


# MetaPlatanus improves metagenome assembly

October 11 2021

## MetaPlatanus: Bettering Metagenomic Sequencing Using a Hybrid Assembler

**Metagenome assembly is essential for generating draft genomes to 'directly' study environmental samples**



Current assembly strategies are limited in addressing heterogeneity in the analyzed metagenome

**MetaPlatanus**

Accurate short reads from direct samples

↓

Basic contig assembly

↓

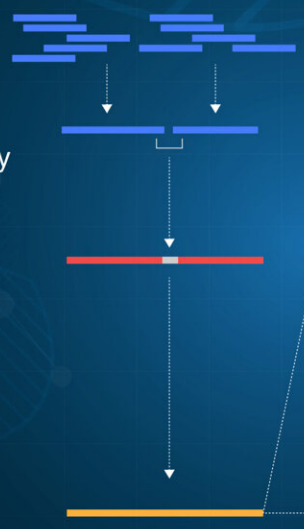
Iterative utilization

+

Binning information

↓

**Species-specific genome assembly**



↑ Contiguity

↓ Interspecific misassemblies


Assembly of intraspecific biologically important genomic elements

Near-complete assembly of high-abundance bacterial genomes

**High quality metagenomic assemblies, combining long-range sequence links and species-specific features, can be achieved using MetaPlatanus**

MetaPlatanus: A Metagenome Assembler that Combines Long-range Sequence Links and Species-specific Features

Kajitani et al. (2021) | *Nucleic Acids Research* | 10.1093/nar/gkab831



東京工業大学  
Tokyo Institute of Technology

Credit: Tokyo Tech

Metagenomics often involves sequencing DNA samples that can only be described as "tricky." Such DNA shows high heterogeneity, which can cause interspecies misassemblies. These misassemblies threaten the very purpose of metagenome sequencing, which is to comprehensively study the gene pool, by generating multiple draft genomes in a given sample. This issue is further complicated by the presence of certain organisms in these samples that cannot be cultured using standard microbiology techniques. How can this cascade of issues be resolved?

Scientists from Tokyo Tech now have the answer. They have developed a novel metagenome assembler called MetaPlatanus, which can generate accurate DNA sequences, including those of uncultured organisms. Their breakthrough findings have been published as a research article in *Nucleic Acids Research*.

MetaPlatanus uses accurate short DNA sequence reads to assemble contigs. Contigs are slightly longer stretches of DNA that are analogous to jigsaw puzzle pieces in the larger genome. The contigs are assembled into larger chromosome-scale scaffolds by repeatedly using inputs like long-range sequence links, species-specific sequence compositions, coverage depth, and binning information.

Explaining the choice of inputs for MetaPlatanus-based scaffold generation, Dr. Rei Kajitani, assistant professor at the School of Life Science and Technology, Tokyo Tech, and the lead scientist of the study, says, "We employ a hybrid metagenome assembly method that not only utilizes the advantages of both short-range and long-range sequence reads, but also compensates for the shortcomings posed by either read lengths, and the sample itself. We have applied binning to link sequences

divided by regions that are hard to assemble, such as repetitive ones. Our approach is novel since the combination of binning and assembly processes have not been implemented as a public tool so far."

Dr. Kajitani and his team left no stone unturned in checking the accuracy of results churned out by MetaPlatanus. In this regard, they performed a process called benchmarking using mock datasets of known bacteria. Not surprisingly, MetaPlatanus gave outputs that were highly contiguous, with very few interspecies misassemblies.

Notably, while testing MetaPlatanus' accuracy with already published human gut data, it additionally assembled many biologically important elements, including coding genes, gene clusters, viral sequences, and over-half bacterial genomes.

Also, compared to other existing tools, only MetaPlatanus was able to perform near-complete assembly of some high-abundance bacterial genomes, while benchmarking using already published human saliva data.

Indeed, Dr. Kajitani and his team appear to have struck metagenomic gold with MetaPlatanus. Excited about the potential applications of MetaPlatanus, he exclaims, "We believe that the metagenome assembler that we have developed at Tokyo Tech could help examine the contexts of sequence elements spreading over the [genome](#), which could have innumerable real-world applications."

**More information:** Rei Kajitani et al, MetaPlatanus: a metagenome assembler that combines long-range sequence links and species-specific features, *Nucleic Acids Research* (2021). [DOI: 10.1093/nar/gkab831](https://doi.org/10.1093/nar/gkab831)

Provided by Tokyo Institute of Technology

Citation: MetaPlatanus improves metagenome assembly (2021, October 11) retrieved 27 April 2024 from <https://phys.org/news/2021-10-metaplatanus-metagenome.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.