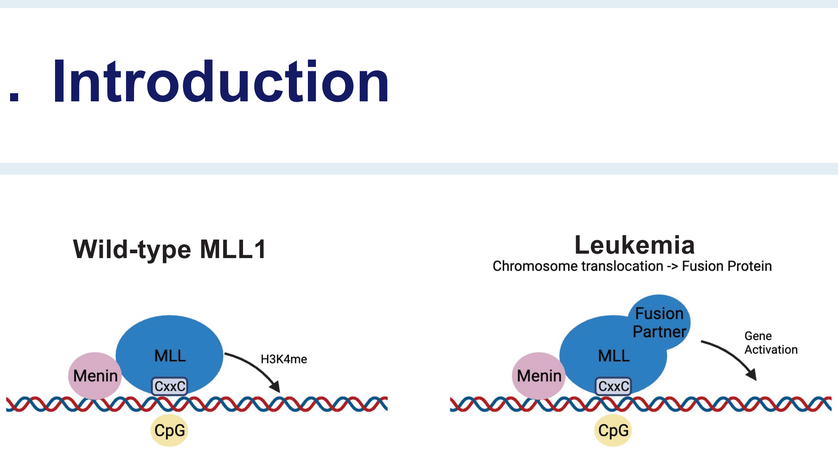
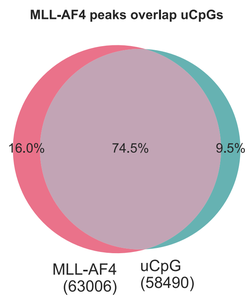
DNA binding specificity of MLL-AF4 in leukemia

Catherine Chahrour, Alastair L. Smith, Thomas A. Milne
MRC Molecular Haematology Unit, MRC Weatherall Institute of Molecular Medicine, University of Oxford

1. Introduction

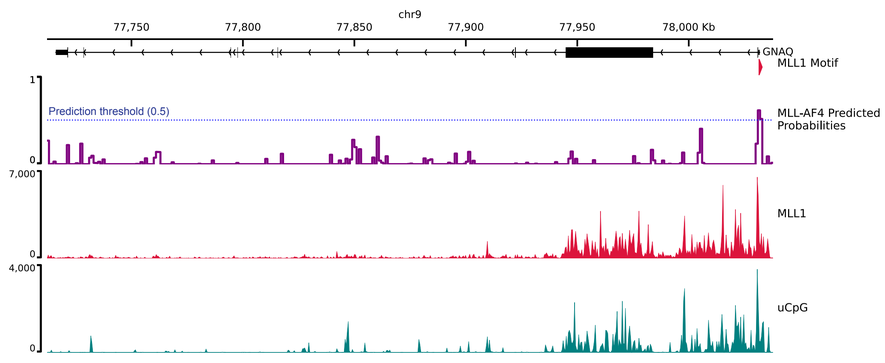


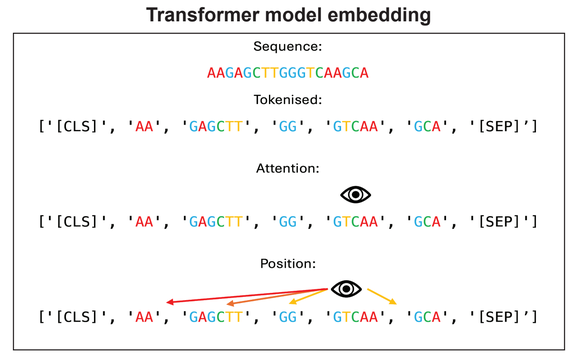
Wild-type MLL1 (KMT2A) and the oncofusion protein MLL-AF4 share an N-terminus that contains a CxxC domain for binding unmethylated CpG dinucleotides essential for the recruitment of MLLwt and MLL-AF4 to specific gene promoters, however, not every CpG island is bound by MLL-AF4 and MLLwt also binds outside of uCpGs[1].

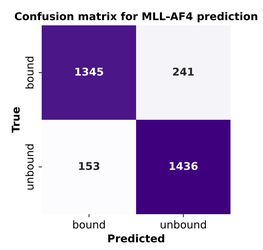


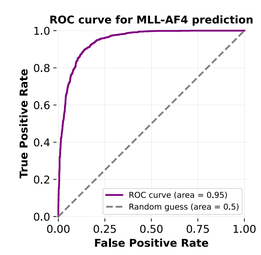
2. MLL-AF4 Binding can be predicted from sequence

To establish if DNA sequence is important for MLL-AF4 binding, a binary classification transformer model[2] was fine-tuned with MLL1 CUT&Tag peaks in MLL-AF4 (SEM) cells. MLL-AF4 binding was predicted with high accuracy (ROC AUC of 0.95).



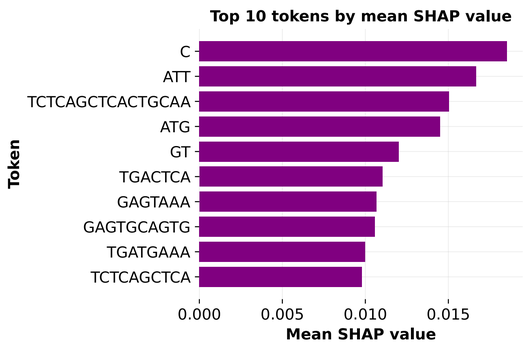


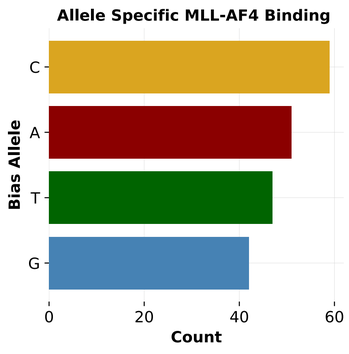




3. MLL-AF4 has a preference for cytosine

Feature extraction from MLL-AF4 binding prediction from DNA sequence showed cytosine alone was the most important token for determining MLL-AF4 bound sequences.

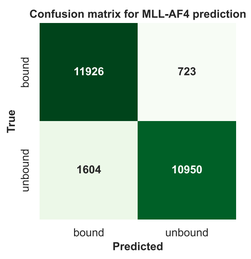


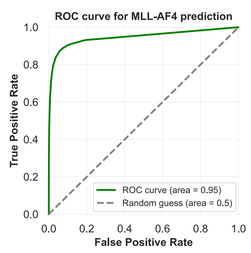


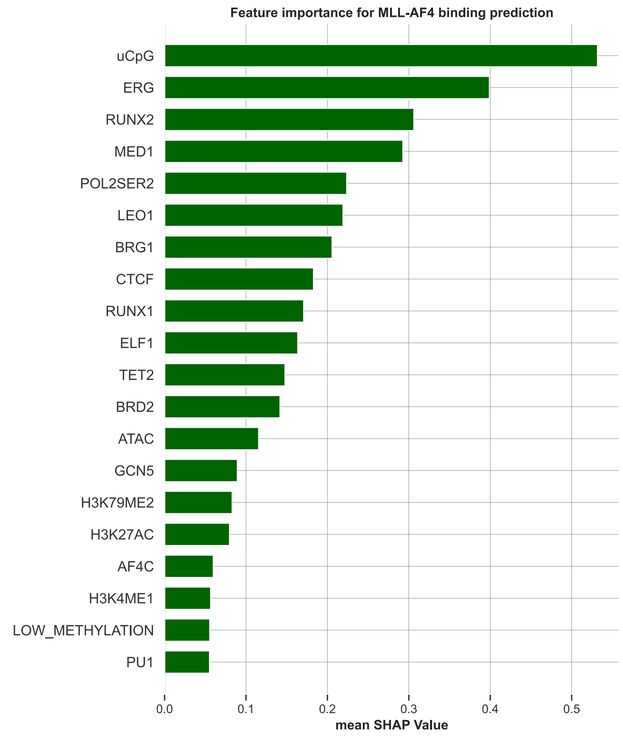
Allelic specific MLL-AF4 binding[3] identified SNPs with enhanced MLL-AF4 binding were biased towards cytosine.

4. Epigenomic landscape influences MLL-AF4 binding

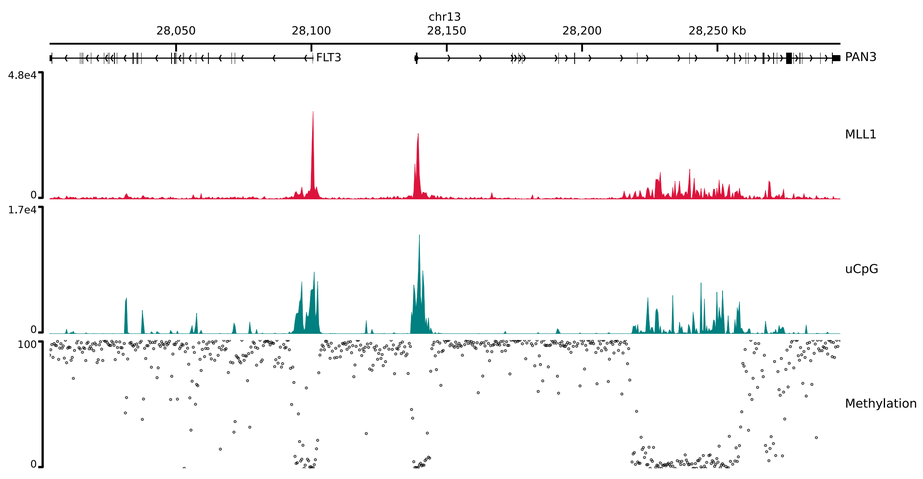
For determining which aspects of the epigenomic landscape are important for MLL-AF4 binding, a GBM classifier was trained to predict MLL-AF4 binding given a panel of 63 features. MLL-AF4 binding was predicted accurately (ROC AUC of 0.95). We confirmed that uCpGs along with ERG and RUNX2 had high importance in determining MLL-AF4 binding.

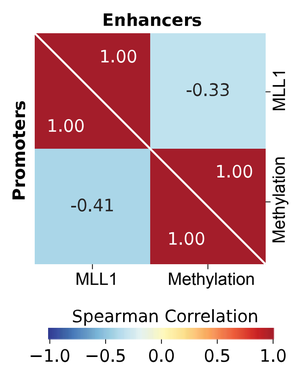






5. Methylation mirrors MLL-AF4 binding

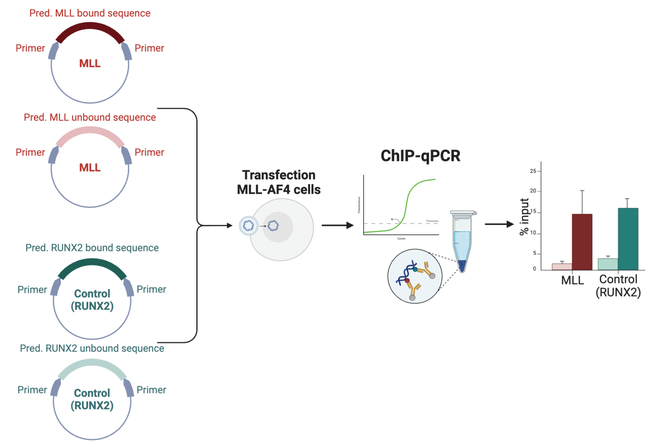




Methylation has an inverse correlation with MLL-AF4 binding, low methylation correlates with MLL-AF4 binding and high methylation anti-correlates with MLL-AF4 binding. This correlation is stronger at both promoters and enhancers.

6. Validation of predicted binding

Our findings enhance the understanding of how specific DNA sequences and the epigenomic landscape contribute to the unique binding patterns of MLL-AF4. Future work will validate predictions of MLL-AF4 binding.



7. References

1. Kerry, J., et al. (2017) *Cell Reports*
2. Zhou, Z., et al. (2024) *ArXiv*
3. de Santiago, I., et al. (2024) *Genome Biology*