TONI GABALDÓN

Biography

Toni Gabaldón is a biochemist and molecular biologist by training (Universities of Valencia and Mainz). After several years working in a molecular biology lab on the regulatory response of yeasts to stress, he was attracted by the emerging fields of genomics and bioinformatics and moved to the comparative genomics group of Martijn Huynen in 2001 (Nijmegen, NL). He obtained his PhD in the Medical Sciences in 2005 (Radbout University Nijmegen), and then moved to the bioinformatics department at CIPF (Valencia) with an EMBO Long Term Fellowship. In September 2008, Gabaldón started his own group at CRG and from 2009 he is also associate professor at the Universitat Pompeu Fabra (UPF), where he teaches bioinformatics and molecular evolution. In 2012 he obtained the highly competitive starting grant from the European Research Council.

The main research interest of Gabaldon's group is to understand the complex relationships between genome sequences and phenotypes and how these two features evolve across species. He generally uses large-scale phylogenetics approaches that allow looking at the evolution of genomes from the perspective of all of their genes, and apply these analyses to a variety of biological questions, including those related to the evolution of fungal pathogens. He has published over 60 articles in international journals including Nature, Science, and PLoS Biology. More information can be found here: http://gabaldonlab.crg.es

Project

European Research Council Starting Grant
Project acronym: NONCODEVOL
Project full title: Evolutionary genomics of long, non-coding RNAs

Overview
Recent genomics analyses have facilitated the discovery of a novel major class of stable transcripts, now called long non-coding RNAs (lncRNAs). A growing number of analyses have implicated lncRNAs in the regulation of gene expression, dosage compensation and imprinting, and there is increasing evidence suggesting the involvement of lncRNAs in various diseases such as cancer. Despite recent advances, however, the role of the large majority of lncRNAs remains unknown and there is current debate on what fraction of lncRNAs may just represent transcriptional noise. Moreover, despite a growing number of lncRNAs catalogues for diverse model species, we lack a proper understanding of how these molecules evolve across genomes. Evolutionary analyses of protein-coding genes have proved tremendously useful in elucidating functional relationships and in understanding how the processes in which they are involved are shaped during evolution. Similar insights may be expected from a proper evolutionary characterization of lncRNAs, although the lack of proper tools and basic knowledge of underlying evolutionary mechanisms are a sizable challenge. Here, I propose to combine state-of-the-art computational and sequencing techniques in order to elucidate what evolutionary mechanisms are shaping this enigmatic component of eukaryotic genomes. The first goal is to enable large-scale phylogenomic analyses of lncRNAs by developing, for these molecules, methodologies that are now standard in the evolutionary analysis of protein-coding genes. The second goal is to explore, at high levels of resolution, the evolutionary dynamics of lncRNAs across selected eukaryotic groups for which novel genome-wide data will be produced experimentally. In particular, I plan to exploit RNAseq experiments and, most importantly, recently developed sequencing techniques that enable obtaining genome-wide footprints of RNA secondary structure. This will allow us to trace the conservation of lncRNA structures across species, information that, in turn, will be instrumental for the informed

development of phylogenetic algorithms. Finally, this dataset will be used to test the impact on lncRNAs evolution of processes such as, among others, purifying and positive selection, gene duplication, co-evolution and gene conversion. The global objective is to understand how lncRNAs evolve across species, and how this relates to the **e**volution of protein-coding genes and the biological processes to which they may be related. Potential outcomes of this research are the discovery of novel evolutionary mechanisms, the development of lncRNA-tailored phylogenetic algorithms, and the prediction of potential functional associations between lncRNAs, proteins, and biological processes.