Scientific goals

The project focuses on two aspects of algal genomics that are being studied in the participating labs.

1. Phylogenomics of red algae: we will use comparative analyses of organelle genomes to advance our knowledge about the timing and nature of the radiations of marine plant life during the last billion years of earth history and study the evolution of the genomes themselves.
2. Genome biology of economically and ecologically important algae: we will sequence and functionally characterize draft genomes of two algal species, one microalga that plays a critical role in the resilience of coral reefs to environmental change and one macroalga that is a important source of agar production.

Algae are critical for the structure and functioning of marine ecosystems and global climate cycles. They are responsible for nearly half of the photosynthesis that takes place on earth, producing the oxygen we breathe and taking up the carbon dioxide we produce. They play a central role in ecosystem functioning and, through the carbon pump and other mechanisms, have major impacts on climate trends. From an anthropocentric perspective, they provide compounds used as food additives and in pharmaceutical and microbiological applications, an industry estimated to be worth US$6 billion annually. Furthermore, algae are starting to be used as biofuels and they provide important ecosystem services in coastal areas.

Despite the importance of algae, many aspects of their evolutionary history as well as the genetic basis of the processes underlying their important functions are not well understood. First, thanks to their ancient age and long and diverse evolutionary history, algae serve as a testbed for a range of evolutionary questions. Red algae are particularly useful in this regard, with a history of more than >1.2 billion years with a significant fossil record, a wide range of functional traits and life cycle characteristics. This gives them many advantages to test hypotheses about the timing and nature of plant evolution prior to the colonization of land by embryophytes (land plants) 400 million years ago. A major obstacle that is currently preventing us to tackle these questions is that the evolutionary relationships among red algae have been difficult to resolve. Our laboratories at UniMelb and USP are using high-throughput sequencing techniques to assemble organelle genomes (chloroplasts and mitochondria) and to reconstruct evolutionary relationships using phylogenomic methods and infer the timeline of evolution with relaxed molecular clocks calibrated with the fossil record. Our data show that our phylogenomic approach is very effective at resolving the rapid radiations that characterize the evolution of the red algae and most other types of organisms. With this project, we aim to bring together the data that is being produced in both groups for joint analysis, in order to answer some long-standing evolutionary questions about marine plant evolution (Goal 1). We will also address the evolutionary dynamics of the organelle genomes, characterizing how genes are reshuffled and transferred to and from other cellular compartments.

In the context of Goal 2, we will develop draft nuclear genomes for two algae of particular ecological and economical interest. The green alga *Ostreobium quekettii* lives inside of the $CaCO_3$ skeletons of corals and plays a key role in the coral reef ecosystem. It facilitates coral survival by transferring a portion of its photosynthate to the coral during bleaching events. It is also the main agent of carbonate dissolution in coral skeletons and contributes greatly to reef bioerosion. As such, *Ostreobium* plays a major yet understudied role in coral reef health and is key to understanding how reefs will develop (or dissolve) under environmental change. Sequencing and characterizing a draft genome for this species is a first step towards understanding the genetic basis of its important functions and forms the foundation of future metagenomic and metatranscriptomic work in experimental settings where temperature and CO_2 levels are adjusted to mimic those expected under global warming and ocean acidification.

The Gracilariaceae algae are of great economic importance due to the various uses of hydrocolloids present in their cell walls. In Brazil, natural populations of Gracilariaceae are abundant, especially in the Northeast coast where these algae are harvested from natural banks and cultivated. The large number of species, broad distribution and the presence of cryptic species contribute to a lot of confusion in the taxonomy of the group. Molecular data have helped to elucidate the phylogenetic relationships of Gracilariaceae, but a more complete dataset is necessary to fully clarify the relationship among the genera. In view of the economic and ecological importance, the problematic taxonomy and the gap in knowlegde of genomic data for the group, and for red algae in general, we are using NGS to obtain the chloroplast and mitochondria genomes for different species of Gracilariaceae. A draft nuclear genome will be assembled for *Gracilaria caudata* an economically important species collected in Brazil.

The combined expertise of Oliveira's and Verbruggen's labs will enable the use of a more broad range of tools for the comparative analyses of those organellar and nuclear genomes. Students from both labs will be trained in different bioinformatic techniques, what is strategic for both teams. A mobility project would allow us to bring together the know-how of both labs and leverage the existing funding on both sides to increase research output and chances of winning future joint funding bids in Brazil and Australia.

Many kinds of evidence support the fact that we are addressing important problems. First, the goals are a good match with stakeholder priorities (listed below). Second, the type of research we propose is highly fundable as illustrated by our active grants on the subject (see under related projects section). Third, algal genomics tends to be high-impact research published in highly ranked journals. Our project addresses the research priorities of the participating institutions and their countries as follows:

- *Australian NRPs*: Within the National Research Priority "An Environmentally Sustainable Australia", our project addresses priority goals "Sustainable use of Australia's biodiversity" and "Responding to climate change and variability". In addition, bioinformatics was singled out as a key enabling capability and skills shortages in the discipline identified as a national vulnerability. This project will deliver national benefit by training postdocs and RHD students in bioinformatics. In doing so, we address the national priority of "Lifting Productivity and Economic Growth", specifically the goals to maximise Australia's competitive advantage in critical sectors, and deliver skills for the new economy.
- *University of Melbourne*: The project fits well with the UniMelb strategic statement in that it addresses relevant problems, promotes internationalisation, and enables research collaboration for mutual benefit. The research goals align particularly well with the goals of the Faculty of Science, as it addresses important questions spanning two of the Faculty's key themes, i.e. "Evolutionary Biology" and "Environment, Earth and Climate Sciences".
- *University of São Paulo*: USP aims to increase internationalisation, including scientific projects and exchange of researchers, including students. USP is also funding the Research Center for Marine Biodiversity (http://npbiomar.cebimar.usp.br/index.php/en/) to which M. Oliveira belongs.
- *FAPESP*: besides promoting internationalisation, FAPESP is committed to increase the quality and impact of the scientific publications produced by Brazilian scientists. Through this collaboration we will work on an innovative project, answering questions of global significance that will lead to publications of higher impact. In terms of research FAPESP has two major programs, Biota (http://www.biota.org.br/) and the Global Climate Change related to the aims of this proposal.
- Brazilian national priorities: as a mega-biodiverse country Brazil aims to record, study and conserve its biodiversity and is financing networks such as the BR-BOL (http://www.brbol.org/). PI Oliveira has an ongoing project financed under the Biota-FAPESP and just finished a project financed by CNPq BR-BOL.Bioinformatics have been considered an strategic field by the Brazilian agencies (such as CAPES and MCT), which have invested in training of students.

Planned activities

1. *Planning workshop* – Melbourne, towards start of the project
    - Present the work being done in both groups and other groups in São Paulo.
    - Narrow in on specific research topics to focus the initial collaboration on. Topics will cover the breadth of goals described above, but we will pinpoint the case studies most suitable to address those goals in a manner that is achievable within the timeframe of this project.
    - Determine which data need to be gathered to answer the specific research questions defined in point 2. Establish strategy to divide tasks between people and between labs and set milestones for completion of this work.
    - Set goals for the continuation of the collaboration beyond this project, and including other partners from São Paulo and UniMelb. Brainstorm about potential funding sources and outline the goals for joint project proposals based on shared interest and expertise.
    - Relevance: This workshop is important to steer the project towards the most opportune case studies, divide tasks, and set milestones.
2. *Production phase* – these actions take place throughout the duration of the project
    - Student exchanges (see team mobility worksheet for details):
    - UniMelb students visiting USP will be trained in genomics and molecular phylogenetics.
    - USP students visiting UniMelb will be trained in evolutionary and computational techniques.
    - Genomic data relevant to the project will be gathered on both sides.
    - PIs will write project proposals for submission to FAPESP and Australian nationally competitive funding bids.
    - Relevance: The production phase is critical because it will facilitate data generation and expansion of the skill sets of students participating in the exchange program.
3. *Assimilation workshop* – São Paulo, near end of the project
    - Present work that has been done on both sides.
    - Finalize data analyses, taking advantage of the shared knowledge of all participants.
    - Group discussion about how to optimize the publication output from the results in terms of papers and target journals. Joint writing of research papers.
    - Group brainstorming about ideas for review papers and "perspectives" papers based on knowledge gained.
    - Set goals for the continuation of the collaboration beyond this project, and including other partners from UniMelb. Brainstorm about potential funding sources and outline the goals for joint project proposals based on shared interest and expertise.
    - Relevance: The assimilation workshop is important as it will permit us to maximize the output from the project by bringing all participants together and using our combined gray matter to write high-impact research and review papers.