Profiling the promoterome of the human brain

Margherita Francescatto[1], Patrizia Rizzu[2], Javi Simon Sanchez[1], Peter Heutink[1]
and the Fantom 5 consortium
1) Genome Biology of neurodegenerative diseases, DZNE Tubingen 72076, Germany
2) Applied Genomics of neurodegenerative diseases, DZNE Tubingen 72076, Germany

Research question and background

The human brain is divided into distinct anatomical districts, characterized by different cellular compositions and functions and interconnected by complex communication networks. Differences in transcription are likely to play a major role in the establishment and maintenance of the morphological and functional differences observed across different brain regions. Aiming to build a complete promoter map to uncover the transcriptional regulatory networks defining the human brain we have coupled Cap analysis of gene expression (CAGE) technology to next generation sequencing. CAGE is a transcriptome exploration technology that captures the 5' end of capped RNA transcripts allowing for the high resolution profiling of transcription start sites in a quantitative and annotation independent manner.

Methods and tissues used

We profiled 15 regions of the human central nervous system, using post mortem tissue from one infant and three aged adult donors (spinal cord, temporal cortex, frontal cortex, parietal cortex, occipital cortex, cerebellum, medulla oblongata, hippocampus, putamen, caudate, thalamus, amygdala, substantia nigra, globus pallidus, locus coeruleus) .
Total RNA was extracted and purified from tissues using the Trizol tissue kit and RNA quality was assessed using the RNA Integrity Number (RIN). CAGE libraries where then prepared for each tissue. All the libraries were sequenced using the Heliscope single molecule sequencer.

Results and conclusion

We created a high-resolution atlas of transcription start sites for 15 anatomical regions of the human central nervous system, using post mortem samples derived from infant and aged adult donors. On the transcriptional level brain is clearly distinguishable from other tissues even if we consider only non-coding genes or expression from genomic regions often described as "genomic dark matter". Preliminary results show a significant fraction of brain-specific transcripts (>15%) that corresponds to loci that map to non-protein coding regions of the genome and lacks experimental characterization. This is an extremely valuable part of our dataset, completely unexplored and unknown, but potentially extremely relevant.
Among these, ~500 are long non –coding RNAs (lncRNAs): very few lncRNAs have been characterized in detail so far but there is increasing evidence of their fundamental involvement in several biological processes, e.g. cellular differentiation and pluripotency and thus it is possible that the brain-specific lncRNAs play an important role in brain region specification and differentiation of neuronal types.