APPENDIX A: THE COMPARISON TABLE OF SOLUTIONS

This appendix contains a table comparing motif discovery solutions. Due to page limitations, the table has not been included in the main text but is available for reference in this appendix.

Method	Advantages	Disadvantages	Time Complexity
MEME [1]	Recognition of complex patterns Additional parameters in specific data Discovering variable-length patterns	High computational complexity Dependency on representative training dataset Sensitivity to parameter tuning	$O(\#Seq^2 \times \lvert Seq\rvert^2)$
GLAM2 [2]	Identification of complex patterns (detecting patterns with and without gaps)	High computational complexity Parameter tuning in analyzing complex dataset	$O(\lvert Seq\rvert \times \#motif^2)$
DREME [3]	Identification of ungapped motifs (particularly rare motifs) Using statistical & inference methods	potentially result of Improper parameter Loss of crucial motifs Weakness in discovering long candidates	$O(\lvert Seq\rvert \times (\#letter\ per\ seq)^{\#motif})$
STREME [5]	Additional parameters in specific data Discovering long motif candidates	Long execution time Input parameters complexity Discovering non-exact candidates	$Uncomputed$
TFEM (This paper)	Execution Time Optimization Considering neighboring conditions Pruning the tree (avoiding redundant outcomes)	Relying on repetitive state High time-complexity in the worst-case	$O((\#letter\ in\ the\ input) \times 20^{\lvert motif\rvert})$
DB-SS [4]	Uncovering hidden insights Understanding the interconnections Predicting future outcomes Enhancing decision-making processes	High Computational complexity Requiring high-quality dataset Sensitivity to changes	$O(function\ of\ nodes\ \&\ edges)$

REFERENCES

[1] T. L. Bailey and C. Elkan, "Fitting a mixture model by expectation maximization to discover motifs in biopolymers.," *Proceedings. Int. Conf. Intell. Syst. Mol. Biol.*, vol. 2, pp. 28–36, 1994.

[2] M. C. Frith, N. F. W. Saunders, B. Kobe, and T. L. Bailey, "Discovering Sequence Motifs with Arbitrary Insertions and Deletions," *PLoS Comput. Biol.*, vol. 4, no. 5, p. e1000071, May 2008.

[3] T. L. Bailey, "DREME: motif discovery in transcription factor ChIP-seq data," *Bioinformatics*, vol. 27, no. 12, pp. 1653–1659, Jun. 2011.

[4] E. Czeizler, T. Hirvola, and K. Karhu, "A Graph-Theoretical Approach for Motif Discovery in Protein Sequences," *IEEE/ACM Trans. Comput. Biol. Bioinforma.*, vol. 14, no. 1, pp. 121–130, Jan. 2017.

[5] T. L. Bailey, "STREME: accurate and versatile sequence motif discovery," *Bioinformatics*, vol. 37, no. 18, pp. 2834–2840, Sep. 2021.