Single-cell sequencing for the discovery of microbial **dark matter** genes

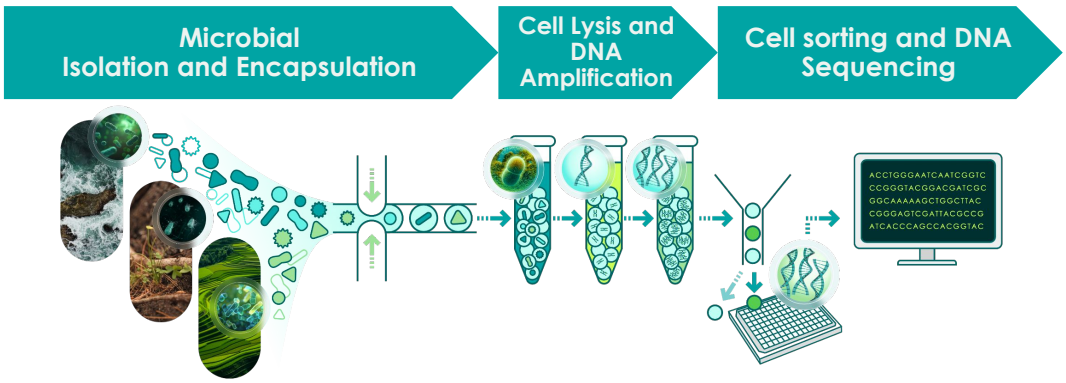
To date, less than 0.001% of all bacterial species have been identified. Technological advances such as 16S RNA sequencing and shotgun metagenomics have advanced our knowledge and discovery of a huge variety of bacterial species. However, these technologies have significant challenges when it comes to the identification of novel strains and species. Here we have developed the only microbial single cell sequencing platform and database for the discovery of not only new species, but also the ability to identify genetically distinct strains and their associated genes.

Key Points

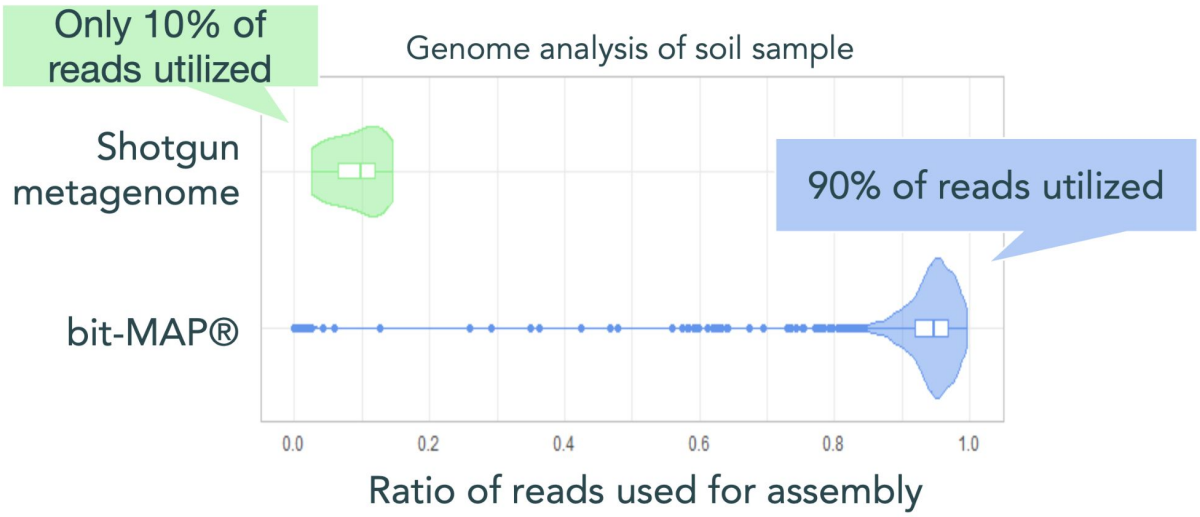
- To date less than 0.001% of all bacterial species have been identified
- Using current technologies and techniques, it would take over 10,000 years to discover them all
- With our single-cell sequencing platform, we can realistically sequence all unknown species within a single generation

bitBiome

Obtain more complete sequences using our bit-MAP® platform

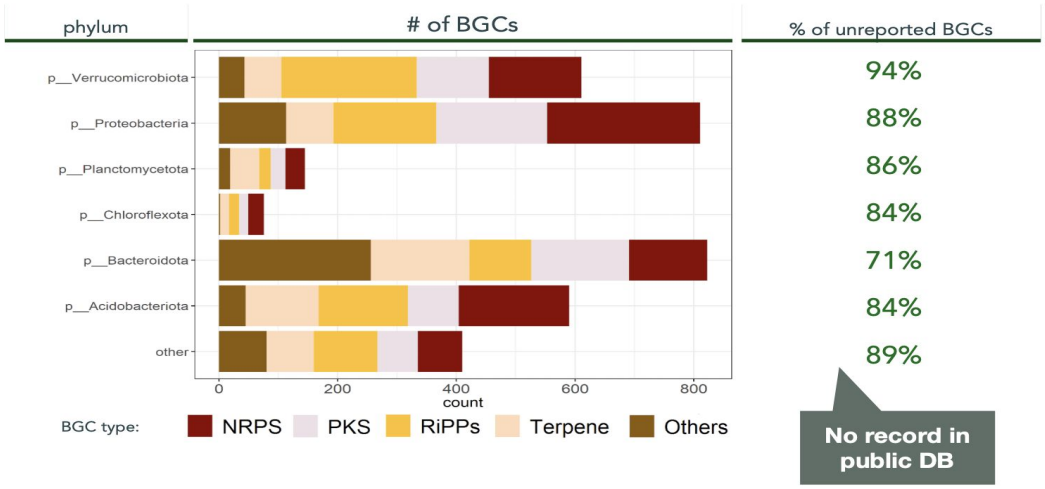


Here at bitBiome we have developed a first-in-class microfluidics based single-cell capture and sequencing platform, bitMAP®. Using this platform, we can isolate individual live cells and encapsulate them where cell lysis and DNA amplification occurs. In this way we have massively parallelized the collection, amplification, and sequencing of individual cells as shown above.



As a result of our unique platform, single-cell sequencing has the benefit of utilizing over 90% of the short reads generated for genome assembly as shown in the figure above. In comparison, metagenomic sequencing is plagued by non-overlapping short reads resulting in the majority of sequences being discarded. A consequence of this limitation is that metagenomic libraries are primarily comprised of short genes, typically 200 amino acids or less. However, many of our important proteins and enzymes are far larger. Therefore, metagenomic libraries are at a significant disadvantage when it comes to the identification of new species, strains, or gene products.

Identification of biosynthetic gene clusters



Of particular interest is the identification of biosynthetic gene clusters. These gene segments represent closely related metabolic pathways and products, often resulting in the identification of novel biologically active molecules. As shown above, due to our ability to utilize nearly all of the short reads generated, our genome assemblies contain large amount of novel sequences, representing genes not found in public databases. These previously unknown gene clusters represent microbial dark matter genes, and represent a huge catalogue of potentially biologically active molecules for use in a wide range of applications and industries.