Nanopore sequencing for reliable outbreak tracing

Mara Lohde[1], Adrian Viehweger[2], Mike Marquet[1], Claudia Stein[1], Oliwia Makarewicz[1], Mathias W. Pletz[1] and Christian Brandt[1]

[1]Institute for Infectious Diseases and Infection Control, Univerisity Hospital Jena, Germany and [2]Institute of Medical Microbiology and Virology, University Hospital Leipzig, Germany

Aim

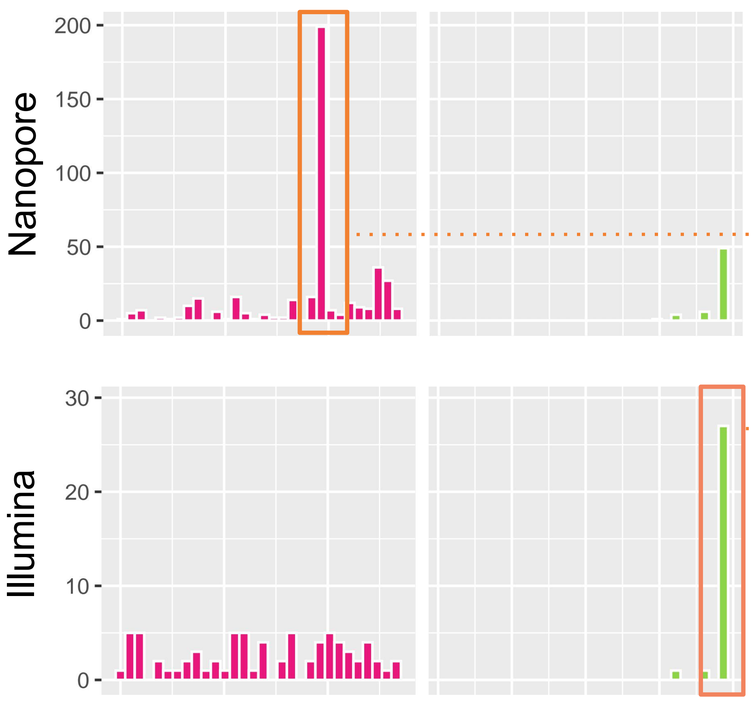
We explored the capabilities of Oxford Nanopore Technology to trace outbreaks in hospitals as an alternative to short-read sequencing (Illumina). The **three-year-long outbreak** caused by *Klebsiella pneumoniae* at the University Hospital Leipzig served as a basis for our study. We sequenced **34 random isolates** among 114 samples from the outbreak using the SQK-NBD112.24 Kit and compared our findings with existing data from short-read sequencing.

Phylogenetic timetree of 114 *Klebsiella pneumoniae* across three years.

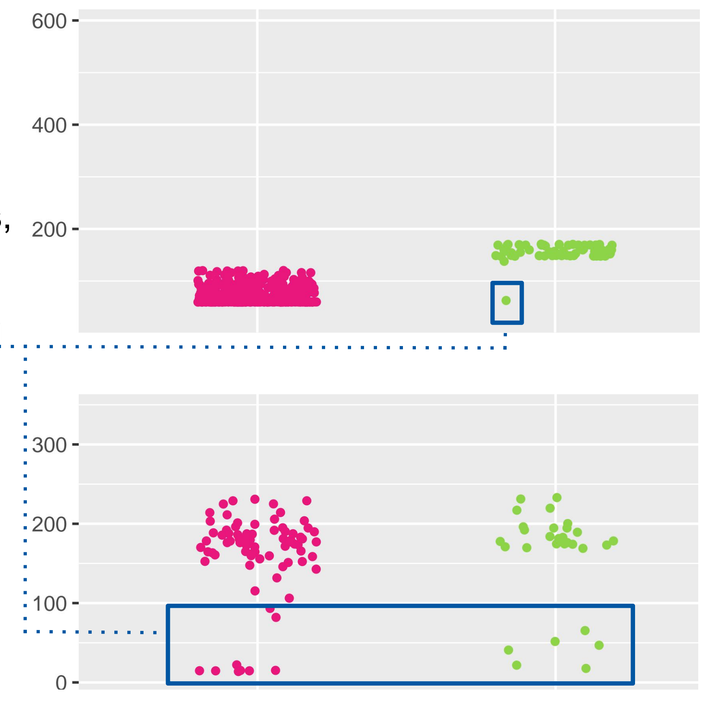
Conclusion

The study demonstrates the potential of Oxford Nanopore Technology as an alternative to short-read sequencing in outbreak tracing. However, some limitations must be noted. The lack of proper core genome alignment tools complicated reliable analysis across multiple samples. The default mapping parameters that Medaka is trained on lead to incorrect mapping of extrachromosomal DNA and introduce false variant calls. Moreover, combining sequencing technologies exposes their respective errors; despite these limitations, Nanopore provides a suitable solution to track outbreaks.

Variant calling - unexpected accumulation and low depth regions calls

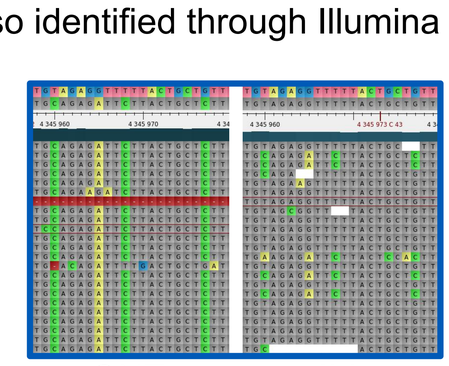


Variant calling was performed with Snippy (v4.6.0) for short-reads and Medaka (v1.7.2) with adjusted quality filters (depth ≥ 30; call quality ≥ 20) for long-reads. We identified unexpected **accumulations of variations** in certain regions, including SNPs, insertions, deletions, and complex variants in both **high and low-coverage regions**. This phenomenon was observed on **chromosome** and **plasmid contigs** and is exemplified with one example but applies to all 34 samples studied. **Both issues highlight mistakes in the variant calls leading to false positives. Nanopore detects more variations than Illumina, which is attributed to reconstruction errors (see below).**

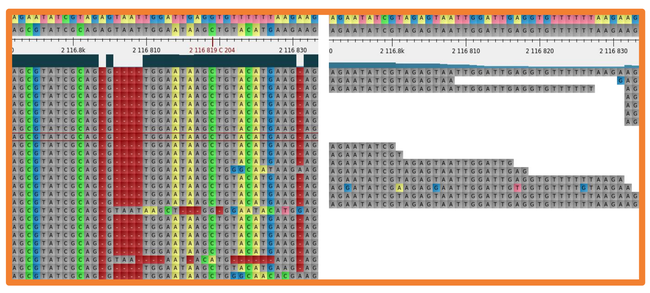


Mapping and reconstruction issues

We identified **false mapping of short-reads in non-coding sites** and **poor reconstruction of short-read assemblies** for Illumina (ILL). This led to false positive structural variation calling for Nanopore (ONT). These false mappings were identified using Snippy, and Medaka for whole genome variant calling; errors were identified in **non-coding sites close to tRNA regions** within Illumina assemblies (left figure) where Nanopore provided clear calls. Illumina genome reconstructions (right figure) led to low-coverage regions that Nanopore interpreted as deletions. Poorly constructed areas were also identified through Illumina read mapping.

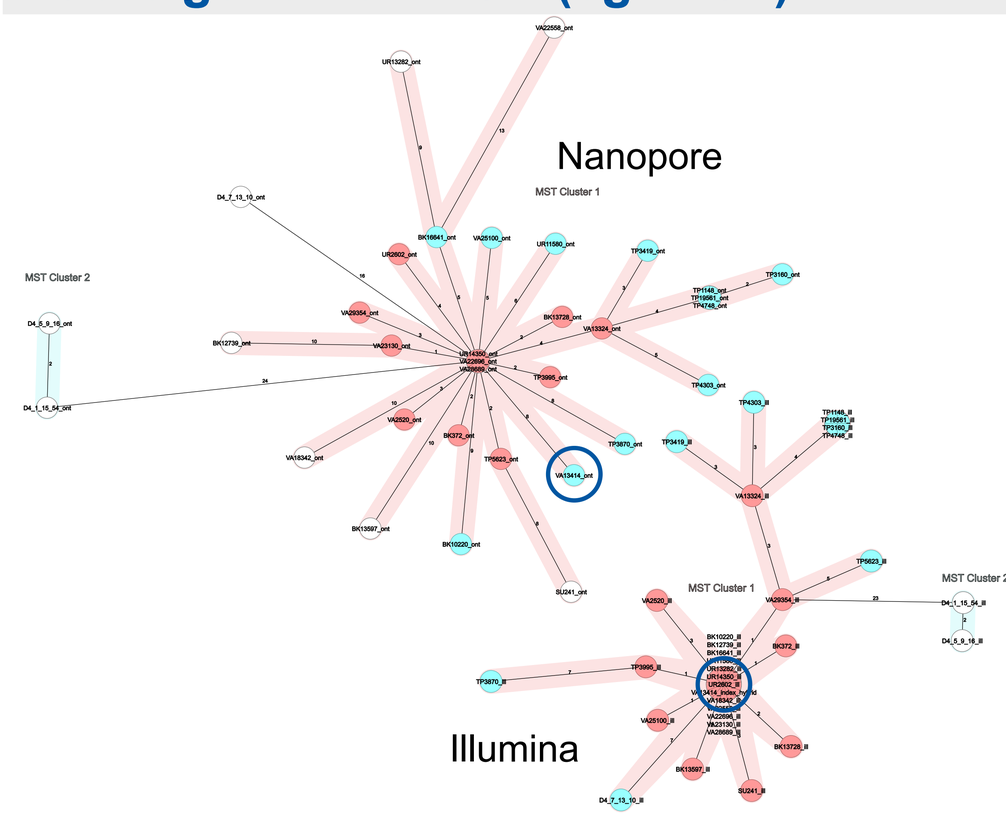


ONT ILL



ONT ILL

Core genome MLST (cgMLST)

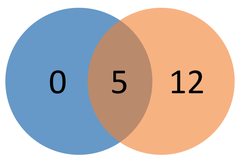


cgMLST was used to analyze all isolates using Nanopore and Illumina sequencing data with corresponding **index samples**. Both technologies showed similar patterns, although slight differences were found in the distances within Nanopore data. Notably, both approaches revealed inconsistencies between the minimum spanning and phylogenetic tree regarding temporal coincidence among isolates (data not shown).

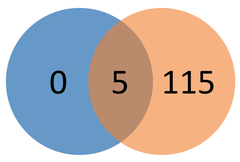
Poor mapping and low quality of short-read reconstruction affected the distance between outbreak isolates, leading to wrong conclusions.

Shared variations inside and outside of coding-sites

inside CDS



outside of CDS



Chromosomal reads were mapped against the illumina hybrid assembled genome of the index patient (“patient zero”). Variations of the chromosome inside and outside of coding sites (CDS) were counted and compared for **Illumina** and **Nanopore** technologies.

Nanopore reveals more SNPs in non-coding sites than Illumina, attributed to the Illumina reconstruction errors highlighted above.

@maralohde