Overview:

Marine bacteria encompass both phototrophic and heterotrophic organisms. Sitting at the base of the food chain these microorganisms underpin the functioning of marine ecosystems, playing a critical role in the biogeochemical cycling of major elements (e.g. carbon, nitrogen, sulphur etc) as primary producers and recyclers, respectively (1). The major bacterial genera comprising phototrophic organisms are Prochlorococcus and Synechococcus, whilst the heterotrophic component commonly includes members of the SAR11 and Roseobacter lineages (2, 3). These bacterial genera comprise the most abundant organisms on planet Earth.

Genomic information is now available for members of each of these major marine bacterial genera. This has highlighted that many of these microbes have undergone genome streamlining with genomes ranging in size from 1.6-2.3 Mb, representing the smallest genomes from free-living organisms known (2, 3). However, even within these ‘minimal’-sized genomes there exists a plethora of genes with no known biological function, so-called ‘hypothetical’ genes. This situation is magnified with the recent explosion in next generation sequencing which has led to an extensive inventory of the global ocean metagenome through projects such as GOS (4) and TARA Oceans (5), but also highlighted the large number of novel protein families contained within these datasets with no known function (6).

Figure 1: Oceans occupy 70% of Earth’s surface encompassing a plethora of microbial life and the genomes contained within. However, many of the gene functions encoded within genomes remain unknown. Using information from large metatranscriptome datasets we will target genes with no known function but that are highly expressed in natural marine systems and interrogate their function using various ‘omic technologies in model marine microbes.

Methodology:

This project will utilise currently available metatranscriptomics datasets (including our own data from natural populations of marine Synechococcus) to identify highly expressed genes encoding hypothetical proteins i.e. with no known function, and that are differentially expressed depending on environmental conditions. Once determined we will identify orthologues in cultured, genetically amenable phototrophic and heterotrophic marine microbes of the genera...
Synechococcus and Roseobacter). This will allow the construction of mutants in these ‘hypothetical’ genes and subsequent physiological, biochemical and molecular characterisation of these mutants. The latter will include the use of various ‘omics’ technologies to compare gene/protein expression in mutants compared to wild type strains. Where appropriate, protein(s) will be over-expressed in E. coli to explore more detailed structural or enzymatic properties. Overall, this project will explore new avenues of gene function in organisms of key environmental relevance that will undoubtedly advance our understanding of microbes that play critical roles in major biogeochemical cycles.

**Training and skills:**

This project will provide excellent training in marine microbiology, bioinformatics, biochemistry and molecular biology using cutting edge ‘omic’ approaches, in the context of revealing the function of novel genes potentially involved in major marine biogeochemical cycles.

CENTA students are required to complete 45 days training throughout their PhD including a 10 day placement. In the first year, students will be trained as a single cohort on environmental science, research methods and core skills. Throughout the PhD, training will progress from core skills sets to master classes specific to CENTA research themes.

**Partners and collaboration (including CASE):** Prof Scanlan has >20 years expertise working with marine cyanobacteria from ecology to genomics, and especially defining nutrient acquisition mechanisms and the role of P in limiting marine production. The student will be assisted by the wide array of scientific expertise already available in the very active Scanlan research lab. In addition the project will greatly benefit from having Dr Yin Chen, a marine microbiologist with extensive molecular and biochemical expertise working with bacterial heterotrophs, and Dr Andrew Millard who has tremendous bioinformatics expertise and will provide specific expertise on interrogating large marine metagenomics and transcriptomics datasets, as co-supervisors on the project.

**Possible timeline:**

**Year 1:** Bioinformatic analysis of marine metatranscriptomic and metagenomic datasets to identify key differentially expressed ‘hypothetical’ genes in model marine phototrophs and heterotrophs.

**Years 2-3:** Construction of gene knockouts in specific hypothetical genes in model marine microbes. Subsequent molecular characterisation of mutants using high throughput transcriptomic/proteomic analyses and/or over-expression of novel proteins in E. coli and biochemical characterisation of purified proteins.

**Further reading:**


**Further details:**

APPLICANTS should possess a BSc or MSc in Microbiology/Biochemistry with a particular interest in molecular biology and bioinformatics techniques. Informal enquiries can be made to Prof Dave Scanlan (d.j.scanlan@warwick.ac.uk). Further details of research in each of the supervisors labs can be obtained from http://www2.warwick.ac.uk/fac/sci/lifesci/people/dscanlan/, http://www2.warwick.ac.uk/fac/med/research/tsm/microinfect/staff/millardlab and http://www2.warwick.ac.uk/fac/sci/lifesci/people/ychen/.