



Master project 2024-2025

Personal Information

Supervisor	Roderic Guigo
Email	roderic.guigo@crg.cat
Institution	Center for Genomic Regulation
Website	https://www.crg.eu/en/programmes-groups/guigo-lab
Group	Roderic Guigo

Project

Computational genomics

Project Title:

The human selenomicrobiome

Keywords:

selenoproteins, microbiome, gene finding,

Summary:

Selenium (Se) is an essential trace element in humans and in organisms throughout the tree of life, including the human microbiota. The main biological form of Se is exerted in form of the non-canonical amino acid selenocysteine (Sec), incorporated in selenoproteins. Sec is analogous to cysteine (Cys), with Se replacing sulfur, and it is generally found in the active site of oxidoreductase enzymes. Sec is inserted co-translationally through a recoding mechanism in which a UGA codon (normally a stop codon) specifies Sec insertion. Specific RNA structures (SECISes, SEC Insertion Sequences) are present on selenoprotein transcripts and act as Sec recoding signals. While the set of selenoproteins in the human genome is well characterized, little is known about the selenoproteins in the human microbiome. The human microbiota is composed of bacteria, archaea, viruses, and microbial eukaryotes that inhabit the human body. These microbial communities interact with the host and play fundamental roles in human health and disease. The aim of the project is to characterize and identify the selenoprotein genes in the human microbiome. Because UGA is normally a stop codon, selenoproteins are poorly predicted in genome sequences. Our group has developed over the years a number of computational tools to precisely annotate selenoprotein genes. We have already some quite advanced preliminary results. Thus, the project is feasible in the time span of six to nine months.

References:

Characterization of Mammalian Selenoproteins, Kryukov et al. Science (2003) Selenoprofiles: profile-based scanning of eukaryotic genome sequences for selenoprotein genes, Mariotti and Guigo, Bioinformatics (2010) Computational identification of the selenocysteine tRNA (tRNA^{Sec}) in genomes, Santesmasses, Mariotti and Guigo, PLoS Computational Biology (2017)

Expected skills:

good programming skills, familiarity with sequence analysis methods, eventually capacity to develop novel methods

Possibility of funding:

To be discussed

Possible continuity with PhD:

To be discussed

