Universitat Pompeu Fabra Barcelona

Master in Bioinformatics for Health Sciences

Master project 2021-2022

Personal Information

Supervisor	Marco Mariotti
Email	marco.mariotti.mm@gmail.com
Institution	Universitat de Barcelona
Website	https://www.mariottigenomicslab.com/
Group	Comparative Genomics and Recoding lab

Project

Computational genomics

Project Title:

Computational genomics of proteins containing the special amino acid selenocysteine

Keywords:

Comparative Genomics; Evolution; Gene analysis; Translation; Recoding

Summary:

Our lab employs comparative genomics to study the mechanisms of gene expression and protein synthesis. We focus in particular on “recoding” events, programmed exceptions to the genetic code (1). A remarkable example of recoding is selenocysteine: this special amino acid is present in human and many other species, but it is not among the canonical 20 residues of the genetic code. Instead, it is encoded by the UGA codon, which is normally a stop, but it is recoded to selenocysteine through a highly regulated “readthrough” mechanism occurring only in specific mRNAs (2). Selenocysteine is found in the catalytic site of specialized enzymes, where it provides enhanced biochemical properties, typically for improved redox catalysis. Due to recoding, the genes encoding for selenocysteine-containing proteins (“selenoproteins”) are often missed or wrongly annotated in genomes, since gene annotation programs only consider the canonical role of UGA as stop (3). Selenoprotein genes in human have various well-known essential functions (4, 5), but a large part of the tree of life remains unexplored in this sense. The student may participate in three projects, related to the interest of the lab: • Development of automated approaches to recognize and correctly annotate selenoprotein genes in nucleotide sequences. In practice, the student will combine and improve programs for gene prediction and RNA motif finding (e.g. see (6, 7)). This is particularly important in context of new species being sequenced at unprecedented speed. • Evolutionary analysis of selenoprotein evolution. The student will apply tools from phylogenetics and sequence analysis to selenoprotein genes from diverse organisms, tracing how the selenocysteine utilization pathways changed throughout lineages (e.g. see (8–10)). • Analysis of selenoprotein function and regulation in disease. In practice, the student will make use of large public datasets of human or mouse data to analyse patterns of selenoprotein expression across tissues and disease conditions (e.g. see (11)).

References:

1. Rodnina,M. V, Korniy,N., Klimova,M., Karki,P., Peng,B.-Z., Senyushkina,T., Belardinelli,R., Maracci,C., Wohlgemuth,I., Samatova,E., et al. (2019) Translational recoding: canonical translation mechanisms reinterpreted. Nucleic Acids Res., 10.1093/nar/gkz783. 2. Labunskyy,V.M., Hatfield,D.L. and Gladyshev,V.N. (2014) Selenoproteins: molecular pathways and physiological roles. Physiol. Rev., 94, 739–77. 3. Santesmasses,D., Mariotti,M. and Gladyshev,V.N. (2020) Bioinformatics of Selenoproteins. Antioxidants Redox Signal., 33, 525–536. 4. Kryukov,G.V., Castellano,S., Novoselov,S.V., Lobanov,A.V., Zehtab,O., Guigo,R. and Gladyshev,V.N. (2003) Characterization of mammalian selenoproteomes. Science (80-.)., 300, 1439. 5. Mariotti,M., Ridge,P.G., Zhang,Y., Lobanov,A. V, Pringle,T.H., Guigo,R., Hatfield,D.L. and Gladyshev,V.N. (2012) Composition and evolution of the vertebrate and mammalian selenoproteomes. PLoS One, 7, e33066. 6. Mariotti,M. and Guigó,R. (2010) Selenoprofiles: Profile-based scanning of eukaryotic genome sequences for selenoprotein genes. Bioinformatics, 26. 7. Mariotti,M., Lobanov,A.V., Guigo,R. and Gladyshev,V.N. (2013) SECISearch3 and Seblastian: New tools for prediction of SECIS elements and selenoproteins. Nucleic Acids Res., 41. 8. Mariotti,M., Santesmasses,D., Capella-Gutierrez,S., Mateo,A., Arnan,C., Johnson,R., D’Aniello,S., Yim,S.H., Gladyshev,V.N., Serras,F., et al. (2015) Evolution of selenophosphate synthetases: Emergence and relocation of function through independent duplications and recurrent subfunctionalization. Genome Res., 25. 9. Mariotti,M., Lobanov,A. V, Manta,B., Santesmasses,D., Bofill,A., Guigó,R., Gabaldón,T. and Gladyshev,V.N. (2016) Lokiarchaeota Marks the Transition between the

Archaeal and Eukaryotic Selenocysteine Encoding Systems. Mol. Biol. Evol., 33, 2441–53. 10. Mariotti,M., Salinas,G., Gabaldón,T. and Gladyshev,V.N. (2019) Utilization of selenocysteine in early-branching fungal phyla. Nat. Microbiol., 4. 11. Avery,J. and Hoffmann,P. (2018) Selenium, Selenoproteins, and Immunity. Nutrients, 10, 1203.

Expected skills::

Python and/or R; Basics of gene evolution and phylogenetics; Expression data analysis.

Possibility of funding::

To be discussed

Possible continuity with PhD: :

To be discussed