Global regulatory transitions at core promoters demarcate the mammalian germline cycle

Nevena Cvetesic

06/11/2020

Abstract

Core promoters integrate regulatory inputs of genes. Global dynamics of promoter usage can reveal systemic changes in how genomic sequence is interpreted by the cell. We previously discovered that early development in fish is associated with a genome-wide shift in how promoters are read: the switch from the oocyte-specific to the somatic-specific promoter usage observed during Zygotic Genome Activation (ZGA) represents one of the most dramatic instances of transcriptional regulatory rewiring known in the animal life cycle. However, the key question of how and when the oocyte-specific pattern is established in the germ line has been hitherto impossible to answer, as well as whether this switch is universal across vertebrate development. Here we report the first analysis of promoter dynamics and code switching in the mammalian germ line. Using Super Low Input Carrier-CAGE (SLIC-CAGE) we show that mouse germline development starts with the somatic promoter code, followed by a prominent switch to the maternal code during follicular oogenesis. The sequence features underlying the shift from somatic to maternal code are conserved across vertebrates, despite large differences in promoter nucleotide compositions. In addition, we also find that the promoters of gonadal germ cells diverge from the canonical somatic transcription initiation. This divergence is distinct from the promoter code used later by developing oocytes and reveals genome-wide promoter remodelling associated with alternative nucleosome positioning during early female and male germline development. Collectively, our findings establish promoter-level regulatory transitions as a central, conserved feature of the vertebrate life cycle.

Package requirements

- Install following R packages from CRAN: dplyr_0.8.4, RColorBrewer_1.1-2, ggseqlogo_0.1, ggplot2_3.2.1, corrplot_0.84

- Install following R packages from Bioconductor: CAGEr_1.20.0, BSgenome_1.52.0, BSgenome.Mmusculus.UCSC.mm10_ rtracklayer_1.44.4, Biostrings_2.52.0, org.Mm.eg.db_3.8.2, TxDb.Mmusculus.UCSC.mm10.knownGene_3.4.7,GenomicFe AnnotationDbi_1.46.1, Biobase_2.44.0, ChIPseeker_1.20.0, TFBSTools_1.22.0, heatmaps_1.8.0, Gviz_1.29.1, GenomicRanges_1.36.1, GenomeInfoDb_1.20.0, IRanges_2.18.3, BiocGenerics_0.30.0, DESeq2_1.24.0, seqPattern_1.16.0

Figure to code map

- Figure 1 C