F-MoDA: Efficient Fourier-based motif discovery in attribution maps

Ofir Yaish[1] **and Yaron Orenstein**[2,3]

[1]School of Electrical and Computer Engineering, Ben-Gurion University of the Negev, Beer Sheva, Israel
[2]Department of Computer Science, Bar-Ilan University, Ramat Gan, Israel
[3]The Mina and Everard Goodman Faculty of Life Sciences, Bar-Ilan University, Ramat Gan, Israel

Deep neural networks have been transforming the field of bioinformatics and computational biology in recent years, especially in genomic-related tasks. Trained neural networks have enabled unprecedented capabilities in predicting molecular and genomic phenotypes. A fundamental step following deep-neural-network training and performance evaluation is the interpretation of the trained neural networks to learn new biology and validate the trained models. Techniques like integrated gradients have been highly successful in local interpretability, i.e., attributing importance to a specific residue in a given DNA, RNA, or amino acid sequence. But, there still remains the challenge of finding the *global* patterns that are shared among many sequences to understand the biological mechanism. Currently, TF-MoDISco is the only available method for this task. However, TF-MoDISco takes hours to run on standard datasets, and it reports many redundant and false motifs. Here, we present F-MoDA (Fourier-based Motif Discovery in Attribution maps), a novel computational method for efficiently and accurately discovering shared sequence motifs in residue-level attribution maps. F-MoDA leverages signal processing techniques and a hierarchical clustering approach to identify recurring regulatory patterns. We evaluated F-MoDA against TF-MoDISco over an established motif-finding benchmark and found that F-MoDA reports motifs that are more similar to the ground truth, in addition to reporting fewer redundant motifs and fewer false motifs. Moreover, F-MoDA runs much faster and uses less memory. We expect F-MoDA to be utilized in many studies applying deep neural networks to genomics data. F-MoDA is publicly available at `https://github.com/OrensteinLab/F-MoDA`.

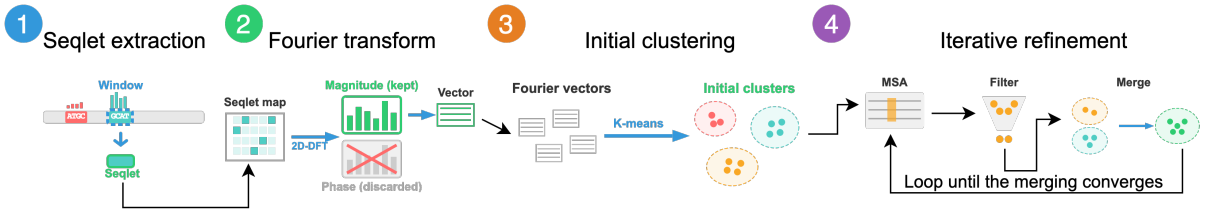

Fig. 1. Overview of F-MoDA workflow for motif discovery in attribution maps. High-attribution subsequences (seqlets) are extracted from attribution maps and transformed into translation-invariant representations using 2D-discrete Fourier transform (DFT). Initial coarse clustering is performed on Fourier magnitude vectors with K-means, followed by iterative refinement with a multiple sequence alignment (MSA)-based hierarchical clustering algorithm.

Introduction

Deep neural networks have revolutionized our ability to understand biological mechanisms, including decoding regulatory DNA sequences and accurately predicting transcription factor (TF) binding, chromatin accessibility, and other epigenomic features (1, 2). Deep neural networks, also termed deep-learning models, have achieved unprecedented prediction performance, but are often criticized for their 'black-box' nature. Interpreting what deep-learning models have learned is critical for discovering new biology and for validating that the models have not learned experimental artifacts or biases (3).

Attribution methods, such as Integrated Gradients (4), DeepLIFT (5), and SHAP (6), achieved great success in assigning importance scores to each residue of a given sequence, producing base-resolution attribution maps that can reveal underlying regulatory and functional features in a given sequence. While there are limitations in gradient-based attribution methods (7), they have become the most established local-interpretability technique in the field of genomics (8–10). Concurrently, efforts have been made to improve the quality and stability of attribution maps during model training, for instance, by employing

exponential activations to learn more robust motif representations (11) or by using Fourier-transform-based attribution priors (12). Yet, interpreting the *global* features associated with many sequences remains a challenge (13).

The most common global-interpretability approach is to extract recurring patterns, termed *motifs*, that contribute significantly to the model's predictions. Traditional motif-discovery methods, such as MEME (14) and STREME (15), are designed to find enriched motifs in raw sequences. As such, these methods are not applicable to finding motifs in attribution maps of deep-learning models. To the best of our knowledge, the only method for the task, TF-MoDISco (16), identifies high-scoring subsequences, termed seqlets, and clusters them into motif representations by comparing both their sequence content and importance scores. However, TF-MoDISco reports many redundant and false motifs, and takes hours to run on standard genomic datasets.

Here, we introduce **F-MoDA** (Fourier-based Motif Discovery in Attribution maps), a fast and accurate method for motif discovery in attribution maps (Figure 1). F-MoDA leverages the translation-invariant properties of the Fourier transform to represent seqlets. F-MoDA efficiently identifies enriched motifs by combining this Fourier-based representation with coarse clustering and a novel iterative procedure for consolidating similar patterns. We demonstrate the effectiveness of F-MoDA on synthetic data and experimental datasets. F-MoDA outperforms TF-MoDISco in finding motifs more similar to the ground truth, reporting fewer redundant motifs, fewer false motifs, and running much faster and using less memory.

Methods

A. F-MoDA workflow. F-MoDA consolidates high-attribution sequence motifs by combining frequency-based representations with iterative refinement, enabling efficient and orientation-consistent motif discovery from model attributions.

A.1. Seqlet extraction and Fourier-based representation. F-MoDA starts with identifying **high-attribution subsequences**, termed seqlets, in the input attribution maps obtained using methods such as Integrated Gradients (4). These seqlets represent regions considered important by the underlying deep-learning model.

Let S be a DNA/RNA sequence of length L, and $A(S) \in \mathbb{R}^{L\times 4}$ be its corresponding attribution map, where each entry $A_{i,j}$ denotes the importance score of nucleotide j (from {A, C, G, T/U}) at position i. F-MoDA extracts seqlets, denoted as s_k, based on the approach used in TF-MoDISco (16). Each seqlet s_k has an associated attribution map $a_k \in \mathbb{R}^{W\times 4}$, where W is the seqlet window length.

To achieve a **translation-invariant representation** of the importance signal within each seqlet, F-MoDA applies a 2D-discrete Fourier transform (DFT) to each seqlet's attribution map a_k:

$$\mathcal{F}(a_k) = \mathrm{DFT}(a_k) \tag{1}$$

F-MoDA then retains only the **magnitude spectrum** $|\mathcal{F}(a_k)|$, discarding the phase information. This ensures that seqlets exhibiting similar attribution patterns, irrespective of positional shifts within the window, will have comparable Fourier magnitude profiles. F-MoDA subsequently flattens these magnitude profiles into vectors $f_k \in \mathbb{R}^D$ ($D = W \times 4$), which serve as the primary feature representation for the initial clustering phase.

A.2. Initial coarse clustering. To efficiently generate initial motif candidates, F-MoDA employs a lightweight clustering algorithm, such as K-means clustering with the Euclidean distance, on the collection of flattened Fourier magnitude vectors $\{f_k\}$:

$$\mathrm{InitialLabels} = \text{K-means}(\{f_k\}, N_c) \tag{2}$$

where N_c is a user-defined number of initial clusters. This step groups seqlets with globally similar attribution signal shapes, providing a coarse partitioning of the data. The primary goal of this stage is to speed up the following steps and to group seqlets for more refined downstream processing.

A.3. Iterative motif refinement and merging. Following the initial coarse clustering, F-MoDA iteratively refines the initial clusters to produce high-quality, consolidated motifs. Each iteration, detailed in Algorithm 1, involves

- Multiple sequence alignment (MSA) of the nucleotide sequences derived from seqlets within each current cluster.
- Generation of a mean-aligned attribution map for each cluster.
- Filtering of seqlets based on their similarity to their cluster's mean attribution map.
- A consolidation pass (Algorithm 2) to merge similar clusters based on attribution-score similarity.

F-MoDA can handle DNA motifs in both their forward and reverse-complement (RC) orientations, conceptually integrated into the similarity assessment and merging logic. However, for clarity, Algorithms 1-2 are presented without RC support.
Let $\mathcal{C}_{\text{initial}}$ be the set of initial clusters. For any set of current clusters $\mathcal{C}$, let $\mathcal{A}'(\mathcal{C})$ denote the set of mean-aligned (and filtered) attribution maps derived from them. If RC support is enabled, the process implicitly tracks an optimal orientation (original or RC) for each cluster as it evolves, and this orientation is used when preparing seqlets for MSA and generating $\mathcal{A}'(\mathcal{C})$.

Algorithm 1 Iterative merging and refinement.

Input: Set of seqlets, initial cluster assignment $\mathcal{C}_{\text{initial}}$, merging threshold θ_{merge}, drop threshold δ.

$\mathcal{C}_{\text{curr}}, \textit{converged} \leftarrow \mathcal{C}_{\text{initial}}, \text{false}$
while not *converged* **do**
  $\mathcal{C}_{\text{prev}} \leftarrow \text{Copy}(\mathcal{C}_{\text{curr}})$
  $\mathcal{A}'_{\text{maps}} \leftarrow \emptyset$ ▷ To store mean maps of current active clusters
  for each cluster $C_j \in \mathcal{C}_{\text{curr}}$ **do**
    Perform MSA on sequences in C_j; generate aligned attribution maps
    Compute mean aligned attribution map $\bar{A}'_j$
    Filter seqlets in C_j against $\bar{A}'_j$ using adaptive drop threshold $\tau_{\text{drop},j}$
    if C_j is not empty after filtering **then**
      Update $\bar{A}'_j$ based on filtered seqlets
      Add $\bar{A}'_j$ to $\mathcal{A}'_{\text{maps}}$
    else
      Remove inactive cluster C_j from $\mathcal{C}_{\text{curr}}$
    end if
  end for
  ▷ Cluster Consolidation Pass using Algorithm 2
  $M_{\text{sim}} \leftarrow \text{ComputeSimilarityMatrix}(\mathcal{A}'_{\text{maps}})$
  $\mathcal{C}_{\text{next}} \leftarrow \text{PerformConsolidationPass}(M_{\text{sim}}, \mathcal{C}_{\text{curr}}, \mathcal{A}'_{\text{maps}}, \theta_{\text{merge}})$
  if $\mathcal{C}_{\text{next}}$ is unchanged from $\mathcal{C}_{\text{prev}}$ **then**
    $\textit{converged} \leftarrow \text{true}$
  else
    $\mathcal{C}_{\text{curr}} \leftarrow \mathcal{C}_{\text{next}}$
  end if
end while
return $\mathcal{C}_{\text{curr}}$

The similarity measure we use in F-MoDA, for both filtering seqlets within clusters and comparing mean cluster profiles, is the normalized 2D cross-correlation. For two mean attribution profiles $\bar{A}'_u$ and $\bar{A}'_v$,

this is:

$$\text{Sim}(\bar{A}'_u, \bar{A}'_v) = \max\left(\frac{\bar{A}'_u}{\|\bar{A}'_u\|_\infty} \star \frac{\bar{A}'_v}{\|\bar{A}'_v\|_\infty}\right) \quad \textbf{(3)}$$

where $\|X\|_\infty$ is the L-infinity norm and $\star$ is a 2D cross-correlation.
Seqlets within each cluster C_j are filtered against their respective mean $\bar{A}'_j$ (Algorithm 1, line 8) using an adaptive drop threshold, $\tau_{\text{drop},j}$. This threshold is established relative to a baseline derived from permuted data. Let Sim_{per} be the mean similarity (calculated using Equation 3) between $\bar{A}'_j$ and versions of its own constituent seqlets whose attribution scores have been randomly permuted along their sequence length. Let $Sim_{\text{clusterMax}}$ be the maximum similarity of any original seqlet in C_j to $\bar{A}'_j$. The adaptive drop threshold is then:

$$\tau_{\text{drop},j} = Sim_{\text{per}} + \delta \times (Sim_{\text{clusterMax}} - Sim_{\text{per}}) \quad \textbf{(4)}$$

where δ is a user-defined drop threshold. Seqlets whose similarity to $\bar{A}'_j$ falls below $\tau_{\text{drop},j}$ are removed. For inter-cluster comparisons, the 'ComputeSimilarityMatrix' step (Algorithm 1, line 16) builds the similarity matrix M_{sim}. When RC support is enabled, the calculation of M_{sim} considers not only the similarity between $\bar{A}'_u$ and $\bar{A}'_v$ in their current orientations but also their similarities when one is reverse-complemented. The matrix M_{sim} then effectively stores the maximum similarity scores achieved for each pair of cluster means and each possible orientation.
The cluster merging logic is executed in a distinct consolidation pass (Algorithm 2) where clusters are iteratively merged. A pair of clusters, C_i and C_j, with the highest similarity score Sim_{max} (obtained from M_{sim}), is merged (including their group clusters) if two conditions are met:

1. The Sim_{max} meets an adaptive similarity threshold:

$$\tau_{\text{merge},ij} = \theta_{\text{merge}} \times \frac{S_{\text{sim}}[i] + S_{\text{sim}}[j]}{2} \quad \textbf{(5)}$$

   where θ_{merge} is a user-defined merging threshold and $S_{\text{sim}}[k]$ is the self-similarity of the mean map $\bar{A}'_k$.

2. The proposed merge must be consistent across all clusters that have already been grouped together. Specifically, if clusters C_i and C_j each belong to larger groups formed by previous merges, then every pair of clusters across these two groups must also satisfy their respective adaptive similarity thresholds (Eq. 5). This requirement is enforced by the `ProposedMergeIsConsistent` condition (Algorithm 2, line 10), which prevents incompatible clusters from being merged.

When RC support is active, the merge decision implicitly uses the optimal orientations determined during the similarity matrix computation. The specific seqlets forming the resulting merged cluster are then accordingly oriented. The merging process continues until no more merges satisfy the criteria. The main loop (Algorithm 1) repeats the entire refinement cycle of MSA, filtering, and consolidation until the set of clusters stabilizes (i.e., no change in consecutive iterations).

A.4. Final motif representation. Upon convergence of the iterative refinement process (Algorithm 1), each final cluster represents a discovered motif. These motifs are presented as **attribution-weighted consensus maps**. These maps are effectively the mean-aligned attribution maps ($\bar{A}'_j$) of the seqlets within each final cluster. These maps can be visualized as sequence logos representing the nucleotide positions and types that contribute positively or negatively to the model's predictions.

B. Benchmarking and evaluation on experimental data. To assess F-MoDA's performance, we designed a comprehensive benchmark based on experimental data. This section details the dataset used, the deep learning model employed for generating attribution maps, and the measures used for evaluation.

Algorithm 2 PerformConsolidationPass.

Input: Similarity matrix M_{sim}, current active clusters C_{active} and their mean maps $\mathcal{A}'_{\text{maps}}$, θ_{merge}.

Initialize *groupMap*: each active cluster is its own group
$S_{\text{sim}} \leftarrow \text{DiagonalElements}(M_{\text{sim}})$ ▷ Self-similarities
$M'_{\text{sim}} \leftarrow M_{\text{sim}} - \text{Diagonal}(S_{\text{sim}})$ ▷ Ignore self-similarities
loop
  $Sim_{max}, (i, j) \leftarrow \max(M'_{\text{sim}}), arg\max(M'_{\text{sim}})$ ▷ Identify best merge candidates
  if $Sim_{max} = 0$ **then**
    Break
  end if
  $\tau_{\text{adaptive}} \leftarrow \theta_{\text{merge}} \times (S_{\text{sim}}[i] + S_{\text{sim}}[j])/2$
  if $Sim_{max} \geq \tau_{\text{adaptive}}$ **and** ProposedMergeIsConsistent($i, j, groupMap, M'_{\text{sim}}, \theta_{\text{merge}}, S_{\text{sim}}$) **then**
    ▷ Consistency ensures overall group coherence after merge
    Update *groupMap*: merge groups of i and j
  end if
  $M'_{\text{sim}}[i, j], M'_{\text{sim}}[j, i] \leftarrow 0, 0$ ▷ Mark candidates as processed
end loop
$\mathcal{C}_{\text{consolidated}} \leftarrow \text{RebuildClustersFromGroupMap}(groupMap, \mathcal{C}_{\text{active}})$
return $\mathcal{C}_{\text{consolidated}}$

B.1. Benchmark dataset. Our benchmarking dataset is based on the one introduced by STREME (15) for evaluating motif-discovery algorithms on ChIP-seq data. This dataset consists of 40 TF ChIP-seq experiments conducted in K562 cells, downloaded from ENCODE (17). A corresponding ground-truth motif is available for each of these experiments, derived either from high-throughput SELEX data for the same TF (18) or, if unavailable, for a member of the same TF family.

B.2. Deep-learning model for attribution map generation. To generate the base-resolution attribution maps required by F-MoDA and TF-MoDISco, we trained a convolutional neural network (CNN), which we denote as `CNN-Classifier`, for each of the 40 ChIP-seq datasets. The `CNN-Classifier` model processes one-hot encoded DNA sequences through two convolutional blocks. The first block consists of a 1D-convolutional layer with 16 filters (kernel size 8) followed by a ReLU activation and a max-pooling layer (kernel size 4). The second block applies a 1D-convolutional layer with 32 filters (kernel size 4), followed by a ReLU activation and a max-pooling layer (kernel size 4). The output from these blocks is then flattened and passed to a fully connected layer with 64 neurons and ReLU activation. Finally, a sigmoid activation function is applied to a single neuron in the output layer for binary classification to predict TF binding.

To test the effect of different architectures on the performance of F-MoDA and TF-MoDISco, we also trained a recurrent neural network based on gated-recurrent units (GRU), which we denote as `RNN-Classifier`. For this model, we replaced the convolutional layers of the `CNN-Classifier` with a bidirectional GRU to capture sequential dependencies along the DNA sequence.

Training procedure: For each of the 40 datasets, positive sequences were derived from the ChIP-seq peak regions, centered and extended to a length of 1000 base pairs. Negative sequences were subsequently generated by randomly selecting either the upstream or downstream flanking region of these positive sequences. We trained the model using a binary cross-entropy loss function and the Adam optimizer with a learning rate of 0.0001. We performed training for a maximum of 10 epochs, with an early-stopping criterion based on validation loss (patience of 5 epochs) to prevent overfitting. The model achieving the best validation loss was saved for subsequent attribution map generation.

Attribution map generation: We used the trained `CNN/RNN-Classifier` model to generate attribution maps for the positive sequences. We employed Integrated Gradients (4) to assign importance

scores to each nucleotide in the input sequences. These attribution maps served as the input for F-MoDA.

B.3. Motif discovery methods configurations. For the experimental benchmark evaluation, we configured F-MoDA as follows: seqlet identification utilized a sliding window of size 10 with a flanking region of 3 on each side, while other seqlet finding parameters adopted the defaults of TF-MoDISco. The initial coarse clustering was performed using K-means with $N_c = 20$. MSA was carried out using MAFFT (19), with a drop threshold (δ) of 0.5 and a merging threshold (θ_{merge}) of 0.7 applied during the iterative refinement and merging stage. We ran TF-MoDISco with seqlet identification parameters matching those of F-MoDA (sliding window of 10, flanking of 3), while all other parameters were at their defaults.

B.4. Evaluation measures. We evaluated the performance of F-MoDA and TF-MoDISco using several measures, comparing the discovered motifs against the ground-truth motifs:

1. **Max-similarity.** For each method, we identify the discovered motif with the highest similarity score (Eq. 3) to the ground-truth motif. To achieve normalized scores across different TFs, the score is normalized by the self-similarity of the relevant TF ground-truth motif. This similarity score reflects the ability of a method to accurately find the ground-truth motif.

2. **Sensitivity.** This measure assesses how many motifs reported by each method are considered a "match" to the ground-truth motif. A motif is counted as matched if its similarity score to the ground truth exceeds a given similarity threshold. We vary this similarity threshold (e.g., from 0.1 to 1.0) and plot the average number of matched motifs per dataset to gauge the sensitivity of a method at different similarity thresholds. To avoid an advantage to methods that report redundant motifs, a maximum count of one is allowed in each experiment.

3. **Specificity.** Complementary to the previous measure, this measure measures the average number of reported motifs per dataset that do not achieve a similarity score greater than or equal to the given similarity threshold compared to the ground truth. This indicates the methods' propensity to report motifs dissimilar to the target, potentially representing false motifs.

4. **Conciseness.** Motif-discovery methods can sometimes report multiple highly similar motifs, which can be redundant. To quantify this, we calculate the pairwise similarity (Eq. 3) between all motifs discovered by a method for a given dataset. We then count the number of pairs whose similarity exceeds a given similarity threshold (e.g., from 0.1 to 1.0). The average count of such redundant pairs across all datasets is reported as a function of the similarity threshold.

These measures collectively provide a comprehensive view of the accuracy, sensitivity, specificity, and conciseness of the motif sets produced by F-MoDA compared to TF-MoDISco.

C. Synthetic data. We generated a synthetic dataset to evaluate the impact of the Fourier transform step in F-MoDA. For this evaluation, we selected three TF motifs (Arid5a, SPI1, and MYOD1) from the JASPAR database (20). We embedded these motifs into background sequences. We generated the background sequences using a 3rd-order Markov chain model trained on human genome data to better mimic natural genomic context.
We trained a model, utilizing the `CNN-Classifier` architecture (as described in Section B.2), to distinguish sequences containing any of the three embedded JASPAR motifs from the generated background-only sequences. The training procedure and the process for generating attribution maps also followed those detailed in Section B.2.

Results

We evaluated the performances of F-MoDA and TF-MoDISco on an established benchmark of 40 TF ChIP-seq experiments from K562 cells, using attribution maps based on trained CNNs

(`CNN-Classifier`, Section B). Moreover, we gauged the computational efficiency (i.e., runtime and memory usage) of F-MoDA compared to TF-MoDISco.

D. F-MoDA discovers more accurate and concise motifs. F-MoDA outperformed TF-MoDISco, producing motifs more similar to the ground-truth motifs while also producing a more focused and less redundant set of motifs. A representative example for the ELF1 ChIP-seq experiment (Figure 2A) shows that while both methods identify the ground truth, F-MoDA generated a single motif with a reported support of 4264 seqlets, whereas TF-MoDISco yields a larger, more redundant set of 32 motifs in total. This conciseness is a general trend: F-MoDA reported an average of 2.33 ± 1.35 motifs per experiment, significantly fewer than TF-MoDISco which reported 21.65 ± 8.12 ($p = 3.53 \times 10^{-8}$, Wilcoxon signed-rank test).

Quantitative comparison of the primary discovered motif to the ground truth (max-similarity measure, Section B.4) shows that F-MoDA achieved significantly higher similarity with an average score of 0.61 ± 0.17 compared to 0.57 ± 0.17 obtained by TF-MoDISco ($p = 0.011$, Wilcoxon signed-rank test; Figure 2B), indicating superior motif quality for F-MoDA. The similarity scores were positively correlated with the prediction performance as gauged by AUC-ROC on a held-out test set. This correlation was statistically significant for F-MoDA ($r = 0.38$, $p = 0.016$) and showed a positive, albeit non-statistically significant, trend for TF-MoDISco ($r = 0.29$, $p = 0.071$).

F-MoDA also demonstrated higher sensitivity in detecting at least one correct motif per dataset across various similarity thresholds (sensitivity measure, Section B.4, Figure 2C). This higher sensitivity, coupled with a significantly lower count of unmatched motifs (specificity measure, Section B.4, Figure 2D), indicates that F-MoDA has a much lower propensity for reporting false motifs compared to TF-MoDISco.

Furthermore, F-MoDA generates substantially fewer redundant motifs (conciseness measure, Section B.4, Figure 2E). This redundancy reduction, as achieved by F-MoDA, is crucial for human and machine interpretability, as it provides a clearer, more focused set of candidate regulatory elements. Collectively, these results indicate that F-MoDA not only produces higher quality motifs and fewer false motifs but also presents them in a more concise and interpretable manner.

As different motif similarity scores have been used in previous studies (16), we also evaluated performance using the continuous Jaccard similarity measure. Prior to computing similarity, we normalized the attribution maps by their maximum peak value. This analysis revealed similar trends as for the cross-correlation similarity, further supporting the robustness of F-MoDA's improvements (Supplementary Figure S1).

E. F-MoDA is robust to training runs and model architecture. To evaluate the robustness of F-MoDA to training runs and model architecture, we conducted two additional analyses. First, we assessed robustness across multiple training runs by training the `CNN-Classifier` model five times and evaluating motif discovery using all four evaluation measures. The results show that F-MoDA is robust across independent training sessions (Supplementary Figure S2A-B) and consistently maintains its superiority compared to TF-MoDISco (Supplementary Figure S2A, C-E).

Second, we examined robustness across model architectures by replacing the CNN with a recurrent neural network based on GRU (`RNN-Classifier`, Section B). In this setting, the max-similarity measure yielded comparable performance between F-MoDA and TF-MoDISco, without statistical significance. However, the results for the remaining evaluation measures were consistent with those obtained using the CNN-based model, further supporting the robustness of F-MoDA across architectures (Supplementary Figure S3).

F. F-MoDA outperforms TF-MoDISco in runtime and memory usage. We compared the runtime and maximum memory usage of F-MoDA and TF-MoDISco as a function of the number of input seqlets, using data from five arbitrary ChIP-seq experiments, excluding the seqlet generation step itself, which is the same in both methods.

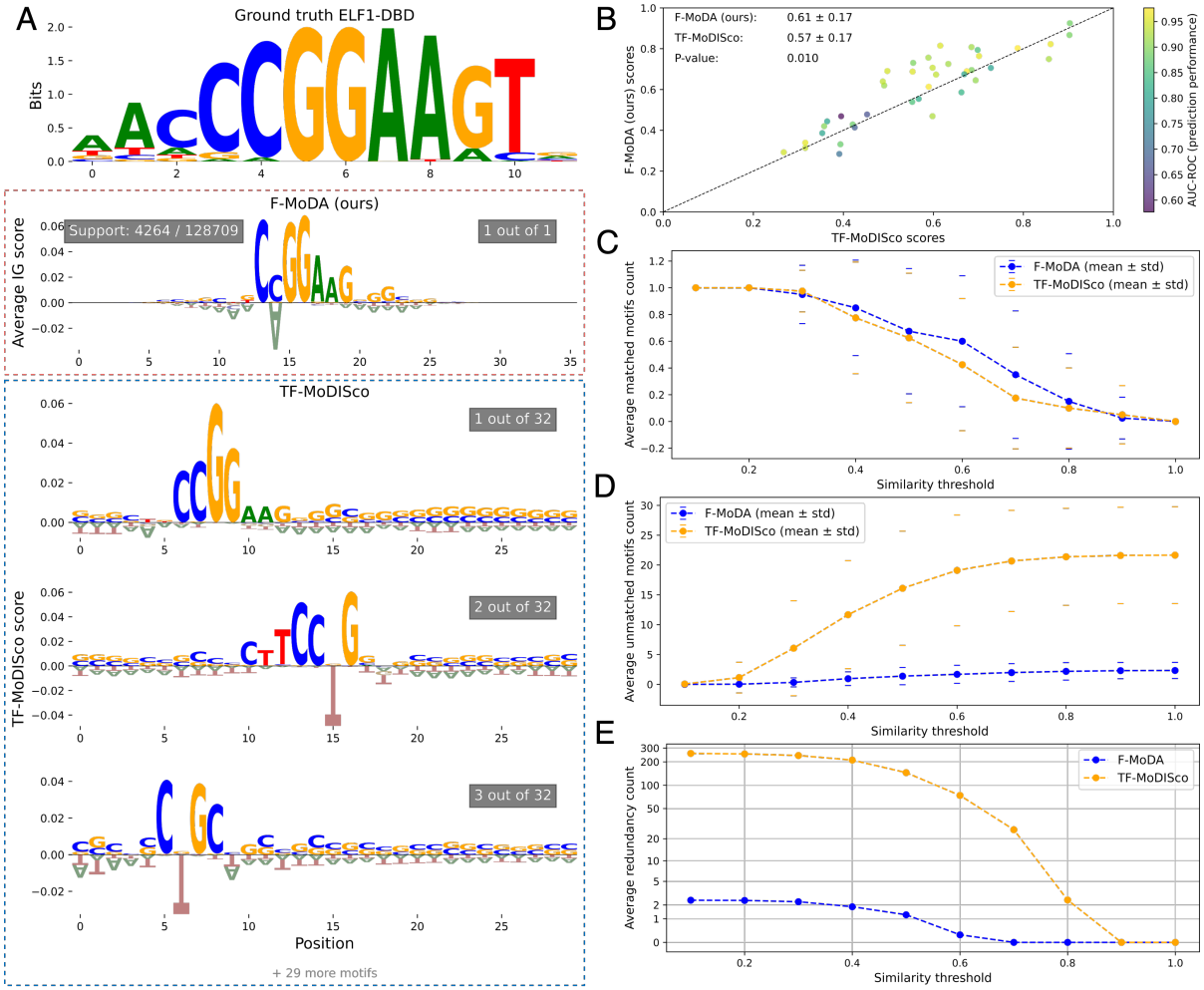

Fig. 2. Performance comparison on benchmark ChIP-seq datasets based on the cross-correlation similarity measure. **(A) Motif discovery results for the ELF1 ChIP-seq experiment.** Motifs identified by F-MoDA and TF-MoDISco in attribution maps generated using Integrated Gradients (IG). F-MoDA also reports a support value indicating the number of seqlets contributing to each motif. **(B) Motif identification quality across 40 ChIP-seq experiments in K562 cells.** Each point represents the maximum similarity score between the discovered motifs and a SELEX-derived reference motif (which is considered the ground truth) for one experiment (max-similarity measure). Point color indicates the model's prediction performance on a held-out test set. Average similarity scores for F-MoDA and TF-MoDISco are also shown. Statistical significance was assessed using the Wilcoxon signed-rank test. **(C-E) Sensitivity, specificity, and conciseness measures.** Average counts per dataset versus similarity threshold for: (C) matched motifs (relative to ground truth; sensitivity measure); (D) unmatched motifs (relative to ground truth; specificity measure); and (E) internal redundancy of reported motifs (conciseness measure).

F-MoDA is much faster than TF-MoDISco across all tested seqlet counts (Figure 3A). The runtime advantage of F-MoDA becomes increasingly pronounced on larger datasets. For instance, with 10^5 seqlets, F-MoDA's mean runtime is 389 ± 52 seconds, compared to $36,000 \pm 3,813$ seconds for TF-MoDISco, an improvement of orders of magnitude. In terms of memory usage, F-MoDA is much more efficient compared to TF-MoDISco (Figure 3B). F-MoDA required less maximum memory across all seqlet counts. This superior scaling with the number of seqlets makes F-MoDA more suitable for interpreting deep neural networks trained on current high-throughput genomic datasets.

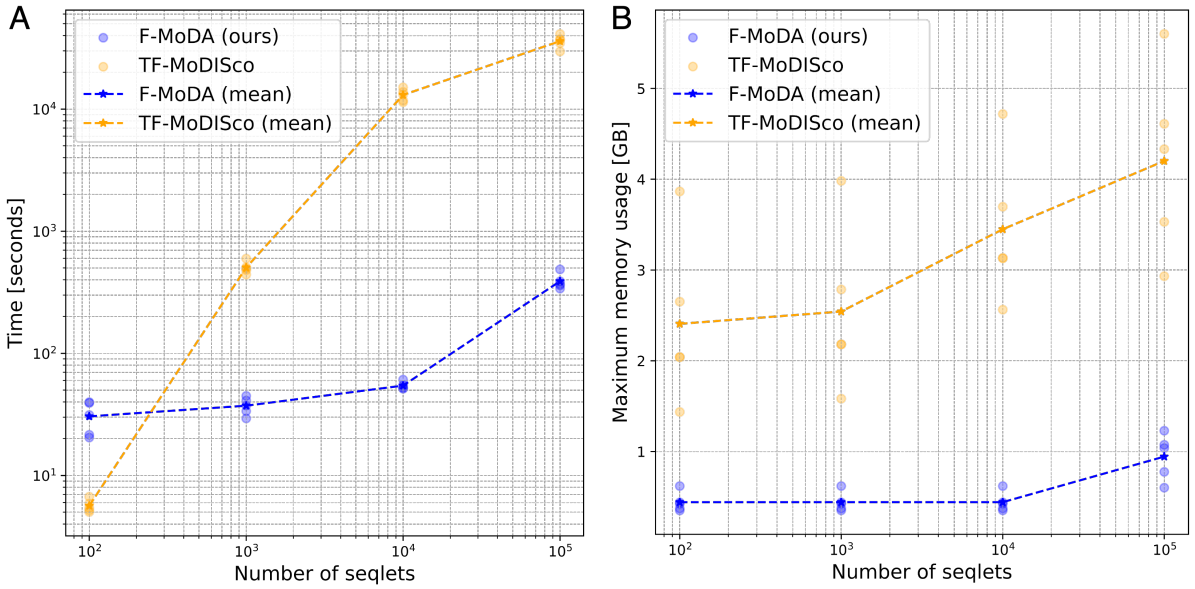

Fig. 3. Comparison of runtime and maximum memory usage. For each seqlet count, five ChIP-seq experiments were used to generate the required number of seqlets. Reported times and maximum memory usage exclude the seqlet generation step, which is the same in both methods.

G. Fourier transform enhances the motif recovery. We assessed the contribution of the Fourier transform component within F-MoDA on the synthetic dataset embedded with Arid5a, SPI1, and MYOD1 JASPAR motifs (Section C). We found that F-MoDA with the Fourier transform discovered the Arid5a motif more clearly and reported fewer false motifs (Figure 4). We note that we ran F-MoDA with the same parameters as in the experimental benchmark (Section B.3), except for a drop threshold (δ) of 0.1, as higher values did not clearly reveal the benefit of the Fourier transform.

We further compared F-MoDA with and without the Fourier transform on the experimental benchmark using all four evaluation measures. We observed a similar trend where incorporating the Fourier transform consistently improved similarity to the ground-truth motifs across different drop thresholds (Supplementary Figure S4).

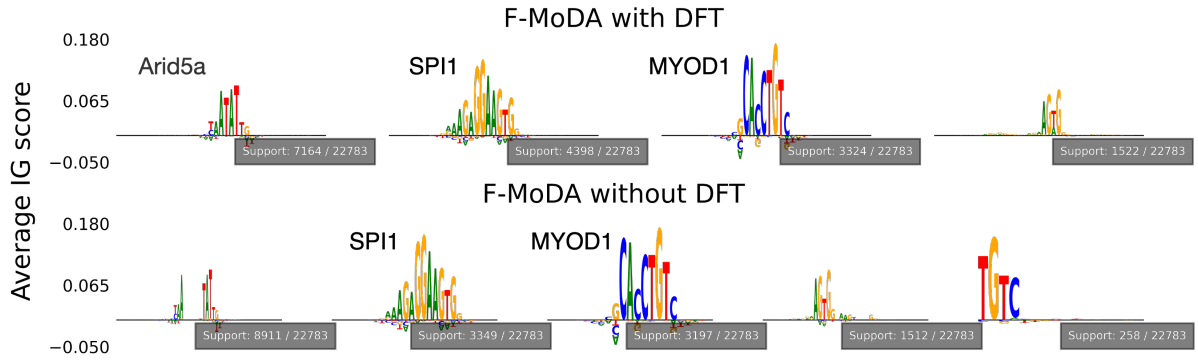

Fig. 4. Impact of Fourier transform on motif discovery in synthetic data. Comparison of motifs discovered by F-MoDA with (top row) and without (bottom row) the initial 2D-discrete Fourier transform (DFT) step on a synthetic dataset. The dataset contains embedded JASPAR motifs of Arid5a, SPI1, and MYOD1. IG denotes Integrated Gradients

Conclusions

While the discovery of enriched motifs from raw genomic data is a foundational problem, the advent of deep learning introduces a distinct, parallel challenge: interpreting the motifs learned by the models themselves. To address this, F-MoDA provides a scalable framework for motif discovery in attribution maps, enabling the efficient identification of learned, biologically relevant motifs. The accuracy, sensitivity, specificity, conciseness, and efficiency of F-MoDA make it well-suited for large-scale genomic analysis pipelines. Key areas for future development include: enhancing seqlet encoding through deep learning to capture richer, potentially phase-inclusive, information beyond the current Fourier-based approach; devising more sophisticated statistical methods for assessing motif significance in seqlet sets beyond simple support values which F-MoDA presents; extending the framework to handle amino-acid sequences, enabling motif discovery in proteomic contexts; evaluating the robustness of F-MoDA to attribution maps generated by different methods, such as DeepLIFT (5) and SHAP (6); and establishing improved, specialized evaluation measures for motifs discovered from attribution maps to create more robust benchmarks.

ACKNOWLEDGEMENTS
Ofir acknowledges the support of the Azrieli PhD fellowship, Kretiman PhD fellowship at Ben-Gurion University, and the cloud-compute credit by the Planning and Budgeting Committee through IUCC.

Bibliography

1. Babak Alipanahi, Andrew Delong, Matthew T Weirauch, and Brendan J Frey. Predicting the sequence specificities of DNA-and RNA-binding proteins by deep learning. *Nature Biotechnology*, 33(8):831–838, 2015.
2. Žiga Avsec, Vikram Agarwal, Daniel Visentin, Joseph R Ledsam, Agnieszka Grabska-Barwinska, Kyle R Taylor, Yannis Assael, John Jumper, Pushmeet Kohli, and David R Kelley. Effective gene expression prediction from sequence by integrating long-range interactions. *Nature Methods*, 18(10):1196–1203, 2021.
3. Gherman Novakovsky, Nick Dexter, Maxwell W Libbrecht, Wyeth W Wasserman, and Sara Mostafavi. Obtaining genetics insights from deep learning via explainable artificial intelligence. *Nature Reviews Genetics*, 24(2):125–137, 2023.
4. Mukund Sundararajan, Ankur Taly, and Qiqi Yan. Axiomatic attribution for deep networks. In *International Conference on Machine Learning*, pages 3319–3328. PMLR, 2017.
5. Avanti Shrikumar, Peyton Greenside, and Anshul Kundaje. Learning important features through propagating activation differences. In *International Conference on Machine Learning*, pages 3145–3153. PMLR, 2017.
6. Scott M Lundberg and Su-In Lee. A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30, 2017.
7. Suraj Srinivas and François Fleuret. Rethinking the role of gradient-based attribution methods for model interpretability. *arXiv preprint arXiv:2006.09128*, 2020.
8. Kelly Cochran, Melody Yin, Anika Mantripragada, Jacob Schreiber, Georgi K Marinov, Sagar R Shah, Haiyuan Yu, JOHN T LIS, and Anshul Kundaje. Dissecting the cis-regulatory syntax of transcription initiation with deep learning. *BioRxiv*, 2024.
9. Sahin Naqvi, Seungsoo Kim, Saman Tabatabaee, Anusri Pampari, Anshul Kundaje, Jonathan K Pritchard, and Joanna Wysocka. Transfer learning reveals sequence determinants of the quantitative response to transcription factor dosage. *Cell Genomics*, 5(3), 2025.
10. Johannes Linder, Divyanshi Srivastava, Han Yuan, Vikram Agarwal, and David R Kelley. Predicting rna-seq coverage from dna sequence as a unifying model of gene regulation. *Nature Genetics*, 57(4):949–961, 2025.
11. Peter K Koo and Matt Ploenzke. Improving representations of genomic sequence motifs in convolutional networks with exponential activations. *Nature Machine Intelligence*, 3(3):258–266, 2021.
12. Alex Tseng, Avanti Shrikumar, and Anshul Kundaje. Fourier-transform-based attribution priors improve the interpretability and stability of deep learning models for genomics. *Advances in Neural Information Processing Systems*, 33:1913–1923, 2020.
13. Valerie Chen, Muyu Yang, Wenbo Cui, Joon Sik Kim, Ameet Talwalkar, and Jian Ma. Applying interpretable machine learning in computational biology—pitfalls, recommendations and opportunities for new developments. *Nature Methods*, 21(8):1454–1461, 2024.
14. Timothy L Bailey, Charles Elkan, et al. Fitting a mixture model by expectation maximization to discover motifs in bipolymers. 1994.
15. Timothy L Bailey. STREME: accurate and versatile sequence motif discovery. *Bioinformatics*, 37(18):2834–2840, 2021.
16. Avanti Shrikumar, Katherine Tian, Žiga Avsec, Anna Shcherbina, Abhimanyu Banerjee, Mahfuza Sharmin, Surag Nair, and Anshul Kundaje. Technical note on transcription factor motif discovery from importance scores (tf-modisco) version 0.5. 6.5. *arXiv preprint arXiv:1811.00416*, 2018.
17. Natalie de Souza. The encode project. *Nature Methods*, 9(11):1046–1046, 2012.
18. Arttu Jolma, Jian Yan, Thomas Whitington, Jarkko Toivonen, Kazuhiro R Nitta, Pasi Rastas, Ekaterina Morgunova, Martin Enge, Mikko Taipale, Gonghong Wei, et al. Dna-binding specificities of human transcription factors. *Cell*, 152(1):327–339, 2013.
19. Kazutaka Katoh, Kazuharu Misawa, Kei-ichi Kuma, and Takashi Miyata. MAFFT: a novel method for rapid multiple sequence alignment based on fast fourier transform. *Nucleic Acids Research*, 30(14):3059–3066, 2002.
20. Oriol Fornes, Jaime A Castro-Mondragon, Aziz Khan, Robin Van der Lee, Xi Zhang, Phillip A Richmond, Bhavi P Modi, Solenne Correard, Marius Gheorghe, Damir Baranašić, et al. JASPAR 2020: update of the open-access database of transcription factor binding profiles. *Nucleic Acids Research*, 48(D1):D87–D92, 2020.