The role of transposable elements in shaping the combinatorial interaction of transcription factors

A. Testori[1], L. Caizzi[2], S. Cutrupi[3], O. Friard[3], M. De Bortoli[1], D. Corà[1] , M. Caselle[4]✉

[1]Center for Molecular Systems Biology, University of Turin, Dept. Oncological Sciences, SP142, Candiolo, Italy
[2]Center for Molecular Systems Biology, University of Turin, Bioindustry Park Silvano Fumero, Colleretto Giacosa, Itlay
[3]Center for Molecular Systems Biology and Dept. of Life Sciences and Systems Biology, University of Turin, Italy
[4]Center for Molecular Systems Biology and Dept. of Physics v. P. Giuria 1 -10125 Turin, University of Turin, Italy

Motivations

In the last few years several studies showed that transposable elements (TEs) in the human genome are significantly associated with transcription factor binding sites (TFBSs) and that in several cases their expansion within the genome led to a substantial rewiring of the regulatory network. Here we suggest another possible role played by TEs in the evolution of the regulatory networks. We discuss a set of evidences supporting the idea that the evolution of particular patterns on combinatorial interactions among Transcription Factors (TFs) was mediated and supported by the expansion of specific classes of TEs in the human genome.

Methods

We tested our conjecture looking, as a case study, at the binding of Estrogen Receptor alpha (ERalpha) to DNA using two chromatin immuno-precipitation sequencing (ChIP-seq) public datasets on MCF7 cell lines corresponding to different modalities of exposure to estrogen. On these datasets we performed a genome-wide analysis of Transposable Elements overlapping ChIP-seq binding peaks and for each of these sequences we performed a genome-wide scan for putative TFBSs employing canonical Positional Weight Matrices (PWMs).

Results

We found a remarkable enrichment of a few well defined types and classes of transposable elements overlapping the ChiP-seq peaks in our two datasets. Among these enriched TEs a prominent role was played by MIR (Mammalian Interspersed Repeats) transposons. These TEs underwent a dramatic expansion at the beginning of the mammalian radiation and then stabilized. We conjecture that the special affinity of ERalpha for the MIR class of TEs could be at the origin of the important role which ERalpha assumed in mammalians. Then, looking at the results of the scan for TFBSs, we found strong enrichment and correlated presence of a few specific TFs. In several cases these TFs were known cofactors of ERalpha, thus supporting the idea of a co-regulatory role of Transcription Factors located within the sane TE. Most of these correlations turned out to be strictly associated to specific classes of TEs thus suggesting the presence of a well defined "transposon code" within the regulatory network. Altogether our results support the idea that transposition events, besides rewiring the network, also played a central role in the emergence and success of combinatorial gene regulation in complex eukaryotes and that the evolution of specific combinations of TFs interactions was actually mediated and driven by the expansion of a few specific classes of Transposable Elements.