To long-read or not to long-read?

ONT long-read technology outperforms Illumina short-read sequencing for plasmid de novo assembly and AMR gene detection

Ryan Cardenas[1], Helen McNeil[1], Andrew Holmes[1], Nick Loman[1,2]
[1]MicrobesNG, [2]University of Birmingham

Background

There are two main approaches used for bacterial whole genome sequencing: long-read (such as **Oxford Nanopore Technologies- ONT**) and short-read (such as **Illumina**).

Long-reads have been key for the assembly of complete bacterial genomes by resolving highly repetitive and structurally complex regions of the genome, where Illumina can struggle.

Illumina has historically been the preferred choice for bacterial genome sequencing, due to its higher accuracy (**up to 99.9%**) compared with **ONT**[1].

However, the recent introduction of the **ONT super accurate (SUP) basecalling model** on the R10.4.1 flow cell has resulted in improved accuracy enabling assemblies using **only ONT data**[1].

- Plasmids, which are mobile genetic elements carried by bacteria, are major drivers of the spread of antimicrobial resistance (AMR) genes.
- **Illumina**-derived assemblies frequently fail to recover plasmids in full due to the presence of repetitive sequences resulting in fragmented assemblies.
- **ONT**'s improvement in accuracy, combined with longer reads, has demonstrated **100% AMR allelic recall**[2], producing longer and more complete assemblies.

Here, we compare the effectiveness of ONT sequencing in resolving complete plasmids and identifying AMR genes to Illumina.

Check out the research!

Methods

185 samples

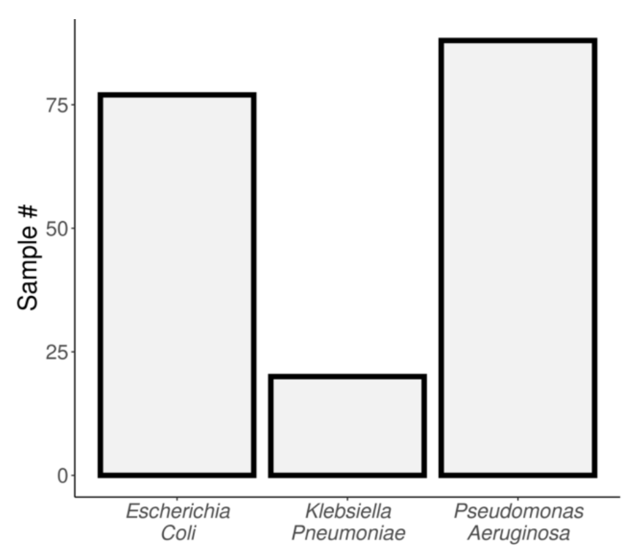

Figure 1 - Bacterial taxa used for comparative plasmid and AMR analysis.

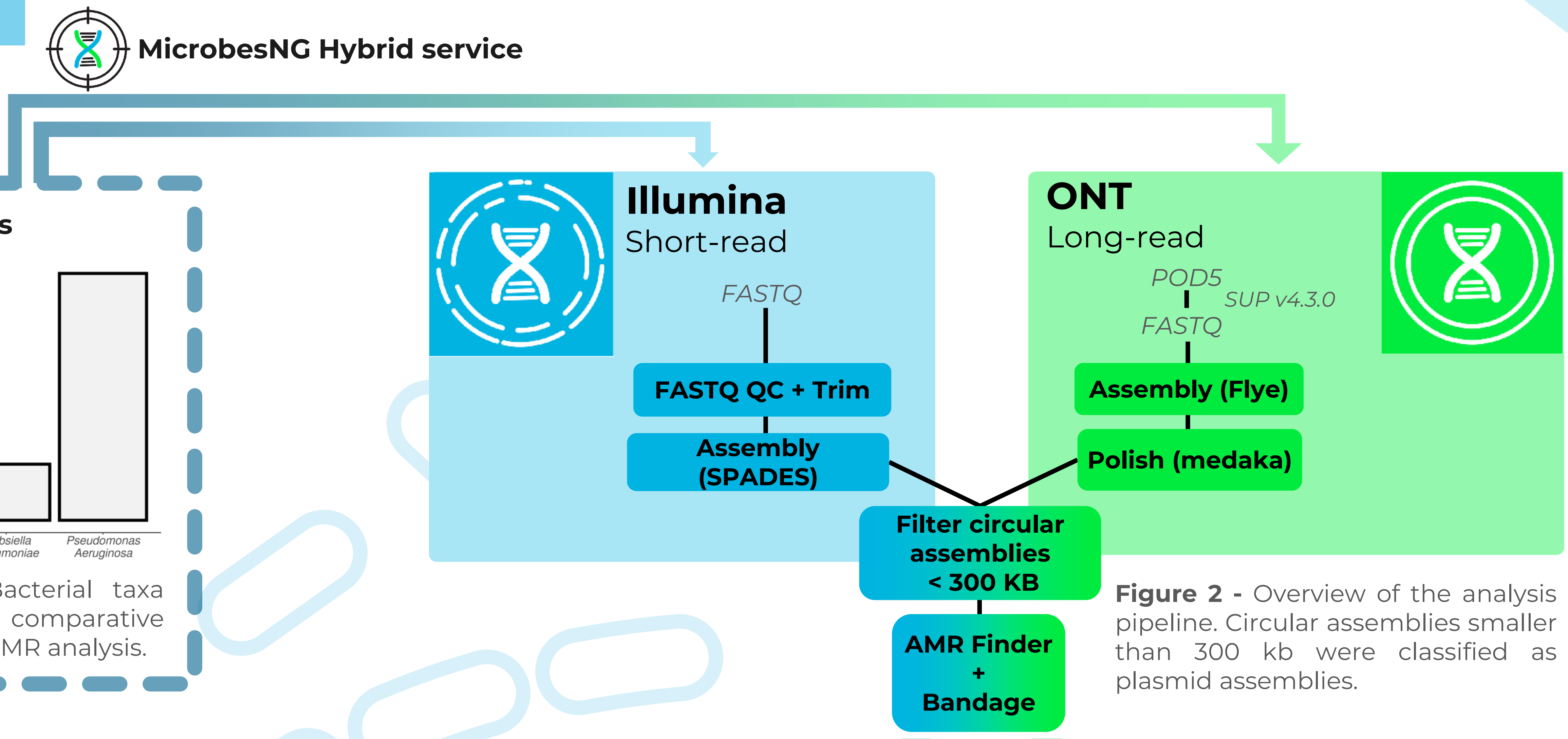

Figure 2 - Overview of the analysis pipeline. Circular assemblies smaller than 300 kb were classified as plasmid assemblies.

Results

We compared assemblies from the same samples, using either **ONT** or **Illumina** sequencing across three different taxa (Figure 1). This allowed the direct comparison of ONT and Illumina technologies for performance in plasmid assembly and AMR gene detection.

There were **more plasmids** assembled for Illumina samples (n=265) compared to ONT (n=162) (Fig 4C.) However, the overall size of ONT plasmid assemblies was **much larger** (Fig 4A and 4C) and **more complete** (Fig 3).

Running AMRfinder on the plasmid assemblies shows a much **higher detection of AMR genes** for ONT assemblies (n=220) compared to Illumina (n=14) (Fig 4B).

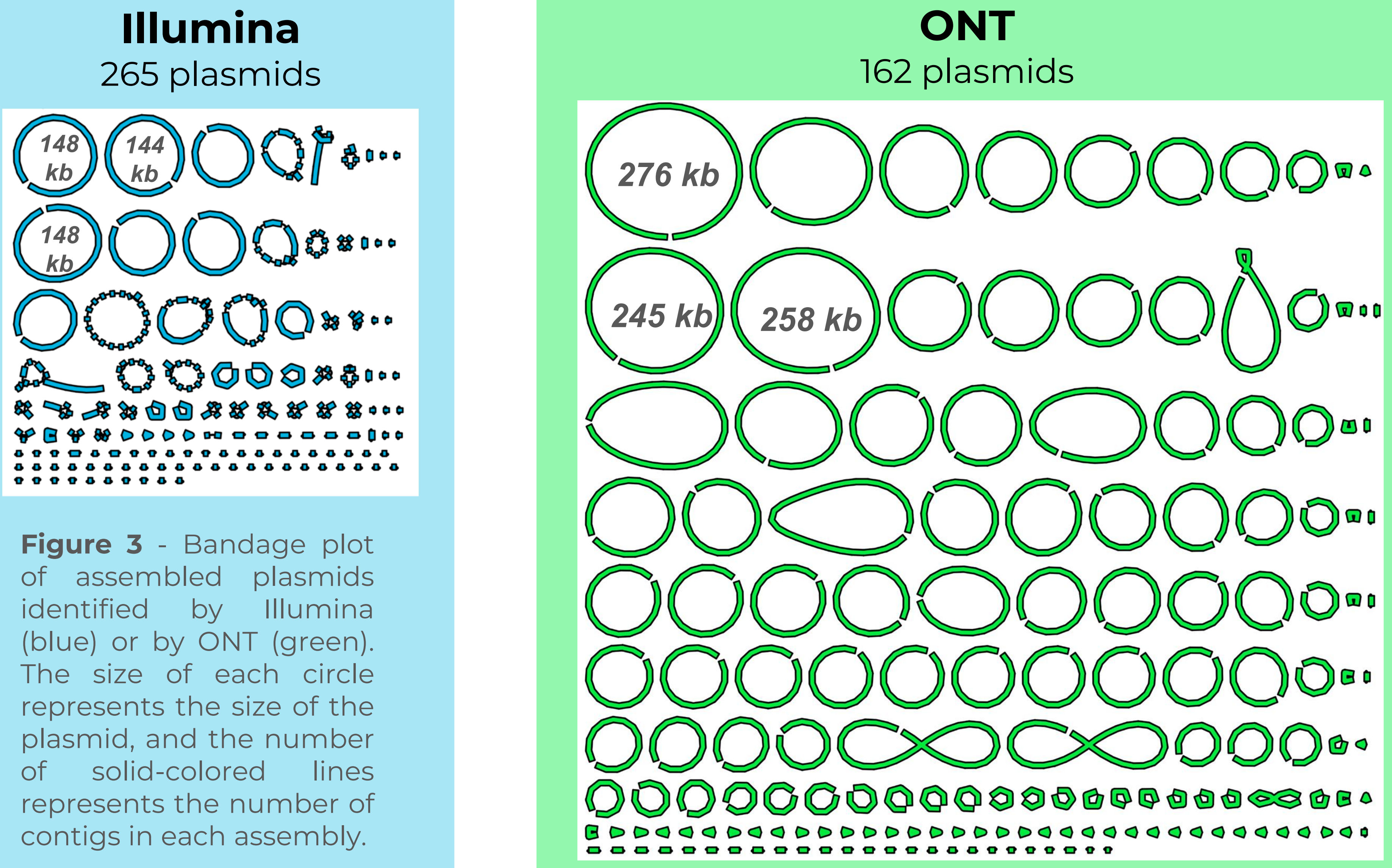

Figure 3 - Bandage plot of assembled plasmids identified by Illumina (blue) or by ONT (green). The size of each circle represents the size of the plasmid, and the number of solid-colored lines represents the number of contigs in each assembly.

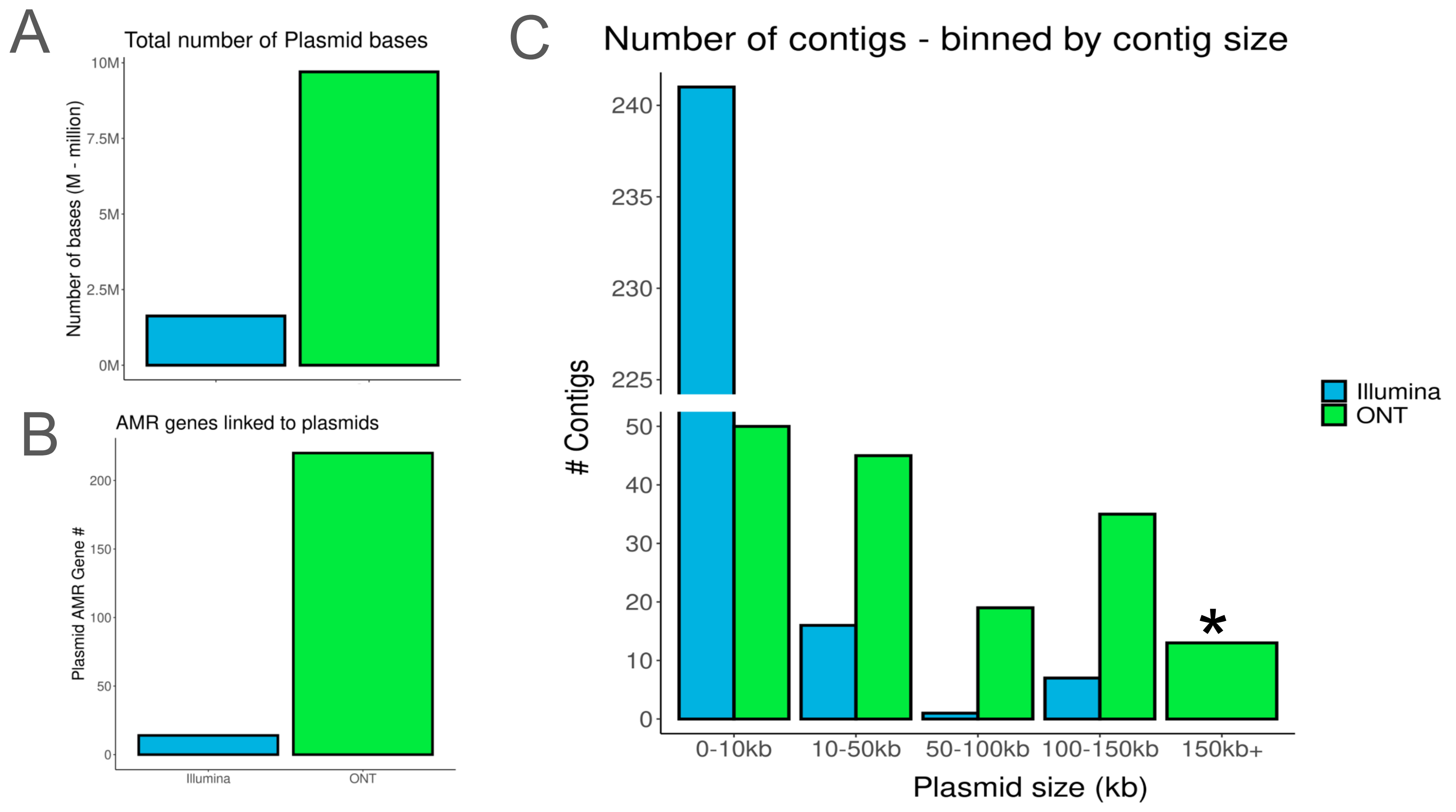

Figure 4 - (A) Total number of bases for assembled plasmids. (B) Total number of AMR genes identified by AMRFinder in assembled plasmids. (C) Number of contigs in different size bins; the largest bin (150 kb) is marked with an asterisk to indicate that only ONT contigs reach this size.

Conclusion

- Understanding the mobilisation of AMR genes in bacterial populations through plasmids is **key** for public health.
- Here we demonstrate that **ONT** resolves a **higher number** of plasmids that are **more complete** compared to paired Illumina samples. As a result, we were able to capture **more plasmid-linked AMR genes.**
- Considering this, MicrobesNG recommend that customers **use our long-read ONT service** for capturing plasmid-linked AMR genes.

Check out our long-read service!