GenoLimit: Genomic adaptations to life limit conditions: Comparative and functional genomics of isolated extremophilic Archaea and Bacteria

Summary

This project proposes to explore the genomic diversity underpinning microbial life in extreme environments. It uses the latest DNA sequencing methods and bioinformatics tools to compare genomes of some extremophiles groups and understand their regulation under extreme conditions.

Project description

Microorganisms occupy a central role in ecosystem function. They efficiently recycle nutrients (Arrigo, 2005), confer enhanced ecological fitness to animal and plant hosts (Backhed et al., 2005; Fan et al., 2012), and alter existing habitats to make them more hospitable to other life (Lavik et al., 2009). They underpin large food webs by utilizing a wide range of energy sources to produce accessible biomass for other organisms' consumption, especially in habitats devoid of light where chemoautotrophs act as the primary producers of organic matter (Dubilier et al., 2008).

Furthermore, microorganisms exclusively occupy the ends of the physicochemical spectrum defining livable habitats, with microbial growth observed in extreme conditions of temperature (Kashefi and Lovley, 2003), salinity (Maignien et al., 2013; Boetius and Joye, 2009) pressure (Zeng et al., 2009), and energy stress (Maignien et al., 2013; Valentine, 2007).

Such a pervasive role of microorganisms can be explained by their vast genetic and metabolic diversity. A long evolutionary process (~3.5 Gyrs), together with the high plasticity of microbial genomes, has led to the development of original metabolic and physiological features allowing microbes to exploit most forms of chemical energy sources. In extreme environments, microorganisms have evolved additional mechanisms to ensure genomic integrity, cell division, and energy conservation in conditions otherwise unable to host life.

This PhD project proposes to explore the evolution, adaptation and ecology of extremophilic Archaea and Bacteria through the study of their genome sequence. Comparative analysis of multiple extrêmophile genomes allows the identification of distinctive genes and metabolic pathways involved in the extremophilic way of life. It provides key data that enables comparison of functional gene composition across organisms and environments, and track evolutionary events involved in environmental adaptations at the population and strain level.
Furthermore, very little is known on the regulation of microbial gene expression in response to extreme environmental changes. Transcriptome analysis will be applied to axenic cultures during controlled in vitro experiments, thus yielding very valuable information regarding functions of differentially expressed genes. This functional genomics approach has the potential to provide novel information regarding the metabolic pathways involved in adaptation to the extremophilic way of life.

Sequencing and assembly of microbial isolate genomes are becoming financially and technically within reach of most laboratories (Tritt et al., 2012). This project will leverage the rich isolate collection of the LM2E by sequencing the genomes and transcriptomes of representative extremophile strains. This data will populate ExtremoBase, a genomic database for the extremophile research community under construction in the LM2E.

Research plan

Task 1: Identification of isolates of interests for comparative and functional genomics.

After a bibliographic work on available/lacking extremophilic genomes from phylogenetic and taxonomic groups, the PhD candidate will select strains with special interest that cover taxonomic or metabolic gaps among extremophilic Archaea and Bacteria. Hyperthermophilic and piezophilic groups from deep-sea hydrothermal and subseafloor environments (e.g. Thermosipho, Clostridium, Thermothoga, Thermus, Pyrococcus, Thermococcus, Archaeoglobus…) are more numerous in our collection and will be considered in priority, but the project remains open to other extremophile groups according PhD candidate interests and strain availability.

Task 2: Genome Sequencing, annotation and comparison

After growth of selected strains at their T and P optima, their genomes will be sequenced using the combination of next-generation high-throughput sequencing platforms, including but not limited to Illumina and PacBio. The PhD student will carry genome assembly and annotation, followed by comparative genomics analysis. The PhD student will partition pan-genomes into core and dispensable gene repertoires, with their respective metabolic routes, and attempt to link genomic variations with ecological data associated with each strain (genomic basis of ecotypes).

Task 3: Functional Genomics and transcriptome response to extreme conditions.

Gradients of temperature, pressure, pH, etc. will be applied to selected strains followed by sampling for RNA sequencing. The PhD student will analyze gene regulation linked with environmental changes by standard statistical methods (R package DeSeq), and identify metabolic response to extreme conditions. The PhD student will compare transcription profiles across multiple strains in order to identify key processes involved in genomic integrity, cell division, and energy conservation at the life limit conditions.
Duration: 1 year

Manuscript #3: “Functional genomics of the group X of extremophilic organisms”

Task 4: PhD manuscript redaction.

Duration: 6 month

Methods description

The review of available extremophilic microorganisms isolates will be carried with the help of major culture collection repositories or references (The LM2E culture collection, BacDive, ExtremeDB, LivingTreeProject). List of available genomes of extremophilic microorganisms will be searched in the NCBI genome repository and the Gold Genome Database.

Genome of selected microbial isolates will be sequenced using Illumina and PacBio sequencing platform. Due to the production of multi-Kbp reads, this later technology has proven to be an efficient way to generate fully assembled circular genomes for microbial organisms at minimal costs.

The PhD student will use automated gene calling and annotation integrated in standard pipelines (Mavromatis et al., 2009) available in the following portals: IMG at the Joint Genome Institute, RAST at Argone National Lab or MicroScope at GenoScope.

In vitro incubations will be used to test the effect of extreme environmental conditions (such as temperature, pressure, and pH) on gene transcription. This PhD project will take advantage of the unique expertise and material available at the LM2E regarding incubation under extreme conditions (high-pressure/temperature fermenters). This includes the recently acquired continuous high pressure (1000 bars), high temperature reactor. Microbial transcriptomes will be sequenced using the RNA-Seq technology on Illumina HiSeq platforms.

Throughout the project, the PhD student will have access to the following bioinformatics resources: public portals (IMG, RAST or MicroScope), the BioGenouest genomic platforms, including the Caparmor Supercomputer at IFREMER, ABIMS at Roscoff marine station and GenOuest in Rennes.

This project will continue and amplify genomic projects undertaken at the LM2E, (Vannier et al., 2011; Lucas et al., 2012; Jun et al., 2011), and will benefit from the high-throughput isolation system “COCAGNE” being developed in this lab.