PhD Project (2014-2017)

Genomic adaptations to life limit conditions: Comparative and functional genomics of 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 expression regulation under extreme conditions.

Project description

Microorganisms occupy a central role in ecosystem function. They efficiently recycle nutrients (*1*), confer enhanced ecological fitness to animal and plant hosts (*2*, *3*), and alter existing habitats to make them more hospitable to other life (*4*). 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 (*5*).

Furthermore, microorganisms exclusively occupy the ends of the physicochemical spectrum defining livable habitats, with microbial growth observed in extreme conditions of temperature (*6*), salinity (*7*, *8*) pressure (*9*), and energy stress (*7*, *10*).

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 analyses of multiple extremophile genomes** allow the identification of distinctive genes and metabolic pathways involved in the extremophilic way of life. Such analyses provide key data that enable comparison of functional genes 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 (*11*). 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.**

The PhD candidate will identify and select strains/isolates of special interest that cover specific taxonomic or metabolic gaps among extremophilic Archaea and Bacteria. Hyperthermophilic and piezophilic groups from deep-sea hydrothermal and subseafloor environements (*e.g. Thermosipho, Clostridum, Thermothoga, Thermus, Pyrococcus, Thermococcus*, *Archaeoglobus*…) are more numerous in our collection and will be considered in priority, however, the project remains open to other extremophile groups according PhD candidate interests and strain availability.

> *Duration: 8 month*
>
> *Manuscript #1: "A Phylogenetic and phylogenomic guide tree across extremophile strains", Mini review*

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 out the assembly and annotation of sequenced genomes, as well as comparative genomic analyses. 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).

> *Duration: 1 year*
>
> *Manuscript #2: "Comparative genomics of the group X of extremophilic organisms"*

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

Selected strains will be subjected to gradients of environmental stress (such as temperature, pressure, pH, etc.) and will be sampled for RNA sequencing. The PhD student will analyze gene regulation linked with environmental changes by standard statistical methods (R package DESeq)(*12*), 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: 4 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 (*13*) 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, (*14–16*), and will benefit from the high-throughput isolation system "COCAGNE" being developed in this lab.

Candidate profile

This project is at the interface of microbial ecology and bioinformatics. We encourage candidates with a MSc. from either backgrounds, and with reciprocal interests to apply for the position. Applications will be evaluated based on experience and records in microbiology, bioreactor operation, microbial ecology, molecular biology and ecology, and bioinformatics. Previous experience with high throughput sequencing datasets, familiarity with UNIX-like environments, and knowledge of at least one scripting language (BASH, Perl, Python, R…) will be appreciated, but not requested if the candidate demonstrates strong motivations to fill this gap.

Funding and supervision

This 3-years PhD project will be supervised by Dr. Loïs Maignien, Associate Professor of EcoGenomics at the IUEM (University Institute for Marine research), and co-supervised by Dr. A. Murat Eren, Assistant Research Scientist in Bioinformatics at the Marine Biological Laboratory in Woods Hole.

IUEM (http://www-iuem.univ-brest.fr/en) is a leading center for Marine Science innovation though its research, teaching and environmental monitoring missions. Its strong multidisciplinary approach spans from research in Human and Social Sciences to the Sciences of the Universe and Life Sciences, putting IUEM at the top of its field nationally. IUEM is situated in the Brest *Technopole* (Technology Centre), where it unites over half of the French scientific community working in the Marine Sciences.
The LM2E has a long-standing record of breakthrough studies on marine extreme ecosystems such as hydrothermal vents, methane cold seeps, and deep subseafloor biosphere.

The MBL's Bay Paul Center in Woods Hole has a pioneering role in using high-throughput sequencing technologies for microbial diversity analysis. See http://www.mbl.edu/jbpc/ for more information on the center, and http://meren.org/research/ for Dr. Eren's research.

The PhD candidate will be hosted primarily at the IUEM in Plouzané, but travel grants will be available to conduct research and education at the Marine Biological Laboratory. The PhD candidate will be enrolled in the doctoral school of marine sciences (http://edsm.univ-brest.fr/fr?set_language=fr). The PhD stipends will be of 1350€ / month for three years, starting between September and December 2014.

Contact

For more information about the project, please contact
Dr. Loïs Maignien

Lois.maignien@univ-brest.fr
Tel +33 290915380
and
Dr. Murat Eren
meren@mbl.edu

We will start reviewing applications immediately until June 3st 2014.

Applications should be submitted through the doctoral school web portal: http://edsm.univ-brest.fr/fr/une-these/les-dossiers-deposes/adaptations-genomiques-aux-conditions-limites-de-la-vie-genomique-comparative-et-fonctionnelle-des-archaea-et-bacteries-extremophiles

Important, only the applications uploaded to this server **before June 3, 2014** will be eligible. Contact Dr. L. Maignien for any further info.

Application material should include a CV, an application letter, a summary of MSc. grads, the name and email of three possible references, as well as any document that could further support your application.

Candidate Interview will take place from June 23 to June 30, either at the IUEM or through videoconferencing.

References

1. K. R. Arrigo, *Nature* **437**, 349–55 (2005).
2. F. Backhed, R. E. Ley, J. L. Sonnenburg, D. A. Peterson, J. I. Gordon, *Science (80-.).* **307**, 1915–1920 (2005).
3. L. Fan *et al.*, **109** (2012), doi:10.1073/pnas.1203287109/-/DCSupplemental.www.pnas.org/cgi/doi/10.1073/pnas.1203287109.
4. G. Lavik *et al.*, *Nature* **457**, 581–4 (2009).
5. N. Dubilier, C. Bergin, C. Lott, *Nat. Rev. Microbiol.* **6**, 725–40 (2008).
6. K. Kashefi, D. R. Lovley, *Science* **301**, 934 (2003).
7. L. Maignien *et al.*, *FEMS Microbiol. Ecol.* **83**, 214–231 (2013).
8. A. Boetius, S. Joye, *Science (80-.).* **324**, 1523–1525 (2009).
9. X. Zeng *et al.*, *ISME J.* **3**, 873–6 (2009).
10. D. L. Valentine, *Nat. Rev. Microbiol.* **5**, 316–323 (2007).
11. A. Tritt, J. a Eisen, M. T. Facciotti, A. E. Darling, *PLoS One* **7**, e42304 (2012).
12. S. Anders, W. Huber, *Genome Biol.* **11**, R106 (2010).
13. K. Mavromatis *et al.*, *Stand. Genomic Sci.* **1**, 63–7 (2009).
14. P. Vannier, V. T. Marteinsson, O. H. Fridjonsson, P. Oger, M. Jebbar, *J. Bacteriol.* **193**, 1481–2 (2011).
15. S. Lucas *et al.*, *J. Bacteriol.* **194**, 5974–5 (2012).
16. X. Jun *et al.*, *J. Bacteriol.* **193**, 4297–8 (2011).