Nanopore-based Amplicon Sequencing on Rapid Detection of Pathogens and AMR genes from Low-biomass Samples in the Hospital Environment

Qing Yang (1,2), Josh Quick (1), Nicholas J. Loman (1), Yu Xia (2)

(1)Institute of Microbiology and Infection, University of Birmingham, UK (2)Southern University of Science and Technology (SUSTech), China

Introduction

1.*Legionella pneumophila* and *Pseudomonas aeruginosa* are the most common bacterium associated with ventilator-associated pneumonia and nosocomial infections. .

2. Investigation of pathogens by genome sequencing has been established as an important tool for outbreak management. Genome sequencing directly from samples (i.e. without isolation and culture) with low biomass remains challenging due to low numbers of genome copies in typical water samples.

3. The portable MinION sequencer (Oxford Nanopore Technologies) has potential for rapid genome sequencing and offline bioinformatic analysis, making it suitable for on-site and real-time detection and lineage assignment and resistance identification of *L. pneumophila* and *P. aeruginosa* directly from environmental samples.

Aims

To improve a multiplex PCR protocol developed before[2] in order to:

1. sensitively and specifically detect the pathogenic *L. pneumophila* from samples containing low number of genome copies.

2. reveal the variations and phylogenetic relationship of the species, therefore providing the evidence for the detection of legionellosis outbreaks.

Methods

Workflow

Sample Acquisition & Sample Filtration
(Lenticle Disks and Water Sample)
↓
DNA Extraction
(QIAamp DNA Mini Kit,
Dneasy Powerwater Kit with 0.22μm PVDF membrane)
↓
Real-Time PCR
↓
Multiplex PCR and Electrophorosis
(Touch Down Program)
↓
Amplicon Cleanup and Quantification
↓
Library Preparation
↓
MinION Sequencing
↓
Bioinformatics

Touch Down PCR Program

L. Pneumophila: 464 primers in 2 pools, for 53 genes

Stage	Cycle number	Denature	Anneal/Extend
1	1	98°C , 30s	
2	10	98°C, 15s	65°C, 5min
3	25	98°C, 15s	65°C, 5min, delta t = - 0.1°C

P. aeruginosa 536 primers in 2 pools, for 103 genes
(7 MLST genes, 3 housekeeping genes, 93 AMR gene clusters)

Stage	Cycle number	Denature	Anneal/Extend
1	1	98°C , 30s	
2	10	98°C, 15s	70°C, 5min
3	25	98°C, 15s	70°C, 5min, delta t = - 0.1°C

Library Preparation

Normalization → End-repair and dA-tailing → Barcode ligation
→ Poor barcoded amplicons → Barcoding adapter ligation
→ Library cleanup and elution → Library loading

Bioinformatics

1. Porechop
(Thorough alignments to effectively find adapters, trim and split reads)

```
porechop -i reads.fastq -b demux --format fastq --require_two_barcodes
```

2.Minimap2
(Mapping positions between reads and reference genomes)

```
minimap2 -a pseudo_ref.fa demux.fastq | samtools view -bS - | samtools sort - > align.bam
```

3.Medaka & Vcffilter
(Variant Calling and Filtration)

```
medaka_variant -f <REFERENCE.fasta> -b <reads.bam> -m r941_flip213
Vcffilter  file.vcf -f  "QC>5"   >  Filtered.vcf
```

4.Parsnp
(Provide core genome phylogeny)

```
parsnp -r <reference_genome> -d <genome_dir> -p <threads>
```

5.FastTree
(Acquire phylogenetic trees with consensus sequences)

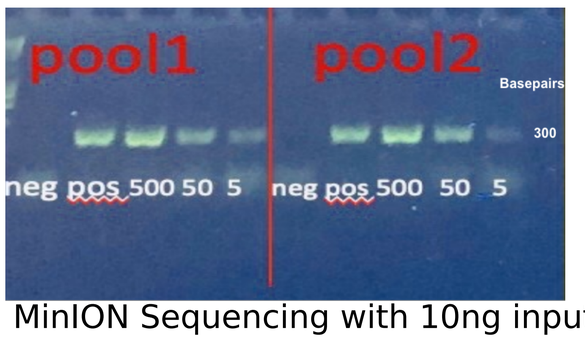
```
FastTree -gtr -nt < consensusread.file > tree_file
```

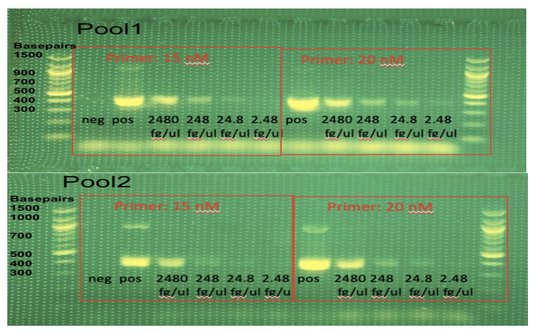
Multiplex scheme for designing target genes

Choose target genes (core genes, easy access) → Acquire target sequences → Align and trim sequences (by clustal omega) → Group homologs and remove identity sequences (by cd-hit est)
→ Run primal scheme on the FASTA file

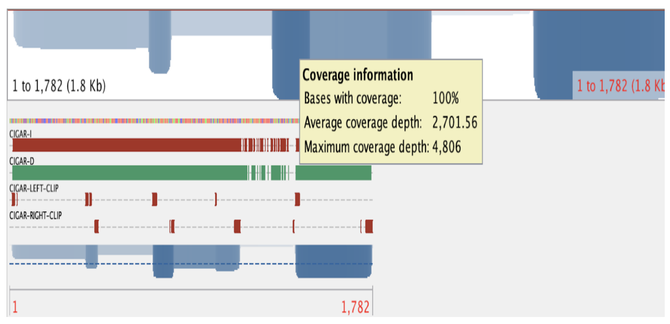
Results: Improved Multiplex PCR Method Works Well with Low Numbers of Genome Copies for MinION Sequencing

Multiplex PCR with DNA templates of:
L.pneumophilaP. aeruginosa as few as 2 copies
as few as 5 copies

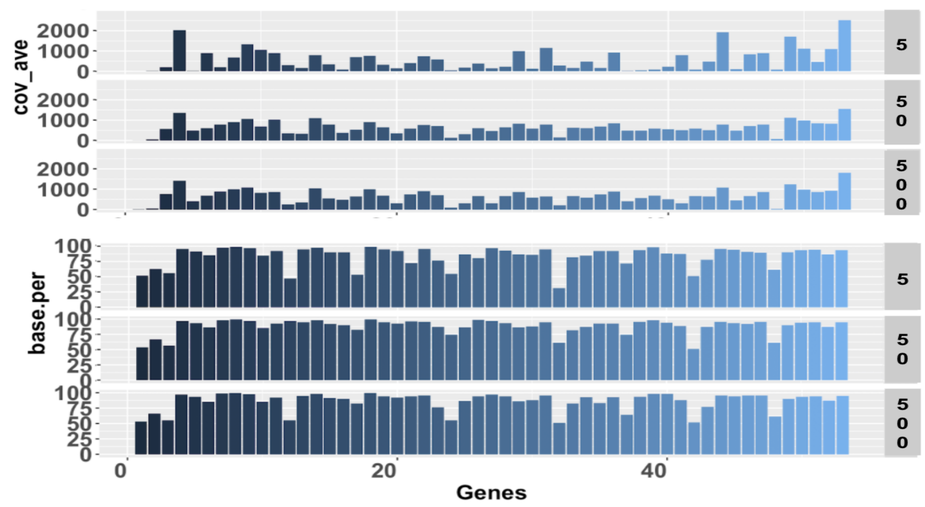




MinION Sequencing with 10ng input DNA of amplicons



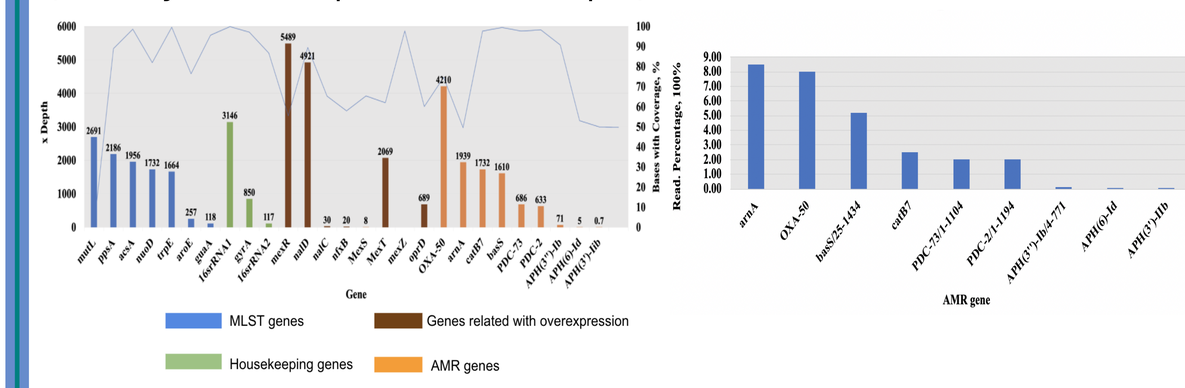
Sequencing covereage information with mapping reads of *L. pneumophila* (500-cell, 50-cell sample and 5-cell samples)



Top: Distribution of average depth across 53 genes of *L. pneumophil*a

Bottom: Depth with coverage at x1 depth across 53 genes of *L. pneumophila*

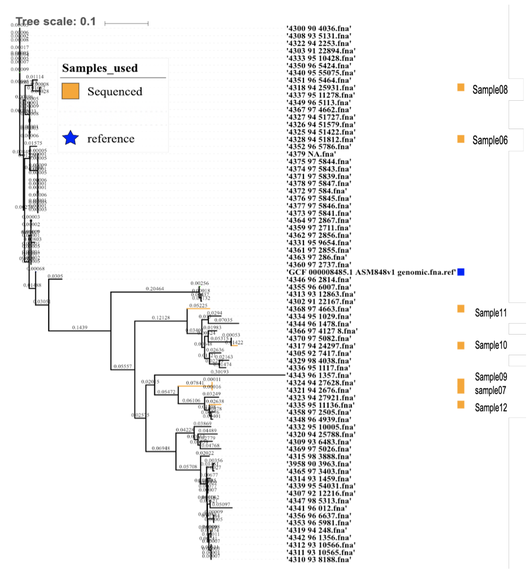
MinION Sequencing with 30ng input DNA of amplicons of *P. aeruginosa* (directly from hospital water sample)



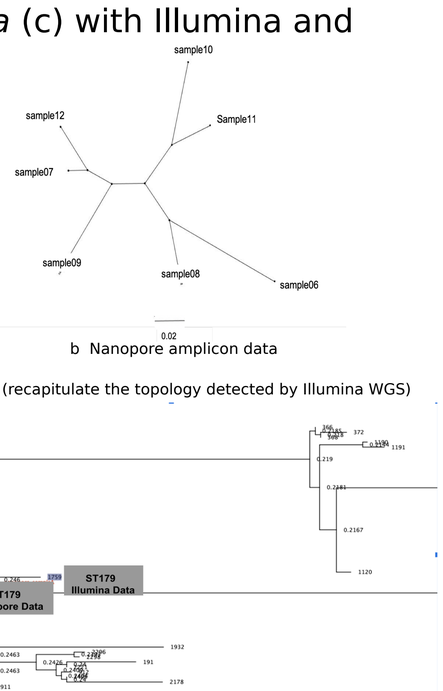
Left: Distribution of average depth, and breadth of depth of *P. aeruginosa* in the water sample

Bottom: AMR genes detected of *P. aeruginosa* in the water sample

Phylogenetic relationship of sequenced DNA samples of *L. pneumophila* (a,b) & MLST determination of *P. aeruginosa* (c) with Illumina and Nanopore amplicon data



a Illumina WGS data



b Nanopore amplicon data

(recapitulate the topology detected by Illumina WGS)

c MLST determination of *P. aeruginosa*

Conclusions

1. The improved protocol could sensitively and specifically detect the pathogenic bacterium *L. pneumophila* and *P. aeruginosa* from water samples containing as few as 5 and 2 genome copies respectively, and reveal the variations, phylogenetic relationship and MLST of the species.
2. Antimicrobial resistance conferred by the horizontal gene transfer and genetic mutation can also be identified by the pipeline.
3. This approach will be of use for rapid risk assessment of contaminated water systems and outbreak investigations.

References

1. Demirjian, A. et al. The Importance of Clinical Surveillance in Detecting Legionnaires' Disease Outbreaks: A Large Outbreak in a Hospital With a Legionella Disinfection System—Pennsylvania, 2011–2012. Clinical Infectious Diseases 60, 1596–1602 (2015).
2. Quick, J. et al. Real-time, portable genome sequencing for Ebola surveillance. Nature 530, 228–232 (2016)
3. Schmidt, K. et al. Identification of bacterial pathogens and antimicrobial resistance directly from clinical urines by nanopore-based metagenomic sequencing. J. Antimicrob. Chemother. 72, 104–114 (2017)