

Report 5: Sequence conservation

Cyril Matthey-Doret

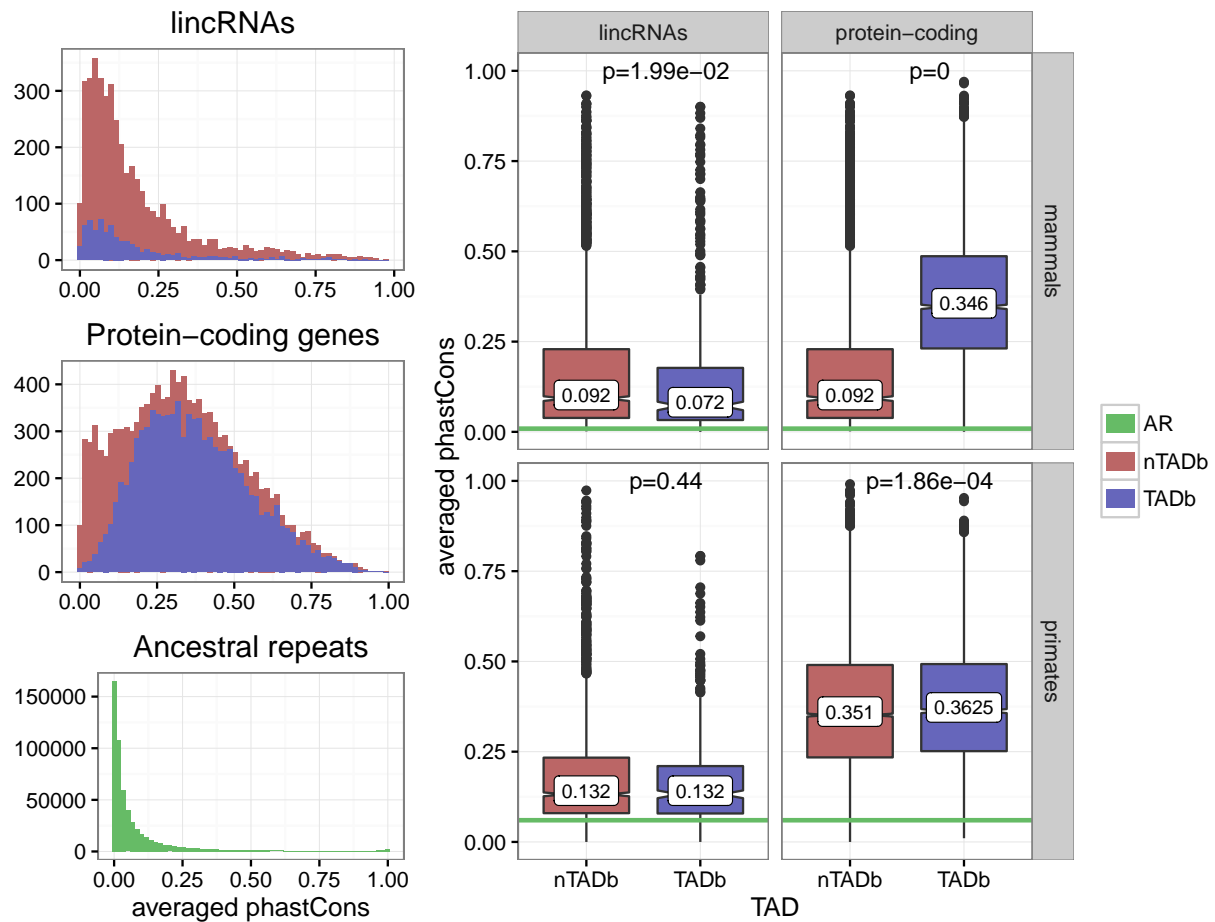
25 octobre 2016

Introduction

Sequence conservation across species can be quantified using **phastCons scores**. Here I compare averaged sequence conservation of TAD bound (TADb) and non-TAD bound (nTADb) lincRNAs and protein-coding genes. Ancestral repeats (AR) are also used as a comparison, since those elements are assumed to be evolving neutrally. Genes that have a higher phastCons score than ancestral repeats are likely to be under purifying selection, while those with a lower score should be under positive selection.

The TAD boundaries used to define if a gene is TAD bound or not are the “flexible” boundaries, described in report 3, section 1.

Comparison of sequence conservation



Conclusion

non-TAD bound lincRNAs seem to be more conserved across mammals than TAD bound ones, however there is no difference in primates. non-TAD bound protein-coding genes on the other hand, are always less conserved. These results indicate a stronger purifying selection for protein-coding genes inside TAD-boundaries, and a somewhat lower purifying selection for lincRNAs inside those boundaries.