Metagenomic approaches are of great utility to understand the taxonomic and functional diversity of natural microbial communities, avoiding the issue of culture bias that can lead to a fragmented picture of this diversity.

One of the problems with the metagenomic paradigm, is the difficulty in assigning particular metabolic functions to specific microbial groups, because the information generated from shallow sequencing of complex microbial communities will only include fragments of individual populations thus complicating the reconstruction of resident microbial groups. This issue can be partially solved by filtration-based fractionation of the microbial communities combined with deep sequencing approaches. Using these practices, the reconstruction of near-complete genomes from the original populations is possible and enables the study of the metabolic potential, ecological role and can provide insight into the genomic heterogeneity that exists in natural microbial populations.

In the present study work, deep metagenomic sequencing of a low complex hypersaline environment (Lake Tyrell, Australia) was employed to generate over 500,000 Sanger and 200,000 454-Titanium reads,. Our goal is to examine the individual populations of this community at different levels of detail, first by looking at the overall phylogenetic and metabolic diversity within the community, to then reconstruct the genomes of the most abundant members and finally evaluate their genetic heterogeneity.

The first approach, a read-based analysis, allowed us to compare the species and metabolic potential of the community as a mixed dataset and to compare our study site with similar hypersaline environments. Using this strategy we identified the unique metabolic and genetic repertoire of the community, and also the identification of novel protein families associated with hypersaline microbial communities.

The second level of analysis allowed reconstruction of the genomes of individual populations members of the microbial community. Using precise filter-based size fractionation of the community samples, combined with innovative binning strategies of the de novo assembled scaffolds, we were able to reconstruct 12 near-complete population genomes. In addition to several genomes representing members of the class Halobacteria, we also recovered two genomes from a previously undescribed taxonomic class of halophilic Archaea. The analysis of the metabolic repertoires for individual populations, allow us to identify gene products that are shared or unique to specific microbial groups, suggesting a partitioning of ecological roles across multiple co-existing, co-evolving microbial groups.

Our final analysis investigated the genetic heterogeneity that exists in these microbial populations. The reconstructed genome assemblies are a composite of different individuals from the populations, allowing us the opportunity to deconstruct these composite assemblies to identify single nucleotide polymorphisms (SNPs) and gene presence/absence dynamics within individual populations. Our preliminary results indicate that some of the dominant members of the community have less genomic heterogeneity than others suggesting different life history strategies among co-existing species. This
approach further provides us with a snapshot of the diversity of the different members of the population, allowing exploration of the different mechanisms that could explain the variability and differences among species within the same environment.

Overall, the combination of read-based and assembly-driven metagenomics provides us with the opportunity to study microbial communities at different levels of detail, from a gene survey of the environment to the deconvolution of the genetic and functional complexity of a natural microbial community.