(\theta) + c_{KL} D_{KL}(\pi_{old} | \pi_\theta)$$

where $c_v = 0.5$ is the value loss coefficient, $c_e = 0.05$ is the entropy bonus coefficient encouraging exploration, and $c_{KL} = 0.01$ is the KL penalty coefficient maintaining proximity to the reference policy.

Value Function. A separate value head was attached to HybriDNA's final hidden layer, consisting of shared feedforward layers followed by prompt-specific output heads. This architecture allows the value function to learn cell-type-specific baseline estimates while sharing representational capacity across prompts.

Experience Replay. To improve sample efficiency and training stability, we maintained an experience replay buffer of 512 trajectories collected from previous PPO iterations. Each trajectory corresponds to one generated 200-bp sequence and stores the full set of (state, action, reward, value, log-probability) tuples required for PPO updates. At each iteration, current on-policy trajectories were combined with replayed trajectories at a 15% replay ratio, sampled with a recency bias (70% drawn from the most recent 25% of the buffer, 30% from older entries).

Analysis of the sequence dynamics during reinforcement learning

Δ K-mer representation of iterative sequences. For each iteration t , the sequence set was converted into a k-mer presence ratio vector x_t , where the value of each k-mer was defined as the fraction of sequences in that iteration in which the k-mer was observed at least once. The Δ k-mer vector was defined as $\Delta x_t = x_t - x_{t-1}$.

Motif scanning and motif match scoring. Motif scanning was performed using HOMER (findMotifs.pl with the -find option) with the Vertebrata known motif collection provided by HOMER and default parameters, applied to both 200-bp AI-generated sequences and variable-length STARR-seq sequences. For each motif hit, HOMER reported a motif match

score (MotifScore), defined as the log-odds score of the motif matrix and higher scores indicate stronger agreement between the matched sequence and the motif model.

Motif matrix definition. For iterative analyses, each iteration consisted of 512 sequences. Motif hits were processed independently for each iteration. Motif match scores below a threshold of 6 were discarded. For each iteration, motif scores were summed across all sequences to generate an iteration-level motif profile. Iteration-level profiles were merged to construct a motif-by-iteration feature matrix, with motifs as rows and iteration indices as columns. Missing values were filled with zero, and columns were ordered according to iteration.

For final generated sequences, motif-based feature matrices were constructed at the sequence level. For each sequence-motif pair, the maximum motif match score was retained. A sequence-by-motif matrix was then generated, with rows representing sequences and columns representing motifs, and missing values filled with zero. For datasets containing more than 10000 sequences, feature matrices were z-score scaled across sequences prior to downstream analysis.

Motif-based stage definition. For each iteration, the sequence set was converted into a motif presence ratio vector m_t , where the value of each motif was defined as the fraction of sequences in that iteration in which the motif was observed at least once.

Non-negative matrix factorization was performed on the motif presence ratio matrix of each task defined as $M = [m_1, m_2, \dots, m_{1000}]^T$. Three components were used for downstream stage definition based on a combined evaluation of the cophenetic correlation coefficient and residual sum of squares (RSS) across candidate factorization ranks. Component weight trajectories were smoothed using a centered moving average. Stage boundaries were defined as the midpoints between adjacent peaks of the smoothed component trajectories.

Construction of the shared sequence grammar space

Construction of the shared dynamic basis. The concatenated $\Delta 6$ -mer matrix was constructed by pooling the first 40 $\Delta 6$ -mer vectors from each of the three tasks and was used to compute the covariance matrix, followed by eigendecomposition. The shared eigenvalue spectrum was converted into an effective rank K based on the entropy of the normalized eigenvalues. The top K covariance eigenvectors defined the shared dynamic basis. Each 6-mer was then represented as

$$f_i = \left[u_{i1}\sqrt{\lambda_1}, \dots, u_{iK}\sqrt{\lambda_K} \right],$$

where u_{ij} denotes the loading of the i -th 6-mer on the j -th shared dynamic axis, and λ_j denotes the corresponding eigenvalue.

Effective rank estimation. The effective rank K of the shared covariance matrix was defined from the entropy of its normalized eigenvalue spectrum. For eigenvalues λ_i , normalized weights were computed as $p_i = \lambda_i / \sum_j \lambda_j$, followed by spectral entropy $H = - \sum_i p_i \log p_i$. The effective rank was then given by $K = \text{ceiling}(\exp(H))$, representing the number of covariance dimensions effectively contributing to the shared variance structure.

Plasticity score estimation. For each 6-mer, the plasticity was defined as an eigenvalue-weighted loading energy:

$$plasticity_i = \sum_{j=1}^K u_{ij}^2 \lambda_j .$$

Nonlinear embedding by UMAP and visualization. The 6-mer feature vectors were z-score standardized and embedded into two dimensions using UMAP (n_neighbors=30, min_dist=0.15, metric="euclidean", random_state=132).

For visualization, plasticity scores were multiplied by a scale factor of 10000. The plasticity and GC fraction of each 6-mer were separately rendered as background landscapes on the

2D grammar space using Gaussian kernel smoothing on a regular grid (GRID_SIZE = 400, SIGMA = 0.04).

Projection of iterative trajectories onto the grammar space. For each iteration, the sequence set was converted into a 6-mer presence ratio vector and normalized to sum to 1. The iteration was then projected into the UMAP space as the weighted centroid of the 6-mer coordinates. The shared initial point was defined as the centroid of the task-specific predicted initial points across the three tasks.

Task-level windowed PCA analysis of sequence update dynamics

Calculation of windowed PC contributions. For each task, PCA was performed once on the task-level Δ 6-mer matrix to define a shared set of dynamic axes across the full trajectory. Ten principal components were retained for downstream analysis based on the average effective rank estimated from 20-iteration windows. A sliding window of 20 iterations with a step size of 5 iterations was then moved across the trajectory.

Within each window w , the root-mean-square (RMS) magnitude of each retained PC score was computed, giving a window-level magnitude vector $r_w \in R_{\geq 0}^{10}$. The RMS magnitudes were then smoothed across adjacent windows using a centered rolling mean with window size 2 and normalized to sum to one, yielding the windowed PC contribution vector

$c_w = (c_{w,1}, \dots, c_{w,10})$, where $c_{w,j}$ denotes the relative contribution of PC j in window w .

Identification of program transition point. Candidate transition points were identified from changes in the smoothed windowed contribution vectors. For transition detection, only the first five retained PCs were used: the vector $c_w^{(5)}$ was defined by restricting c_w to PC1-PC5 in the global PCA ordering and renormalizing it to sum to one. A mutation score was then computed for each window as

$$m(w) = z\left(JSD\left(c_w^{(5)}, c_{w-1}^{(5)}\right)\right) + 0.35 \cdot z\left(\left|c_{w,1}^{(5)} - c_{w-1,1}^{(5)}\right|\right),$$

where $z(\cdot)$ denotes z-score normalization across windows and JSD denotes Jensen-Shannon divergence.

Candidate peaks were taken as local maxima of $m(w)$ above the 75th percentile and separated by at least 80 iterations. A candidate peak was retained as program transition point only if it passed an additional stability filter based on the full 10-dimensional contribution vectors:

letting

$$c_b^{-pre} = \frac{1}{2} \sum_{u=w_b-2}^{w_b-1} c_u, \quad c_b^{-post} = \frac{1}{3} \sum_{u=w_b}^{w_b+2} c_u, \quad c_b^{-pers} = \frac{1}{5} \sum_{u=w_b}^{w_b+4} c_u.$$

The peak was retained only if

$$JSD\left(c_b^{-pre}, c_b^{-post}\right) \geq 0.005, \quad JSD\left(c_b^{-pre}, c_b^{-pers}\right) \geq 0.005, \quad JSD\left(c_b^{-post} - c_b^{-pre}, 0\right) \geq 0.03$$

Identification of the initial program and transition-associated update programs. In addition to transition-derived programs, an initial program P_0 was defined from the first window by selecting the top three PCs with the largest entries in c_0 . For each transition point b , a transition-associated update program P_b was defined from the positive contribution change across the transition,

$$\delta c_b = \max\left(c_b^{-post} - c_b^{-pre}, 0\right)$$

The three PCs with the largest entries in δc_b were selected as the sparse signature of P_b .

Their normalized values were used as program weights $a_{b,j}$, such that $\sum_{j \in P_b} a_{b,j} = 1$. If δc_b

contained no positive entries, the signature was instead taken from the top three PCs in

$$c_b^{-post}.$$

Relative activity of initial and transition-associated programs. For each program P_b , a raw window-level score was first computed as the weighted sum of PC contributions,

$$s_b(w) = \sum_{j \in P_b} a_{b,j} c_{w,j}.$$

To reflect delayed onset after a transition, transition-associated programs were multiplied by a smooth activation gate defined from the window-center iteration relative to the corresponding transition point, using a clipped smooth-step function over a transition scale of 24 iterations; the initial program P_0 used a constant gate of 1. The gated score was then rectified to nonnegative values, transformed by a power of 1.85, and smoothed on the active segment using a Savitzky–Golay filter (target window length 141, polynomial order 3). Finally, the activities of all programs were renormalized within each window to sum to one, yielding the relative program activity

$$a_b(w) = \frac{\hat{s}_b(w)}{\sum_b \hat{s}_b(w)},$$

where $\hat{s}_b(w)$ denotes the gated and smoothed nonnegative score for program b .

Projection of initial and transition-associated programs into 6-mer space. For each program P_b , a sparse program vector $g_b \in R^{10}$ was first defined in PC space, with $g_{b,j} = a_{b,j}$ for $j \in P_b$ and $g_{b,j} = 0$ otherwise. Let $L \in R^{10 \times 4096}$ denote the PCA loading matrix learned from the standardized $\Delta 6$ -mer matrix. The program-specific 6-mer tendency was then obtained by direct projection of the sparse program vector onto the 6-mer loadings as

$$\tau_b^{scaled} = g_b L.$$

Empirical two-sided p-values, Benjamini–Hochberg adjusted q-values, and rank statistics were calculated across all 6-mers using τ_b^{scaled} as the primary tendency score.

Projection of initial and transition-associated programs into motif space. Each motif PWM was first converted into a strand-collapsed 6-mer probability distribution. For a motif m , probabilities were averaged across all overlapping 6-bp windows of the PWM. Forward and reverse-complement 6-mer distributions were then merged by elementwise maximum and renormalized, yielding a motif-specific 6-mer distribution $d_m = (d_{m,1}, \dots, d_{m,4096})$. The value of motif tendency of motif m for program was defined as the expectation of the program's 6-mer tendency under this motif-induced distribution,

$$\eta_{b,m} = \sum_k d_{m,k} \tau_{b,k}^{scaled}.$$

Here, the sign of $\eta_{b,m}$ indicates whether the motif aligns more strongly with the positive or negative side of the program tendency axis.

To assign a direction more directly related to post-transition motif dynamics, motif tendency was further direction-corrected using the observed post-iteration motif change. For each program transition point b , the motif dynamic of post 10-iteration was used to determine the final direction. If $|\Delta_{b,m}^{post10}|$ exceeded the threshold defined as the median absolute post-iteration change across the motif set, the motif tendency score was reoriented by the sign of the observed post-branch change,

$$\hat{\eta}_{b,m} = \text{sign}(\Delta_{b,m}^{post10}) |\eta_{b,m}|.$$

Motifs with $\left|\hat{\eta}_{b,m}\right| < 0.001$ were discarded, and empirical two-sided p-values,

Benjamini–Hochberg adjusted q-values, and rank statistics were calculated from $\hat{\eta}_{b,m}$.

Motif-based visualization in the generated sequences

UMAP embedding and density-based clustering. Motif feature matrices were constructed as described above and used as input for dimensionality reduction and clustering analyses. Low-dimensional representations were computed using UMAP as implemented in the *uwot* v0.2.3 R package. A fixed random seed (123) was applied to ensure reproducibility. UMAP was performed with PCA initialization (`init = "pca"`) and approximate nearest-neighbor search using Annoy (`nn_method = "annoy"`). Dataset-specific parameters were applied to accommodate differences in motif space density and sample size. For HepG2-specific sequences, UMAP was performed with `n_neighbors = 40`, `min_dist = 0.05`, and `n_epochs = 800`; for K562-specific, `n_neighbors = 40`, `min_dist = 0.01`, and `n_epochs = 1000`; for SK-N-SH, `n_neighbors = 60`, `min_dist = 0.05`, and `n_epochs = 800`. All embeddings were generated in two dimensions.

Cluster-level motif heatmap visualization. Clustering was performed on the resulting UMAP coordinates using HDBSCAN [58] to identify density-based groups while allowing noise points. The `minPts` parameter was adjusted per dataset (HepG2: 80; K562: 30; SK-N-SH: 110) to reflect differences in local point density. Points labeled as noise were excluded from downstream analyses, and scatter plots were generated using the `ggplot2` v4.0.1 R package.

To visualize motif usage patterns across clusters, heatmaps were generated from the original motif feature matrix. Motif scores were joined with cluster assignments, and only non-noise samples were retained. The feature matrix was transposed to motif-by-sample format, and motifs present in fewer than 600 samples were filtered out. Cluster annotations derived from HDBSCAN labels were added as column annotations. Heatmaps were rendered using `ComplexHeatmap` [63], with cluster identity indicated by a categorical color scale.

Comparison with endogenous CREs

Endogenous CREs were obtained from reproducible HepG2 and K562 STARR-seq peaks processed by Lee et al. using datasets ENCFF047LDJ and ENCFF045TVA, respectively. For each cell type, endogenous CREs and AI-generated candidate CREs were represented using motif-based sequence feature matrices as described above and analyzed jointly.

Dimensionality reduction by UMAP. Dimensionality reduction was performed by applying PCA to the combined feature matrix, retaining the first 30 principal components.

Two-dimensional embeddings were generated using UMAP on the PCA transformed features with the following parameters: metric = "cosine", init = "spectral", nn_method = "annoy", n_components = 2, n_epochs = 1000, random seed 213 and n_neighbors = 30 for HepG2 and 50 for K562.

Two-step clustering and motif heatmap visualization. Two-step clustering was performed on the two-dimensional UMAP embeddings. First, K-means clustering was applied with 20000 centers (nstart = 80, iter.max = 100) to partition the embedding space. Second, HDBSCAN was run on the K-means cluster centers (minPts = 50 for HepG2 and 70 for K562) to identify dense regions [58]. The resulting HDBSCAN cluster assignments were then mapped back to the motif-based feature matrices.

For each non-noise cluster, up to 60 endogenous sequences were randomly sampled with replacement using a fixed random seed (set.seed = 668), and all AI-generated sequences were retained. The motif-based feature matrix was processed using Seurat v5.3.0 to identify the 300 most variable motifs (selection.method = "vst") [64]. Motifs present in fewer than 200 sequences were excluded and visualized as heatmap using the ComplexHeatmap [63].

MPRA library design

Negative control. Negative control sequences were derived from GM12878 ATAC-seq open chromatin regions. Multiple GM12878 ATAC-seq BED files were first intersected to identify

regions consistently accessible across datasets, and an aggregated ATAC-seq score (sum_score) was calculated for each region by summing the peak scores across all GM12878 ATAC-seq datasets. Candidate regions were then stringently filtered by excluding any intervals overlapping ATAC-seq or DNase-seq peaks from K562, HepG2, or SK-N-SH, and regions shorter than 200 bp were removed. The remaining regions were ranked in descending order first by sum_score per base pair and then by sum_score. From the top 100 ranked regions, 50 were randomly sampled (random_state = 42). In addition, 50 regions were randomly sampled from an intermediate ranking range spanning ranks 1001 to 5000 (random_state = 42). These two subsets were combined, and each interval was subsequently partitioned into multiple overlapping 200 bp fragments using an adaptive tiling strategy. The number of fragments was determined by the interval length, and fragment start positions were evenly spaced to ensure complete coverage of the original interval while maintaining a fixed fragment length. All sequences were extracted from the hg38 reference genome. Sequences containing ambiguous bases were excluded. Genomic specificity was assessed using the UCSC Genome Browser BLAT search. Sequences were retained only if they showed a single full-length 200 bp perfect match in the human genome, with no additional alignments longer than 30 bp.

Library for cell-type-specific CREs validation. This library comprised 1000 constructs, including 50 negative control sequences, 450 AI-generated candidate CREs (150 per cell type), and reverse complements of all sequences. The 50 negative control sequences were randomly sampled without replacement from the negative control library using a fixed random seed (431). AI-generated CREs were selected based on predicted regulatory activity scores generated by the Malinois model as described above. Sequences were first grouped according to their UMAP-derived clusters. Within each cluster, sequences were ranked using a combined criterion that jointly considered i) the predicted activity score in the target cell type and ii) a cell type specificity score, defined as the difference between the predicted activity in the target cell type and the sum of predicted activities in the two non-target cell

types. The combined rank was computed as the sum of the individual ranks for these two metrics, and the top 30 sequences were selected from each cluster. Clusters were then compared based on the mean predicted activity score and mean cell-type–specificity score of their selected sequences. A cluster-level combined rank was derived from these two mean values, and the top five non-noise clusters were retained, yielding a total of 150 AI-generated CREs per cell type.

Library for comparison with endogenous. The MPRA library comprised 1,000 constructs, including 400 HepG2 endogenous CREs, 400 AI-generated HepG2-specific candidate CREs, and 200 negative control sequences which were randomly sampled without replacement from the negative control library using a fixed random seed (431). Endogenous CREs identified from HepG2 STARR-seq data were obtained from reproducible HepG2 STARR-seq peaks reported by Lee et al. Peaks shorter than 200 bp were excluded, and each remaining peak was converted into fixed-length 200 bp fragments by generating up to 100 overlapping windows with uniformly spaced start positions to ensure complete coverage of the original interval. These fragments were first scored using boda2 [36], and the top 5000 fragments with the highest MinGap for HepG2 as the target cell type were retained. The retained fragments were subsequently ranked by chromatin accessibility predicted by ChromBPNet using the mean predicted value across the prediction region. Fragments were selected under the additional constraint that selected fragments were separated by at least 20 fragment indices within the same peak. Based on these criteria, the top 400 fragments were retained as representative HepG2 endogenous CREs. AI-generated HepG2-specific candidate CREs were filtered in two steps. The top 1,000 sequences were first selected based on MinGap scores predicted by the boda2 framework using the Malinois model. These sequences were subsequently ranked by chromatin accessibility predicted by ChromBPNet, and the top 400 CREs were retained.

Amplification of custom oligonucleotide pools

The oligonucleotide library, synthesized by Twist Bioscience, was cloned into the lentiviral MPRA vector following PCR amplification. Forward (Oligo-F: 5'–AACCGGTGGCCACTTCATTT–3') and reverse (Oligo-R: 5'–CCCTCTAGTGACCTTGACAGC–3') primers, complementary to the predefined 5' and 3' adaptor sequences of the oligonucleotide pool, were used for amplification. PCR amplification was performed using Q5 High-Fidelity 2× Master Mix (New England Biolabs, M0492S) under the following conditions: initial denaturation at 98°C for 30 s, followed by 13 cycles of 98°C for 10 s, 58°C for 10 s, and 72°C for 20 s, with a final extension at 72°C for 2 min. The amplified products were resolved on a 1.5% agarose gel, and the fragment of the expected length was extracted and purified using MagPure Gel Extraction Kit (Magen, D2110) for ligation into the linearized lentiviral backbone.

Lentiviral MPRA library construction

The lentiviral MPRA backbone was based on plasmid pLS-Scel (Addgene #137725) [65]. This vector was modified to introduce an hPGK promoter-driven mCherry cassette downstream of the original reporter and to replace the Scel cloning site with dual BfuAI restriction sites. The modified vector was linearized using BfuAI (New England Biolabs, R0701S) in NEBuffer™ r3.1. The linearized product was separated on a 1% agarose gel, and the backbone fragment of the correct size was gel-extracted and purified using MagPure Gel Extraction Kit (Magen, D2110).

The final MPRA vector contains two expression cassettes: 1) an upstream reporter cassette, in which oligonucleotide libraries are inserted between flanking adaptor sequences within a minimal promoter-driven transcription unit. This unit consists of the minimal promoter, the oligo cloning site, an EGFP sequence fused to a PEST destabilization domain, and a bovine growth hormone polyadenylation signal (bGHpA); 2) a downstream constitutive expression cassette consisting of the hPGK promoter driving mCherry, which serves as a transduction and normalization control. Purified, double-stranded oligonucleotide pool PCR products (250

bp) were cloned into the linearized backbone using a 2× Seamless Cloning Mix (Biomed, CL117-01). For cloning, six independent 20 µL ligation reactions were assembled, each containing 330 ng of linearized backbone and 50 ng of insert DNA. Reactions were incubated at 50°C for 60 min, pooled, and co-precipitated with 4 µg of glycogen (Roche, 10901393001) using isopropanol. The pelleted DNA was resuspended in 5 µL of nuclease-free water. Two separate transformations were performed using Endura™ ElectroCompetent Cells (Biosearch Technologies, 60242), each with 2.5 µL of the purified ligation product. Following transformation, cells were recovered in 2 mL of SOB medium at 37°C for 1 h. The two cultures were combined and plated across forty 15-cm LB agar plates. After overnight growth, colonies were harvested, and plasmid DNA was extracted using a QIAGEN Plasmid Maxi Kit (QIAGEN, 12362).

Plasmid library quality control

The plasmid library was deep-sequenced for quality control. We designed sequencing primers flanking the oligo insert region, compatible with the Illumina TruSeq system and containing unique dual indices (UDIs) to enable multiplexing (see oligonucleotides).

A two-step PCR strategy was employed. In the first step, the plasmid library was amplified for 10 cycles using the TruSeq adapter primers under conditions otherwise identical to those described for oligonucleotide pool amplification. The PCR product was purified. In the second step, the purified product was used as a template for an 8-cycle amplification with the UDI-indexed primers. The final PCR product was purified and subjected to paired-end sequencing on an Illumina NovaSeq 6000 platform at Annoroad (Beijing, China) to assess library complexity and sequence representation.

Cell culture

HepG2, K562, and HEK293T cell lines were kindly provided by Dr. Ting Han's laboratory, while SK-N-SH (ATCC HTB-11) cells were purchased from the American Type Culture

Collection (ATCC). HepG2, K562, and HEK293T cells were cultured in Dulbecco's Modified Eagle Medium (Gibco, 11965092). SK-N-SH cells were maintained in Minimum Essential Medium (Gibco, 11095080). All media were supplemented with 10% fetal bovine serum (GeminiBio, 900-108-500) and 1% penicillin–streptomycin (Gibco, 15140122). All cells were maintained at 37 °C in a humidified incubator with 5% CO₂.

Lentiviral vector production

293T cells were seeded in 10 cm dishes and grown to approximately 70% confluence prior to transfection. Lentiviral vectors were produced using a second-generation packaging system consisting of psPAX2 (Addgene, 12260) and pMD2.G (Addgene, 12259). For transient transfection, 10 µg of plasmid library, 7.5 µg of psPAX2, and 2.5 µg of pMD2.G were co-transfected into 293T cells using Lipofectamine 3000 (Thermo Fisher Scientific, L3000075) according to the manufacturer's instructions. Eight hours after transfection, the culture medium was replaced with fresh complete medium. Viral supernatants were collected 48 hours post-transfection, clarified by filtration through a 0.45 µm filter, and used for downstream transduction.

Lentiviral transduction

For suspension cells, K562 cells (1×10^6) were resuspended in 10 mL of filtered lentiviral supernatant supplemented with 10 µg/mL polybrene (Yeasten, 40804ES). Cells were subjected to spinfection by centrifugation at $1,500 \times g$ for 1 h at room temperature, followed by incubation at 37 °C for 12 h. For adherent cell lines, HepG2 and SK-N-SH cells were seeded in 6-well plates one day prior to transduction to reach approximately 70% confluence at the time of infection, with 5 mL of filtered lentiviral supernatant added per well. Polybrene was added at a final concentration of 8 µg/mL for HepG2 cells and 5 µg/mL for SK-N-SH cells. HepG2 cells were subjected to spinfection at $1,000 \times g$ for 30 min, whereas SK-N-SH cells were spininfected at $500 \times g$ for 1 h. All cells were incubated with lentiviral supernatant for 12 h post-infection, after which the medium was replaced with fresh complete culture

medium. K562 and HepG2 cells were passaged and expanded at 48 h post-infection. For all lentiviral MPRA experiments, the multiplicity of infection (MOI) was maintained below 0.3 to ensure predominantly single-copy integration. Transduced populations were enriched by flow cytometric sorting at 96 h after infection.

Flow cytometric sorting and data analysis

Post-transduction cells were washed with DPBS (Thermo Fisher Scientific, 14190144). Adherent HepG2 and SK-N-SH cells were detached using Trypsin-EDTA (Thermo Fisher Scientific, 25200114) at 37 °C (HepG2: 5 min; SK-N-SH: 2 min). Digestion was terminated by adding complete culture medium, and cells were collected by centrifugation and resuspended in MACS buffer (DPBS supplemented with 2% FBS and 2 mM EDTA). Cells were sorted on a BD FACSAria™ III Cell Sorter (BD Biosciences). Untransduced cells were included in each experiment as gating controls. The sorting strategy was as follows: debris was excluded based on FSC/SSC parameters, DAPI-negative cells were gated as viable cells, singlets were selected, and mCherry-positive cells were gated. Within the mCherry-positive population, cells were further subdivided into four equal bins according to the EGFP/mCherry fluorescence ratio. Approximately 1×10^6 cells were sorted and collected from each bin for downstream Illumina sequencing. Flow cytometry data were analyzed using FlowJo software (v10.8.1, BD Biosciences).

Sequencing library preparation from sorted MPRA cell populations

Genomic DNA from each of the four sorted bins was extracted using the QIAamp DNA Mini Kit (QIAGEN, 51304) following the manufacturer's instructions.

A two-step PCR was performed using primer sets identical to those used in plasmid library quality control. In the first PCR, 2 µg genomic DNA per 50 µL reaction was amplified with TruSeq adapter primers for 20 cycles. Products were purified using NovoNGS® DNA Clean Beads (Novoprotein, N240), and fragments of 200–300 bp were selected using a 0.9× bead

ratio. In the second PCR, purified first-round products served as templates (2 ng per 50 μ L reaction) and were amplified with UDI-indexed primers for 10 cycles. PCR products were size selected on a 1% agarose gel, and correctly sized bands were excised and purified using the MagPure Gel Extraction Kit (Magen, D2110). Final libraries were subjected to paired-end sequencing on an Illumina NovaSeq 6000 platform at Annoroad (Beijing, China) to assess sequence distribution.

Supplementary Figures

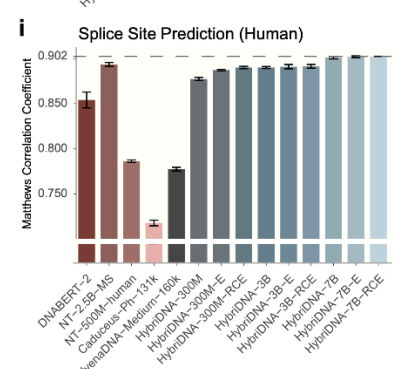
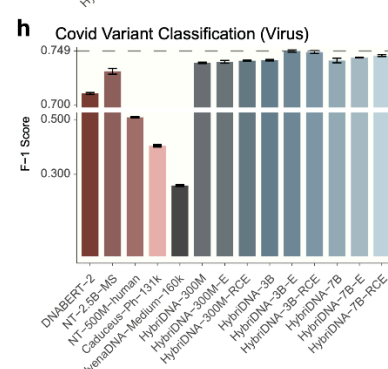
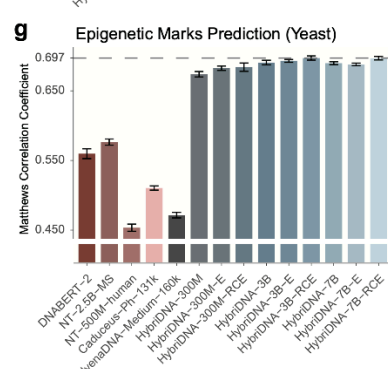
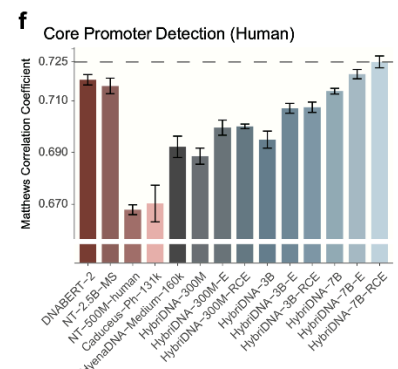
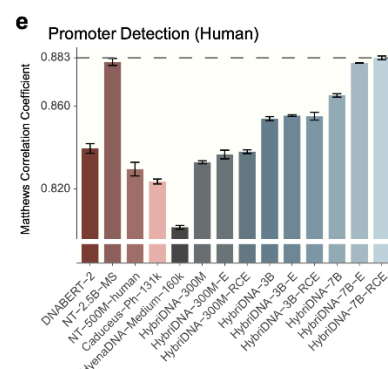
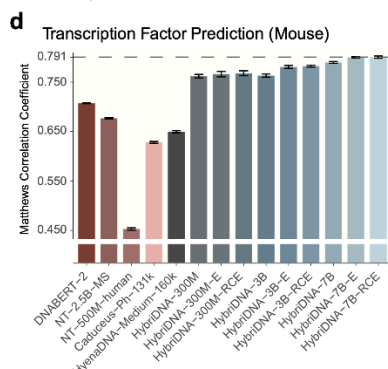
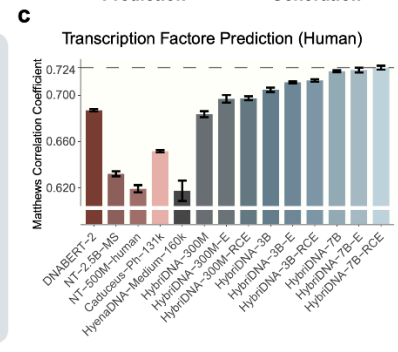
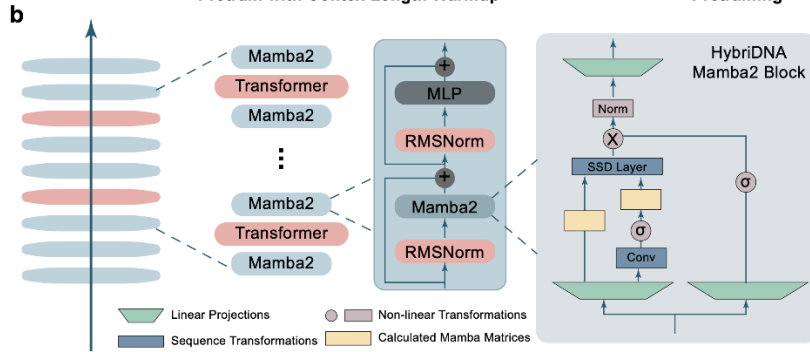
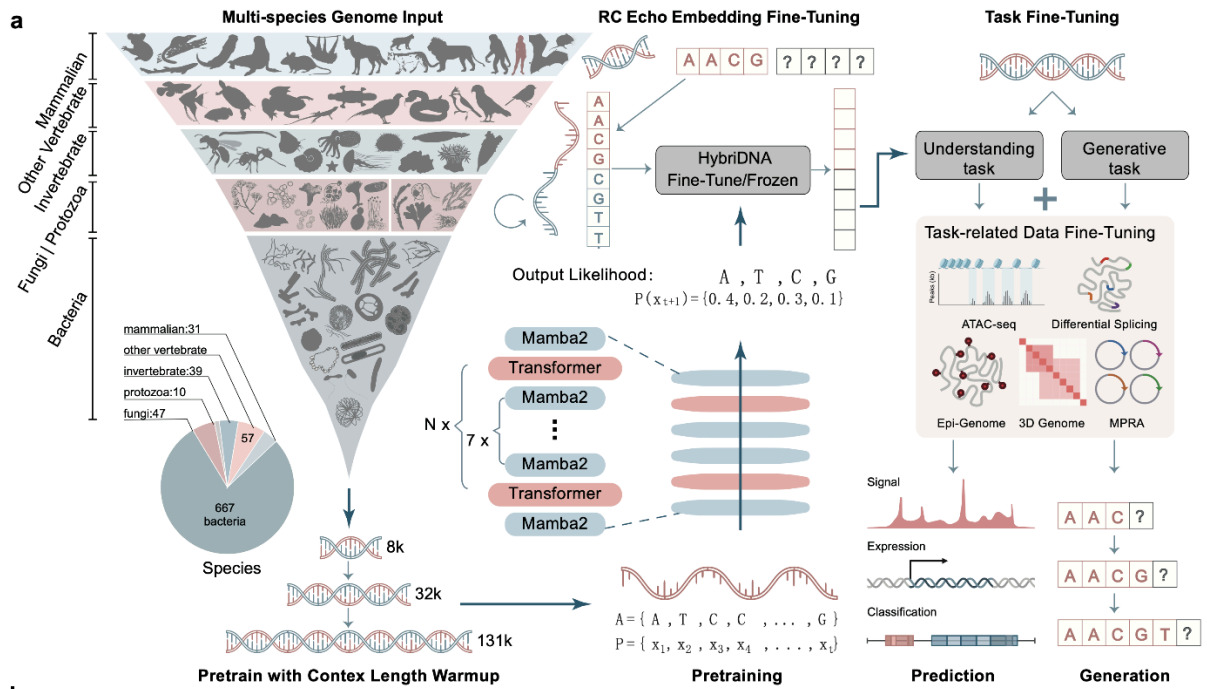


Figure S1. Optimized HybriDNA architecture and benchmarking on genome understanding tasks, related to Figure 1.

(a) Schematic of the HybriDNA framework, including multi-species pretraining with context-length warmup, reverse-complement echo fine-tuning, and adaptation to predictive and generative tasks. (b) Detailed architecture of the HybriDNA Mamba2 block. (c–i) Performance of genomic language models on Genome Understanding Evaluation (GUE) tasks. (c) Transcription factor prediction in human. (d) Transcription factor prediction in mouse. (e) Promoter detection in human. (f) Core promoter detection in human. (g) Epigenetic mark prediction in yeast. (h) COVID variant classification. (i) Splice-site prediction in human. Models compared include DNABERT-2, NT-2.5B-MS, NT-500M-human, Caduceus-Ph-131k, HyenaDNA-Medium-160k, and HybriDNA variants. Matthews correlation coefficient (MCC) was used in (c–g) and (i), and F1 score was used in (h). E, Echo Embedding; RCE, Reverse Complement Echo Embedding. Error Bar represents S.D. of three independent runs with random seeds.

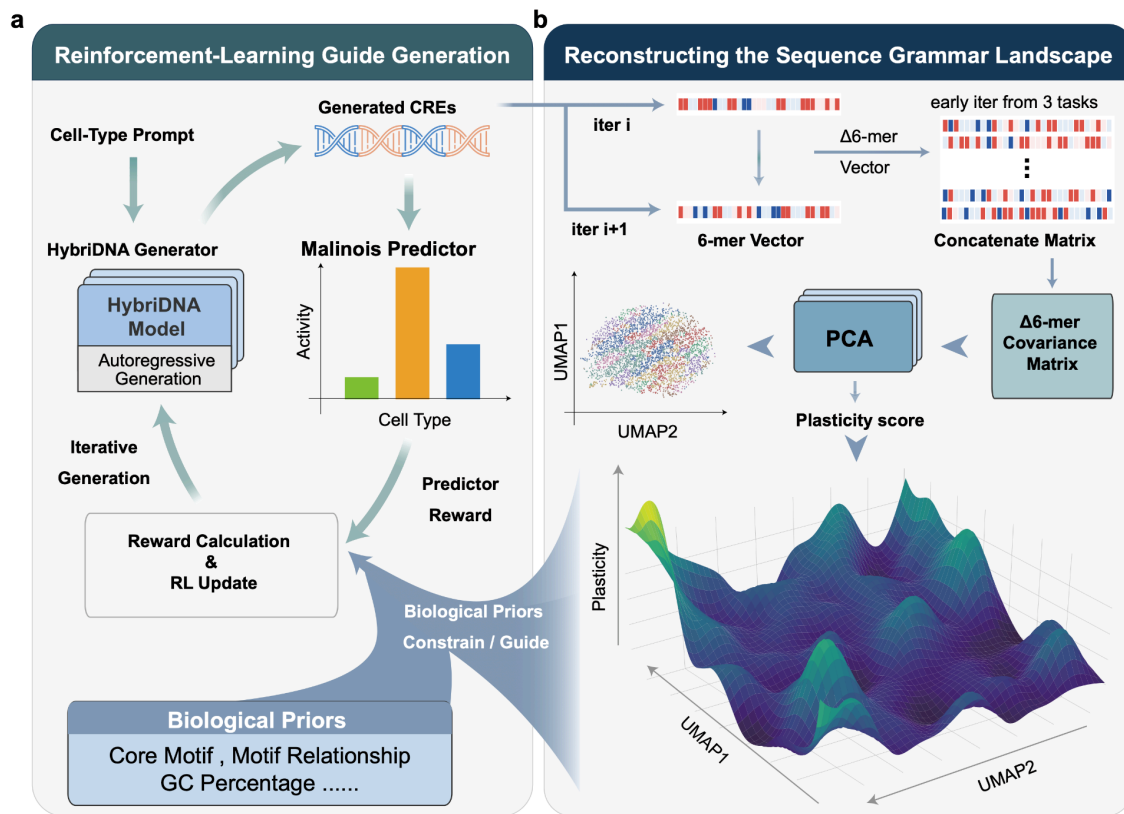


Figure S2. Detailed overview of the GO-CRE framework, related to Figure 2.

(a) Closed-loop reinforcement learning (RL) for cell-type-specific CRE generation, in which candidate sequences are evaluated using predictor-derived rewards and biological priors. Malinois predictor is used to predict the cell-type-specificity of the sequences

(b) Reconstruction of the sequence-grammar landscape from changes in 6-mer prevalence between RL iterations ($\Delta 6\text{-mer}$), followed by deconvolution of grammar features and their feedback into subsequent rounds of optimization.

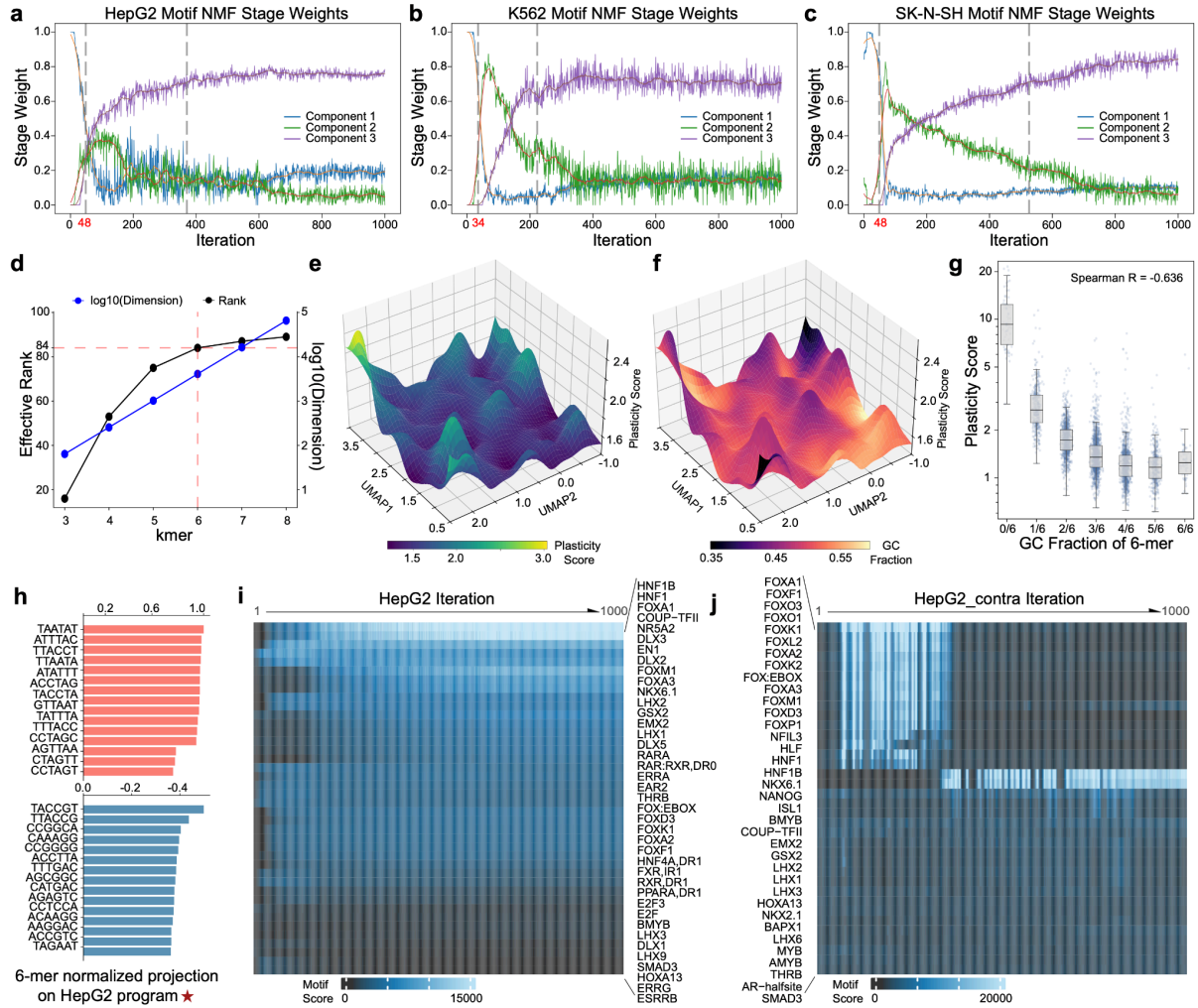


Figure S3. Phase dynamics and optimization trajectories in reconstructed sequence-grammar space, related to Figure 2.

(a–c) Non-negative matrix factorization (NMF) of generated sequences' motif presence across 1,000 optimization iterations for HepG2 (a), K562 (b), and SK-N-SH (c). Three latent components are shown. Gray dashed lines indicate putative phase boundaries, defined as the midpoints between adjacent smoothed component peaks. (d) Effect of increasing k-mer length on the dimensionality of sequence-grammar space. Black line indicates effective rank, and blue line indicates the dimensionality of sequence k-mer distributions. (e–f) Gaussian-smoothed 6-mer grammar landscape colored by plasticity score (e) or GC fraction (f). Details of generating the grammar landscape are provided in Methods. (g) Relationship between plasticity score and GC fraction across all 6-mers. Each point represents one 6-mer (Spearman's $\rho = -0.636$). (h) Top positively and negatively weighted 6-mers associated with

the first major HepG2 update program at iteration 100. The x-axis denotes the normalized projection of each 6-mer onto this program direction. Positive and negative values indicate alignment with and opposition to this direction, respectively, while larger absolute values indicate stronger associations. Scores were scaled by the largest absolute projection across all 6-mers. (i-j) Heatmap of motif match scores across HepG2 contra (i) and HepG2 (j) RL iterations. Columns represent iterations and rows represent transcription factor motifs. For panels a-c, i-j, the x axis indicates RL iteration (1-1,000), increasing from left to right; y axes indicate the corresponding plotted features.

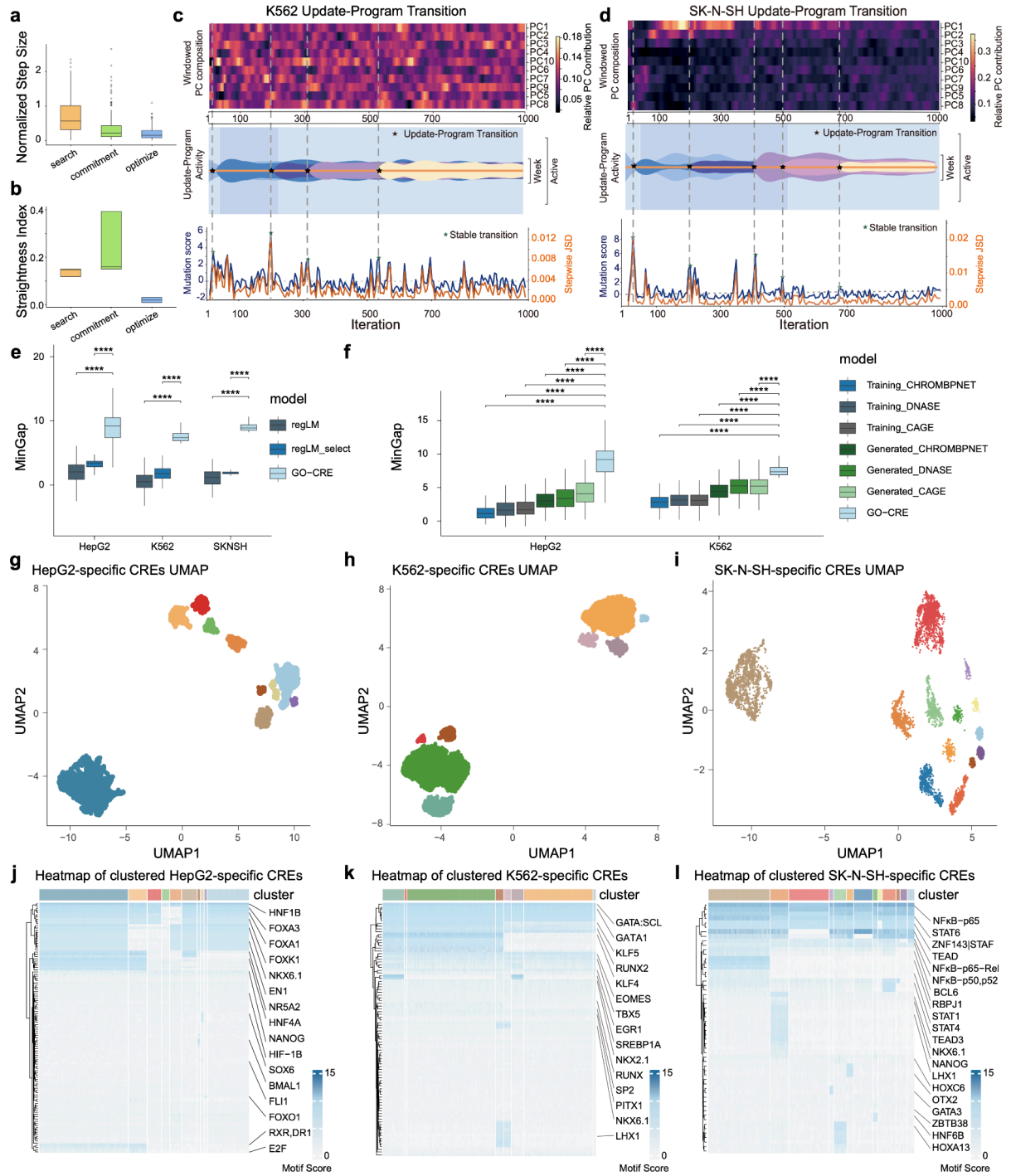


Figure S4. Dissection of generated CREs for iterations and cell-type-specificity, related to Figure 3.

(a–b) Trajectory statistics across three representative phases: search, commitment, and optimization. (a) Step size normalized to iteration 2. (b) Straightness index, defined as net displacement divided by total path length within each phase. (c–d) Update-program transitions during K562 (c) and SK-N-SH (d) CRE generation process, including windowed

principal component composition, relative principal component contributions, update-program activity, mutation score (blue), and stepwise Jensen–Shannon divergence (orange). Stars indicate emergence of stable transition points. Other plot feature representations are the same as Figure 2f-2g. (e) Comparison of MinGap scores of GO-CRE-generated CREs, regLM-generated CREs and 100 selected regLM-generated CREs across HepG2, K562, and SK-N-SH. Statistical significance was assessed using two-sided Wilcoxon rank-sum tests. GO-CRE was compared to each control group within each cell type. **** $p < 0.0001$. (f) Comparison of MinGap scores among GO-CRE generated CREs, DNA-Diffusion generated CREs, and DNA-Diffusion training CREs in HepG2 and K562. For DNA-Diffusion-generated and training CREs, the top 200 sequences were selected based on ChromBPNet, DNase, or CAGE scores. Statistical significance was assessed using two-sided Wilcoxon rank-sum tests. GO-CRE was compared to each control group within each cell type. **** $p < 0.0001$. (g–i) Individual UMAP visualization of 9,000 post-evolved generated CREs from HepG2 (g), K562 (h), and SK-N-SH (i). (j–l) Heatmap of motif scores across representative transcription factor motifs in designed CREs from HepG2 (j), K562 (k), and SK-N-SH (l). Columns represent individual CREs ordered by hierarchical clustering, excluding the noise cluster ($n = 7,898$, $8,778$ and $7,042$ for HepG2, K562 and SK-N-SH). Rows correspond to distinct motifs ($n = 107$, 91 and 55 for HepG2, K562 and SK-N-SH). Top annotation bars represent color code matched subclusters from g–i.

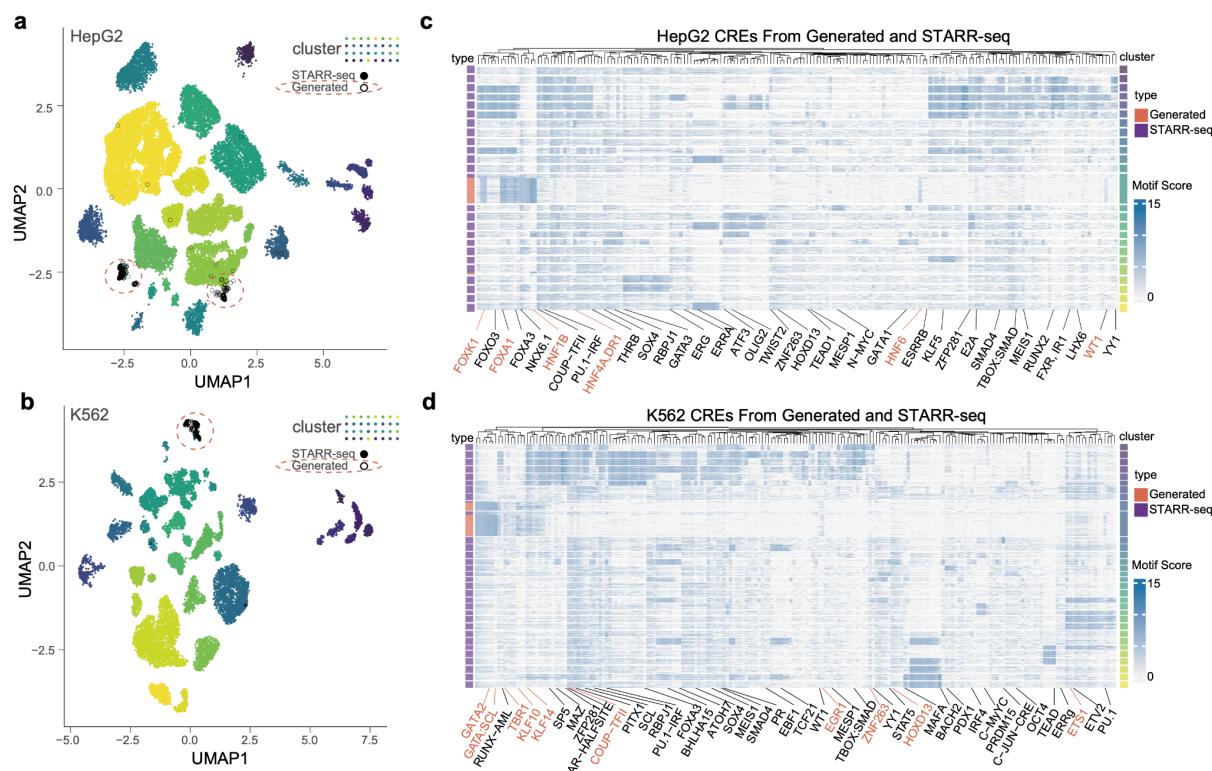


Figure S5. Comparison of generated and endogenous CREs, related to Figure 4.

(a–b) Generated CREs projected onto the endogenous CRE UMAP space in HepG2 (a) and K562 (b) cells. Endogenous CREs defined by STARR-seq are clustered into multiple subtypes indicated in the plots. Projected generated CREs are marked as blank circles.

(c–d) Heatmap of motif scores across representative transcription factor motifs in post-evolved CREs from HepG2 (c) and K562 (d). Rows represent individual CREs ordered by hierarchical clustering. Left-side annotation bars represent CRE origin from endogenous (purple) or generated (red). Right-side annotation bars represent color-coded subclusters matched to a–b.

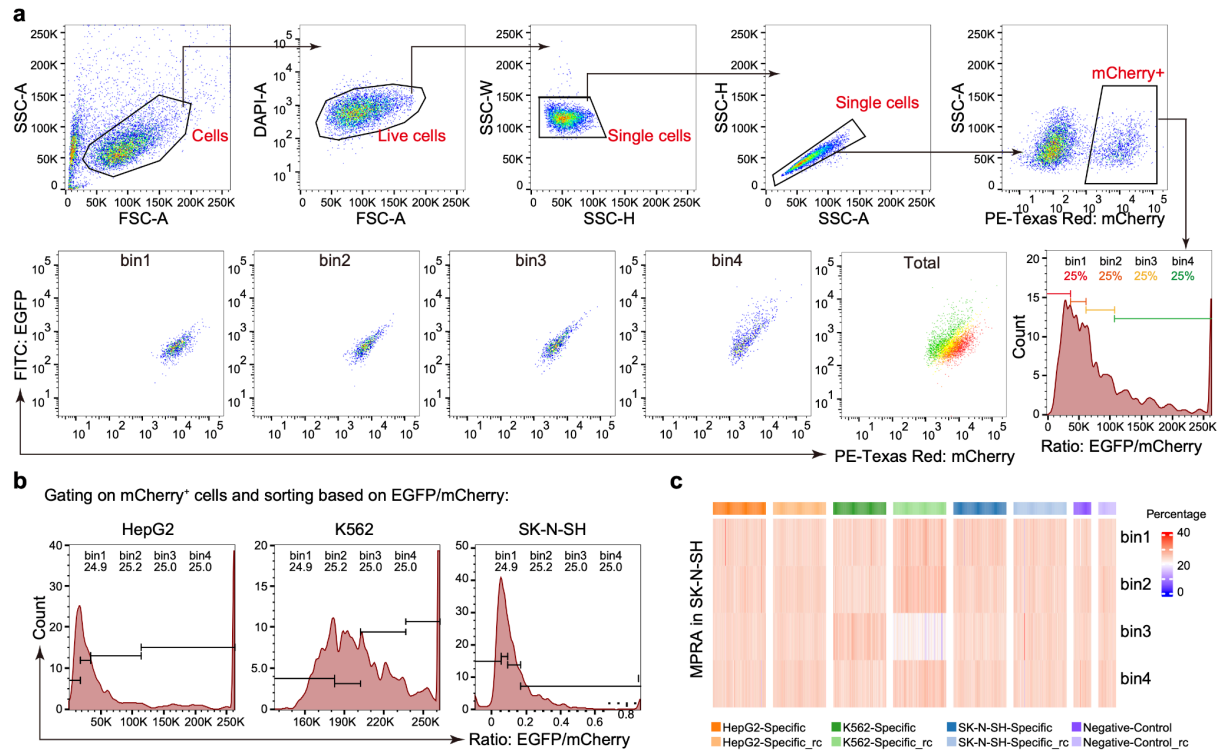


Figure S6. LentiMPRA flow cytometric sort gating strategy and SK-N-SH LentiMPRA profiling, related to Figure 5.

(a) Representative gating strategy for sorting LentiMPRA reporter cells. Cells were gated sequentially on intact events, DAPI-negative live cells, singlets, and mCherry-positive cells. The mCherry-positive population was then divided into four bins according to the EGFP/mCherry fluorescence ratio, with bins 1–4 containing approximately equal fractions of cells and corresponding to increasing enhancer activity. (b) Representative histograms of EGFP/mCherry fluorescence ratios in mCherry-positive HepG2, K562, and SK-N-SH cells. Bin boundaries used for sorting and the fraction of cells in each bin are indicated. (c) Heatmap of CRE enrichment across bins 1–4 in SK-N-SH cells. Columns represent individual CREs, and rows represent sorted bins of increasing EGFP/mCherry ratio as defined in (a). Top annotation bars indicate CRE class. RC, reverse-complemented sequence. Data represent 2 independent screens.

Supplementary Tables

Table S1: MPRA datasets for fine-tuning

Biosample	Accession
HepG2	ENCSR045VPN, ENCSR167HAR, ENCSR222QJT, ENCSR354FXV, ENCSR447RLV, ENCSR463DWN, ENCSR554UKK, ENCSR694VUM, ENCSR712VZT, ENCSR736XUK, ENCSR811QES, ENCSR833BYO, ENCSR843APF, ENCSR847SNQ, ENCSR979TSK, ENCSR632EPR, ENCSR208DGH, ENCSR839LDC, ENCSR215YXA, ENCSR463IRX, ENCSR405QCT, ENCSR255GVC, ENCSR049RIU, ENCSR022GQD, ENCSR509EUX, ENCSR202CHZ, ENCSR787MOJ, ENCSR002UCW, ENCSR760GMX, ENCSR080JVM, ENCSR478ESZ, ENCSR881ERG, ENCSR661ZUO, ENCSR069RCF
K562	ENCSR003TJR, ENCSR135MRI, ENCSR274WZF, ENCSR326HQU, ENCSR341GVP, ENCSR363XER, ENCSR653LWA, ENCSR697PNZ, ENCSR753ZRI, ENCSR820RPI, ENCSR857JFG, ENCSR909JVP, ENCSR971BEX, ENCSR971PLA, ENCSR997GVD, ENCSR460LZI, ENCSR382BVV, ENCSR203UFY
SK-N-SH	ENCSR087VPA, ENCSR386XIB, ENCSR505LQA, ENCSR541FWB, ENCSR562GGR, ENCSR563TTU, ENCSR574GPT, ENCSR579QGV, ENCSR623DJZ, ENCSR742AAP, ENCSR833GXJ, ENCSR856SQG, ENCSR959NUX

Table S2: Detailed RL training hyperparameters.

Hyperparameter	Value
Policy learning rate	5.00E-05
Value learning rate	8.00E-05
Off-Target Penalty (λ_{off})	1.0 -> 5.0 (linear)
PPO clip ratio (ϵ)	0.2
Value loss coefficient (c_v)	0.5
Entropy coefficient (c_e)	0.05
KL penalty coefficient (c_{KL})	0.01
Target KL divergence	0.04
Discount factor (γ)	0.99
GAE parameter (λ)	0.95
Batch size	64
Mini-batch size	32
PPO epochs per iteration	3
Total training iterations	1000
Warmup iterations	50
Sequences per iteration	512

Nucleus Sampling (top-p)	0.9
Experience replay buffer size	512
Experience replay ratio	0.15

Table S3: Datasets used in the paper.

Datasets	Source	Identifier
STARR-seq in HepG2 cells	ENCODE	ENCFF047LDJ
ATAC-seq in HepG2 cells	ENCODE	ENCFF439EIO, ENCFF913MQB
DNase-seq in HepG2 cells	ENCODE	ENCFF401PML, ENCFF419JVG, ENCFF978TLN
STARR-seq in K562 cells	ENCODE	ENCFF045TVA
ATAC-seq in K562 cells	ENCODE	ENCFF333TAT, ENCFF948AFM
DNase-seq in K562 cells	ENCODE	ENCFF021UTP, ENCFF103YEL, ENCFF201KEN
ATAC-seq in SK-N-SH cells	ENCODE	ENCFF716JUM, ENCFF144HUM
DNase-seq in SK-N-SH cells	ENCODE	ENCFF528WUL, ENCFF730KYY
ATAC-seq in GM12878 cells	ENCODE	ENCFF470YYO, ENCFF748UZH

ATAC-seq ChromBPNet model for K562	ENCODE	ENCFF984RAF
ATAC-seq ChromBPNet model for HepG2	ENCODE	ENCFF137WCM
Massively parallel reporter assay (MPRA) data generated in this study	GEO	GSE330487

Table S4: The MPRA backbone and pooled oligonucleotide library used in the paper.

Resources	Source
pLS-MPRA-Lenti backbone	This study
Pooled oligonucleotide library (MPRA, 250 bp)	Twist Bioscience

Table S5: Bacterial and virus strains resources used in the paper.

RESOURCE	SOURCE	IDENTIFIER
Endura™ ElectroCompetent Cells	Biosearch Technologies	Cat# 60242
TransStbl3 Chemically Competent Cell	TransGen	Cat# CD521-01
pMD2.G (second generation lentiviral plasmids)	Addgene	Cat# 12259 RRID:Addgene_12259
psPAX2 (second generation lentiviral plasmids)	Addgene	Cat# 12260 RRID:Addgene_12260

Table S6: Oligonucleotides used in the paper.

Oligonucleotides	Source
Truseq-A-F: 5'-3' ACACTCTTTCCCTACACGACGCTCTTCCGATCTAAccggTGGCcacttCATtt	GenScript
Truseq-A-R: 5'-3' TGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCCCTCTAGTGACCTtG ACAGC	GenScript
Truseq-B-F-UDI002 AATGATACGGCGACCAACCGAGATCTACACGATATCGAACACTCTTTCCC TACACGACG	GenScript
Truseq-B-R-UDI002 CAAGCAGAAGACGGCATACGAGATGGTTATAAGTGACTGGAGTTCAGA CGTGTGC	GenScript
Truseq-B-F-UDI003 AATGATACGGCGACCAACCGAGATCTACACCGCAGACGACACTCTTTCC CTACACGACG	GenScript
Truseq-B-R-UDI003 CAAGCAGAAGACGGCATACGAGATCCAAGTCCGTGACTGGAGTTCAG ACGTGTGC	GenScript
Truseq-B-F-UDI004 AATGATACGGCGACCAACCGAGATCTACACTATGAGTAACACTCTTTCCC TACACGACG	GenScript

Truseq-B-R-UDI004 CAAGCAGAAGACGGCATAACGAGATTTGGACTTGTGACTGGAGTTCAGA CGTGTGC	GenScript
Truseq-B-F-UDI005 AATGATACGGCGACCAACCGAGATCTACACAGGTGCGTACACTCTTTCC CTACACGACG	GenScript
Truseq-B-R-UDI005 CAAGCAGAAGACGGCATAACGAGATCAGTGGATGTGACTGGAGTTCAG ACGTGTGC	GenScript

Table S7: Chemicals and reagents used in this study

RESOURCE	SOURCE	IDENTIFIER
DMEM, high glucose	Thermo Fisher Scientific	Cat# 11965092
MEM	Thermo Fisher Scientific	Cat# 11095080
Fetal Bovine Serum	GeminiBio	Cat# 900-108-500
Penicillin-Streptomycin (10,000 U/mL)	Thermo Fisher Scientific	Cat# 15140122
Trypsin-EDTA (0.25%)	Thermo Fisher Scientific	Cat# 25200114
OPTI-MEM	Thermo Fisher Scientific	Cat# 31985088
DPBS, no calcium, no magnesium	Thermo Fisher Scientific	Cat# 14190144

Lipofectamine™ 3000 Transfection Reagent	Thermo Fisher Scientific	Cat# L3000075
Polybrene	Yeasen	Cat# 40804ES
BfuAI	New England Biolabs	Cat# R0701S
Q5 High-Fidelity 2× Master Mix	New England Biolabs	Cat# M0492S
Glycogen	Roche	Cat# 10901393001
DAPI	Thermo Fisher Scientific	Cat# D1306

Table S8. Other resources used in the paper.

RESOURCE	SOURCE	IDENTIFIER
EndoFree Plasmid Maxi Kit	Qiagen	Cat# 12362
HiPure Gel Pure Micro Kit	Magen	Cat# D2110
QIAamp DNA Mini Kit	Qiagen	Cat# 51304
NovoNGS® DNA Clean Beads	Novoprotein	Cat# N240