14th International Multiconference
Bioinformatics of Genome Regulation and Structure / Biology

PREDICTION AND ANNOTATION OF ALTERNATIVE TRANSCRIPTION STARTS IN THE CHICKEN GENOME

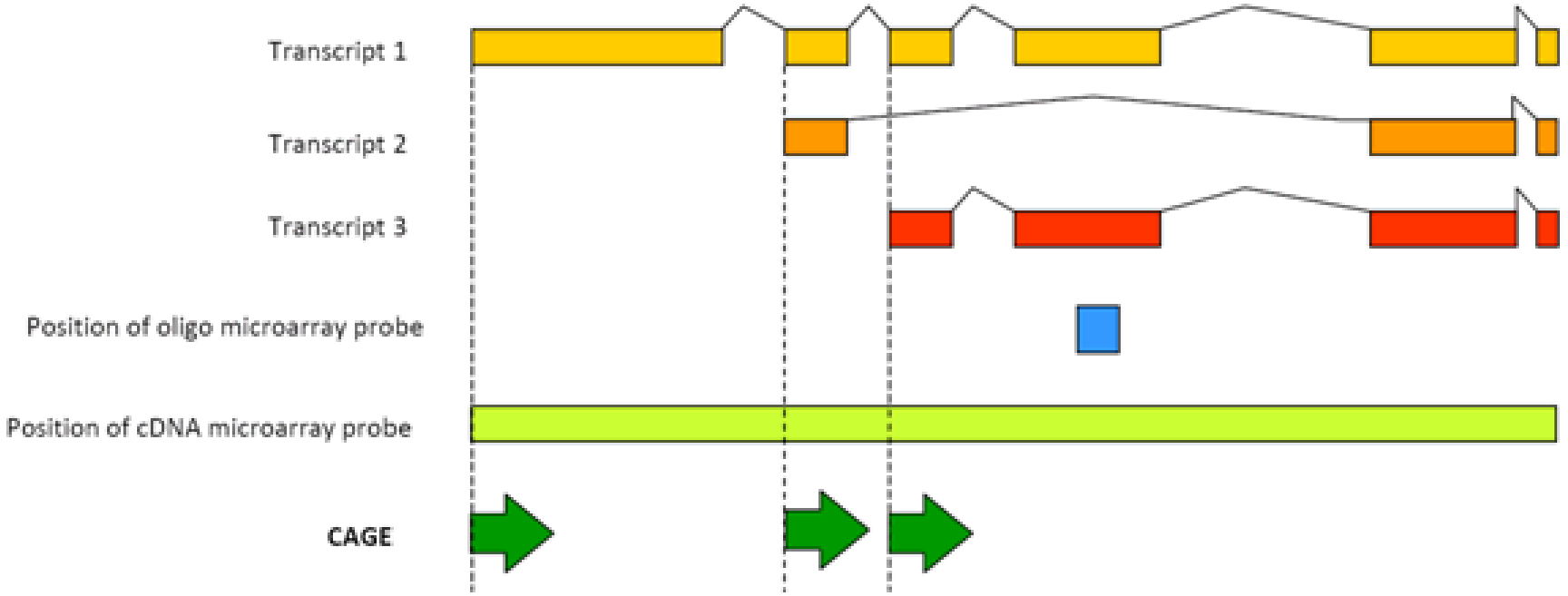
Grushina V., Pintus S.
Sirius University of Science and Technology

Sirius
2024

Introduction

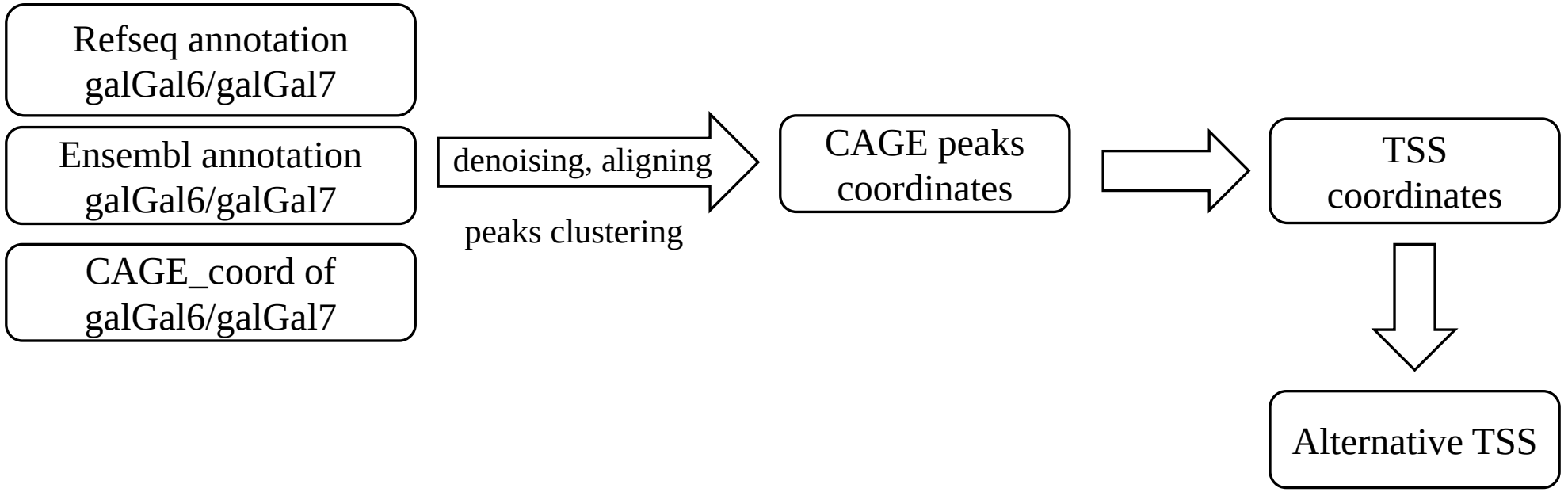
The analysis of promoter regions of genes often reveals differences from the existing annotation, which can be caused both by the insufficient accuracy of the existing annotation and the existence of transcripts previously unknown and activated under certain non-standard conditions [1, 2].

The data from CAGE experiments would make it possible to better identify the initial positions of such transcripts, TSS and the events of the promoter shift.



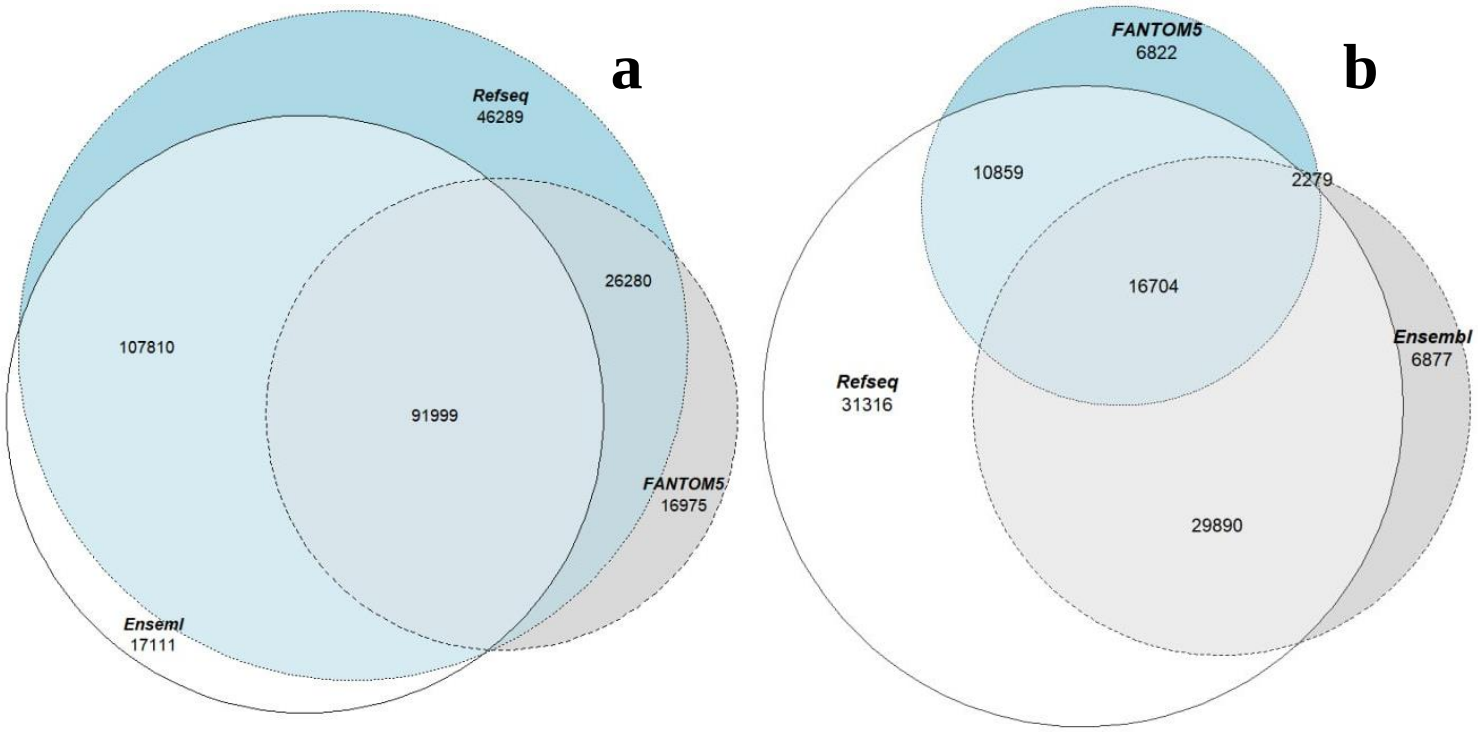
Methods

The galGal6 and galGal7 genome assemblies and Ensemble and Refseq genome annotations were used as a reference for comparing and validating TSS coordinates. STAR was used to align the CAGE reads against the reference genome, then the CAGE tag peaks were clustered with DPI1 tool. The TSS coordinates and alternative transcription sites (promoter shift events) were identified with the CageFightR package.



Results

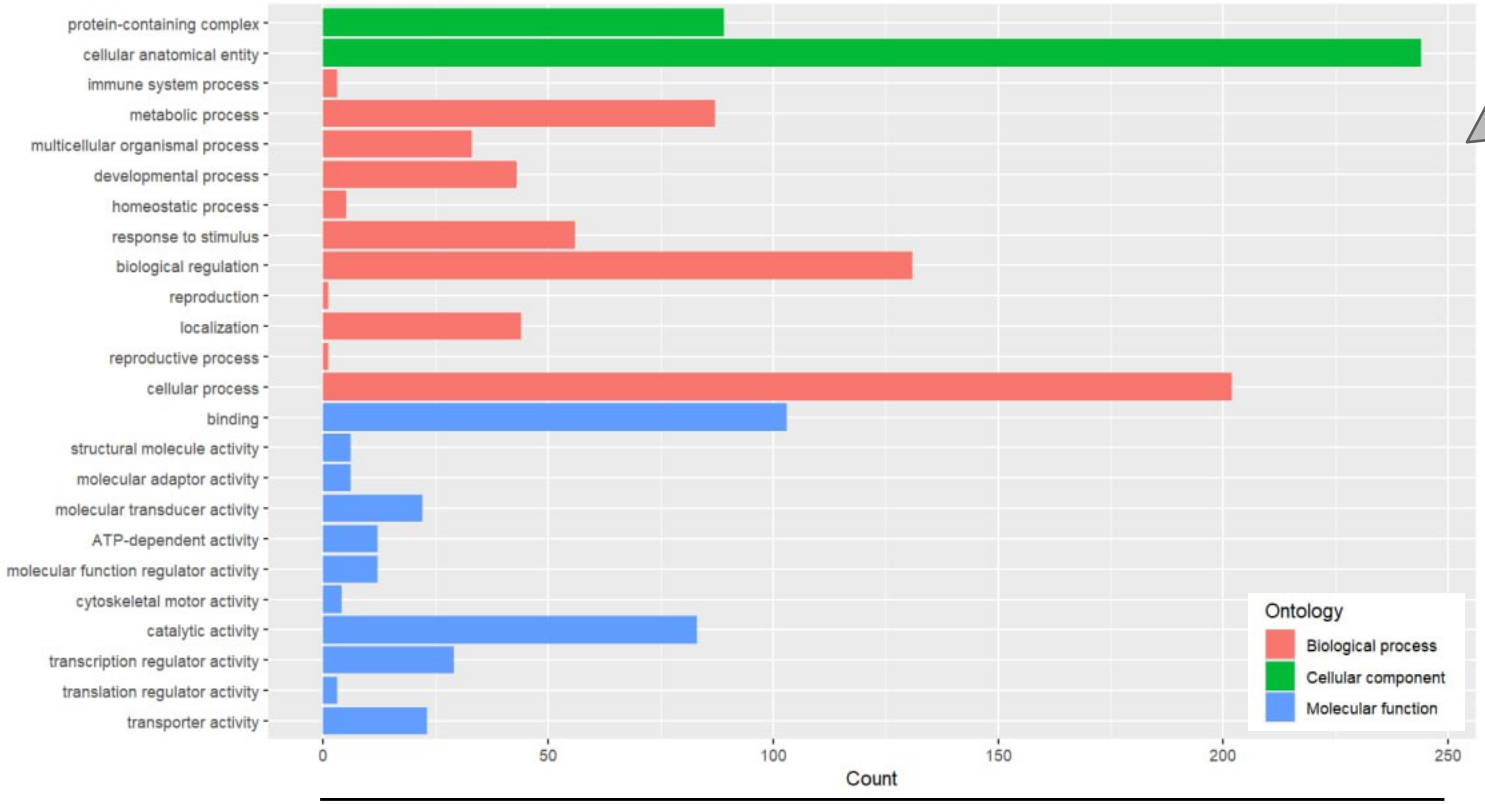
Differences were found between the predicted transcription starts in FANTOM5 and Ensembl and Refseq genome annotations of the chicken genome version 6, the re-aligned readings of FANTOM5 to the 7th version genome and Ensembl and Refseq genome annotations of the chicken genome version 7.



TSS of the chicken version 6
(a) and version 7 (b)

Discussion

2456 alternative TSS
276 genes



GO enrichment analysis

Was identified and annotated **2456** statistically significant alternative transcription starts in the chicken genome (putative promoter shifts) based on data from the FANTOM5 project and new experimental CAGE data.

Conclusion

In addition to the major transcription starts annotated in Refseq and Ensembl reference annotations, CAGE experiments reveal additional TSS. Our results suggest that some genes are characterized by the phenomenon of transcription start shifts that manifest under specific conditions.

References

1. Link V.M., Duttke S.H., Chun H.B., Holtman I.R., Westin E., Hoeksema M.A., Abe Y., Skola D., Romanoski C.E., Tao J., et al. Analysis of Genetically Diverse Macrophages Reveals Local and Domain-wide Mechanisms that Control Transcription Factor Binding and Function. Cell. 2018:173: 1796–1809. doi 10.1016/j.cell.2018.04.018
2. Warren D Anderson, Fabiana M Duarte, Mete Civelek, Michael J Guertin, Defining data-driven primary transcript annotations with primaryTranscriptAnnotation in R. Bioinformatics. 2020:36(9):2926–2928. doi 10.1093/bioinformatics/btaa011