Extracting closed genomes from comammox enrichment cultures using Nanopore sequencing

Laura N. Wenzel, Jeroen Frank, Maartje A.H.J. van Kessel, Mike S.M. Jetten, Sebastian Lücker

Department of Microbiology, IWWR, Radboud University, Heyendaalseweg 135, 6525 AJ Nijmegen, the Netherlands.

BACKGROUND

Nitrification is a step within the nitrogen cycle (Figure 1) in which ammonium is oxidized over nitrite to nitrate. Until recently this process was considered to be a two-step reaction performed by two distinct groups of organisms, ammonia oxidizing microorganisms and nitrite-oxidizing bacteria. The recently discovered group of complete ammonium oxidizers (comammox) are capable of performing complete nitrification on their own [1,2]. The lack of comammox pure cultures and the resulting lack of closed, high quality genomes is hindering the study of comammox bacteria. Combining the advantages of long read Nanopore sequencing and high accuracy Illumina read data, this study aims on obtaining closed high quality comammox genomes from enrichment cultures.

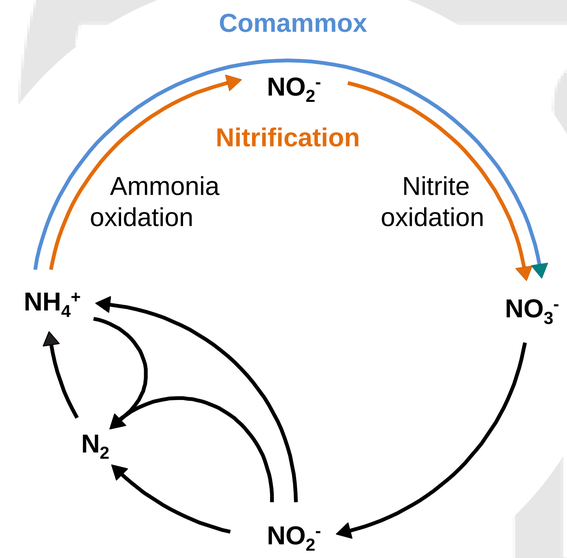

Figure 1: Illustration of the nitrogen cycle including the complete nitrification performed by comammox bacteria.

METAGENOME ASSEMBLY

While both the OLC-pipeline (Fig. 2a) and the Scaffolding-pipeline (Fig. 2b) produced closed contigs of the *N. nitrosa* genome, the scaffolding-pipeline resulted in lower error rates and a higher amount of identified single copy marker genes. In comparison to the *N. nitrosa* reference genome, rearrangements were found in the generated contig, likely originating from misassembles of the original Illumina-based reference genome (Fig. 4). Furthermore we *de novo* sequenced another enrichment culture, resulting in the assembly of the closed genome of a novel comammox *Nitrospira* species without high similarity to previously published genomes.

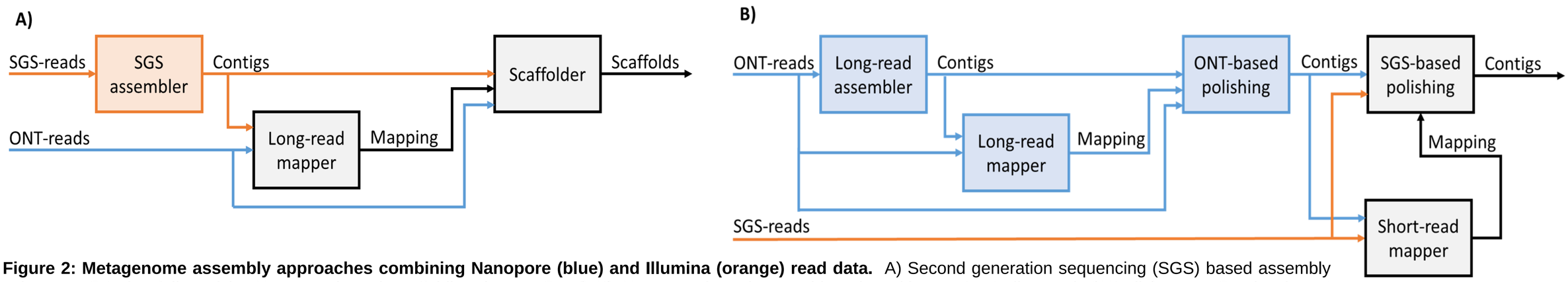

Figure 2: Metagenome assembly approaches combining Nanopore (blue) and Illumina (orange) read data. A) Second generation sequencing (SGS) based assembly using e.g. SPAdes followed by Nanopore-based scaffolding (e.g. npScarf) B) Nanopore based assembly using either a De-Bruijn graph (ABruijn) or an Overlap layout consensus based assembler (e.g. Canu) followed by optional long read polishing (e.g. Racon) and subsequent SGS-based polishing (e.g. Pilon).

COMPOSITION OF THE ENRICHMENT CULTURE

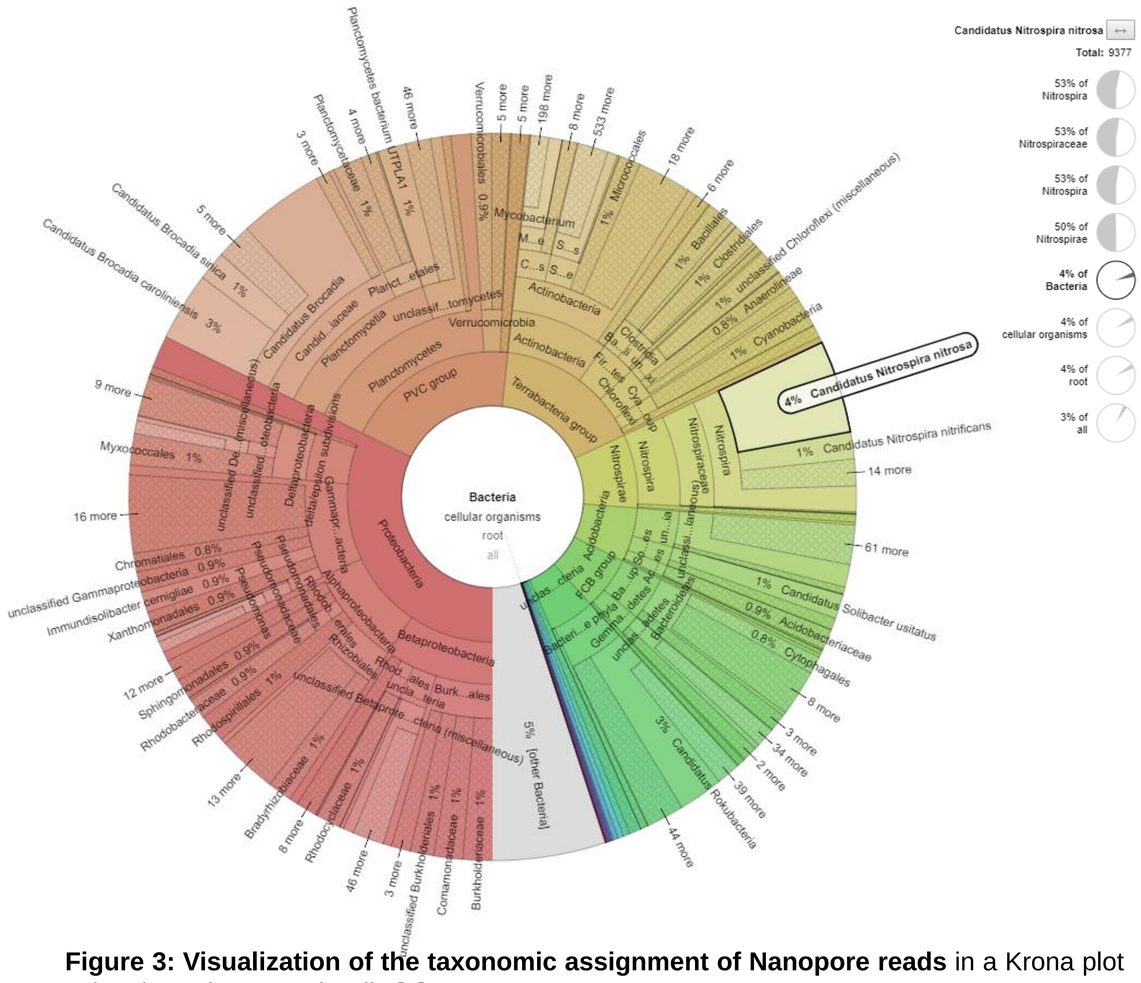

Figure 3: Visualization of the taxonomic assignment of Nanopore reads in a Krona plot using the software tool Kaiju [3]**.**

REARRANGEMENTS IN CLOSED CONTIG

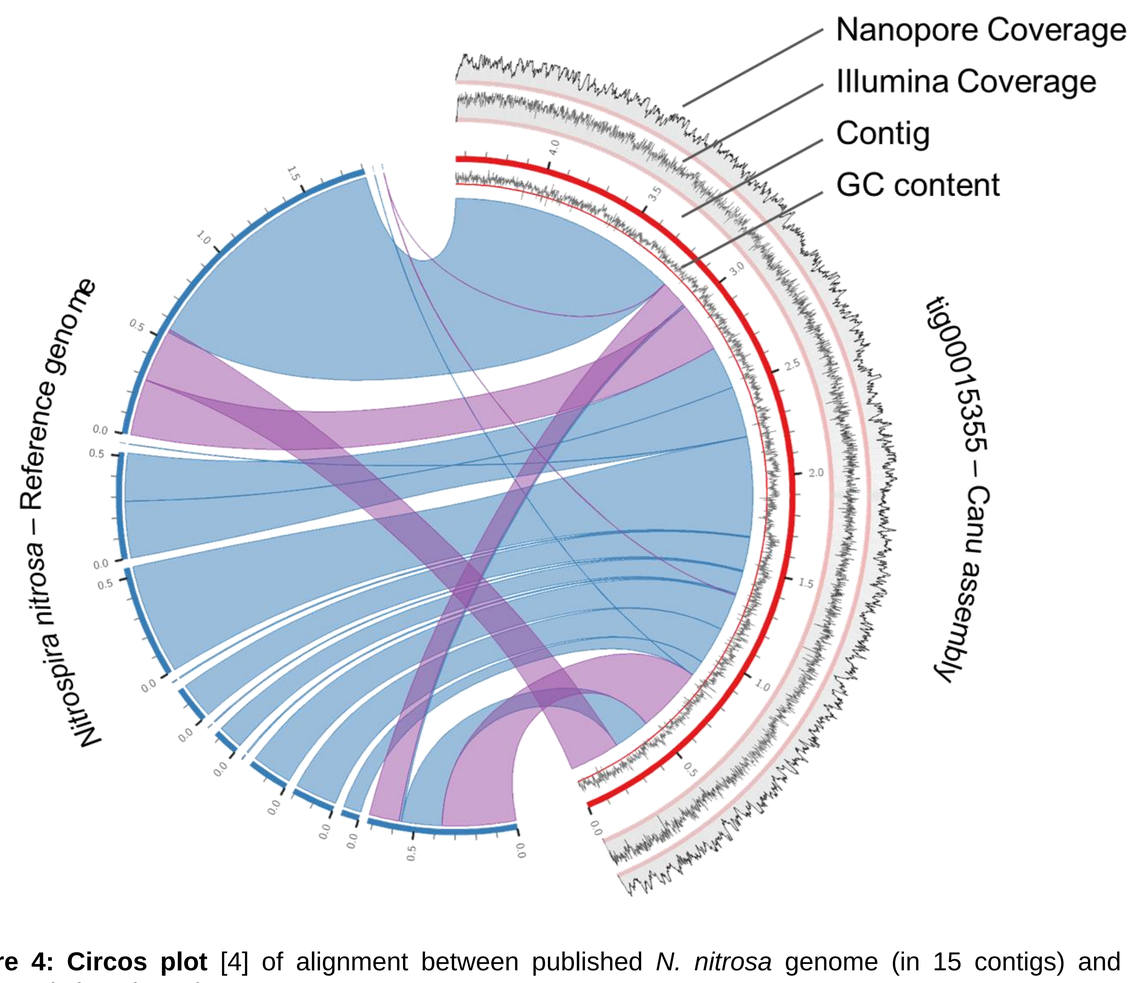

Figure 4: Circos plot [4] of alignment between published *N. nitrosa* genome (in 15 contigs) and generated closed contig.

CONCLUSION

The tested hybrid approaches produces more continuous metagenome assemblies than possible before. We were able to obtain closed genomes of the two most dominant *Nitrospira* species from two enrichment cultures. Both genomes contained all necessary enzymes for complete nitrification. The combination of long but error prone Nanopore reads with short Illumina reads with high accuracy proved to be a good method to obtain high quality assemblies.

REFERENCES

[1] van Kessel *et al.* (2015). *Nature* **528**: 555–559
[2] Daims *et al.* (2015). *Nature* **528**: 504–509
[3] Menzel *et al*. (2016). Nature Communications **7**: 11257
[4] Krzywinski *et al*. (2009). Genome research **19**: 1639-45