Fig. S1 Characteristics of chromatin loop anchors across diverse species.Box plots showing the distribution of **a** chromatin loop anchor lengths (log10 scale) and **b** inter-anchor distances (log10 scale) for representative species across different eukaryotic groups.

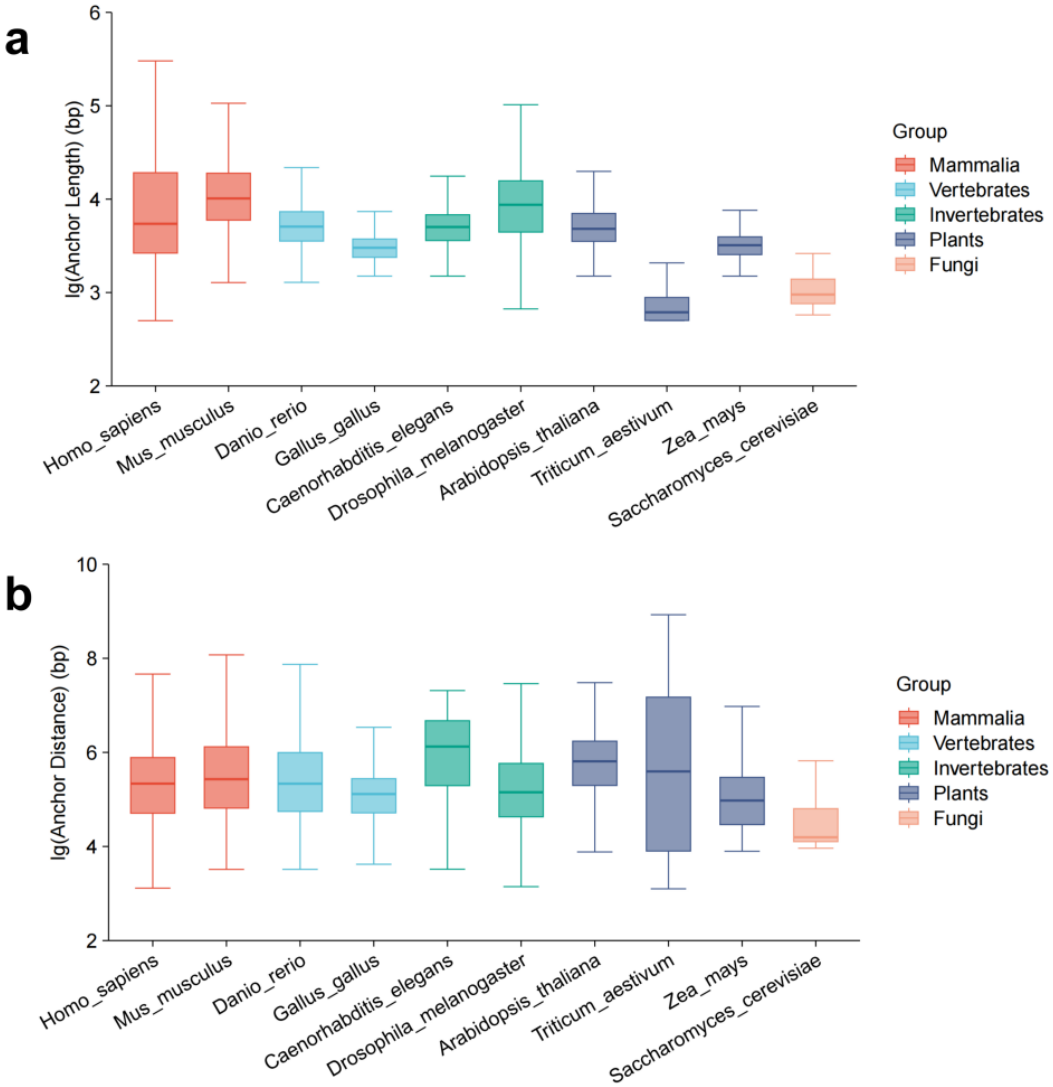

Fig. S2 Generalization capabilities of CLAMP across proteins, cell types, and species. Heatmaps showing pairwise prediction performance (MCC) for generalization experiments. **a** Cross-protein generalization: Models trained on loops mediated by one protein (e.g., CTCF in GM12878) tested on loops mediated by another protein (e.g., RAD21 in GM12878) within the same cell type context. **b** Cross-cell type generalization: Models trained in one cell type (e.g., CTCF in GM12878) tested in another cell type (e.g., CTCF in Hela-S3) for the same protein and species. **c** Cross-species generalization: Models trained in one species (e.g., CTCF in Homo sapiens) tested in another species (e.g., CTCF in Mus musculus) for the same protein.

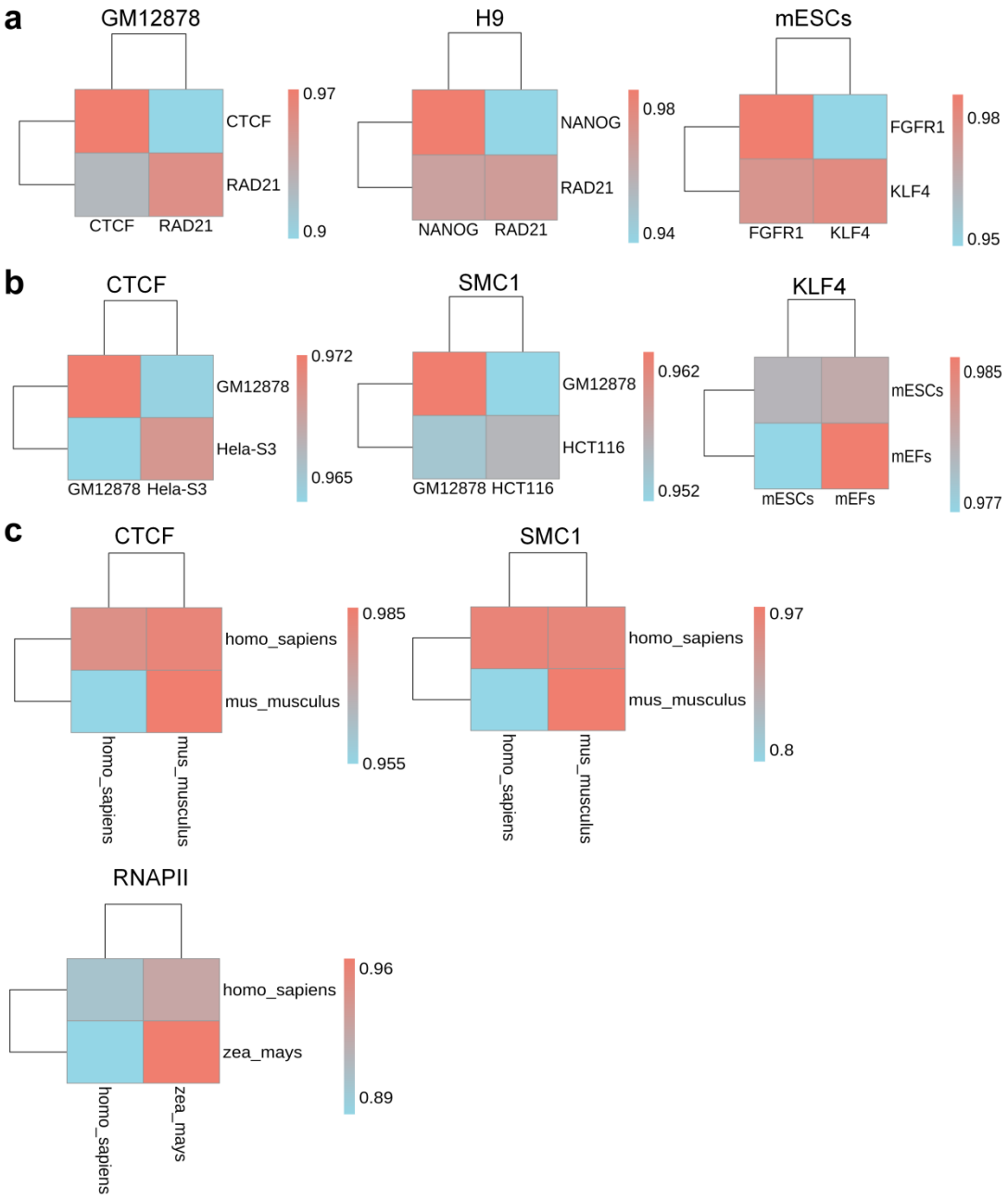

Fig. S3 Intersection analysis of variable epigenetic signal experiment. This analysis revealed minimal overlap between chromatin loops for prediction from different cell types, confirming the cell-type-specificity predictions.

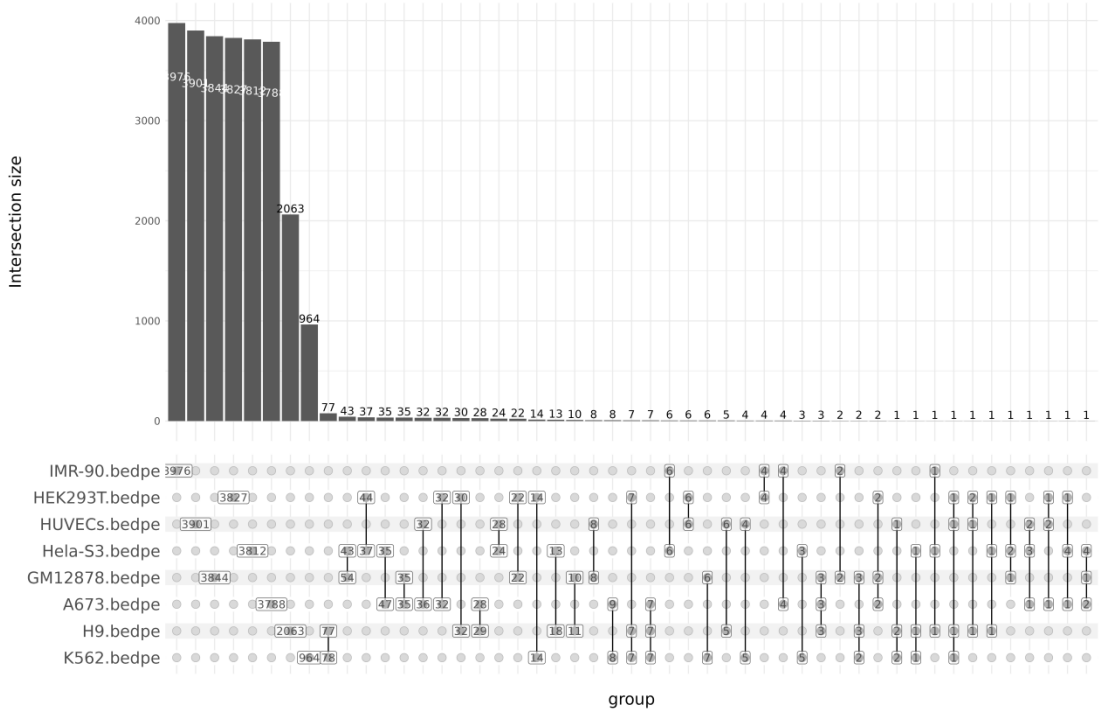

Fig. S4 Cross-species generalization using variable epigenetic signals. Bar chart showing the AUROC for CLAMP's performance when predicting chromatin loops in different target species.

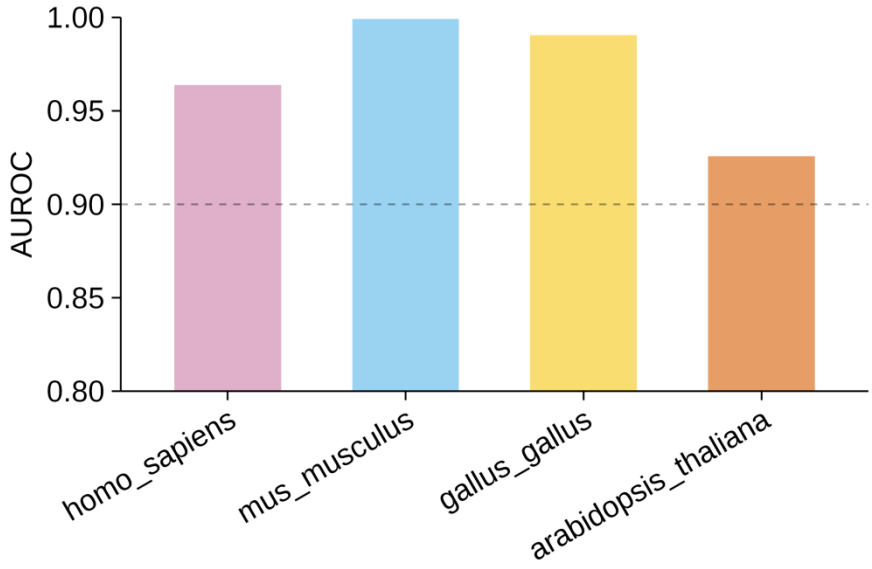

Fig. S5 Performance comparison of CLAMP and other methods on Bacon dataset.

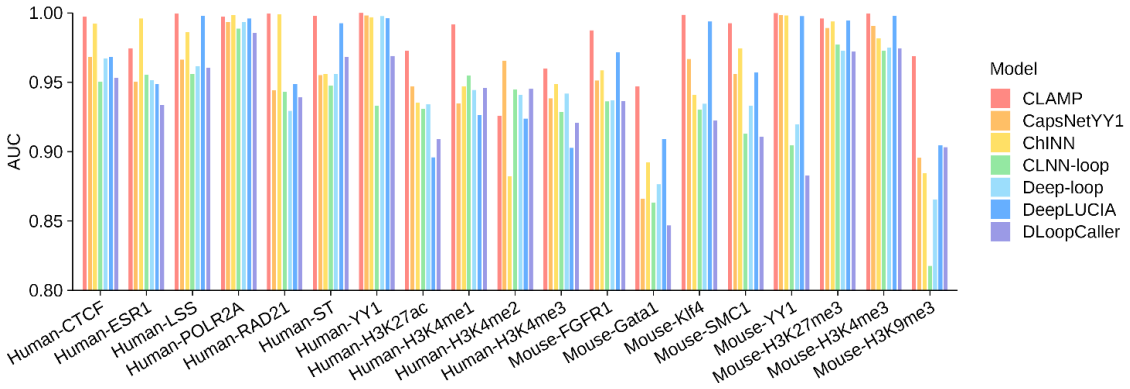

Fig. S6 Performance comparison of CLAMP and other methods on the CLAMP dataset. Performance (AUC) below 0.5 is not displayed.

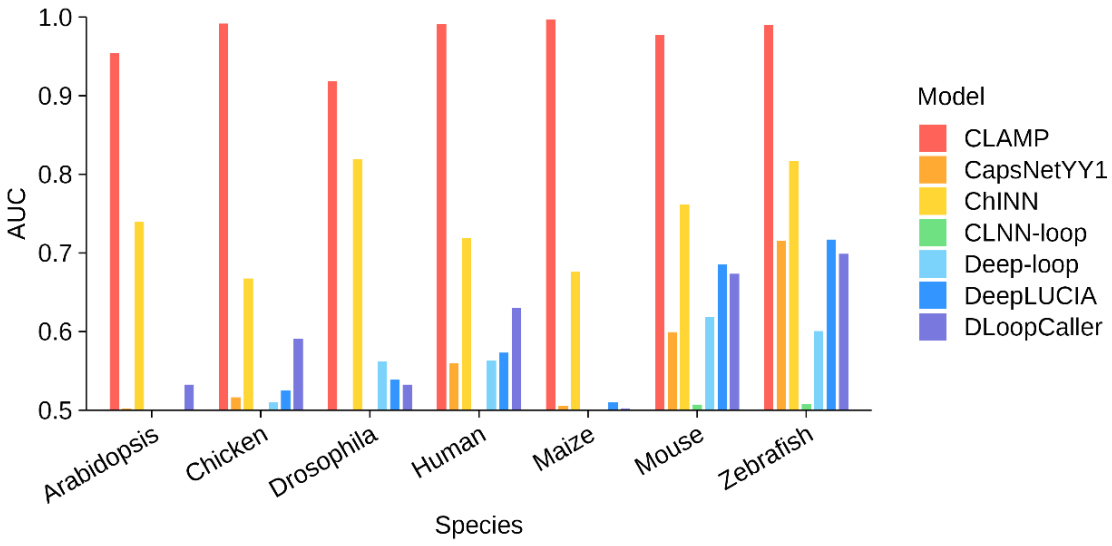

Fig. S7 Comparative performance for CLAMP's pre-trained representations on diverse genomic tasks. Bar plots comparing the (Top-1) accuracy of CLAMP against other DNA foundation models (DNABERT, DNABERT-2, GENA-LM, NT-v2, HyenaDNA, Caduceus) and a baseline CNN on seven distinct genomic tasks.

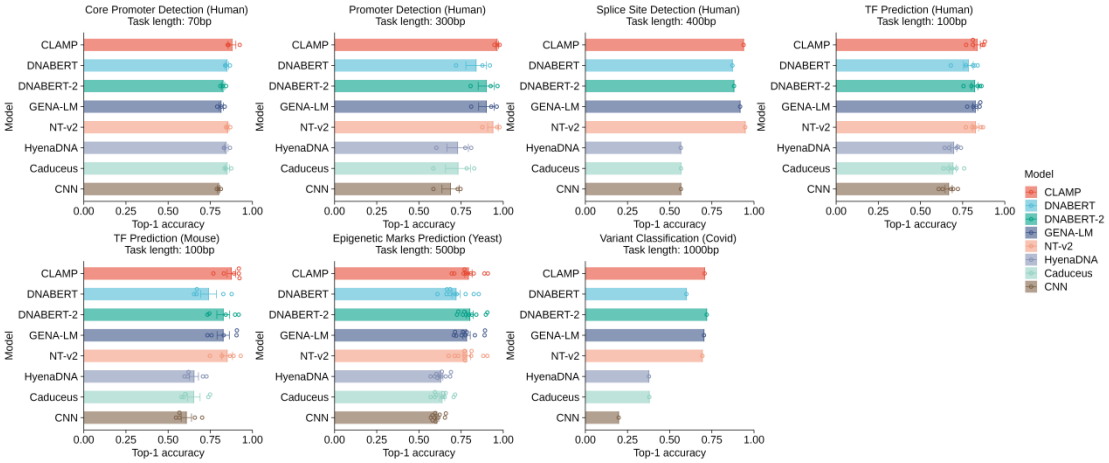

Fig. S8 CLAMP attention scores highlight biologically relevant genomic regions. Visualization of attention scores from the final layer of the pre-trained CLAMP model. **a** Zebrafish H3K4me3 peak region (chr23: 43157750-43158250). **b** Chicken RNAPII peak region (chr33: 2890948-2891148).

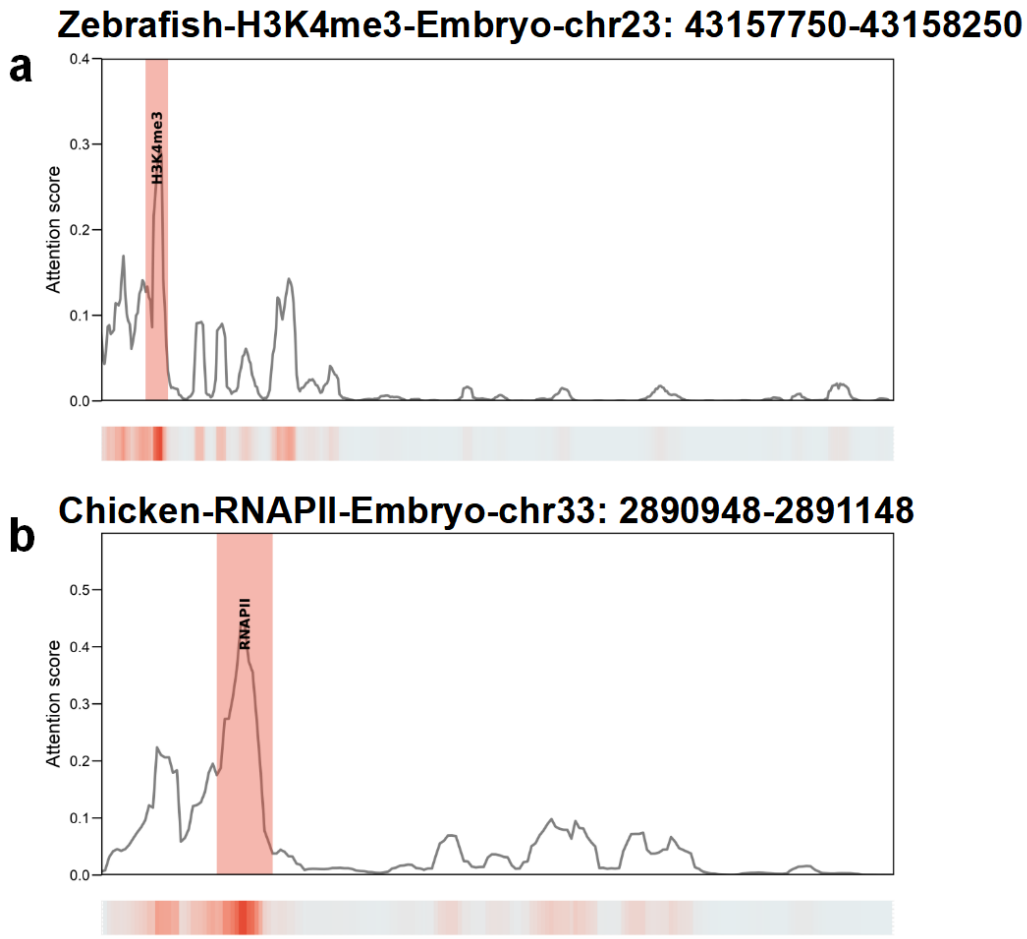

Fig. S9 CoVE interpretability analysis on YY1-mediated loops in mESC dataset. **a** Correlation heatmap using CoVE method. **b** Feature stability test on top 10 features of signal feature. **c** Feature stability test on top 10 features of distance feature.

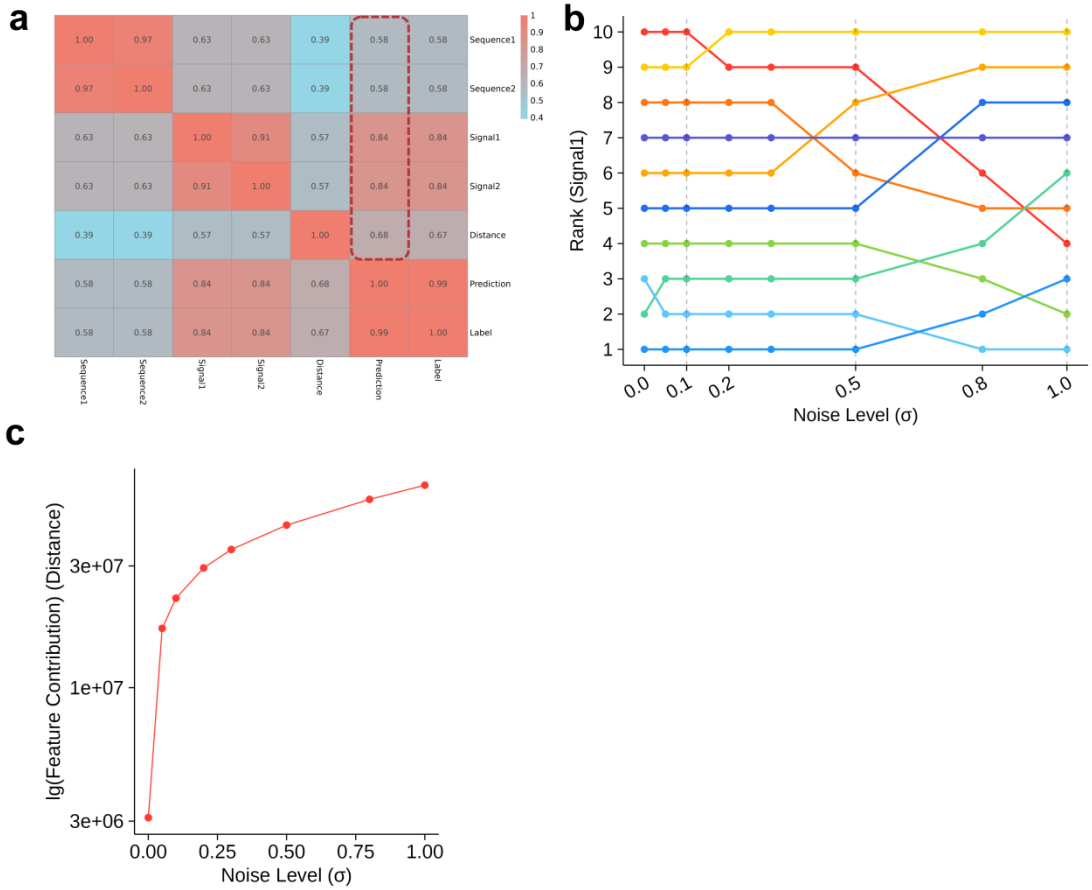

Fig. S10 Relationship between CLAMP prediction scores and chromatin loop interaction intensity. **a** Bar chart showing the CLAMP prediction score for chromatin loops binned by their interaction intensity, measured as PET count from ChIA-PET experiments. **b** Relationship between CLAMP prediction scores and their quantile. The dashed line highlights the score (0.99) corresponding to the 0.4 quantile (top 60% of predictions).

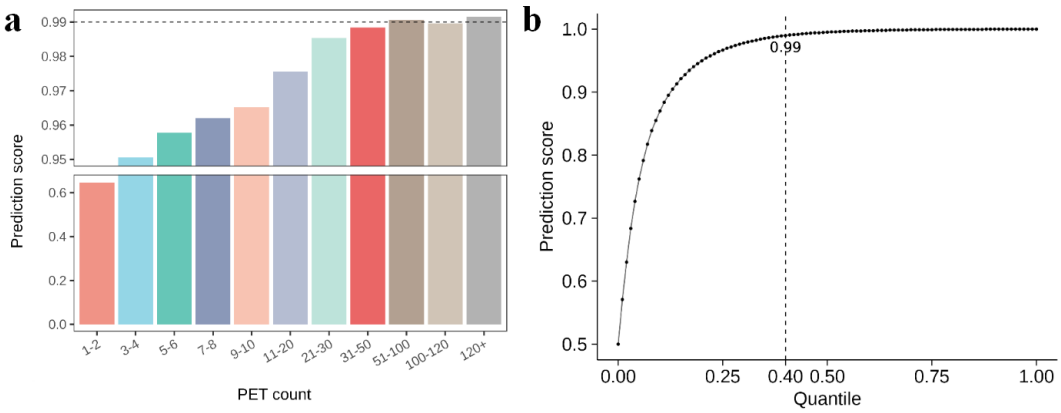

Fig. S11 Functional enrichment analysis of genes associated with high-confidence CLAMP-predicted loops. Pathway clustering networks illustrating enriched KEGG pathways for genes located at the anchors of high-confidence CLAMP-predicted loops. **a** Analysis for human triple-negative breast cancer (TNBC) cells. **b** Analysis for 20-HE treated Drosophila melanogaster embryonic cells. Nodes represent significantly enriched pathways, with node size potentially related to gene count and colour indicating functional clusters or enrichment scores. Edges represent gene overlap between pathways. Feature labels indicate key genes driving enrichment in specific clusters.

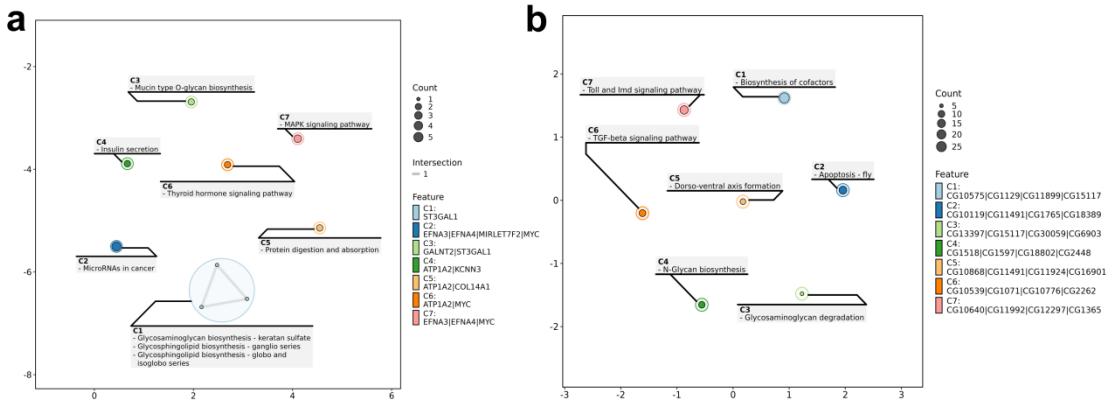

Fig. S12 Characteristics of the cross-species chromatin accessibility dataset used for CLAMP pre-training. **a** Bar chart showing the number of ATAC-seq peaks (log10 scale) included for each of the seven species (bovine, chicken, human, medaka, mouse, zebrafish) in the pre-training dataset. **b** Box plots comparing the distribution of ATAC-seq peak region lengths (bp) across the seven species. **c** Stacked bar chart showing the proportion of peaks annotated to different genomic features for each species.

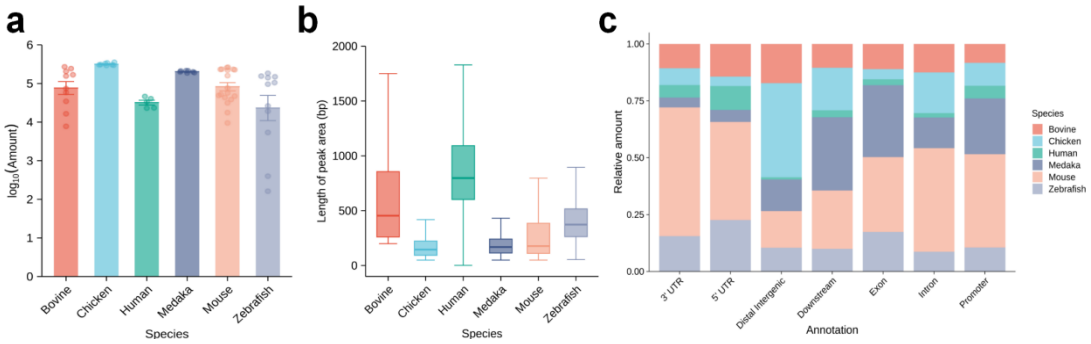

Fig. S13 Impact of input sequence length and tokenizer strategy on CLAMP performance. **a** Violin plots showing MCC for CLAMP performance on downstream genomic tasks using different input sequence lengths extracted around peak regions (200bp, 500bp, 1000bp, 1500bp). **b**, Violin plots showing MCC for CLAMP performance using different tokenizer strategies (Byte Pair Encoding (BPE), 3-mer, 4-mer, 5-mer, 6-mer).

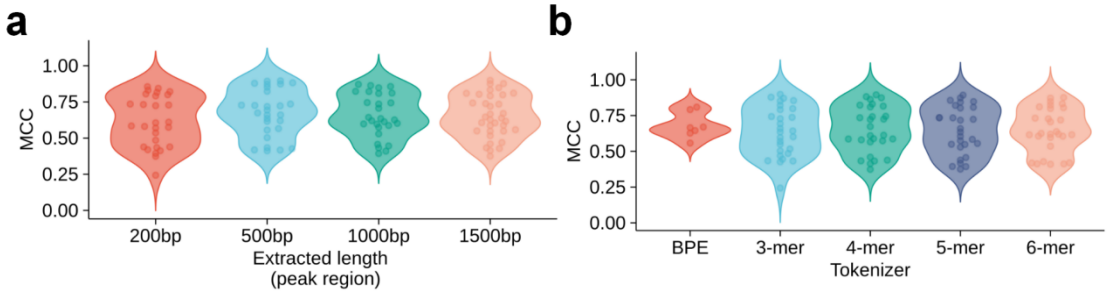

Fig. S14 CLAMP performance across diverse genomic tasks using varied input lengths and tokenizers.

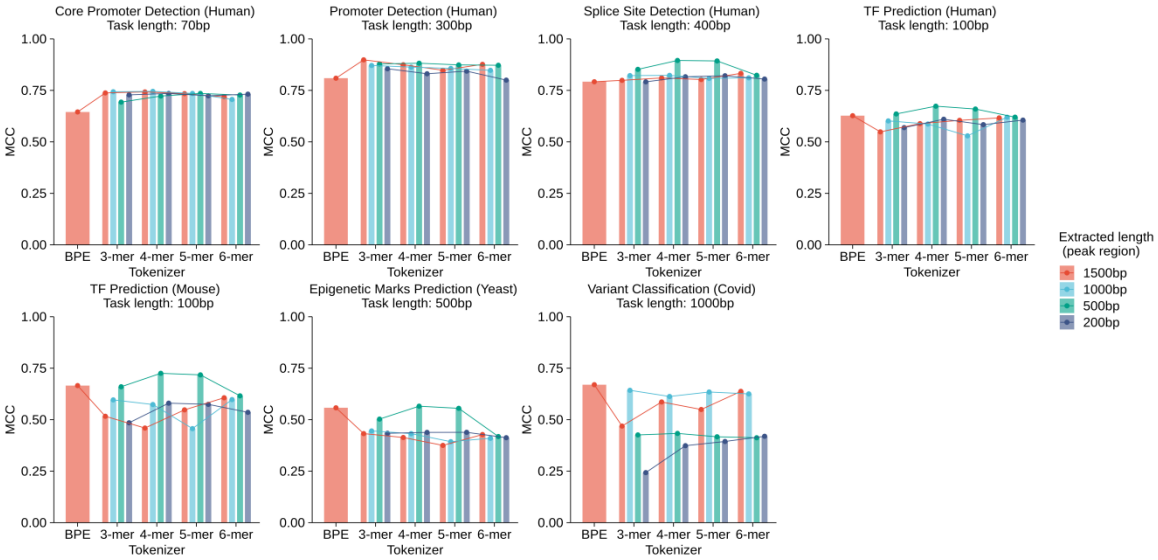

Fig. S15 Ablation study of the pre-training stage on evolutionarily distant species. Bar chart comparing the performance (AUC) of the standard pre-trained CLAMP model (Pretrained) against a version with a randomly initialized sequence encoder (Non-pretrained, Xavier initialization). Both models were fine-tuned and evaluated on species not included in the pre-training dataset, grouped into three major lineages: Invertebrates, Plants, and Fungi.

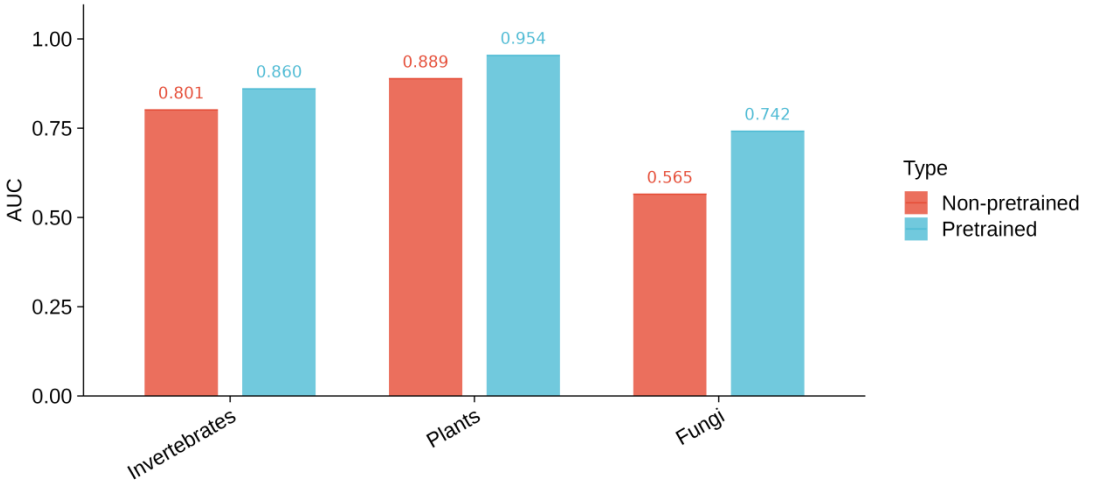

Fig. S16 CLAMP performance stratified by chromatin loop distance across different species groups. Bar chart showing the Matthews correlation coefficient (MCC) for CLAMP's predictions, with chromatin loops categorized by interaction distance: short-range (<50 kb), medium-range (50 kb–10 Mb), and long-range (>10 Mb). The results indicate that CLAMP performs well on short-range and medium-range interactions across all lineages but shows a marked decrease in performance for predicting long-range interactions.

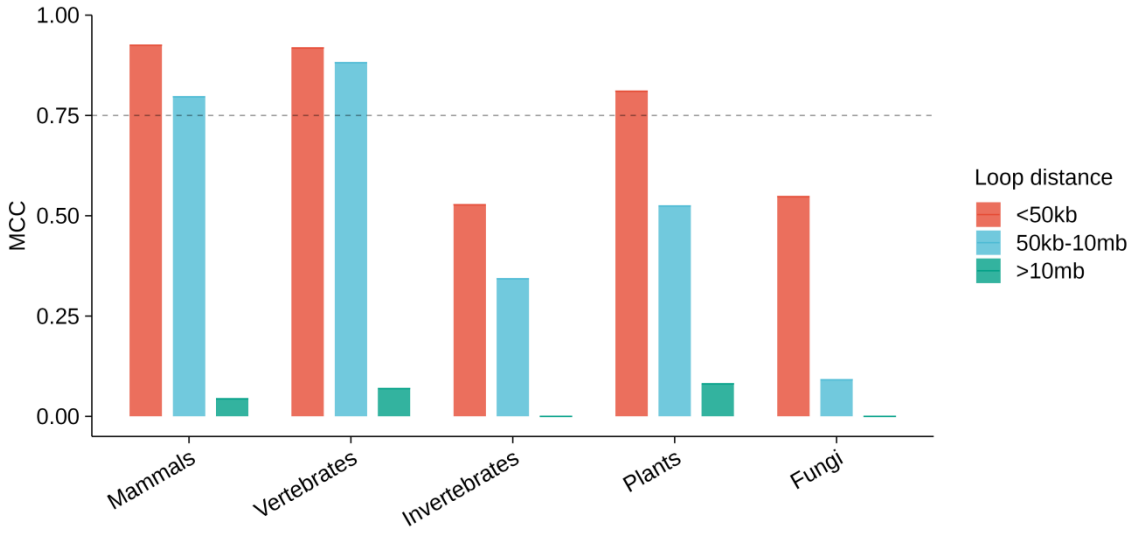

Fig. S17 Impact of different negative sampling strategies on CLAMP's generalization performance. Bar chart comparing the Matthews correlation coefficient (MCC) for models trained on the human CTCF-A673 dataset using three different negative sampling strategies: Adversarialness, Random, and Distal (>10 Mb). Generalization performance was evaluated on the test set across five different human cell lines.