Slide #2

The first sequencing method was developed by Sanger. We broke DNA up randomly into numerous small segments, using the chain termination method to obtain reads, that overlap. The principle is as follows: in each of the 4 tubes the polymerase stops chance, since “bad” nucleotide is incorporated. As a result, we have a lot of DNA fragments with different lengths, corresponding the position of each nucleotide. So, the sequencing is done by synthesis. Sanger method is very expensive and long, so it isn’t widely used nowadays.

Slide #3

In principle, it is the same, but it is much more effective in terms of speed and cost. It is massively parallel, sequencing millions of fragments simultaneously per turn. The only problem is that we have short read lengths. Also, Illumina sequenator is a huge machine, which are not present in every laboratory.

Slide #4

We have the membrane and the pore in it. We apply some voltage across the membrane inducing an electric field that drives charged particles, in this case our DNA, into motion, the principle is the same as in gel electrophoresis. DNA passes through the membrane, and we measure the current density of the pair, which depends on which base is in the nanopore. Compared to sequencing by synthesis, we get low accuracy, but long read lengths. This made it possible for DNA sequencing to be carried out almost anywhere, even in remote areas with limited resources.

Slide #5

In this project, we already had Illumina reads of Staphylococcus aureus strains from the hospital outbreak in Tallinn. It was already sequenced with Illumina, but to assemble the genome we sequenced the same genome with Nanopore, because, as you remember, Illumina reads are quite short. We performed the hybrid assembly of this data and and then annotated mobile elements in this genome (such elements - for example, plasmids, which can provide resistance of bacteria to antibiotics and other features).

Slide #6

The whole sequencing process takes place in a small palm-sized box! There is an electrical device and a flow cell, which is used for several tens of hours and can be replaced. So we just connect it to the computer, special software checks is all ok, then we prepare all the needed mixes: containing DNA library and primig mix. It feels like a chemistry set for kids, but is it real science :) When everything is ready, we leave this device connected to the computer, for example, overnight, and our DNA is sequenced.

Slide #7

To decrypt the obtained data, we need to translate a sequence of electrical signals into a sequence of nucleotides: this process is called basecalling. We used Guppy software for this purpose. Guppy is a neural network based basecaller that in addition to base calling also performs filtering of low quality reads and clipping of Oxford Nanopore adapters, so it was suitable for our project.

Slide #8

To combine the Illumina and the Nanopore data we used a program called SPAdes.

Slide #10

To annotate mobile elements we used the MOB-suite. We found the plasmid containing the genes with the following products: RepL protein, which helps the plasmid to replicate rRNA methyltransferase ErmC methylase leader peptide ErmCL, which is needed for the ErmC induction From the literature we found that the ErmC is resistant to macrolide antibiotics. So we concluded that this plasmid provides resistance to macrolides of the strain that we have sequenced