

# Comparative genomics to identify regulatory regions: an example from the PAX8 gene

Valeria Di Dato<sup>(1)</sup>, Roberto Di Lauro<sup>(2)</sup> and Maria Luisa Chiusano<sup>(3)</sup>

<sup>(1)</sup>Stazione Zoologica “Anton Dohrn”, villa Comunale I, Napoli

v.didato@szn.it, dilauro@szn.it

<sup>(2)</sup>Dipartimento di Genetica, Biologia Generale e Molecolare, Università “Federico II”, Mezzocannone 8, 80134 Napoli

chiusano@unina.it

**Keywords:** Comparative genomics, promoter analysis, transcription factor

Comparisons between human and rodent non coding sequences are widely used for the identification of highly conserved sequences that could suggest functional implications [1], [2]. In particular, intergenomic comparisons are rapidly evolving for investigations on regulatory regions involved in promoter activity [3], [4], [5]. Moreover, the efficacy of such comparisons for the identification of functional regulatory elements, can be of help also in the study on the evolutionary dynamics of promoter sequences [6].

We are conducting computational analyses, based on comparative genomics between *Homo sapiens* and *Mus musculus*, on regions of at list 200kb spanning the entire genomic locus of genes involved in tyroid differentiation, to understand their expression mechanisms and regulation. A preliminary study on the PAX8 gene was supported by experimental analysis. The analysis resulted in the identification of 91 conserved regions of which 35 located at the 5' of the gene were chosen to start the experimental analysis. They were tested for functional implications in PAX8 promoter activity leading to the identification of tyroid specific regulatory regions.

The results of the current analysis provide experimental evidences that in turn have three fundamental perspectives: to help the clarification of the mechanisms of regulation and expression of the genes investigated; to improve the computational methodology proposed and strengthen its predictive power; to validate the computational approaches for the analysis of transcription factor binding sites, giving more hints to understand their organization and the pattern of evolution in regulatory sequences.

## References

[1] R.C. Hardison, J. Oeltjen, and W. Miller, Long human-mouse sequence alignments reveal novel regulatory elements: a reason to sequence the mouse genome. *Genome Res.* 7:959–966, 1997

[2] N. Jareborg, E. Birney, and R. Durbin, Comparative analysis of noncoding regions of 77 orthologous mouse and human gene pairs. *Genome Res.* 9:815–824, 1999

[3] J. Y. Leung, F. E. McKenzie, A. M. Ugliodoro, P. O. Flores-Villanueva and B. C. Sorkin, Identification of phylogenetic footprints in primate tumor necrosis factor-alpha promoters, *Proc. Natl. Acad. Sci. USA* 97:6614–6618, 2000

[4] W. Wasserman, M. Palumbo, W. Thompson, J.W. Fickett and C. E. Lawrence, Human-mouse genome comparisons to locate regulatory sites, *Nat. Genet.* 26:225–228, 2000

[5] L. Florea, M. Li, C. Riemer, B. Giardine, W. Miller et al., Validating computer programs for functional genomics in gene regulatory regions, *Curr. Genomics* 1:11–27, 2000.

[6] E. T. Dermitzakis and A. G. Clark, Evolution of transcription factor binding sites in Mammalian gene regulatory regions: conservation and turnover, *Mol. Biol. Evol.* 19(7):1114–1121, 2002