Oxford Nanopore Technologies

Genomic insights into the biology of complex microbiomes with Oxford Nanopore metagenomic sequencing

Oxford Nanopore reads of unrestricted length enable highly contiguous metagenomic assemblies and accurate annotation, delivering strain-resolved biological insights into complex real-world microbial communities

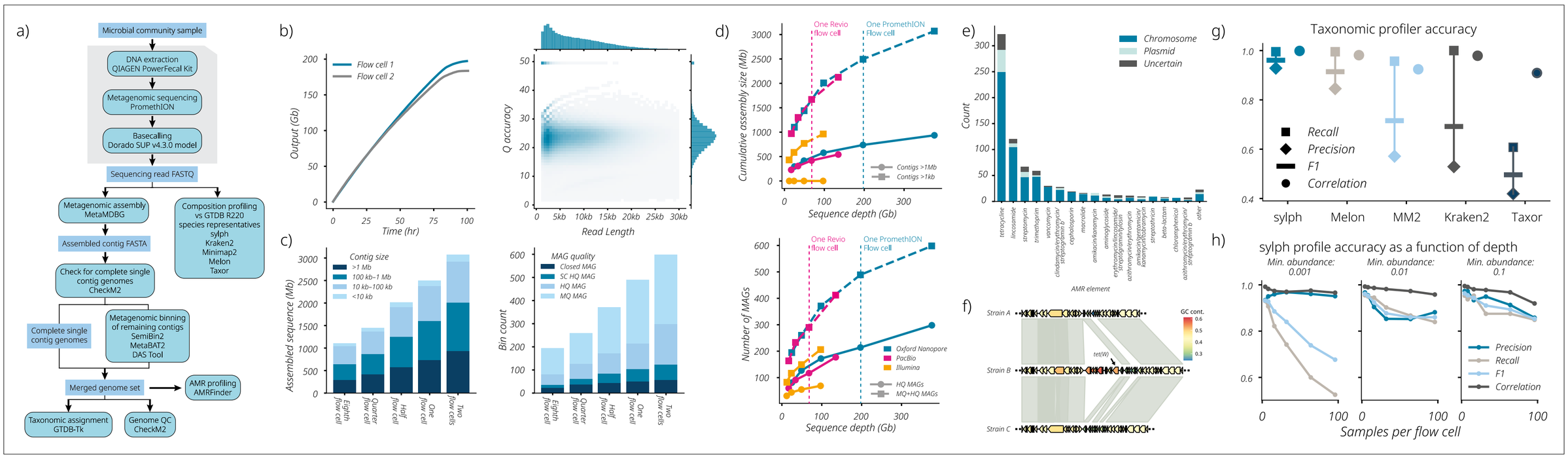
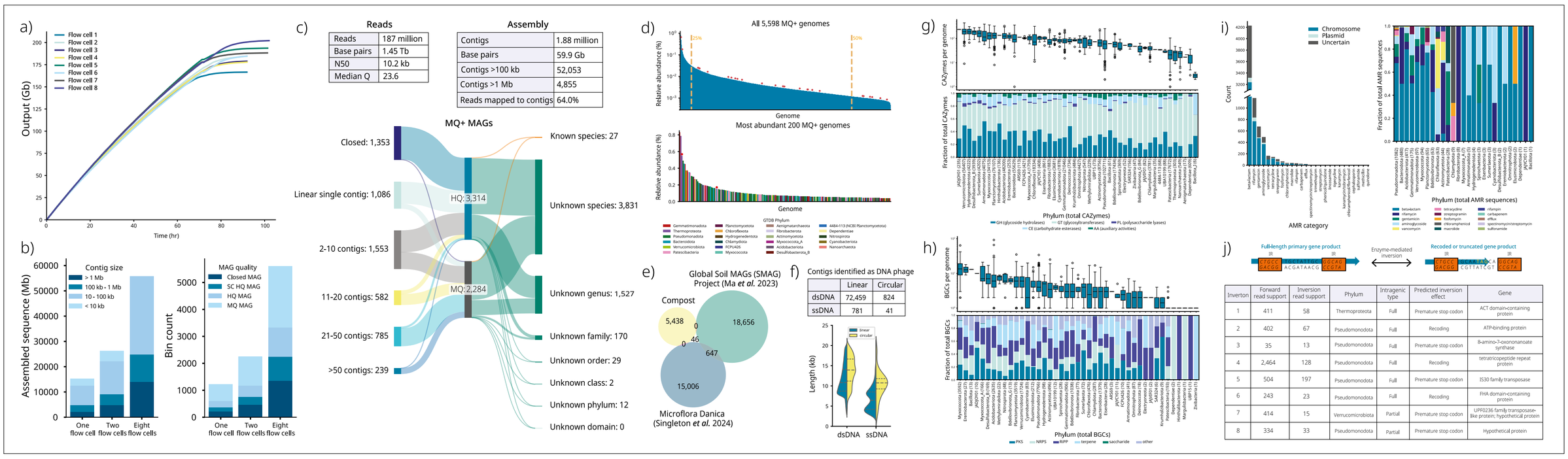


Fig. 1 a) Experimental workflow, b) sequencing output/accuracy, c) assembly/binning performance, d) platform comparison, e) AMR genes, f) strain alignment, g) profiler accuracy, and h) sylph accuracy by depth.

Accurate and complete *de novo* genome assembly and taxonomic profiling of complex metagenomic communities using Oxford Nanopore long reads with MetaMDBG and sylph

High-accuracy reads of unrestricted length from Oxford Nanopore Technologies unlock deep insights into metagenomic communities through *de novo* genome assembly and/or read-based profiling. To assess the performance of Oxford Nanopore sequencing for complex metagenomic samples, the ZymoBIOMICS Fecal Reference sample was sequenced and analysed using state-of-the art metagenomic tools (**Fig. 1a**). Sequencing yielded almost 200 Gb of high accuracy data per flow cell (**Fig. 1b**). Assembly with MetaMDBG resulted in hundreds of large contigs, and binning recovered hundreds of medium-or-greater quality metagenome-assembled genomes (MAGs), including 298 high-quality MAGs (HQ MAGs) (**Fig. 1c**). Multiplexing samples still recovers dozens of HQ MAGs (**Fig. 1c**). Comparison to other platforms demonstrated the superior performance for both assembly contiguity and MAG recovery on a per-gigabase (Illumina) or per-flow cell (PacBio HiFi) basis (**Fig. 1d**). MAGs can be surveyed for antimicrobial resistance (AMR) genes on genome contigs and plasmids (**Fig. 1e**), and strain-level variance in AMR can be identified (**Fig. 1f**). Analysis of taxonomic profiling tools showed high recall and abundance correlation with an assembly-based ground truth (for a subset of reads mapping to 204 MAGs with clear species assignment), but only sylph and Melon showed high precision in this analysis (**Fig. 1g**). For all but the lowest abundance species, sylph showed high accuracy even at low read depths / when sequencing many samples per flow cell (**Fig. 1h**).



Inverton	Forward read support	Inversion read support	Phylum	Intragenic type	Predicted inversion effect	Gene
1	411	58	Thermoproteota	Full	Premature stop codon	ACT domain-containing protein
2	402	67	Pseudomonodota	Full	Recoding	ATP-binding protein
3	35	13	Pseudomonodota	Full	Premature stop codon	8-amino-7-oxononanoate synthase
4	2,464	128	Pseudomonodota	Full	Recoding	tetratricopeptide repeat protein
5	504	197	Pseudomonodota	Full	Premature stop codon	IS30 family transposase
6	243	23	Pseudomonodota	Full	Recoding	FHA domain-containing protein
7	414	15	Verrucomicrobiota	Partial	Premature stop codon	UPF0236 family transposase-like protein; hypothetical protein
8	334	33	Pseudomonodota	Partial	Premature stop codon	Hypothetical protein

Fig. 2 a) Sequencing output, b) assembly and MAGs by depth, c) statistics for full dataset, d) species rank abundance curve, e) MAG overlap, f) phage contigs, g) CAZymes, h) BGCs, i) AMR genes, and j) invertons.

Highly contiguous assembly and high-quality genome recovery from a mature compost microbiome reveals high taxonomic diversity, functional potential, and intraspecific genomic variation

Metagenomic DNA from a mature compost sample was sequenced on eight PromethION™ Flow Cells (**Fig. 2a**). *De novo* assembly and binning of MAGs (following the bioinformatics workflow as depicted in **Fig. 1a**) show that assembled sequence and MAG recovery increases as a function of sequencing depth (**Fig. 2b**). Altogether, sequencing yielded 1.48 Tb of data which assembled to almost 60 Gb of contig sequence, with a 64% read mapping rate (**Fig. 1c**). Out of 5,598 MAGs of medium or high quality (MQ+), 2,439 were comprised of a single contig (**Fig. 2c**). These MQ+ MAGs are highly divergent at the species level from reference sequences, shown by the fact that only 27 were taxonomically assigned at the species level by GTDB-Tk (GTDB version r220; **Fig. 2c** and **Fig. 2d**, red dots) and only 46 species were shared with two recent large-scale soil MAG catalogues (**Fig. 2e**). Over 70,000 contigs were identified as DNA phage by VirSorter2, including over 800 circular contigs (**Fig. 2f**). The set of bacterial MQ+ MAGs encode a large number of carbohydrate active enzymes (CAZymes, identified with dbCAN3; **Fig. 2g**) and biosynthetic gene clusters (BGCs, identified with antiSMASH) (**Fig. 2h**). AMR sequences were identified and their locations mapped to chromosome vs. plasmid (**Fig. 2i**). Long reads allowed us to identify invertons, loci flanked by inverted repeats that can be inverted to generate gene variation within a clonal population (**Fig. 2j**). Among HQ MAGs at >0.1% relative abundance, the PhaVa tool found eight intragenic invertons, including one within an archaeon.

PO_1242(EN)_V3_19May2026