Whole-genome sequencing for syphilis, gonorrhoea and shigella epidemiology and control

David Eyre

Robertson Fellow and Honorary Consultant in Infection
Big Data Institute, University of Oxford
June 2019

david.eyre@bdi.ox.ac.uk
What can whole-genome sequencing tell us?

- Reconstruct historical transmission events at a population scale
  - When did Syphilis come to Europe?
- Track the spread of infections in particular populations
  - Shigella as a cause of colitis in MSM
- Track the spread of antibiotic resistance
  - Identifying the likely origin of cases of MDR/XDR gonorrhoea
- Individual transmission tracking
  - Understand local transmission networks, e.g. evidence for serosorting or associative selectivity
  - Enhanced partner notification
- Better diagnostics?
Decoding pathogen DNA: Sequencing

Culture (or direct from clinical sample)

Extract DNA

Sequencing £50-100 per sample

Analyse the sequence data

Analysis is typically priced at £50-100 per sample.
Pathogen DNA

Bacterial DNA consists of one to a few million letters of genetic code – A, C, G or T

If infection spreads between two people the DNA of the bacteria is likely to be identical or very similar

This DNA allows pathogens to be put into similar groups

Phylogenetic tree:
Two theories of how Syphilis came to Europe

Columbian theory
• Brought back to Europe by Christopher Columbus’ crew
• Spread to Italy by the French during the siege of Naples in 1495 (via Spanish mercenaries)

Pre-Columbian
• Existed before - descriptions in keeping with tertiary syphilis in ancient Greece
• Skeletons consistent with congenital syphilis from Pompeii

Can whole-genome sequencing tell us the answer?
Treponema pallidum: DNA capture

- First genome sequenced in 1998, but only a handful of genomes until 2016...
- Sequencing possible direct from clinical samples in primary and secondary syphilis using RNA ‘baits’ → allows selective capture of low concentrations of pathogen DNA
Historical perspective

- 70 clinical samples 2012-13 and 18 historic samples from 1912 onwards

Common ancestor of modern Syphilis dates back to 17-19th century (cf. 1495 reports in Europe)

Most circulating strains worldwide are from the SS14 lineage
- Emerged 1950s to 1970s
- Most azithromycin resistant

Nichols limited mostly to N America ≤ 1980s

0.75 SNPs / genome / year

Lineages named after reference genomes

NATURE MICROBIOLOGY 2, 16245 (2016)
Success of the SS14 lineage likely multifactorial
• WGS of 73 samples from UK and USA + 49 previous sequences

Multiple mutations have emerged that confer macrolide resistance
Not all lineages azithromycin resistant

Genomic epidemiology of syphilis reveals independent emergence of macrolide resistance across multiple circulating lineages
Shigellosis as a sexually transmitted infection

• Outbreaks of colitis due to *Shigella flexneri* serotype 3a reported from 1970s, large UK outbreak in 2009 onwards in MSM

• *Shigella flexneri* serotype 3a from UK 2004-2013 sequenced together with isolates from 29 countries
Cases associated with isolates from Asia or travel to Asia

Cases associated with isolates from Africa or travel to Africa

No travel to high risk areas for foodborne acquisition
MSM-outbreak lineage found predominant in men without a history of recent travel; of those with data on sexuality 40/41 MSM had MSM-outbreak lineage.
**Antibiotic resistance**

<table>
<thead>
<tr>
<th>Associated resistances</th>
<th>In all isolates</th>
<th>In common isolates from 2009 onward</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Shigella resistance locus multidrug resistance element</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bla*&lt;sub&gt;CTX-M-1&lt;/sub&gt;</td>
<td>β-lactams</td>
<td></td>
</tr>
<tr>
<td>catA1</td>
<td>Chloramphenicol</td>
<td></td>
</tr>
<tr>
<td>aadA1</td>
<td>Aminoglycosides</td>
<td></td>
</tr>
<tr>
<td>tet(B)</td>
<td>Tetracyclines</td>
<td></td>
</tr>
<tr>
<td><strong>pKSR100 (conjugative R-plasmid)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>erm(B)</td>
<td>Macrolides (erythromycin)</td>
<td></td>
</tr>
<tr>
<td>mph(A)</td>
<td>Macrolides (azithromycin)</td>
<td></td>
</tr>
<tr>
<td>bla*&lt;sub&gt;SEMI&lt;/sub&gt;</td>
<td>β-lactams</td>
<td></td>
</tr>
<tr>
<td><strong>pKSR100 integron</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dfrA17</td>
<td>Trimethoprim</td>
<td></td>
</tr>
<tr>
<td>sul1</td>
<td>Sulfonamides</td>
<td></td>
</tr>
<tr>
<td>aadA5</td>
<td>Aminoglycosides</td>
<td></td>
</tr>
<tr>
<td><strong>pCERC1 (R-plasmid)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dfrA14</td>
<td>Trimethoprim</td>
<td></td>
</tr>
<tr>
<td>sul2</td>
<td>Sulfonamides</td>
<td></td>
</tr>
<tr>
<td>strA</td>
<td>Aminoglycosides</td>
<td></td>
</tr>
<tr>
<td>strB</td>
<td>Aminoglycosides</td>
<td></td>
</tr>
</tbody>
</table>

MSM = men who have sex with men.

*Table 3: Antibiotic resistance genes and associated resistances on mobile genetic elements in the MSM-associated outbreak lineage*
Also associated with other serotypes, e.g. 2a and *S. sonnei*
**Shigella sonnei**

- In 2015, *Shigella sonnei*, isolated in 4 men in London
  - Macrolide resistant, but also ESBL
- pKRS100-like plasmid found (same as *S. flexneri* 3a outbreak), with additional \( \text{bla}_{\text{CTX-M-27}} \) gene
- Look back over WGS of all *S. sonnei* sent to PHE identified 9 patients within 5 SNPs, 7 identified as MSM, but no epidemiological links found between them
Tracking XDR gonorrhoea

• March 2018 – **XDR gonorrhoea** diagnosed
  • Ceftriaxone resistant and high-level azithromycin resistance
  • UK resident heterosexual male
  • Recent sexual contact with female in Thailand

• Key **public health questions**
  • **Where did this strain come from?** → is it circulating elsewhere?
  • **How has it become resistant?**
  • **What has this strain evolved from? What is its potential to spread?**
How do we use WGS to identify person-to-person transmission?

• Context from previous studies: 1407 sequences obtained from Brighton from 1061 genetically distinct infections

• How many genetic differences do we expect between recently transmitted genomes?
  • SNPs, single nucleotide polymorphisms

Whole-genome sequencing to determine transmission of *Neisseria gonorrhoeae*: an observational study

*Dilrini De Silva*, Joanna Peters*, Kevin Cole, Michelle J Cole, Fiona Cresswell, Gillian Dean, Jayshree Dave, Daniel Rh Thomas, Kirsty Foster, Alison Waldram, Daniel J Wilson, Xavier Didelot, Yonatan H Grad, Derrick W Crook, Tim E A Peto, A Sarah Walker, John Paul†, David W Eyre†

Within-sample diversity:
12-14 colonies picked from 6 randomly chosen samples

Minimal within sample diversity, a sweep of growth sequenced for remaining samples
Anatomical site diversity:
Opportunistic sequencing of all patients with samples from more than one site

13% of patients had distinct strains at different anatomical sites (206 pairs of samples obtained within \( \leq 30 \) days)

Transmission pairs
30 samples from 15 contact pairs

Both used together to estimate within-host diversity
Evolutionary rates:
Very little chronic infection, so estimated from time-scaled phylogenies

Mutation rate:
3.55 SNPs / genome / year
(95% credibility interval 3.27–3.83)
Transmission nomogram

Number of SNPs expected between isolates linked by direct or indirect transmission
Figure 3: Proportion of Brighton gonorrhoea infections genetically linked to another sampled case.
Evidence of mixed HIV sero-status transmission clusters

Peters et al, 2017, STI
Gonorrhoea in east London

Clustering seen by reported sexual orientation, but not HIV status or ethnicity

Results do not support assortative selectivity as an explanation for infection rate differences between ethnic groups

Dave et al, submitted
Detecting antibiotic resistance mechanisms

- 681 samples sequenced from UK, USA, Canada, and WHO reference collections
- MICs determined by gold standard agar dilution in national surveillance
- Catalogue of common resistance determinants identified

Overall predict MIC to nearest doubling dilution 54% of assays and to within 1 doubling dilution 93% and to within 2 doubling dilutions 98%

Eyre et al, 2018 JAC
UK XDR case vs. previous ceftriaxone resistant cases

Eyre et al, 2018 Eurosurveillance
What about resistance?

Ceftriaxone resistance from the most successful penA gene variant to date, azithromycin from the most common high-level resistance mutation

<table>
<thead>
<tr>
<th>Gene</th>
<th>Variant</th>
<th>Mechanism</th>
<th>Antimicrobials affected</th>
</tr>
</thead>
<tbody>
<tr>
<td>23S rRNA</td>
<td>A2059G, 4 copies</td>
<td>Decreased macrolide binding to 50S ribosome</td>
<td>AZM</td>
</tr>
<tr>
<td>penA</td>
<td>FC428 mosaic penA - 100% identity</td>
<td>Reduced β-lactam acylation of penicillin binding protein (PBP) 2</td>
<td>CRO, PEN</td>
</tr>
<tr>
<td>penB</td>
<td>G120K, A121D</td>
<td>Reduced influx through PorB1b</td>
<td>CRO, PEN, TET</td>
</tr>
<tr>
<td>mtrR</td>
<td>G45D, Promoter deletion</td>
<td>Over-expression of MtrCDE efflux pump resulting in increased efflux</td>
<td>AZM, CRO, PEN, TET</td>
</tr>
<tr>
<td>ponA</td>
<td>L421P</td>
<td>Reduced β-lactam acylation of PBP1</td>
<td>PEN</td>
</tr>
<tr>
<td>tetM</td>
<td>Gene presence</td>
<td>Prevents tetracycline binding to the 30S ribosome</td>
<td>TET</td>
</tr>
<tr>
<td>rpsJ</td>
<td>V57M</td>
<td>Reduced affinity of 30S ribosome for tetracycline</td>
<td>TET</td>
</tr>
<tr>
<td>gyrA</td>
<td>S91F, D95A</td>
<td>Reduced quinolone binding to DNA gyrase</td>
<td>CIP</td>
</tr>
<tr>
<td>parC</td>
<td>S87R</td>
<td>Reduced quinolone binding to topoisomerase IV</td>
<td>CIP</td>
</tr>
</tbody>
</table>

AZM: azithromycin; CIP: ciprofloxacin; CRO: ceftriaxone; PEN: benzylpenicillin; TET: tetracycline.
Comparing the UK isolate to all 7,812 sequences available for *N. gonorrhoeae*

- Rapid search to find 100 most closely related
- Precision tree for these 100
A further development...

Two Australian residents
• A male who reported sexual contact with a female in South East Asia
• A female who had no travel history outside of Australia
Both had ceftriaxone-resistant gonorrhoea with high-level azithromycin resistance
Australian cases identical and 1 SNP different to UK case

Jennison et al, 2019, Eurosurveillance
European transmission of ceftriaxone resistant gonorrhoea October – December 2018

- Two female patients developed gonorrhoea following contact with UK-resident men from the same sexual network linked to travel to Ibiza, Spain
- Ceftriaxone resistant, intermediate susceptible to azithromycin

Both cases were genetically identical – suggesting recent common source
Part of the most successful ceftriaxone resistant lineage to date, FC428 clone
• No direct linkage to previous cases
Potential impacts

• Able to confirm likely transmission of ceftriaxone resistant gonorrhoea in Europe → public health response and raised public awareness

• Data from all cases from the FC428 clone suggest South East Asia or China, is the likely reservoir → need for enhanced surveillance in region

• Highlights the need for multifactorial interventions – new drugs, ?vaccine, access to diagnostics, surveillance, test of cure, effective partner notification
What about sequencing as a routine diagnostic tool?

• Potential to
  • Detect presence of infection
  • Detect antibiotic resistance determinants
  • Enhance partner notification
Leeds real-time sequencing study

• Illumina based sequencing of positive cultures
• Investigate the utility of WGS for routine use alongside partner notification
• Aiming to return final report based by 14 day test of cure visit
WGS enables links to be made between cases not revealed by existing partner notification

- Cluster of 6 MSM
- 4 are identical isolates
- Only one link made on the basis of partner notification

![Diagram showing dates and links]

Link confirmed by routine PN and WGS

WGS linked

13 SNPs from nearest neighbor
WGS suggest the presence of unsampled common sources

Cluster containing 3 verified couples and 3 heterosexual females without matches
Direct from sample sequencing

Extract DNA

Sequence 20 mins → 8 hours

Bioinformatics

Analysis → Final report, using database of antibiotic resistance determinants and previous sequences
Direct from sample sequencing

ONT sequencing direct from a patient urine sample:
Urine sample A

Proof of principle:
Resistance determinants identified within 8-12 hours
Conclusions

• Whole-genome sequencing can inform about historical and recent transmission events at a global and national scale

• Whole-genome sequencing can identify antibiotic resistant determinants and track their spread

• There is a challenge to know how best to use whole-genome sequencing in the clinic
  • It may be very helpful to identify antibiotic resistance
  • Genomic links allow the reach of partner notification to be potentially quantified
  • Active question about how to approach genomically linked cases not linked by routine partner notification
Acknowledgements

GU Medicine, Oxford
- Emily Lord, Natasha Regisford-Reimmer, Anne Edwards

Microbiology, Oxford
- Monique Andersson, Markus Morgan, Robert Newnham

University of Oxford
- Teresa Street, Nicholas Sanderson, Kevin Chau, Leanne Barker, Derrick Crook, Tim Peto

PHE
- Gwenda Hughes, Nick Phin, Michelle Cole, Helen Fifer

WHO coordinating centre
- Daniel Golparian, Magnus Unemo

Australia
- Amy Jennison, David Whiley, Monica Lahra

London
- Jayshree Dave

Brighton
- John Paul, Jo Peters, Keven Cole

Leeds
- Ling Yuan Kong, Mark Wilcox, Ines Moura
- Janet Wilson and rest of GU medicine team