Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

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

Mingqian Ma
Carnegie Mellon University
Wanjuan Bu
Peking University
Guoqing Liu
Microsoft Research (United Kingdom)
Yuxuan Liu
National Institute of Biological Sciences, Beijing
Sizhen Liu
Peking University
Zhen Zhao
National Institute of Biological Sciences, Beijing
Shijie Yao
Peking University
Qingru Hua
University of Rochester
Yujie Zhang
Peking University
Cuiting Zhong
Peking University
Haitao Huang
Hunan University
Pan Deng
Zhongguancun Academy
Peiran Jin
Zhongguancun Academy
Qijin Yin
Zhongguancun Academy
Chuan Cao
Zhongguancun Academy
Haiguang Liu

Zhongguancun Academy

Mo Xu

National Institute of Biological Sciences, Beijing

Yuan He

Zhongguancun Academy

Tao Qin

Zhongguancun Academy

Zeyu Chen

zeyuchen@hsc.pku.edu.cn

Peking University

Research Article

Keywords: cis-regulatory element design, genomic language models, reinforcement learning.

Posted Date: September 23rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10913378/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background

Designing *cis*-regulatory elements (CREs) with cell-type-specific activity is critical for precise gene and cell therapies. Genomic language models (gLMs) have shown potential for CRE generation, but their optimization is often treated as a black box, limiting understanding of how regulatory grammar emerges and how unsuccessful generation trajectories can be corrected.

Results

We present Guided Optimization of C*is*-Regulatory Elements (GO-CRE), an interpretable deep learning framework for cell-type-specific CRE generation. GO-CRE is built on HybriDNA, a hybrid Transformer-Mamba2 gLM that supports iterative sequence optimization under practical computational constraints. GO-CRE then runs reinforcement learning (RL) and reconstructs sequence-grammar trajectories using deconvolved k-mer and motif features. The trajectories exhibit three phases - search, commitment, and optimization - each associated with the acquisition of distinct biological features. In HepG2, trajectory analysis identified a low-complexity polyG trap that motivated an updated reinforcement learning policy, and redirected optimization toward higher activity and specificity. In HepG2 and K562, GO-CRE generated diverse, cell-type-specific CREs with compact lineage-associated grammars, including HNF1B-FOXA1-HNF4A programs in HepG2 and GATA/RUNX-associated programs in K562. Lentiviral massively parallel reporter assays validated the cell-type-specific activity of the generated CREs in both cell types, and showed that HepG2 designs had higher average activity than endogenous CREs.

Conclusions

GO-CRE integrates efficient gLM-based generation, sequence grammar trajectory reconstruction, and biologically guided reward shaping. It links iterative sequence changes to regulatory grammar and feeds interpretable features back into reward design, and thus enables the design of diverse, experimentally validated, cell-type-specific CREs.

Background

Cis-regulatory elements (CREs) are central regulators of context-specific gene expression, shaping cell fate decisions, cell state transitions, and cell type specification [1–4]. Genome-wide chromatin profiling has revealed broad associations between CRE states and gene expression [5, 6], whereas genetic and epigenetic perturbation studies have provided causal evidence for how specific CREs control gene expression in a context-dependent manner [7–15]. Recent advances in gLMs have expanded our ability to decode regulatory information in DNA across organisms and cell types, particularly for understanding CRE function and regulation. Many of these models use encoder-based architectures that are well suited

to capturing contextual information within DNA sequences, and several highly interpretable models have provided base-resolution insights into regulatory mechanisms [16–30]. These advances have transformed our ability to interpret endogenous regulatory sequences. However, the inverse problem, the *de novo* design of CREs with context-specific or user-defined functions, remains substantially less mature.

CRE design has important applications in functional genomics and programmable gene control, particularly when endogenous regulatory repertoires are limited, insufficiently selective, or poorly suited to therapeutic delivery. Recent approaches have addressed this challenge through either experimental directed evolution in mammalian cells or *in silico* generative design [9, 30, 31]. *In situ* directed-evolution strategies couple mutagenesis with functional selection in native cellular environments, but remain constrained by limited mutational diversity, locus throughput and scalability [31–33]. *In silico* generative models, including convolutional neural networks, diffusion models and autoregressive gLMs, can explore substantially larger sequence spaces and have generated DNA sequences with cell-type-specific regulatory activity [34–37]. More recently, iterative deep-learning design has integrated model-guided sequence generation, experimental testing and model refinement, yielding a second generation of human enhancers with improved cell-type specificity and a more compact motif vocabulary than endogenous enhancers [38]. However, how regulatory grammar emerges, when optimization trajectories commit to a productive regulatory program, and why they become trapped in particular cellular contexts remain unclear.

Here, we present Guided Optimization of C*is*-Regulatory Elements (GO-CRE), an interpretable reinforcement learning framework for CRE design. GO-CRE reconstructs optimization trajectories in sequence-grammar space and examines the underlying biological features and events to reveal how regulatory features emerge during reinforcement learning. These trajectory-level insights enable unproductive optimization to be diagnosed and redirected through biologically informed reward shaping. Productive trajectories in K562 and HepG2 converge on distinct cell-type-associated regulatory programs. Lentiviral massively parallel reporter assays (MPRA) validate the cell-type-specific activity of the designed CREs in both cell types. Together, GO-CRE transforms synthetic CRE design from endpoint optimization into a trajectory-resolved, interpretable and steerable process for understanding and controlling regulatory sequence generation.

Results

HybriDNA combines computational efficiency with strong genomic performance for reinforcement-learning-based CRE design

To dissect and optimize synthetic CRE generation using reward-based feedback, we introduced an interpretable and tunable reinforcement learning framework for context-specific regulatory sequence design (Fig. 1a; see Methods). Since iterative optimization requires repeated sequence sampling and policy updates, the framework requires a generative model that can operate under practical GPU memory

constraints. We thus used HybriDNA, a decoder-only gLM that interleaves Mamba2 selective state-space blocks with Transformer attention layers at a 7:1 ratio, leveraging the sub-quadratic scaling of Mamba2 to process long genomic contexts efficiently while preserving nucleotide resolution via periodic attention layers. The model was pre-trained using a multispecies corpus comprising approximately 160 billion nucleotides from 816 species, and validated on 35 held-out species [39–41].

We introduced reverse-complement echo (RCE) embeddings, enabling each base-level representation to incorporate information from the full input sequence [42]. We then collected cell-type-specific MPRA data and fine-tuned HybriDNA to improve its capacity to generate cell-type-specific CREs [5, 36] (Fig. 1b; Additional file 1: Figures S1a and S1b; see Methods).

We first evaluated the computational efficiency for HybriDNA. At a matched model size of approximately 300 million parameters and under identical single-GPU settings, HybriDNA achieved higher training throughput and lower peak memory consumption than both a pure Transformer and StripedHyena across all tested context lengths. These advantages became more pronounced as sequence length increased (Figs. 1c and 1d). We next benchmarked computational efficiency during RL optimization. Under matched proximal policy optimization (PPO) settings on a single NVIDIA A100 80-GB GPU, HybriDNA-7B completed full-batch PPO optimization, whereas full-batch Evo 2-7B exceeded the memory limit. The per-iteration runtime of HybriDNA-7B was comparable to that of Evo 2-1B, whereas HybriDNA-300M and HybriDNA-3B were faster than Evo 2-1B (Figs. 1e and 1f; see Methods).

We then evaluated HybriDNA's capacity for predicting genomic sequences' features and biological functions (see Methods). On the BEND benchmark [43], HybriDNA-7B achieved strong performance across chromatin accessibility, histone modification and CpG methylation prediction tasks, matching or outperforming established gLMs including DNABERT-2, Nucleotide Transformer, HyenaDNA and Caduceus variants (Fig. 1g) [16, 21, 23, 25]. HybriDNA also performed competitively on the Genome Understanding Evaluation (GUE) tasks [16], including transcription-factor binding, promoter and core-promoter detection, epigenetic-mark prediction, viral-variant classification, and splice-site recognition (Additional file 1: Figures S1c-1i; see Methods).

GO-CRE reconstructs and redirects sequence-grammar trajectories during reinforcement learning

To extend the efficient RL capacity of HybriDNA toward cell-type-specific CRE design, we developed Guided Optimization of C*is*-Regulatory Elements (GO-CRE), a closed-loop framework that reconstructs sequence-grammar trajectories during RL and uses them to steer sequence optimization (Additional file 1: Figure S2a and Figure S2b; see Methods). This design makes CRE generation interpretable and tunable by feeding deconvolved grammar features back into biologically informed reward shaping. GO-CRE uses an MPRA-finetuned HybriDNA model as the sequence generator and a pretrained regulatory activity predictor from CODA [36] to score the cell-type-specific regulatory activity for each candidate sequence (see Methods).

To interpret the optimization process, we deconvolved each RL iteration into k-mer and motif-level features and reconstructed the sequence-grammar trajectories (Figs. 2a, 2b). We first defined a grammar space using principal component analysis (PCA) on the first 40 Δ6-mer profiles, and then reconstructed each trajectory by projecting the 6-mer profiles from each RL iteration into this space (Additional file 1: Figures S3a, S3b, S3c and S3d; see Methods). We then mapped sequence properties onto the reconstructed trajectories. In HepG2, we observed that the trajectory became trapped in a high-GC local region (HepG2 contra, Fig. 2b). Besides, we observed that GC content was correlated with how rapidly the generated sequence evolved, with low-GC sequences exhibiting faster evolution, or higher plasticity, than GC-rich sequences (Figs. 2a, 2b; Additional file 1: Figures S3e, S3f and S3g; see Methods). Along the trajectories, the generated sequences were accompanied by continuous accumulation of polyG features (HepG2 contra, Fig. 2c).

Based on this trajectory-derived failure mode, we introduced a polyG penalty into the RL reward function and repeated optimization for HepG2 (Additional file 1: Figure S2a and Figure S2b; see Methods). This intervention redirected the trajectory away from the low-complexity regime, accelerated improvement in MinGap score, and led to a more stable and higher reward plateau (Figs. 2d). These results demonstrate that the failure mode identified from sequence-grammar trajectories can be fed back into reward design to redirect optimization (Figs. 2a, 2b).

We then examined the sequence-grammar trajectory of the polyG-penalized HepG2 optimization process to understand the underlying biological events. We observed three phases of the optimization process, which we termed search, commitment, and optimization. During the search phase, generated sequences broadly explored the grammar space. As reward-aligned features accumulated, the trajectory entered a commitment phase marked by directional movement across grammar barriers and enrichment of target-associated features. As the reward approached saturation, the RL process entered the optimization phase, during which sequences continued to diversify locally while retaining their core grammar (Figs. 2a and 2b).

To quantify transitions among these phases, we decomposed windowed Δ6-mer profiles into sequence-update programs (see Methods). Three major shifts occurred: i) at iteration 100, corresponding to the search-to-commitment transition, ii) at iteration 290, when a new program emerged and the commitment-associated program remained predominant and iii) at iteration 640, when a newly dominant program marked completion of the commitment-to-optimization transition (Figs. 2e and 2f). Compared with the unpenalized HepG2 trajectory, the polyG-penalized trajectory showed more stable phase transitions, with newly emerged programs attaining sustained dominance, particularly at the search-to-commitment transition. This pattern indicates that suppressing the polyG trap promoted stable commitment to productive sequence-update programs (Figs. 2f and 2g; see Methods).

We next mapped motif-level biological changes onto these trajectory transitions (Fig. 2h; Additional file 1: Figure S3h; see Methods). In the unpenalized HepG2 contra trajectory, motifs from two major liver-lineage transcription factor (TF) families, FOXA and HNF, exhibited temporally offset waves of

enrichment (Fig. 2i, Additional file 1: Figure S3i). By contrast, polyG-penalized optimization progressively acquired lineage-associated regulatory features, with HNF1B emerging as a dominant motif, FOXA-family retained, and a second wave of HNF1B enrichment at iteration 100, coinciding with the transition (Figs. 2k and Additional file 1: Figure S3j).

Together, these analyses establish sequence-grammar trajectories as an interpretable feedback layer within GO-CRE. These trajectories reveal phase-specific regulatory events, diagnose a low-complexity failure mode, and enable reward-based redirection toward productive cell-type-associated regulatory programs.

GO-CRE generates cell-type-specific CREs with distinct regulatory programs

We next reconstructed sequence-grammar trajectories during reinforcement learning for K562 and SK-N-SH, in addition to HepG2. The trajectories differed in step size and directional straightness across the search, commitment, and optimization phases (Figs. 3a and 3b, Additional file 1: Figure S4a and S4b). The K562 and HepG2 trajectories progressed through search, commit and optimization, whereas the SK-N-SH trajectory remained largely confined to a high-GC region after the search phase (Figs. 3a and 3b). As MinGap scores increased in the three cell types, sequence diversity approached a stable level during the optimization phase (Figs. 3c and 3d; see Methods). The HepG2 and K562 generated CREs stabilized at diversity levels comparable to Malinois-filtered natural sequences, suggesting an efficient generation process can maintain sequence diversity.

We next examined the biological events along the trajectories. Similar to HepG2, both K562 and SK-N-SH acquired TF motif features accompanied by phase-associated transition points (Figs. 3e, 3f, Additional file 1: Figures S4c and S4d). The K562 trajectory progressively acquired motifs associated with hematopoietic regulators, including GATA1- and RUNX-associated motifs (Fig. 3e). Consistently, the K562 Δ6-mer trajectory exhibited multiple stable transitions between update programs, indicating the progressive acquisition of lineage-associated features (Additional file 1: Figure S4c). By contrast, SK-N-SH showed weak convergence toward neural-lineage grammar, yet strong enrichment of NF-κB, inflammatory-related and broadly active transcription-factor programs (Fig. 3f, Additional file 1: Figure S4d).

We next dissected the final generated CREs. The designs were segregated by target cell type while maintaining substantial within-group diversity (Fig. 3g). They also achieved higher predicted regulatory activity than DNase-hypersensitive natural regions, Malinois-filtered natural CREs, and Malinois-generated CREs across the three cell lines (Fig. 3h). GO-CRE designs also achieved higher predicted scores than CREs generated by regLM across all three cell lines and than DNA-Diffusion designs in K562 and HepG2 (Additional file 1: Figures S4e and S4f). The generated CREs displayed distinct, cell-line-specific TF motif profiles, further supporting the cell-type-specific generation capacity of GO-CRE (Fig. 3i, Additional file 1: Figures S4g-S4l). Core motifs were broadly represented within each target-specific design set, including FOXA1 and HNF1B in HepG2 and GATA1 in K562, whereas other motifs, such as

RARA in HepG2, were restricted to specific subclusters (Fig. 3i). These results indicate that GO-CRE captures core lineage requirements while avoiding collapse onto a single repeated sequence family.

Together, these results show that GO-CRE is able to generate diverse, cell-type-specific CRE with biologically meaningful regulatory features. HepG2 and K562 designs showed strong predicted activity and lineage-aligned motif enrichment, whereas SK-N-SH retained sequence diversity but showed weaker convergence toward a neural-lineage regulatory program.

GO-CRE-generated CREs encode compact lineage-associated motif programs

To understand the regulatory logic encoded by GO-CRE-generated CREs, we examined their chromatin accessibility and motif architecture. ChromBPNet predicted that designed CREs were more accessible in matched than unmatched cell types, indicating their cell-type-specific chromatin accessibility (Figs. 4a and 4b; see Methods) [17]. We further observed that, whereas endogenous CREs showed broad motif distributions, GO-CRE-generated CREs were concentrated within one or two dominant motif clusters, suggesting convergence on shared motif grammars (Additional file 1: Figures S5a-S5d). Comparison with STARR-seq-identified endogenous CREs [44] further showed a greater prevalence of lineage-aligned motifs among generated CREs (Figs. 4c and 4d; see Methods). HepG2 designs were enriched for HNF1B, HNF4A, FOXA1, and other hepatocyte-associated motifs, whereas K562 designs were enriched for motifs associated with hematopoietic regulators, including GATA1/2, RUNX2, and EOMES (Figs. 4c and 4d). These results indicate that GO-CRE distilled compact lineage-associated motif programs rather than recapitulating the full heterogeneity of endogenous CREs.

We next asked whether this convergence extended beyond motif abundance to higher-order motif syntax (see Methods). In HepG2, distances from HNF4A or HNF1B to a central FOXA1 motif exhibited sharper and more recurrent peaks in generated CREs than in endogenous CREs, with orientation-specific spacing observed on both sides of FOXA1 (Figs. 4e and 4f). Joint analysis of the two distances revealed a concentrated HNF1B-FOXA1-HNF4A configuration in generated sequences, compared with the more diffuse organization of endogenous CREs (Fig. 4g).

Trajectory analysis showed that this higher-order motif organization emerged progressively during optimization. Across representative iterations 29, 99, and 249, the joint HNF1B-FOXA1-HNF4A distance distribution became increasingly structured, while the full trajectory showed continued consolidation of the motif configuration (Fig. 4h, Additional file 2: Video S1). Representative endpoint designs encoded compact arrangements of the three motifs with recurrent spacing and orientation (Fig. 4i). Together, these results show that GO-CRE-generated sequences exhibited cell-type-specific accessibility and encoded more compact lineage-associated motif syntax than matched endogenous CREs.

LentiMPRA validates the cell-type-specific activity of generated CREs

We next used a fluorescence-based lentiviral massively parallel reporter assay (MPRA) to experimentally evaluate the activity and cell-type-specificity of GO-CRE-generated CREs. We designed a dual-reporter

vector in which hPGK-driven mCherry served as an internal normalization control for variation in lentiviral delivery and genomic integration, whereas candidate CREs were cloned upstream of a minimal promoter-EGFP cassette to report enhancer activity. Four days after transduction, mCherry-positive cells were sorted into four bins according to the EGFP/mCherry ratio, representing increasing levels of CRE activity (Fig. 5a; Additional file 1: Figures S6a and S6b; see Methods).

We then profiled the CRE libraries designed for HepG2, K562, and SK-N-SH, together with reverse-complemented sequences and negative controls (see Methods). In K562, K562-target CREs were preferentially enriched in high-activity bins and achieved higher activity scores than off-target sequences and their reverse-complemented counterparts (Figs. 5b and 5d; see Methods). Similarly, in HepG2, HepG2-targeted CREs were concentrated in bins with the highest activity and outperformed off-target sequences and controls (Figs. 5c and 5e). By contrast, SK-N-SH-targeted CREs did not show clear target-selective activity, consistent with weak convergence toward neural-lineage regulatory grammar observed along their optimization trajectory (Additional file 1: Figure S6c). Clonal validation of individual HepG2-generated CREs further confirmed strong HepG2-specific reporter activation relative to a random-sequence negative control (Fig. 5f).

We next compared designed CREs targeting HepG2 with endogenous HepG2 CREs identified by STARR-seq. On average, generated CREs showed higher LentiMPRA activity than endogenous CREs, although the top-performing sequences in the two groups achieved comparable scores (Figs. 5g and 5h). Despite their similar maximal activities, the underlying motif organizations differed: the highest-ranking generated CREs encoded compact HNF1B-FOXA1-HNF4A-like syntax, whereas the highest-ranking endogenous CREs displayed more heterogeneous motif combinations (Fig. 5i). Together, these results indicate that GO-CRE can generate diverse, cell-type-specific CREs that are functionally competitive with endogenous regulatory elements while converging on compact, lineage-aligned motif programs.

Discussion

Cis-regulatory elements design provides an approach to program context-specific gene expression for both basic research and gene and cell therapy applications [7, 45, 46]. Generative models can expand the accessible design space beyond endogenous regulatory repertoires, particularly for cellular states that are under-represented in existing datasets. However, most sequence design approaches optimize primarily for endpoint performance. Although predictive models can rank candidate sequences, they provide limited insight into how regulatory features emerge during optimization, why trajectories succeed or fail, or how biological knowledge can be used to redirect unproductive searches. Consequently, score-driven optimization may exploit low-complexity or non-physiological sequence features that satisfy the predictor without establishing a coherent regulatory program.

GO-CRE addresses this limitation by treating CRE design as an observable and steerable process. Enabled by the computational efficiency of HybriDNA, GO-CRE supports iterative sequence generation while reconstructing changes in sequence grammar throughout optimization. Productive trajectories

resolve into search, commitment, and optimization phases, with broad exploration progressively narrowing toward stable regulatory programs. The combination of iterative sequence variation, functional selection, staged feature acquisition, and reward-dependent trajectory redirection makes GO-CRE conceptually analogous to *in silico* directed evolution. Under this framework, the objective extends beyond selecting a functional endpoint, to understanding and controlling the path through which that endpoint is reached.

These trajectories reveal that successful CRE optimization is shaped by the path through sequence-grammar space. Productive optimization was associated with coordinated transitions among sequence-update programs and convergence toward lineage-aligned regulatory features. The polyG intervention in HepG2 provides proof of principle that trajectory-derived failure modes can be identified during optimization and converted into actionable constraints. Thus, GO-CRE functions as both a generative engine and a diagnostic framework for CRE generation.

The generated CREs suggest that synthetic and endogenous regulatory elements may achieve comparable regulatory activity through different grammatical solutions. GO-CRE designs retained sequence diversity while concentrating lineage-associated motifs into compact regulatory programs. In HepG2, the progressive emergence of an HNF1B-FOXA1-HNF4A organization suggests that higher-order motif syntax can be progressively assembled during iterative optimization. The earlier emergence of FOXA-family motifs relative to HNF1B was consistent with established roles of FOXA factors as pioneer regulators and HNF factors in hepatocyte identity [47, 48]. This correspondence is intriguing, although it should not be interpreted as evidence that the optimization trajectory recapitulates liver development.

LentiMPRA confirmed cell-type-specific activity for K562 and HepG2 designs and showed that the most active synthetic and endogenous CREs achieved comparable reporter output despite distinct motif organizations. The lack of clear cell-type-specific activity in SK-N-SH designs further indicates that trajectory analysis may help identify cellular contexts in which the available training data, predictor, or reward formulation is insufficient for productive design. However, reporter assays measure autonomous regulatory activity outside native genomic and chromatin contexts. Future work should therefore test the designs at endogenous loci in primary cells and in vivo to incorporate chromatin state and transcriptional information.

More broadly, extending GO-CRE beyond reporter-based optimization will require objectives that reflect endogenous chromatin, genomic context, and stimulus-responsive states, because reporter-active and native regulatory grammars may not fully overlap. Matched functional and chromatin datasets, particularly state-resolved accessibility and transcriptional measurements, will be important for training such objectives. The weaker performance in SK-N-SH further highlights the need for richer cell-state-resolved data and biological priors, especially for heterogeneous or underrepresented cellular systems [49, 50]. Search strategies that preserve multiple candidate trajectories may also reduce premature convergence on a single dominant solution. Together, these extensions could move trajectory-resolved

CRE design from autonomous reporter activity toward predictable regulatory control in native cellular and therapeutic contexts.

Conclusions

We established GO-CRE as a framework for making cell-type-specific CRE design interpretable and steerable. By reconstructing the trajectories through which regulatory grammar emerges, GO-CRE enables unproductive optimization paths to be diagnosed and redirected. The resulting CREs exhibited experimentally validated cell-type-specific activity in K562 and HepG2. By shifting CRE generation from black-box sequence optimization to trajectory-aware design, GO-CRE establishes a foundation for controllable CRE engineering and systematic analysis of the regulatory grammar dynamics during gLM optimization.

Methods

HybriDNA Architecture and Model Configurations

HybriDNA is a decoder-only autoregressive DNA language model in which Mamba2 selective state-space model (SSM) blocks and Transformer attention blocks are interleaved at a 7:1 ratio. Mamba2 blocks provide sub-quadratic O(NL) complexity, where N is the SSM state dimension and L is sequence length, while the periodic Transformer blocks use causal masking, RMSNorm, grouped-query attention and a feed-forward multilayer perceptron to retain high-fidelity nucleotide interactions. The first layer is a Mamba2 block, and positional information is represented through the recurrent SSM dynamics without explicit positional embeddings. Each Mamba2 block follows the State-Space Duality formulation with input-dependent parameter generation [51, 52]. The 7:1 hybrid ratio was empirically selected over a pure-Mamba2 baseline based on lower pretraining loss at matched parameter count, consistent with recent observations across hybrid language models [39–41].

Three model versions were trained and evaluated. HybriDNA-300M comprised 24 layers with hidden dimension 1,024, intermediate dimension 2,048, 32 attention heads, eight key-value heads and head dimension 64. HybriDNA-3B comprises 16 layers with hidden dimension 4,096, intermediate dimension 8,192, 128 attention heads, eight key-value heads and head dimension 64. HybriDNA-7B comprised 32 layers with hidden dimension 4,096, intermediate dimension 8,192, 128 attention heads, eight key-value heads and head dimension 64. HybriDNA-300M was used for SFT and CRE optimization; the larger variants were included in scaling and computational-efficiency analyses.

Model Pre-training, Fine-tuning and Benchmarking

Multispecies Pretraining. The pretraining corpus was curated from the Nucleotide Transformer collection derived from NCBI RefSeq [53]. It contained approximately 160 billion nucleotides from 816 training species: 647 bacterial species (16.5 billion bp), 44 fungal species (2.0 billion bp), 37 invertebrate species

(19.9 billion bp), nine protozoan species (0.45 billion bp), 28 mammalian vertebrate species (65.2 billion bp) and 51 non-mammalian vertebrate species (57.4 billion bp). An additional 35 species comprising 13.25 billion bp were held out for validation. Sequences were tokenized at single-nucleotide resolution using A, C, G and T. Long-context capacity was introduced through a three-stage curriculum: 8,192-token sequences for 500,000 steps at approximately 0.5 million tokens per batch; 32,768-token sequences for a further 10,000 steps; and 131,072-token sequences for a final 10,000 steps. Models were trained using next-token prediction and AdamW ($\beta_1 = 0.9,\ \beta_2 = 0.95,\ \epsilon =$1e^{-8}, weight decay 0.1). Learning rates were 1 × 10^{-3}, 6 × 10^{-4} and 1 × 10^{-4} for the 300M, 3B and 7B models, respectively, with 2,000 warm-up steps followed by cosine decay. Training used BF16 mixed precision, DeepSpeed ZeRO Stage 1 and FlashAttention 2 [54].

Reverse-Complement Echo Embeddings. To include reverse-complement echo embeddings, an input sequence x of length L was concatenated with either a copy of itself, $x_{echo} = [x; x]$, or its reverse complement, $x_{echo} = [x;\ RC(x)]$. RC(x) was obtained by A↔T and C↔G substitution followed by sequence reversal [42]. The decoder processed the resulting 2L-token input autoregressively. Hidden states at positions L + 1 to 2L could therefore condition on the complete first copy. Representations were obtained by mean pooling hidden states from the echoed half and were passed to a task-specific classification head. The twofold increase in input length produced an approximately 1.5-fold increase in fine-tuning wall-clock time because only the periodically interleaved Transformer layers incurred quadratic cost.

Supervised Fine-Tuning on MPRA Data. We curated 200-bp CREs for K562, HepG2 and SK-N-SH cells using MPRA datasets from ENCODE [5] and the CODA framework [36] (Additional file 1: Table S1). Malinois predicted activity in all three cell types. Within each cell type, sequences in the upper activity quartile were labeled high activity and those in the lower quartile were labeled low activity. A sequence was designated cell-type-specific when it was high activity in exactly one cell type and low activity in the other two. Cell-type prompt tokens ([K562], [HepG2] and [SKNSH]) were added to the HybriDNA vocabulary and prepended to each training sequence. The pretrained HybriDNA-300M model was fine-tuned for three epochs using an 80/20 train-validation split, and the checkpoint with the lowest validation loss was used to initialize all subsequent RL tasks. SFT was performed on one NVIDIA A100 GPU with 80 GB memory and required approximately five hours.

Benchmark Evaluation. BEND [43] presents a series of biologically meaningful tasks for genomics foundation models derived from a series of studies. We selected the three short-range supervised tasks: Chromatin Accessibility, Histone Modification, and CpG Methylation. For the BEND benchmark, pretrained embeddings were frozen and a two-layer convolutional classifier was trained for 100 epochs with learning rate 3×10^{-3} and linear decay to zero. Mean-pooled hidden states were used for standard autoregressive models, and hidden states from the echoed half were pooled for echo variants. Model selection was based on the lowest validation loss, and mean AUROC was reported on the held-out test set.

Genome Understanding Evaluation (GUE) [16] comprised 28 datasets spanning promoter and core-promoter detection, splice-site prediction, human and mouse transcription-factor binding, yeast epigenetic-mark prediction and COVID variant classification. Full fine-tuning used Adam with learning rate 5×10^{-5}, batch size 32 and weight decay 0.01. Standard, repetition-echo and reverse-complement-echo representations were evaluated using a linear classification head. Matthews correlation coefficient was reported for all tasks except COVID variant classification, for which F1 score was used. Results represent mean ± standard deviation over three random seeds. *Computational Efficiency Analysis.* Training throughput and peak GPU memory were measured for approximately 300-million-parameter HybriDNA, Transformer and StripedHyena models on one NVIDIA A100 GPU with 80 GB memory. Throughput used the largest batch size that fit memory, whereas peak memory was measured at batch size one. For the RL benchmark, HybriDNA-300M, HybriDNA-3B and HybriDNA-7B were compared with Evo 2-1B and Evo 2-7B under identical PPO settings: 512 sequences per iteration, batch size 64, mini-batch size 32 and generation length 200 bp. Full-batch Evo 2-7B exceeded the 80-GB memory limit and was additionally evaluated at three-quarters of the original batch size.

Reinforcement Learning Process

We used the HybriDNA-300M model as the generative backbone for cell-type-specific CRE design given its high efficiency and strong performance (Fig. 1g). The SFT-initialized generator was optimized using Proximal Policy Optimization (PPO) [55]. Each generated 200-bp sequence was padded with the flanking MPRA vector context expected by Malinois. For a sequence *x* designed to target cell type *t*, the MinGap reward was defined as $R_{raw}(x,\ t) = S_t\left(x\right) -\ \lambda_{\ off} max_{c \neq\ t} S_c\left(x\right)$, where $S_t\left(x\right)$ denotes the Malinois-predicted activity in the target cell type and $max_{c \neq\ t} S_c\left(x\right)$ denotes the maximum predicted activity across all off-target cell types. The off-target penalty coefficient $\lambda_{\ off}$ was increased linearly from 1.0 at iteration 0 to 5.0 at iteration 1,000. Quality-based reward shaping was applied to accelerate convergence. We further introduced a trajectory-informed polyG penalty to penalize the low-complexity polyG trap identified by trajectory analysis of the HepG2 run. The reward was defined as $R_{raw,\ corr}(x) = R_{raw}\left(x\right) -\ P_{polyG}[x]$, where $P_{polyG}[x] = max[0,\ L_G\left(x\right) - 5,\ \ L_C\left(x\right) - 5]$, and $L_G\left(x\right)$ denotes the longest consecutive run of G in sequence x and $L_C\left(x\right)$ denotes the longest consecutive run of C in sequence x to account for the reverse complement sequence of polyG. Further details are provided in Additional file 1: Supplementary Methods and Additional file 1: Table S2.

Analysis of the sequence dynamics during reinforcement learning

Representations of the generated sequences. Sequence feature dynamics were characterized using complementary k-mer and TF motif representations. At each reinforcement learning iteration, *K-mer* prevalence was defined as the fraction of generated sequences containing each k-mer, and changes between consecutive iterations *ΔK-mer* were used to quantify shifts in sequence composition. Motifs

were identified using HOMER v5.1 with the Vertebrata known motif collection [56]. Motif match scores were aggregated to construct iteration-level profiles, whereas sequence-level profiles retained the strongest match for each sequence-motif pair. To identify major phases of sequence optimization, we applied non-negative matrix factorization to motif-presence of the generated sequences across iterations and defined stage boundaries from the smoothed temporal weights of three components. Additional details are provided in Additional file 1: Supplementary Methods.

Construction of the shared sequence grammar space. To construct a shared representation of early sequence-optimization dynamics, we pooled the first 40 Δ6-mer vectors from each of the three cell-type-specific tasks and performed eigendecomposition of the resulting 6-mer covariance matrix. The number of retained dynamic axes was determined from the spectral effective rank of the covariance matrix. Each 6-mer was represented by its loadings on these shared axes, and its plasticity was quantified as the eigenvalue-weighted energy of its loadings. Standardized 6-mer representations were embedded into two dimensions using UMAP [57], and the distributions of plasticity and GC content were visualized as smoothed landscapes across the resulting sequence-grammar space. Additional details are provided in Additional file 1: Supplementary Methods.

Task-level windowed PCA analysis of sequence update dynamics. To characterize dynamic sequence-grammar programs during optimization, we performed task-specific PCA on the full Δ6-mer trajectory and projected sliding windows of iterations onto the resulting dynamic axes. Window-level principal component (PC) contributions were quantified from the root mean square (RMS) magnitude of PC scores across 20-iteration windows and normalized to represent relative dynamic activity. Program transition points were identified based on abrupt changes in PC contribution profiles using Jensen-Shannon divergence and additional stability filtering. Initial and transition-associated programs were defined from sparse combinations of dominant PCs and their temporal changes. Program activities were quantified by gated temporal activation and mapped back into sequence space through PCA loadings, enabling identification of associated 6-mer and TF motif signatures. Additional details are provided in Additional file 1: Supplementary Methods.

Analysis of the Generated Sequences

Motif-based clustering and visualization. To characterize sequence diversity and motif usage patterns among the generated sequences, motif feature matrices were embedded into a low-dimensional space using UMAP [57] and clustered using HDBSCAN [58], enabling identification of density-based sequence groups without predefining the number of clusters. Cluster-specific motif usage patterns were visualized using heatmaps generated from the original motif feature matrices [63, 64]. Additional details are provided in Additional file 1: Supplementary Methods.

Normalized 6-mer Vendi scores. Sequence diversity was quantified using the Vendi score based on sequence-level 6-mer composition. For each sequence, 6-mer frequencies were L2-normalized, and pairwise cosine similarities were used to construct a similarity kernel matrix. The Vendi score was

calculated from the entropy of the normalized non-negative eigenvalues of this matrix and further normalized by the smaller of the number of sequences and the 6-mer feature dimensionality (4,096). Scores were calculated independently for each reinforcement-learning iteration and reference sequence set.

Comparison with endogenous CREs. Endogenous CREs were obtained from reproducible HepG2 and K562 STARR-seq peaks processed using datasets ENCFF047LDJ and ENCFF045TVA, respectively [44] (Additional file 1: Table S3). For each cell type, endogenous CREs and the designed candidate CREs were represented using motif-based sequence feature matrices as described above and analyzed jointly.

ChromBPNet prediction. Chromatin accessibility of 200-bp candidate CREs was evaluated using ChromBPNet with pretrained ATAC-seq models for HepG2 (ENCFF137WCM) and K562 cells (ENCFF984RAF). For ChromBPNet prediction, each 200-bp candidate sequence was embedded at a fixed position within a standardized 2,114 bp input sequence. Specifically, candidate sequences replaced the 200-bp segment corresponding to chr16: 28,930,800 – 28,930,999 within a fixed 2,114 bp genomic window spanning chr16:28,929,800 – 28,931,913. Predictions were generated using five independently trained no-bias ChromBPNet models per cell line, with averaged outputs used as the final chromatin accessibility profiles. For each sequence, the mean predicted accessibility within the insertion window (positions 470–570 of the model output) was extracted and used to represent the chromatin accessibility of the corresponding CRE.

Motif configuration analysis. Motif syntax analysis was performed for three selected motifs. Candidate CREs containing all three motifs were extracted, and motif positions were represented by the offset reported by HOMER. For each analysis, one motif was designated as the reference, and each occurrence of it was treated as the coordinate origin. The relative positions of the other two motifs with respect to this origin were calculated, forming a set of three-motif position triplets. The resulting two-dimensional distributions of motif triplets were visualized as 3D bar plots using *persp()* in R.

MPRA library design

To evaluate the regulatory activity of the designed CREs, we designed MPRA libraries containing the designed sequences, endogenous CRE controls, and genomic negative controls. Negative control sequences were derived from GM12878 open chromatin regions and filtered to exclude accessibility signals from the tested cell types (K562, HepG2, and SK-N-SH), minimizing the likelihood of endogenous enhancer activity in the validation contexts.

For cell-type-specific CRE validation, the MPRA library consisted of the designed candidate CREs selected from multiple sequence-grammar clusters identified during model optimization. Candidate sequences were ranked based on both predicted regulatory activity and cell-type specificity, and representative sequences from the highest-performing clusters were selected for experimental validation. Reverse-complement sequences were included to evaluate strand effects.

For comparison with endogenous regulatory elements, we constructed an MPRA library containing HepG2 endogenous CREs identified from STARR-seq, designed HepG2-specific CREs, and negative controls. Endogenous CRE candidates were selected based on STARR-seq activity, predicted chromatin accessibility, and sequence diversity. Designed CRE candidates were selected using a two-stage prioritization strategy based on predicted regulatory activity and ChromBPNet-predicted accessibility. Detailed library construction and sequence selection procedures are provided in Additional file 1: Supplementary Methods.

Lentiviral MPRA library construction and quality control

Synthetic oligonucleotide libraries containing the designed CRE sequences were cloned into a lentiviral MPRA vector following PCR amplification. The MPRA backbone was derived from pLS-SceI and modified to include an hPGK promoter-driven mCherry expression cassette as a transduction and normalization control, while retaining a minimal promoter-driven destabilized EGFP reporter cassette for measuring CRE activity [65]. The cloning site was replaced with dual BfuAI sites to enable insertion of the synthetic oligonucleotide library. The pooled oligonucleotide library can be found in Additional file 1: Table S4.

Amplified double-stranded oligonucleotide pools were cloned into the linearized MPRA backbone using seamless cloning, followed by transformation into high-complexity electrocompetent bacteria to generate plasmid libraries (Additional file 1: Table S5). The final MPRA vector contained two expression cassettes: an upstream CRE-responsive reporter consisting of the inserted regulatory sequence, minimal promoter, destabilized EGFP fused to a PEST domain, and bGH polyadenylation signal; and a downstream constitutive mCherry cassette used for normalization of transduction efficiency.

Plasmid libraries were subjected to deep sequencing before lentiviral production to evaluate library complexity and sequence representation. Insert regions were amplified using TruSeq-compatible primers followed by UDI indexing and paired-end sequencing on an Illumina NovaSeq 6000 platform. Detailed cloning procedures, primer sequences, and library quality-control procedures are provided in Additional file 1: Supplementary Methods, and oligonucleotide and primer sequences are listed in Additional file 1: Table S6.

Lentiviral MPRA screening and fluorescence-based activity measurement

HepG2, K562, and SK-N-SH cells were used for lentiviral MPRA screening of the designed CRE libraries. Lentiviral particles were produced in 293T cells using a second-generation packaging system and used to transduce target cell lines at a multiplicity of infection (MOI) below 0.3 to minimize multiple integrations (Additional file 1: Table S5). Following transduction, reporter activity was quantified based on the ratio of destabilized EGFP fluorescence normalized by constitutive mCherry expression.

At 96h after infection, transduced cells were enriched and sorted by flow cytometry. After gating for viable singlets and mCherry-positive cells, populations were separated into four bins according to the EGFP/mCherry fluorescence ratio, representing different levels of CRE-driven reporter activity. Approximately 10^6 cells from each fluorescence bin were collected for downstream sequencing.

Genomic DNA from sorted populations was extracted, and integrated MPRA constructs were amplified using indexed sequencing libraries and quantified by paired-end Illumina sequencing. Sequence-level activity scores were calculated based on the relative enrichment of each construct across fluorescence bins as described above. Detailed cell culture conditions, lentiviral production, sorting strategy, and sequencing library preparation procedures are provided in Additional file 1: Supplementary Methods.

Chemicals, culture reagents, enzymes, and other experimental reagents used for LentiMPRA construction and screening are listed in Additional file 1: Table S7; DNA extraction and purification resources are listed in Additional file 1: Table S8.

Analysis of MPRA sequencing reads

Raw MPRA sequencing reads were first adapter-trimmed using cutadapt v5.1 [59] and merged using PEAR v0.9.6 [60] to generate contiguous sequences for alignment. For libraries containing reverse complement (RC) sequences, separate Bowtie2 v2.5.4 [61] indices were built for the forward and RC sequences. Merged reads were aligned independently to both indices, and only forward-aligned reads were retained to generate counts represented forward and RC sequences respectively. For libraries without RC sequences, a single Bowtie2 index was built and reads were aligned once to extract counts. BAM files were sorted and indexed using samtools v1.22.1 [62], and read counts for each sequence were tabulated. For each sequence, activity scores were computed as a weighted sum of binned signal counts. Specifically, four bins (bin1-bin4) were combined as:

$$score = 3 \times\ bin4 + bin3 - bin2 - 3 \times\ bin1$$

The resulting score was then standardized using z-score normalization, yielding the normalized activity score. For experiments with multiple replicates, normalized scores were averaged across replicates to obtain the final activity score for each sequence.

Use of generative AI and AI-assisted technologies

During the preparation of this manuscript, the authors used Codex to assist with code review, and GPT-5.4 and GPT-5.5 Pro to assist with language editing and manuscript refinement. All outputs generated with these tools were reviewed and edited by the authors, who take full responsibility for the content of the published article.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Funding

Z.C. was supported by the “U35 Pei Miao Funding Program” of Changping District, Beijing (No. CHPU35202501001), Beijing Natural Science Foundation (No.5264034, No. Z231100004823015), Noncommunicable Chronic Diseases-National Science and Technology Major Project (No. 2023ZD0501204), Open Funding for National Center for Protein Sciences at Peking University (No. KF-202501). Y.H. was supported by Beijing Natural Science Foundation (No. Z260011), National Natural Science Foundation of China (No. 32600792), Zhongguancun Academy (No. XTS0010).

Author Contribution

M.M., W.B., and Z.C. conceived the study; M.M. and G.L. supported the use of HybriDNA with input from P.D., Q.Y., C.C., H.L., and T.Q.; M.M., W.B., and Z.C. optimized the RC-echo embedding strategy and designed the GO-CRE reinforcement learning framework and reward shaping; W.B. performed bioinformatic analyses and scoring with support from Z.C. and Y.H.; Y.L., S.L., Z.Z., and S.Y. performed wet-lab experiments with support from Z.C. and M.X.; P.D. and P.J. provided AI modeling and computational assistance with support from H.L. and T.Q.; Z.C., W.B., M.M., Y.H., Y.L., and S.L. wrote and edited the manuscript with assistance from Q.H., Y.Z., C.Z., P.D., Q.Y., M.X., and T.Q.

Acknowledgement

We thank Ting Han’s lab at NIBS for providing HepG2, K562 and 293T cell lines, Salvador Casaní-Galdón, Yanjing Liu, Shuang Li, Ruochi Zhang, and Bradley E. Bernstein for initial idea shaping, and Jiecong Lin, Bohan Li, Han Yuan, Zhijian Li, Xi Dawn Chen and Fadi Najm for helpful discussions.

Data Availability

MPRA data generated from this paper is publicly available in GEO under accession GSE330487

Code availability

The implementation of the HybriDNA model, the GO-CRE framework, scripts for sequence grammar landscape analysis, and code for MPRA experimental data processing are available at https://github.com/zeyuChenLab/GO-CRE.git.

Competing interests

The authors declare no competing interests.

References

1. Wittkopp P, Kalay G. Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. Nat Rev Genet. 2012;13:59–69.
2. Preissl S, Gaulton KJ, Ren B. Characterizing cis-regulatory elements using single-cell epigenomics. Nat Rev Genet. 2023;24(1):21–43.
3. Andersson R, Sandelin A. Determinants of enhancer and promoter activities of regulatory elements. Nat Rev Genet. 2020;21(2):71–87.
4. Nasser J, Bergman DT, Fulco CP, Guckelberger P, Doughty BR, Patwardhan TA, et al. Genome-wide enhancer maps link risk variants to disease genes. Nature. 2021;593(7858):238–43.
5. ENCODE Project Consortium, Moore JE, Purcaro MJ, Pratt HE, Epstein CB, Shoresh N, et al. Expanded encyclopaedias of DNA elements in the human and mouse genomes. Nature. 2020;583(7818):699–710.
6. Stunnenberg HG, International Human Epigenome Consortium, Hirst M. The International Human Epigenome Consortium: A Blueprint for Scientific Collaboration and Discovery. Cell. 2016;167:1145–9.
7. Martyn GE, Montgomery MT, Jones H, Guo K, Doughty BR, Linder J, et al. Rewriting regulatory DNA to dissect and reprogram gene expression. Cell. 2025;188(12):3349–e6623.
8. Martin-Rufino JD, Castano N, Pang M, Grody EI, Joubran S, Caulier A, et al. Massively parallel base editing to map variant effects in human hematopoiesis. Cell. 2023;186(11):2456–e7424.
9. Sanjana NE, Wright J, Zheng K, Shalem O, Fontanillas P, Joung J, et al. High-resolution interrogation of functional elements in the noncoding genome. Science. 2016;353(6307):1545–9.
10. Gilbert LA, Larson MH, Morsut L, Liu Z, Brar GA, Torres SE, et al. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. Cell. 2013;154(2):442–51.
11. Ren X, Yang H, Nierenberg JL, Sun Y, Chen J, Beaman C, et al. High-throughput PRIME-editing screens identify functional DNA variants in the human genome. Mol Cell. 2023;83(24):4633–e46459.
12. Diao Y, Fang R, Li B, Meng Z, Yu J, Qiu Y, et al. A tiling-deletion-based genetic screen for cis-regulatory element identification in mammalian cells. Nat Methods. 2017;14(6):629–35.
13. Chen Z, Javed N, Moore M, Wu J, Sun G, Vinyard M, et al. Integrative dissection of gene regulatory elements at base resolution. Cell Genom. 2023;3(6):100318.

14. Fulco CP, Munschauer M, Anyisi R, Mockus TE, Erdin S, Kamitaki N, Schenone M, Macdonald JR, Macosko EZ, Doench JA, et al. Systematic mapping of functional enhancer–promoter connections with CRISPR interference. Science. 2016;354:769–73.
15. Korkmaz G, Lopes R, Ugalde AP, Nevedomskaya E, Han R, Myacheva K, Zwart W, Elkon R, Agami R. Functional genetic screens for enhancer elements in the human genome using CRISPR-Cas9. Nat Biotechnol. 2016;34:192–8.
16. Zhou Z et al. vol. DNABERT-2: efficient foundation model and benchmark for multi-species genomes. In International Conference on Learning Representations (2024), 41642–41665 (2024).
17. Pampari A, Shcherbina A, Kvon EZ et al. ChromBPNet: bias factorized, base-resolution deep learning models of chromatin accessibility reveal cis-regulatory sequence syntax, transcription factor footprints and regulatory variants. Preprint bioRxiv 202412.25.630221; doi: https://doi.org/10.1101/2024.12.25.630221
18. Sanabria M, Hirsch J, Joubert PM, Poetsch AR. DNA language model GROVER learns sequence context in the human genome. Nat Mach Intell. 2024;6(8):911–23.
19. Linder J, Srivastava D, Yuan H, Agarwal V, Kelley DR. Predicting RNA-seq coverage from DNA sequence as a unifying model of gene regulation. Nat Genet. 2025;57(4):949–61.
20. Yuan H, Linder J, Kelley DR. Parameter-efficient fine-tuning enables scalable transfer of regulatory sequence models to novel contexts. Genome Biol. 2026;27:60.
21. Nguyen E, Poli M, Faizi M, Thomas A, Wornow M, Birch-Sykes C, Massaroli S, Patel A, Rabideau C, Bengio Y et al. HyenaDNA: Long-Range Genomic Sequence Modeling at Single Nucleotide Resolution. In: *Advances in Neural Information Processing Systems*. 2023.
22. Wang A, Li J, Dong H, Xu B, Yin Q, Xu Y, et al. Omnireg-gpt: a high-efficiency foundation model for comprehensive genomic sequence understanding. Nat Commun. 2025;16(1):10139.
23. Dalla-Torre H, Gonzalez L, Mendoza-Revilla J, Lopez Carranza N, Grzywaczewski AH, Oteri F, et al. Nucleotide Transformer: building and evaluating robust foundation models for human genomics. Nat Methods. 2025;22(2):287–97.
24. Nguyen E, et al. Sequence modeling and design from molecular to genome scale with Evo. Science. 2024;386(6723):eado9336.
25. Schiff Y, Kao CH, Gokaslan A, Dao T, Gu A, Kuleshov V. Caduceus: Bi-Directional Equivariant Long-Range DNA Sequence Modeling. Proc Mach Learn Res. 2024;235:43632–48.
26. Benegas G, Kang J, Einhorn Y, Song YS. A DNA language model based on multispecies alignment predicts the effects of genome-wide variants. Nat Biotechnol. 2025;43:1960–5.
27. Fishman V, Kuratov Y, Shmelev A, Petrov M, Penzar D, Shepelin D, et al. GENA-LM: a family of open-source foundational DNA language models for long sequences. Nucleic Acids Res Volume. January 2025;53(2):gkae1310.
28. Avsec Ž, Agarwal V, Visentin D, Ledsam JR, Grabska-Barwinska A, Taylor KR, et al. Effective gene expression prediction from sequence by integrating long-range interactions. Nat Methods. 2021;18(10):1196–203.

29. Brixi G, Durrant MG, Ku J, et al. Genome modeling and design across all domains of life with Evo 2. Nature. 2026;652(8112):1349–61.
30. Taskiran II, Spanier KI, Dickmänken H, et al. Cell-type-directed design of synthetic enhancers. Nature. 2024;626(7997):212–20.
31. Chen H, Liu S, Padula S, Lesman D, Griswold K, Lin A, et al. Efficient, continuous mutagenesis in human cells using a pseudo-random DNA editor. Nat Biotechnol. 2020;38(2):165–8.
32. Chen XD, Chen Z, Wythes G, Zhang Y, Orr BC, Sun G, et al. Helicase-assisted continuous editing for programmable mutagenesis of endogenous genomes. Science. 2024;386(6718):eadn5876.
33. Hess GT, Frésard L, Han K, Lee CH, Li A, Cimprich KA, et al. Directed evolution using dCas9-targeted somatic hypermutation in mammalian cells. Nat Methods. 2016;13(12):1036–42.
34. DaSilva LF, Senan S, Kribelbauer-Swietek JF, Patel ZM, Louis LK, Reddy AJ, et al. Designing synthetic regulatory elements using the generative AI framework DNA-Diffusion. Nat Genet. 2026;58:180–94.
35. Lal A, Garfield D, Biancalani T, Eraslan G. Designing realistic regulatory DNA with autoregressive language models. Genome Res. 2024;34(9):1411–20.
36. Gosai SJ, Castro RI, Fuentes N, Butts JC, Mouri K, Alasoadura M, et al. Machine-guided design of cell-type-targeting Cis-regulatory elements. Nature. 2024;634:1211–20.
37. Kempynck N, De Winter S, Blaauw CH, Konstantakos V, Ekşi EC, Dieltiens S, et al. CREsted: modeling genomic and synthetic cell-type-specific enhancers across tissues and species. Nat Methods. 2026;23:946–59.
38. Yin C, Castillo-Hair S, Byeon GW, Bromley P, Meuleman W, Seelig G. Iterative deep learning design of human enhancers exploits condensed sequence grammar to achieve cell-type specificity. Cell Syst. 2025;16(7):101302.
39. Ma M, Liu G, Cao C et al. HybriDNA: a hybrid transformer-mamba2 long-range DNA language model. arXiv:2502.10807. Available from: http://dx.doi.org/10.48550/arXiv.2502.10807
40. Nvidia, Blakeman A, Basant A, Khattar A, Renduchintala A, Bercovich A et al. Nemotron-H: A family of accurate and efficient hybrid Mamba-Transformer models. arXiv:2504.03624 [cs.CL]. 2025. Available from: http://arxiv.org/abs/2504.03624
41. Lieber O, Lenz B, Bata H, Cohen G, Osin J, Dalmedigos I et al. Jamba: A Hybrid Transformer-Mamba Language Model. In: The Thirteenth International Conference on Learning Representations. 2025.
42. Springer JM, Kotha S, Fried D, Neubig G, Raghunathan A. Repetition Improves Language Model Embeddings. In: The Thirteenth International Conference on Learning Representations. 2025. Available from: https://doi.org/10.48550/arXiv.2402.15449
43. Marin FI, Teufel F, Horlacher M et al. BEND: benchmarking DNA language models on biologically meaningful tasks. In: The Twelfth International Conference on Learning Representations, 2024.
44. Lee D, Shi M, Moran J, et al. STARRPeaker: uniform processing and accurate identification of STARR-seq active regions. Genome Biol. 2020;21(1):298.

45. Chen AXY, Yap KM, Kim JS, Sek K, Huang YK, Dunbar PA, et al. Rewiring endogenous genes in CAR T cells for tumour-restricted payload delivery. Nature. 2025;644(8075):241–51.
46. Koeber U, Matjusaitis M, Alfazema N, et al. Synthetic super-enhancers enable precision viral immunotherapy. Nature. 2026;653:232–41.
47. Zaret KS. Pioneer transcription factors initiating gene network changes. Annu Rev Genet. 2020;54(1):367–85.
48. Tachmatzidi EC, Galanopoulou O, Talianidis I. Transcription control of liver development. Cells. 2021;10(8):2026.
49. Thirant C, Peltier A, Durand S, Kramdi A, Louis-Brennetot C, Pierre-Eugène C, et al. Reversible transitions between noradrenergic and mesenchymal tumor identities define cell plasticity in neuroblastoma. Nat Commun. 2023;14(1):2575.
50. Xu Y, Lou D, Chen P, Li G, Usoskin D, Pan J, et al. Single-cell MultiOmics and spatial transcriptomics demonstrate neuroblastoma developmental plasticity. Dev Cell. 2025;60(17):2248–e226311.
51. Dao T, Gu A. Transformers are SSMs: Generalized models and efficient algorithms through structured state space duality. In: The Forty-first International Conference on Machine Learning. 2024.
52. Glorioso P, Anthony Q, Tokpanov Y, Whittington J, Pilault J, Ibrahim A et al. Zamba: A Compact 7B SSM Hybrid Model. arXiv:2405.16712 [cs.LG]. 2024. Available from: http://arxiv.org/abs/2405.16712
53. O'Leary NA, Wright MW, Brister JR, Ciufo S, Haddad D, McVeigh R, et al. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic Acids Res. 2016;44(D1):D733–45.
54. Dao T. FlashAttention-2: Faster attention with better parallelism and work partitioning. In: The Twelfth International Conference on Learning Representations. 2024.
55. Schulman J, Wolski F, Dhariwal P, Radford A, Klimov O. Proximal Policy Optimization Algorithms. arXiv:1707.06347 [cs.LG]. 2017. Available from: http://arxiv.org/abs/1707.06347
56. Heinz S, Benner C, Spann N, et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell. 2010;38(4):576–89.
57. McInnes L, et al. UMAP: Uniform Manifold Approximation and Projection. J Open Source Softw. 2018;3(29):861.
58. Hahsler M, Piekenbrock M, Doran D. dbscan: Fast Density-Based Clustering with R. J Stat Soft. 2019;91(1):1–30.
59. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J. 2011;17:10–2.
60. Zhang J, Kobert K, Flouri T, Stamatakis A. PEAR: a fast and accurate Illumina Paired-End reAd mergeR. Bioinformatics. 2014;30(5):614–20.

61. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9(4):357–9.
62. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25(16):2078–9.
63. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics. 2016;32(18):2847–9.
64. Hao Y, Stuart T, Kowalski MH, et al. Dictionary learning for integrative, multimodal and scalable single-cell analysis. Nat Biotechnol. 2024;42(2):293–304.
65. Gordon MG, Inoue F, Martin B, Schubach M, Agarwal V, Whalen S, et al. lentiMPRA and MPRAflow for high-throughput functional characterization of gene regulatory elements. Nat Protoc. 2020;15(8):2387–412.

Figures

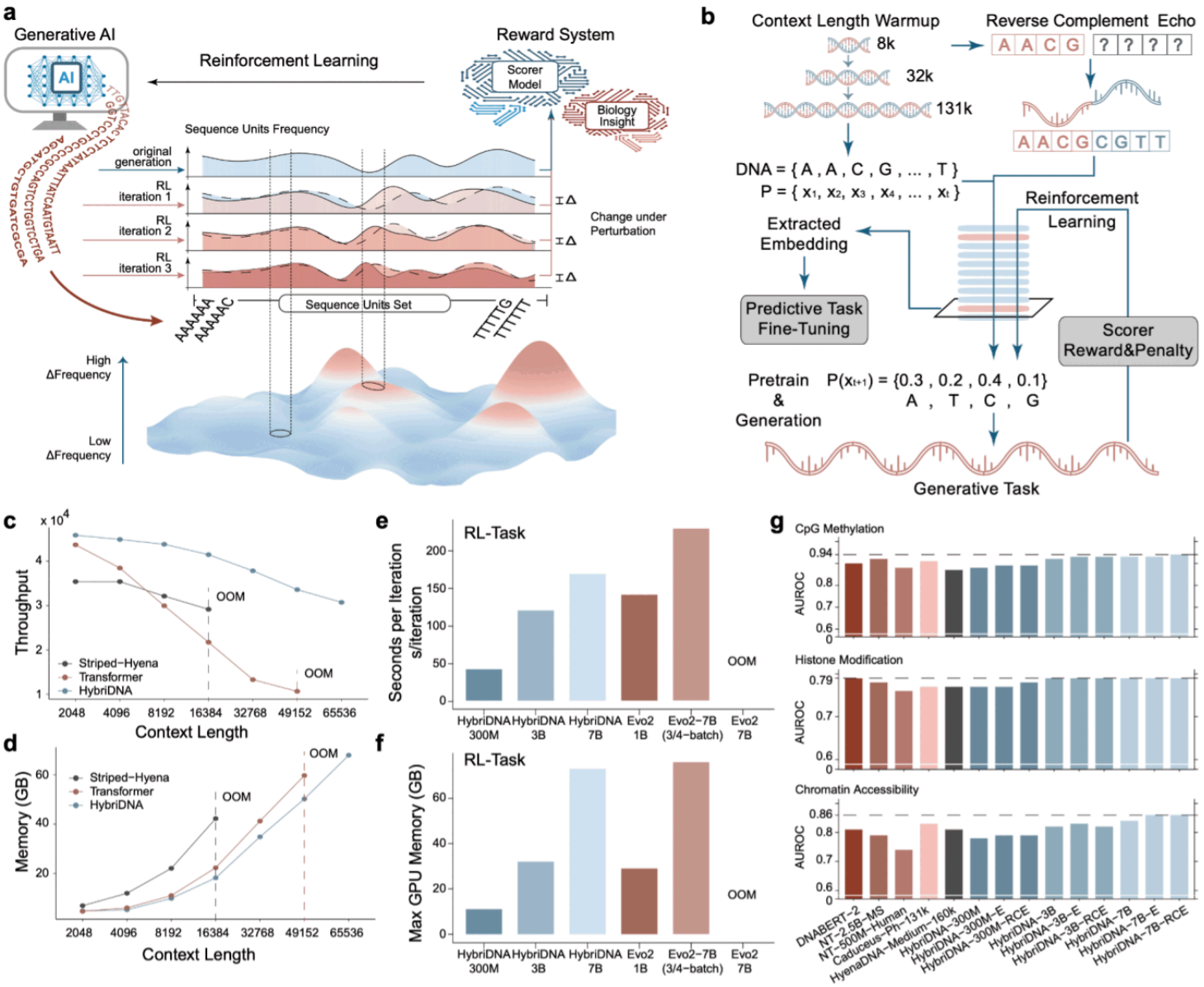
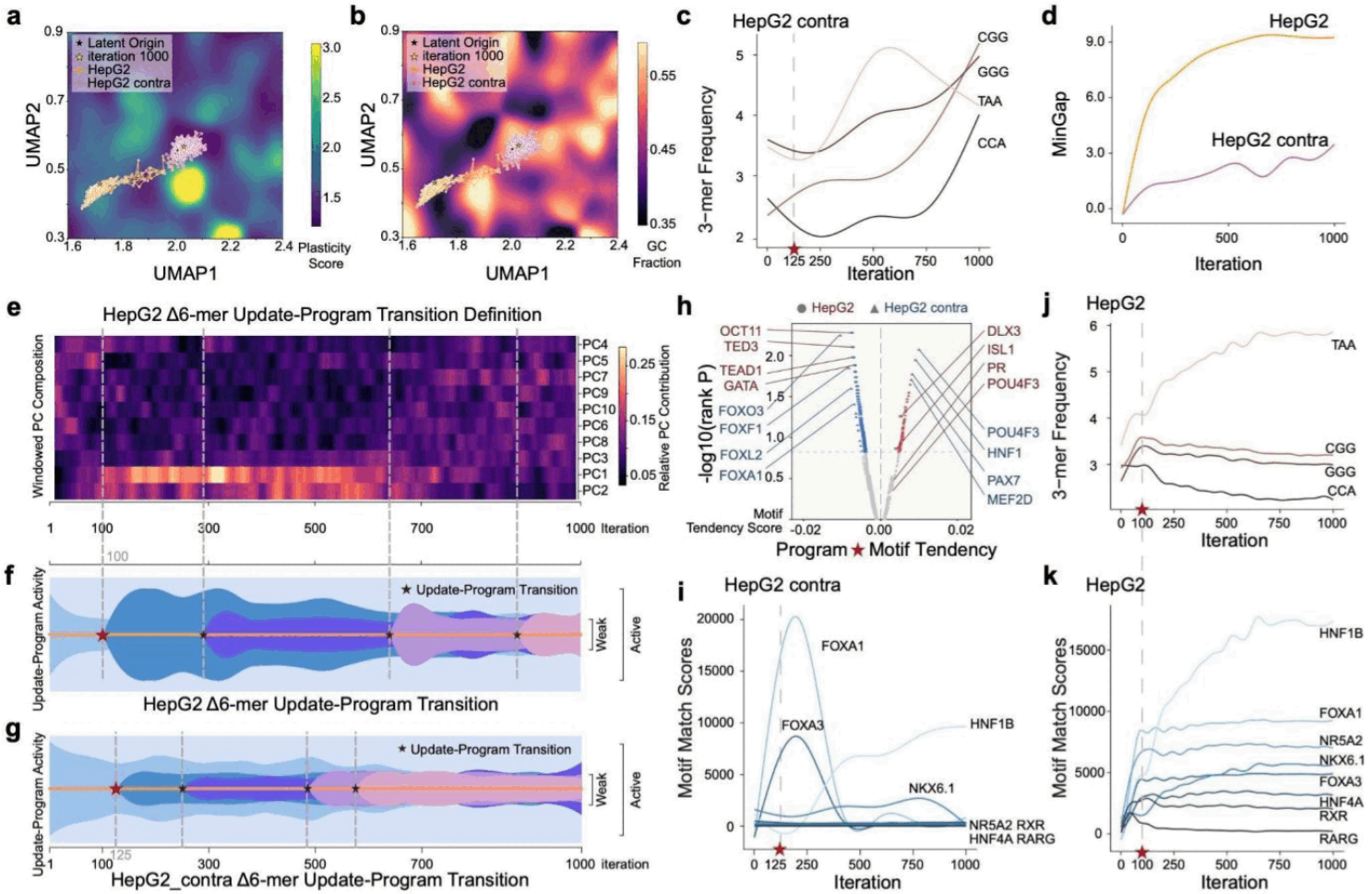


Figure 1



HybriDNA provides an efficient generative backbone for reinforcement-learning-based CRE optimization. **a.** Overview of the GO-CRE framework. A generative DNA language model proposes candidate CREs, a predictive model evaluates their cell-type-specific regulatory activity, and reward feedback combined with biological priors, guides iterative sequence optimization. The trajectory is reconstructed for interpretation. **b.** Overview of the HybriDNA training and optimization workflow. A hybrid Transformer-Mamba2 architecture is pretrained at single nucleotide resolution with DNA sequences of 8 kb, 32 kb, and 131 kb with RCE embedding. For CRE generation, the pretrained model is further optimized through RL. **c, d.** Throughput (tokens per second) (**c**) and GPU memory usage (**d**) for HybriDNA, a pure Transformer, and StripedHyena across input context lengths. All models were matched at approximately 300 million parameters. OOM, out of memory. Dashed line indicates the maximum context length reached before OOM. **e, f.** Iteration time (**e**) and maximum GPU memory usage (**f**) during RL optimization of HybriDNA and Evo 2 on a single NVIDIA A100 80-GB GPU. HybriDNA was tested at 300M, 3B and 7B parameters, and Evo 2 tested at 1B and 7B parameters at 3/4 of the full batch. Full-batch Evo 2-7B exceeded the 80-GB memory limit and is shown as OOM. **g.** Performance of gLMs on the BEND benchmark, measured by area under the receiver operating characteristic curve (AUROC) for chromatin-accessibility, histone-modification, and CpG-methylation prediction tasks. The vertical black dashed lines indicate the best performance for each task.

Figure 2

GO-CRE reconstructs interpretable and tunable sequence-grammar trajectories. **a, b.** Optimization trajectories of polyG-penalized HepG2 sequences (HepG2) and unpenalized controls (HepG2 contra) projected onto the reconstructed sequence-grammar space, colored by plasticity score (**a**) or GC fraction (**b**). Stars mark final iteration 1,000. **c.** Representative 3-mer frequency trajectories during unpenalized HepG2 optimization. The dashed line marks iteration 125. **d.** MinGap reward trajectories across RL iterations for HepG2 and HepG2 contra. **e.** Windowed principal component (PC) composition used to define Δ6-mer update programs across the HepG2 trajectory; relative contributions of the retained PCs are shown at right. **f, g.** Activities of inferred sequence-update programs across RL iterations for HepG2 (**f**) and the unpenalized HepG2 control(**g**). Each colored ribbon represents one update program, and its width indicates the program's relative activity, defined by the alignment between the current Δ6-mer update composition and the program-specific weighted PC signature. Labels from "Weak" to "Active" indicate increasing representation of this signature in the current sequence-update dynamics. Stars and dashed lines mark the emergence of newly inferred update programs. Red stars denote the first major program transition, at iteration 100 for HepG2 (**f**) and iteration 125 for unpenalized HepG2 control (**g**). **h.** Corrected motif tendency scores associated with the first transition-associated program in HepG2 and HepG2 contra. The y-axis denotes -log10(rank P), where rank P is derived from the ranking of absolute tendency scores; a value of 1 approximately denotes the top 10% of motifs. **i.**Representative motif match score trajectories during unpenalized HepG2 optimization. The red star and dashed line mark iteration 125. **j, k.**Representative 3-mer frequency (**j**) and motif match score (**k**) trajectories during HepG2 optimization. Dashed lines mark iteration 100, the first major program transition. For (**c-g**) and (**i-k**), the x-axis shows RL iteration 1 - 1,000 from left to right, and the y-axis shows the corresponding feature.

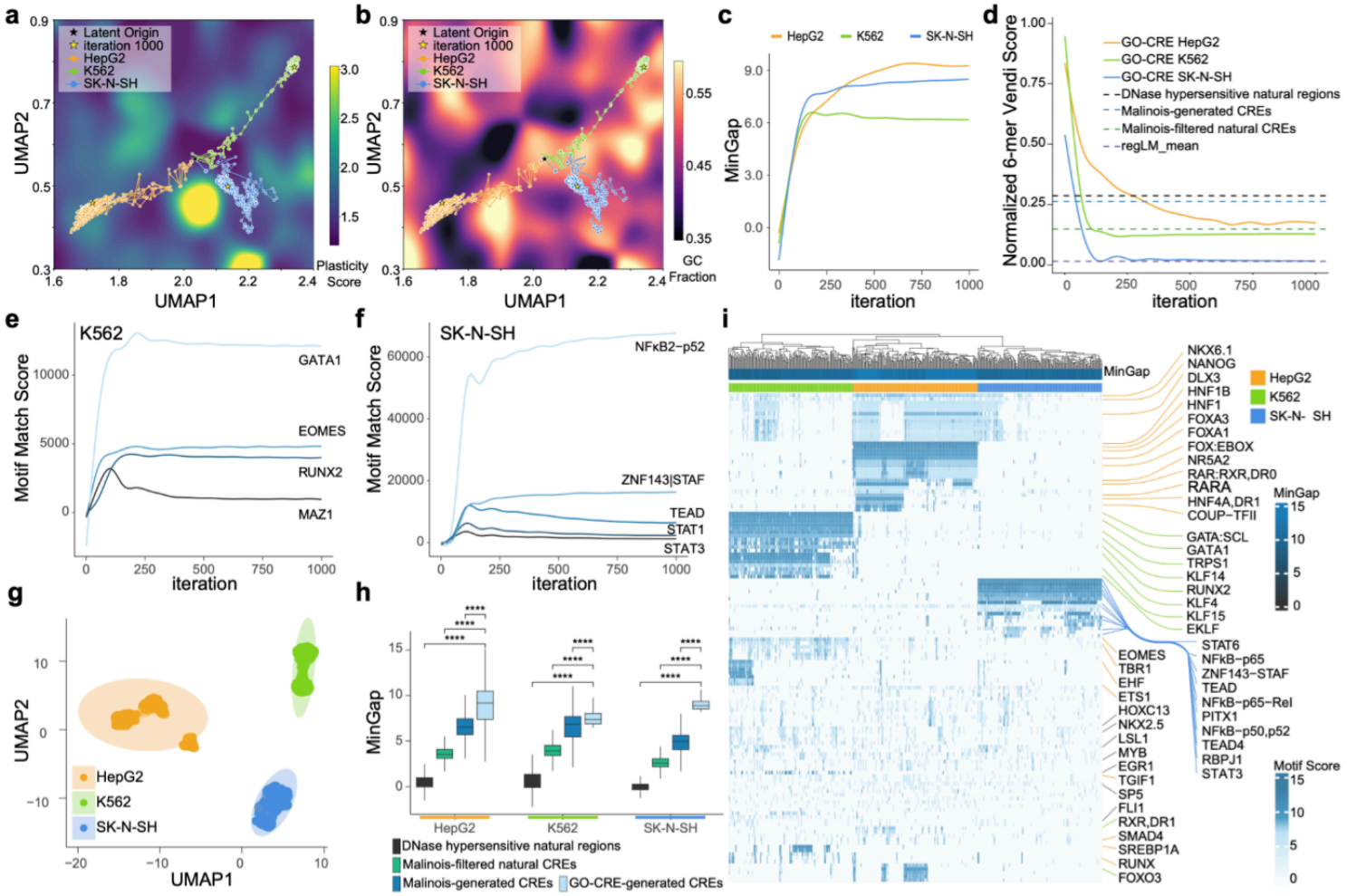


Figure 3

Cell-type-specific optimization converges on distinct regulatory programs. **a, b.** RL trajectories projected onto the reconstructed sequence-grammar space, colored by plasticity score (**a**) or GC fraction (**b**). The origin represents the centroid of iteration 1 across tasks. Trajectories over 1,000 RL iterations are shown for HepG2, K562 and SK-N-SH. Stars mark the final iteration 1,000. **c**. MinGap trajectories for HepG2, K562 and SK-N-SH. **d**. Normalized 6-mer Vendi scores during GO-CRE optimization of the three cell types. Dashed lines show mean normalized scores across the three cell types for DNase hypersensitive natural regions (black), Malinois-generated CREs (blue), Malinois-filtered natural CREs (green), and regLM-generated CREs (purple). **e, f**. Representative motif match score trajectories during K562 (**e**) and SK-N-SH (**f**) optimization. **g**. UMAP visualization of final CREs generated for the three cell types (9,000 sequences per cell type). **h**. MinGap scores of GO-CRE-generated CREs, DNase hypersensitive natural regions, Malinois-filtered natural CREs, and Malinois-generated CREs across the three cell types. GO-CRE was compared with each control within each cell type using two-sided Wilcoxon rank-sum tests. **** $p < 0.0001$. **i**. Heatmap of representative TF motif match scores in final CREs generated for the three cell types. Columns represent generated CREs ordered by hierarchical clustering (n=150 per cell type), and rows represent motifs (n=132). For each cell type, the top 30 sequences were selected from each of the five highest-ranked MinGap clusters. Top annotation bars indicate each sequence's MinGap score and

target cell type. For (**c-f**), the x-axis denotes RL iterations 1 - 1,000, and the y-axis denotes the corresponding feature.

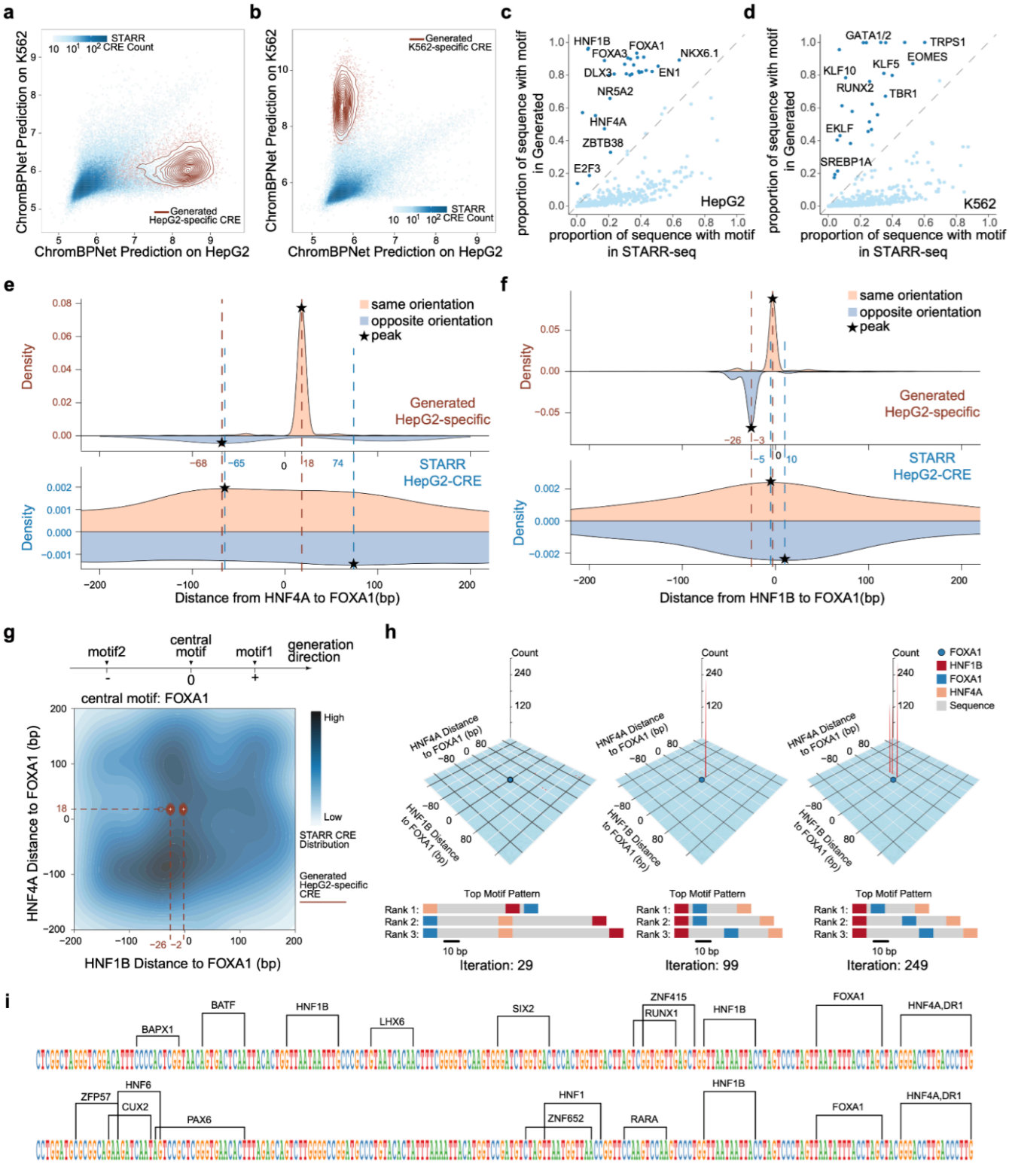


Figure 4

Generated and endogenous CREs encode distinct motif organizations. **a, b**. Density projections of generated HepG2-specific (**a**) and K562-specific (**b**) CREs in ChromBPNet prediction space, defined by

predicted accessibility in HepG2 (x-axis) and K562 (y-axis). Endogenous STARR-seq CREs are shown as background density. **c, d**. Motif prevalence in generated versus endogenous STARR-seq CREs from HepG2 (**c**) and K562 (**d**). Each point represents a motif, with the proportions of generated and endogenous CREs containing it on the two axes. The dashed line denotes y=x. **e, f**. Density plots of motif spacing centered on FOXA1 in generated HepG2-specific and endogenous HepG2 STARR-seq CREs, showing distances from HNF4A to FOXA1 (**e**) and from HNF1B to FOXA1 (**f**). Negative and positive x-axis values indicate that the HNF motif is located 5' and 3' of FOXA1. Densities above and below zero indicate same and opposite motif orientations. Asterisks mark peaks; red and blue dashed lines indicate same- and opposite- oriented peaks. **g**. Two-dimensional density of HNF1B-to-FOXA1 and HNF4A-to-FOXA1 distances in endogenous HepG2 STARR-seq CREs and generated HepG2-specific CREs. Red dashed lines project the density peak onto both axes. **h.** Emergence of HNF1B-FOXA1-HNF4A organization during HepG2 optimization. Distance distribution and top-ranked motif arrangements are shown at representative iterations 29, 99, and 249. The horizontal axes show HNF1B-to-FOXA1 and HNF4A-to-FOXA1 distances, with FOXA1 at the origin. The vertical axis shows the number of three-motif configurations at each distance combination. Schematics show the highest-ranked arrangements at each iteration. The full trajectory is shown in Additional file 2: Video S1. **i**. Motif organization of two representative generated HepG2-specific CREs.

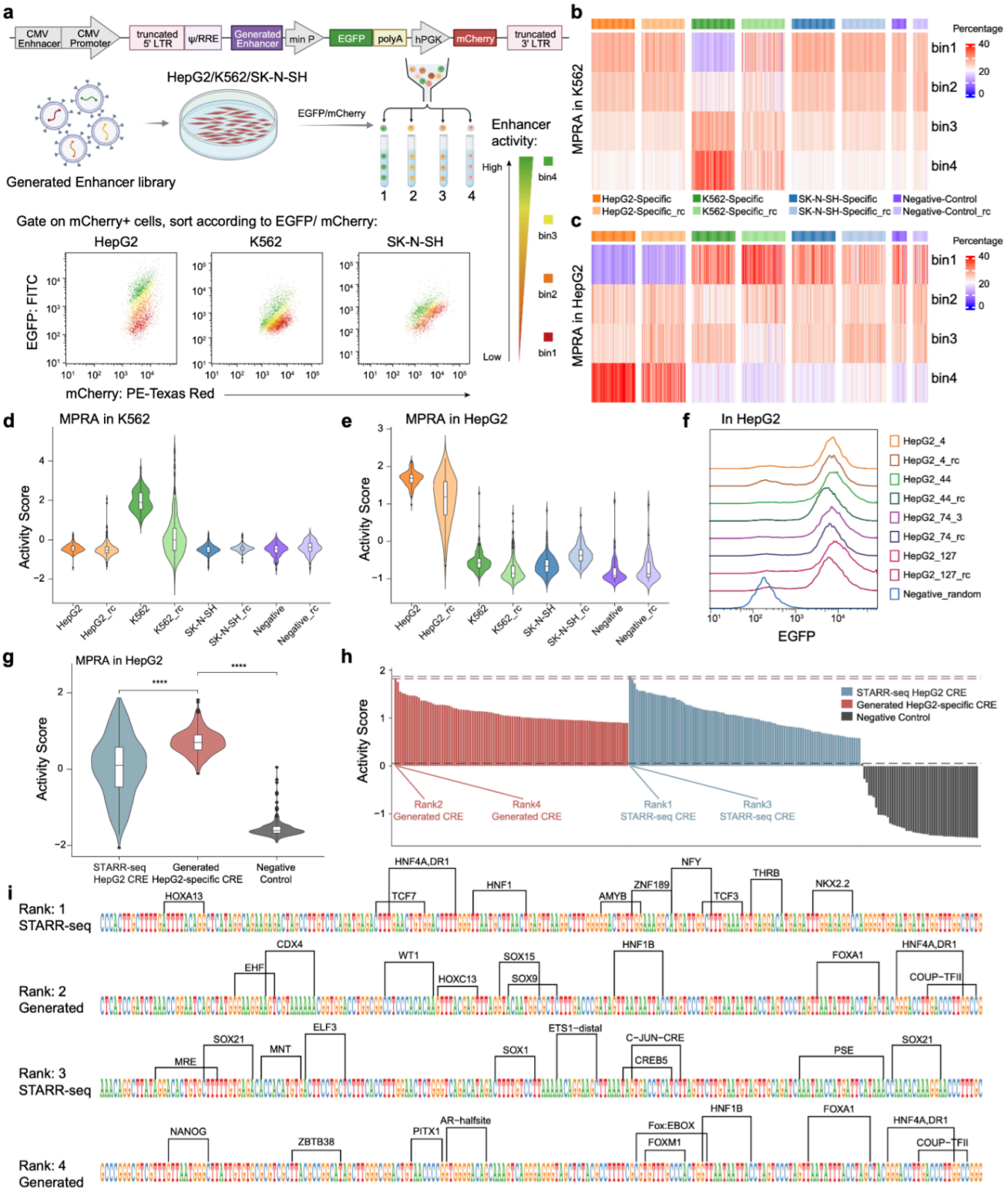


Figure 5

LentiMPRA validates the cell-type-specific activity of generated CREs. **a**. Schematic of the fluorescence-based LentiMPRA workflow. Candidate CRE libraries were cloned upstream of a minimal promoter-EGFP reporter in a dual-color lentiviral vector containing hPGK-mCherry as an internal normalization control. Libraries were transduced into HepG2, K562, or SK-N-SH cells. Four days after transduction, mCherry-positive cells were sorted into four bins by EGFP/mCherry ratio (see Additional file 1: Figure S6), and CRE

abundance in each bin was quantified by amplicon sequencing. **b, c**. Heatmaps of CRE enrichment across bins 1-4 in K562 (**b**) and HepG2 (**c**) cells. Columns represent CREs, rows represent sorted bins of increasing EGFP/mCherry ratio, and top annotation bars indicate CRE class. **d, e**. LentiMPRA activity scores for CREs tested in K562 (**d**) and HepG2 (**e**) cells. **f**. Representative EGFP fluorescence histograms from individual validation of HepG2-generated CREs. **g**. LentiMPRA activity scores in HepG2 for endogenous STARR-seq CREs and generated HepG2-specific CREs selected by ChromBPNet. **** $P < 0.0001$. **h**. Ranked LentiMPRA activity scores for the top 25% endogenous and generated HepG2 CREs. Red, blue, and gray indicate generated CREs, endogenous CREs, and negative controls. Dashed lines indicate the top-ranked generated and endogenous CREs. Sequences shown in (**i**) are marked. **i**. TF motif organization of the four highest-ranking endogenous and generated HepG2 CREs in (**h**). In (**b-h**), the suffix “rc” denotes a reverse-complemented sequence. Negative controls were derived from GM12878-accessible regions lacking accessibility in the target cell type. Data represent 2, 3, 2, and 3 independent experiments for HepG2, K562, SK-N-SH and the generated-versus-endogenous HepG2 comparison.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile1supplementarymethodsfiguresandtables.pdf