Reconstruction of plasmids and MGEs from NGS and SMRT Seq

Anna Sheppard
IMMEM XII, 19th September 2019
Disclosures

• None
Deaths attributable to AMR every year compared to other major causes of death

AMR in 2050
10 million

Tetanus
60,000

Road traffic accidents
1.2 million

Cancer
8.2 million

Measles
130,000

Cholera
100,000 – 120,000

Diarrhoeal disease
1.4 million

Diabetes
1.5 million

AMR now
700,000
(low estimate)

O'Neill Review on Antimicrobial Resistance, 2014

ANTIBIOTIC RESISTANCE:
THE GLOBAL THREAT

SUPER RESISTANT BACTERIA:
PROBLEM TODAY,
CRISIS TOMORROW

In INDIA, over 58,000 babies
died in one year as a result of infection with super-resistant bacteria usually passed on from their mothers

In the EUROPEAN UNION,
antibiotic resistance causes
25,000 deaths per year and 2.5m extra hospital days

In THAILAND,
antibiotic resistance causes
38,000+ deaths per year and 3.2m hospital days

In the UNITED STATES,
antibiotic resistance causes
23,000+ deaths per year and >2.0m illnesses

CARBAPENEM-RESISTANT
ENTEROBACTERIACEAE

9,000 DRUG RESISTANT
INFECTIONS PER YEAR

1,400 CARBAPENEM-RESISTANT
KLEBSIELLA PNEUMONIAE

1,400 CARBAPENEM-RESISTANT
E. COLI

CRE HAVE BECOME RESISTANT TO ALL
OR NEARLY ALL AVAILABLE ANTIBIOTICS

THREAT LEVEL
URGENT

CDC Antibiotic Resistance Initiative Fact Sheet, 2014

CDC Antibiotic Resistance Threats in the United States, 2013
Resistance genes associated with MGEs

How can we use WGS to better understand resistance dissemination?
Approaches for plasmid reconstruction from WGS

- Sequencing technology
  - Short-read (typically Illumina) – 100s of bp
  - Long-read (Pacific Biosciences / Oxford Nanopore) – 10s of kb

- De novo assembly
  - Short-read: plasmids generally fragmented
    - Fragmentation due to repeats
    - Repeats include MGEs involved in transposition, recombination etc
  - Long-read: can generate complete, closed genomes

- Non-assembly approaches
  - Plasmid “presence”: query with reference plasmid(s)
    - Doesn’t account for structural continuity
    - OK if queried structure is stable; not ideal for resistance plasmids
**Klebsiella pneumoniae**
carbapenememase (KPC)

- Ambler Class A serine β-lactamase
- Hydrolyzes all β-lactams; penicillins, extended spectrum cephalasporins, aztreonam and carbapenems
- Found across various *Enterobacteriaceae* species
- $bla_{KPC}$ is usually found within Tn4401, which undergoes replicative transposition with 5 bp target site duplication (TSD)

Long-term multi-species KPC outbreak

Perirectal surveillance begins

- P. mirabilis
- Kluyvera sp.
- C. farmeri
- M. morganii
- C. youngae
- C. amalonaticus
- R. planticola
- E. asburiae
- K. intermedia
- Enterobacter sp.
- Citrobacter sp.
- R. ornitholytica
- Proteus sp.
- E. aerogenes
- Aeromonas
- E. coli
- K. oxytoca
- S. marcescens
- C. freundii
- K. pneumoniae
- E. cloacae
Long-term multi-species KPC outbreak

281 isolates from 182 patients (2007-2012)
- Illumina: all isolates
- PacBio: select isolates

→ High diversity in species, strains and plasmids carrying $bla_{KPC}$
Long-term multi-species KPC outbreak

Multiple levels of $bla_{KPC}$ mobility:

- Klebsiella pneumoniae carbapenemase (KPC) resistance gene
- Mobile transposon Tn4401
- Tn4401 nested within mobile Tn2-like element
- Transposon located within mobile extra-chromosomal plasmid DNA
- Bacterial cell with mobile plasmid and immobile host chromosome (not to scale)

Transposition of Tn4401 into different plasmids

Proposed mobilization of Tn4401 between Tn2-like elements in different plasmids, via homologous recombination

Plasmid movement via conjugation into different bacterial strains and species

Mobilisation of $bla_{KPC}$ via:
- Tn4401 transposition
- Recombination involving Tn2-like elements
- Plasmid conjugation

Difficult to infer transmission

Can we find ways to track $bla_{KPC}$ through multiple patients, strains and plasmids?
How to track $bla_{KPC}$?

- Cannot reconstruct plasmids from short-read WGS!
- Long-read WGS?
- “Transposon outbreak”
  - Tn4401 is largely conserved
- Can we better utilise existing short-read WGS by focusing on Tn4401 as the unit of transmission?
TETyper: a bioinformatic pipeline for typing transposable elements (TEs)

• Inputs:
  • Short-read WGS data for a single isolate
  • TE reference sequence (e.g. Tn4401)

• Outputs:
  • Deletions within TE
  • SNVs within TE
  • Flanking genetic context(s) of TE (~5-30 bp)

$bla_{KPC}$ outbreak: application of TETyper to Tn4401

  - 660 isolates from 384 patients
  - 606 isolates from the hospital environment
Tn4401 SNVs and associations with host species

![Graph showing the distribution of Tn4401 SNVs across different species.](image)

**Species**
- Serratia marcescens
- Enterobacter sp.
- Klebsiella pneumoniae
- Citrobacter sp.
- Klebsiella oxytoca
- Aeromonas hydrophila
- Raoultella ornitholytica
- Klyuyvera intermedia
- Escherichia coli
- Unknown

**Tn4401 SNVs**
- None
- bla<sub>KPC-2</sub>
- C8015T
- C8015Y
- bla<sub>KPC-3</sub>
- bla<sub>KPC-2</sub> + bla<sub>KPC-3</sub>
- T6800C, C7509G, T7917G
- A420G, C8015T
- C8015T, T9621C
- Other

© ESCMID eLibrary by author
### T9663C mutation in Tn4401

- 310 isolates with T9663C, 14 non-Serratia (12 pat, 2 env)
- 5/6 patients also have KPC-Serratia
  - Evidence for within-patient HGT of $\text{bla}_{KPC}$ from *Serratia* to other species

<table>
<thead>
<tr>
<th>Isolate</th>
<th>Patient</th>
<th>Date</th>
<th>Species</th>
<th>Tn4401 structure</th>
<th>Tn4401 SNPs</th>
<th>Left flanks</th>
<th>Right flanks</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAV1601</td>
<td>280</td>
<td>2012-04</td>
<td>Aeromonas hydrophila</td>
<td>Tn4401b</td>
<td>T9663C</td>
<td>TTTTT</td>
<td>TTTTT</td>
</tr>
<tr>
<td>CAV1773</td>
<td>280</td>
<td>2012-04</td>
<td>Serratia marcescens</td>
<td>Tn4401b</td>
<td>none</td>
<td>GTTCT</td>
<td>GTTCT</td>
</tr>
<tr>
<td>CAV1602</td>
<td>280</td>
<td>2012-04</td>
<td>Serratia marcescens</td>
<td>Tn4401b</td>
<td>none</td>
<td>GTTCT</td>
<td>GTTCT</td>
</tr>
<tr>
<td>CAV1606</td>
<td>280</td>
<td>2012-04</td>
<td>Serratia marcescens</td>
<td>Tn4401b</td>
<td>none</td>
<td>GTTCT</td>
<td>GTTCT</td>
</tr>
<tr>
<td>CAVp65</td>
<td>294</td>
<td>2014-08</td>
<td>Serratia marcescens</td>
<td>Tn4401b</td>
<td>T9663C</td>
<td>TTTTT</td>
<td>TTTTT</td>
</tr>
<tr>
<td>CAVp80</td>
<td>294</td>
<td>2014-09</td>
<td>Morganella morgani</td>
<td>Tn4401b</td>
<td>C8015T</td>
<td>TTTTT</td>
<td>GTTCT</td>
</tr>
<tr>
<td>CAVp91</td>
<td>294</td>
<td>2014-10</td>
<td>Escherichia coli</td>
<td>Tn4401 truncD</td>
<td>C8015T,T9663C</td>
<td>TTTTT</td>
<td>none</td>
</tr>
<tr>
<td>CAVp269</td>
<td>302</td>
<td>2015-06</td>
<td>Serratia marcescens</td>
<td>Tn4401b</td>
<td>T9663C</td>
<td>TTTTT</td>
<td>TTTTT</td>
</tr>
<tr>
<td>CAVp270</td>
<td>302</td>
<td>2015-06</td>
<td>Serratia marcescens</td>
<td>Tn4401b</td>
<td>T9663C</td>
<td>TTTTT</td>
<td>TTTTT</td>
</tr>
<tr>
<td>CAVp296</td>
<td>302</td>
<td>2015-10</td>
<td>Klebsiella pneumoniae</td>
<td>Tn4401b</td>
<td>T9663C</td>
<td>TTTTT</td>
<td>TTTTT</td>
</tr>
<tr>
<td>CAVp360</td>
<td>302</td>
<td>2016-12</td>
<td>Klebsiella pneumoniae</td>
<td>Tn4401b</td>
<td>T9663C</td>
<td>TTTTT</td>
<td>TTTTT</td>
</tr>
<tr>
<td>CAVp361</td>
<td>302</td>
<td>2016-12</td>
<td>Escherichia coli</td>
<td>Tn4401b</td>
<td>T9663C</td>
<td>TTTTT</td>
<td>TTTTT</td>
</tr>
<tr>
<td>CAVp374</td>
<td>302</td>
<td>2017-03</td>
<td>Citrobacter freundii complex</td>
<td>Tn4401b</td>
<td>T9663C</td>
<td>TTTTT</td>
<td>TTTTT</td>
</tr>
<tr>
<td>CAVp223</td>
<td>326</td>
<td>2015-01</td>
<td>Klebsiella aerogenes</td>
<td>Tn4401b</td>
<td>T9663C</td>
<td>TTTTT</td>
<td>TTTTT</td>
</tr>
<tr>
<td>CAVp66</td>
<td>333</td>
<td>2014-08</td>
<td>Serratia marcescens</td>
<td>Tn4401b</td>
<td>T9663C</td>
<td>TTTTT</td>
<td>TTTTT</td>
</tr>
<tr>
<td>CAVp83</td>
<td>333</td>
<td>2014-09</td>
<td>Klebsiella aerogenes</td>
<td>Tn4401b</td>
<td>T9663C</td>
<td>TTTTT</td>
<td>TTTTT</td>
</tr>
<tr>
<td>CAVp98</td>
<td>338</td>
<td>2014-11</td>
<td>Serratia marcescens</td>
<td>Tn4401b</td>
<td>T9663C</td>
<td>ATTAC,TTTTT</td>
<td>ATTAC,TTTTT</td>
</tr>
<tr>
<td>CAVp100</td>
<td>338</td>
<td>2014-11</td>
<td>Escherichia coli</td>
<td>Tn4401b</td>
<td>T9663C</td>
<td>TTTTT</td>
<td>TTTTT</td>
</tr>
<tr>
<td>CAVp101</td>
<td>338</td>
<td>2014-11</td>
<td>Escherichia coli</td>
<td>Tn4401b</td>
<td>T9663C</td>
<td>TTTTT</td>
<td>TTTTT</td>
</tr>
<tr>
<td>CAVp102</td>
<td>338</td>
<td>2014-11</td>
<td>Escherichia coli</td>
<td>Tn4401b</td>
<td>T9663C</td>
<td>TTTTT</td>
<td>TTTTT</td>
</tr>
<tr>
<td>CAVp99</td>
<td>338</td>
<td>2014-11</td>
<td>Escherichia coli</td>
<td>Tn4401b</td>
<td>T9663C</td>
<td>TTTTT</td>
<td>TTTTT</td>
</tr>
<tr>
<td>CAVp212</td>
<td>338</td>
<td>2015-01</td>
<td>Serratia marcescens</td>
<td>Tn4401b</td>
<td>T9663C</td>
<td>ATTAC,TTTTT</td>
<td>ATTAC,TTTTT</td>
</tr>
</tbody>
</table>
Long-read sequencing

• Enables generation of high-quality finished genomes
  • Typically achieved using hybrid assembly (short- and long-read) – Unicycler (Wick et al 2017)
  • Ideally resulting genome is “closed”, i.e. single contig per plasmid / chromosome, with all contigs circular

• Two examples of insights from long-read sequencing in $bla_{KPC}$ outbreak:
  • Klebsiella quasipneumoniae
  • Tn5403 association with Tn4401
**Klebsiella quasipneumoniae**

- 23 KpIIA isolates from $\text{bla}_{KPC}$ outbreak

<table>
<thead>
<tr>
<th>Isolate</th>
<th>Patient</th>
<th>Room</th>
<th>Year</th>
<th>Tn4401 structure</th>
<th>Tn4401 SNVs</th>
<th>Tn4401 left flanks</th>
<th>Tn4401 right flanks</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAV1360</td>
<td>1</td>
<td></td>
<td>2009</td>
<td>Tn4401b</td>
<td>none</td>
<td>GT</td>
<td>CT</td>
</tr>
<tr>
<td>CAVp203</td>
<td>2</td>
<td></td>
<td>2013</td>
<td>Tn4401b</td>
<td>C8015T</td>
<td>GTCTCTCT</td>
<td>AATAA</td>
</tr>
<tr>
<td>CAV2013</td>
<td>2</td>
<td></td>
<td>2013</td>
<td>Tn4401b</td>
<td>C8015T</td>
<td>GTCTCT</td>
<td>GTCTCT</td>
</tr>
<tr>
<td>CAVp26</td>
<td>2</td>
<td></td>
<td>2014</td>
<td>Tn4401b</td>
<td>C8015T</td>
<td>GTCTCT</td>
<td>GTCTCT</td>
</tr>
<tr>
<td>CAVp20</td>
<td>2</td>
<td></td>
<td>2014</td>
<td>Tn4401b</td>
<td>C8015T</td>
<td>GTCTCT TAGTG</td>
<td>GTCTCTTAGTG</td>
</tr>
<tr>
<td>CAVp103</td>
<td>2</td>
<td></td>
<td>2014</td>
<td>Tn4401b</td>
<td>C8015Y</td>
<td>GTCTCT</td>
<td>TTTTT</td>
</tr>
<tr>
<td>CAVp72</td>
<td>2</td>
<td></td>
<td>2014</td>
<td>Tn4401b</td>
<td>C8015T</td>
<td>TAGTG</td>
<td>TAGTG</td>
</tr>
<tr>
<td>CAVp64</td>
<td>2</td>
<td></td>
<td>2014</td>
<td>Tn4401b</td>
<td>C8015T</td>
<td>TTTTT</td>
<td>TTTTT</td>
</tr>
<tr>
<td>CAVp67</td>
<td>3</td>
<td></td>
<td>2014</td>
<td>Tn4401b</td>
<td>C8015T</td>
<td>GTCTCT</td>
<td>GTCTCT</td>
</tr>
<tr>
<td>CAVp275</td>
<td>4</td>
<td></td>
<td>2015</td>
<td>Tn4401b</td>
<td>C8015T</td>
<td>TTTTT</td>
<td>TTTTT</td>
</tr>
<tr>
<td>CAV2244</td>
<td>A</td>
<td></td>
<td>2014</td>
<td>Tn4401b</td>
<td>C8015T</td>
<td>TTTTT</td>
<td>TTTTT</td>
</tr>
<tr>
<td>CAV2279</td>
<td>B</td>
<td></td>
<td>2014</td>
<td>Tn4401b</td>
<td>none</td>
<td>AACTTACATA</td>
<td>GTCT</td>
</tr>
<tr>
<td>CAV1947</td>
<td>C</td>
<td></td>
<td>2014</td>
<td>Tn4401b</td>
<td>C8015T</td>
<td>T9663Y</td>
<td>ACAATAGCT</td>
</tr>
<tr>
<td>CAV1945</td>
<td>C</td>
<td></td>
<td>2014</td>
<td>ΔT7075-7153</td>
<td>C8015T</td>
<td>GTCTCT</td>
<td>GTCTCT</td>
</tr>
<tr>
<td>CAV2018</td>
<td>C</td>
<td></td>
<td>2014</td>
<td>Tn4401b</td>
<td>T9663C</td>
<td>GTCTCT</td>
<td>GTCTCT</td>
</tr>
<tr>
<td>CAV2397</td>
<td>C</td>
<td></td>
<td>2014</td>
<td>Tn4401b</td>
<td>C8015T</td>
<td>GTCTCT</td>
<td>GTCTCT</td>
</tr>
<tr>
<td>CAV1964</td>
<td>C</td>
<td></td>
<td>2014</td>
<td>Tn4401b</td>
<td>C8015Y</td>
<td>GTCTCT TAGTG</td>
<td>GTCTCTTAGTG</td>
</tr>
<tr>
<td>CAV2019</td>
<td>D</td>
<td></td>
<td>2014</td>
<td>Tn4401b</td>
<td>T9663C</td>
<td>TAGTGTTTT</td>
<td>TAGTGTTTT</td>
</tr>
<tr>
<td>CAV2697</td>
<td>E</td>
<td></td>
<td>2014</td>
<td>Tn4401b</td>
<td>C8015T</td>
<td>TAGTG</td>
<td>TAGTG</td>
</tr>
<tr>
<td>CAV2957</td>
<td>F</td>
<td></td>
<td>2015</td>
<td>Tn4401b</td>
<td>C8015T</td>
<td>TTTTT</td>
<td>TTTTT</td>
</tr>
<tr>
<td>CAV2983</td>
<td>G</td>
<td></td>
<td>2015</td>
<td>Tn4401b</td>
<td>C8015T</td>
<td>TTTTT</td>
<td>TTTTT</td>
</tr>
<tr>
<td>CAV2984</td>
<td>G</td>
<td></td>
<td>2015</td>
<td>Tn4401b</td>
<td>C8015T</td>
<td>TTTTT</td>
<td>TTTTT</td>
</tr>
<tr>
<td>CAV3444</td>
<td>G</td>
<td></td>
<td>2016</td>
<td>Tn4401b</td>
<td>C8015T</td>
<td>TTTTT</td>
<td>TTTTT</td>
</tr>
</tbody>
</table>
Genetic variation in KpIIA isolates
Long-read sequencing

CAV2013 (*K. quasipneumoniae*)

- RepA
- IncU
- IncX5

CAV1761 (*S. marcescens*)

- IncU
- IncX5

CAV1947 (*K. quasipneumoniae*)

- RepA
- IncX5

CAV2018 (*K. quasipneumoniae*)

- RepA
- IncX5

Pat 2:

- Plasmid transfer
- Homologous recombination

16 kb deletion adjacent to Tn4401

6 kb deletion truncating Tn4401

BlasKPC-2

Plasmid loss

**Env:**

- Tn4401 transposition

Tn5403 association with Tn4401

- Novel Tn4401 structure identified in a *Raoultella planticola* isolate from the hospital environment:

![Diagram of Tn4401 and Tn5403 structures]

- Prevalence across *bla*KPC outbreak?
- Queried short-read data for junction sequences
  - Left junction present in 109/1266 (9%) isolates
  - Right junction present in 90 isolates
Flanking genetic contexts of Tn4401::Tn5403

- TETyper on Tn4401 with 30 bp flank length
  - From 109 isolates: 32 distinct left flanks, 25 right flanks
  - Query flanks against known sequence databases (ISFinder)

- 7/32 left flanks match Tn5403!
### Isolate | Year | Species | Tn4401 structure | Tn4401 left SNVs | Tn4401 left flanks | Tn4401 right SNVs | Tn4401 right flanks
--- | --- | --- | --- | --- | --- | --- | ---
CAV1663 | 2012 | Klebsiella pneumoniae | CAV4041 | none | none | none | none
Cavp50 | 2014 | Klebsiella pneumoniae | CAV9999-10006 | none | none | none | none
Cavp7 | 2014 | Klebsiella pneumoniae | CAV9999-10006 | none | none | none | none
Cavp104 | 2014 | Klebsiella pneumoniae | CAV9999-10006 | none | none | none | none
Cavp123 | 2013 | Klebsiella pneumoniae | CAV4041 | none | none | none | none
Cavp124 | 2013 | Klebsiella pneumoniae | CAV9999-10006 | none | none | none | none
Cavp1661 | 2013 | Klebsiella pneumoniae | CAV4041 | none | none | none | none
Cavp205 | 2013 | Klebsiella pneumoniae | CAV9999-10006 | none | none | none | none
Cav1921 | 2014 | Raoultella planticola | CAV4041 | none | none | none | none
Cav1988 | 2015 | Raoultella planticola | CAV4041 | none | none | none | none
Cav2800 | 2015 | Raoultella planticola | CAV4041 | none | none | none | none
Cav2339 | 2015 | Raoultella planticola | CAV4041 | none | none | none | none
Cav1978 | 2013 | Raoultella planticola | CAV4041 | none | none | none | none
Cav1779 | 2013 | Raoultella planticola | CAV4041 | none | none | none | none
Cav1780 | 2013 | Raoultella planticola | CAV4041 | none | none | none | none
Cav1792 | 2013 | Raoultella planticola | CAV4041 | none | none | none | none
Cav1800 | 2013 | Raoultella planticola | CAV4041 | none | none | none | none
Cav1831 | 2013 | Raoultella planticola | CAV4041 | none | none | none | none
Cav1832 | 2013 | Raoultella planticola | CAV4041 | none | none | none | none
Cav1836 | 2013 | Raoultella planticola | CAV4041 | none | none | none | none
Cav1913 | 2013 | Raoultella planticola | CAV4041 | none | none | none | none
Cav1813 | 2013 | Raoultella planticola | CAV4041 | none | none | none | none
Cav1814 | 2013 | Raoultella planticola | CAV4041 | none | none | none | none
Cav1815 | 2013 | Raoultella planticola | CAV4041 | none | none | none | none
Cav1844 | 2013 | Raoultella planticola | CAV4041 | none | none | none | none
Cav1847 | 2013 | Raoultella planticola | CAV4041 | none | none | none | none
Cav1848 | 2013 | Raoultella planticola | CAV4041 | none | none | none | none
Cav1861 | 2013 | Raoultella planticola | CAV4041 | none | none | none | none
Cav1867 | 2013 | Raoultella planticola | CAV4041 | none | none | none | none
Cav1868 | 2013 | Raoultella planticola | CAV4041 | none | none | none | none
Cav1885 | 2013 | Raoultella planticola | CAV4041 | none | none | none | none
Cav1866 | 2013 | Raoultella planticola | CAV4041 | none | none | none | none
Cav1892 | 2013 | Raoultella planticola | CAV4041 | none | none | none | none
Cav1902 | 2013 | Raoultella planticola | CAV4041 | none | none | none | none
Cav1936 | 2013 | Raoultella planticola | CAV4041 | none | none | none | none
Cav1948 | 2014 | Raoultella planticola | CAV4041 | none | none | none | none
Cav1956 | 2014 | Raoultella planticola | CAV4041 | none | none | none | none
Cav1958 | 2014 | Raoultella planticola | CAV4041 | none | none | none | none
Cav1987 | 2014 | Raoultella planticola | CAV4041 | none | none | none | none
Cav2685 | 2014 | Raoultella planticola | CAV4041 | none | none | none | none
Cav2688 | 2014 | Raoultella planticola | CAV4041 | none | none | none | none
Cav278 | 2014 | Raoultella planticola | CAV4041 | none | none | none | none
Cav4574 | 2016 | Raoultella planticola | CAV4041 | none | none | none | none
Cav6020 | 2017 | Raoultella planticola | CAV4041 | none | none | none | none
Cav1819 | 2013 | Serratia marcescens | Tn4401b | CAV9963 | Tn4401 left flanks | CAV9963 | Tn4401 right flanks

### 47 isolates with Tn4401 flank matching Tn5403
### Isolate Year Species Tn\textsubscript{4401} structure Tn\textsubscript{4401} SNVs Tn\textsubscript{4401} left flanks Tn\textsubscript{4401} right flanks
<table>
<thead>
<tr>
<th>Isolate</th>
<th>Year</th>
<th>Species</th>
<th>Tn\textsubscript{4401}</th>
<th>Tn\textsubscript{4401}</th>
<th>left flanks</th>
<th>right flanks</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAV1817</td>
<td>2013</td>
<td>Raoultella planticola</td>
<td>Tn\textsubscript{4401b}</td>
<td>C8015T</td>
<td>GCTGGCGCTGGATGATGATTTGGACGTTCT</td>
<td>GTTCTGGACGGCCATGCTGGCCGTTATCATC</td>
</tr>
<tr>
<td>CAV1958</td>
<td>2014</td>
<td>Raoultella planticola</td>
<td>Tn\textsubscript{4401b}</td>
<td>C8015T</td>
<td>GCTGGCGCTGGATGATGATTTGGACGTTCT</td>
<td>GTTCTGGACGGCCATGCTGGCCGTTATCATC</td>
</tr>
<tr>
<td>CA\textsubscript{V}p50</td>
<td>2014</td>
<td>Klebsiella pneumoniae</td>
<td>(\Delta)9299-10006</td>
<td>C8015T</td>
<td>GGCCCGGAATATCGTAAGGCACAAAGTTGGACGTTCT</td>
<td>none</td>
</tr>
<tr>
<td>CAV1921</td>
<td>2014</td>
<td>Raoultella planticola</td>
<td>Tn\textsubscript{4401b}</td>
<td>C8015T</td>
<td>ACCGACAGATAAAACGATGAGCAAGATCCT</td>
<td>GTTCTGGACGGCCATGCTGGCCGTTATCATC</td>
</tr>
</tbody>
</table>

**CAV1817 (R. planticola)**

- pKPC\_CAV1817
  - 72,181 bp Inc\textsubscript{N}
  - 183,032 bp Inc\textsubscript{FiB}
  - 94,581 bp Inc\textsubscript{FiA}/Inc\textsubscript{FiII}

**CAV1958 (R. planticola)**

- pKPC\_CAV1958-72
  - 72,181 bp Inc\textsubscript{N}
  - 183,030 bp Inc\textsubscript{FiB}
  - 94,581 bp Inc\textsubscript{FiA}/Inc\textsubscript{FiII}

### Long-read sequencing

- Plasmid with Tn\textsubscript{5403}
- Integration of Tn\textsubscript{4401}::Tn\textsubscript{5403} within Tn\textsubscript{5403}
- Recombination involving downstream Tn\textsubscript{5403} fragment
- Deletion of intervening region

- E.g. pKPC\_CAV1958-20

© ESCMID eLibrary by author
### Long-read sequencing

<table>
<thead>
<tr>
<th>Isolate</th>
<th>Year</th>
<th>Species</th>
<th>Tn4401 structure</th>
<th>Tn4401 SNVs</th>
<th>Tn4401 left flanks</th>
<th>Tn4401 right flanks</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAV1817</td>
<td>2013</td>
<td>Raoultella planticola</td>
<td>Tn4401b</td>
<td>C8015T</td>
<td>GCTGGCGCTGGATGATGCTTTCT</td>
<td>TGTCTGGACGCCATGCTGGCCAATC</td>
</tr>
<tr>
<td>CAV1958</td>
<td>2014</td>
<td>Raoultella planticola</td>
<td>Tn4401b</td>
<td>C8015T</td>
<td>GCTGGCGCTGGATGATGCTTTCT</td>
<td>TGTCTGGACGCCATGCTGGCCAATC</td>
</tr>
<tr>
<td>CAVp50</td>
<td>2014</td>
<td>Klebsiella pneumoniae</td>
<td>Δ929-10006</td>
<td>C8015T</td>
<td>GCCGCGGATATCGGACACAAAGTCCA</td>
<td>none</td>
</tr>
<tr>
<td>CAV1921</td>
<td>2014</td>
<td>Raoultella planticola</td>
<td>Tn4401b</td>
<td>C8015T</td>
<td>GCTGGCGCTGGATGATGCTTTCT</td>
<td>TGTCTGGACGCCATGCTGGCCAATC</td>
</tr>
</tbody>
</table>

#### CAV1817 (R. planticola)
- **pKPC_CAV1817**: 72,181 bp IncN
- **pKPC_CAV1817::Tn5403**: 13,675 bp
- **Tn4401b**: 10,006 bp
- **Tn5403**: 13,675 bp

#### CAV1958 (R. planticola)
- **pKPC_CAV1958**: 72,181 bp IncN
- **pKPC_CAV1958::Tn5403**: 13,675 bp
- **Tn4401b**: 10,006 bp
- **Tn5403**: 13,675 bp

#### CAVp50 (K. pneumoniae)
- **pKPC_CAVp50**: 12,072 bp NT
- **pKPC_CAVp50::Tn5403**: 13,675 bp
- **Tn4401b**: 10,006 bp
- **Tn5403**: 13,675 bp

© ESCMID eLibrary by author
### Table

<table>
<thead>
<tr>
<th>Isolate</th>
<th>Year</th>
<th>Species</th>
<th>Tn4401 structure</th>
<th>Tn4401 left SNVs</th>
<th>Tn4401 left flanks</th>
<th>Tn4401 right flanks</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAV1817</td>
<td>2013</td>
<td>Raoultella planticola</td>
<td>Tn4401b</td>
<td>C8015T</td>
<td>GCTGGCGCTGGATGATGATGGACGTTC</td>
<td>GTTCTGGACGCGATGCTGGCCGTTATCATC</td>
</tr>
<tr>
<td>CAV1958</td>
<td>2014</td>
<td>Raoultella planticola</td>
<td>Tn4401b</td>
<td>C8015T</td>
<td>GCTGGCGCTGGATGATGATGGACGTTC</td>
<td>GTTCTGGACGCGATGCTGGCCGTTATCATC</td>
</tr>
<tr>
<td>CAVp50</td>
<td>2014</td>
<td>Klebsiella pneumoniae</td>
<td>A9299-10006</td>
<td>C8015T</td>
<td>GGCCCGATATCGTAAGGCACAAAGTTGA</td>
<td>none</td>
</tr>
<tr>
<td>CAV1921</td>
<td>2014</td>
<td>Raoultella planticola</td>
<td>Tn4401b</td>
<td>C8015T</td>
<td>ACCTGGCGCTGGATGATGATGGACGTTC</td>
<td>ACCTGGCGCTGGATGATGATGGACGTTC</td>
</tr>
</tbody>
</table>

**CAV1817 (R. planticola)**

- pKPC_CAV1817 72,181 bp IncN
- 8756 bp Col440I/Col440II
- 3,223 bp Col440I
- 183,032 bp IncFiB

**CAV1958 (R. planticola)**

- pKPC_CAV1958-72 72,181 bp IncN
- 3,223 bp Col440I
- 183,030 bp IncFiB

**CAV1921 (R. planticola)**

- pKPC_CAV1921 72,181 bp IncN
- 94,581 bp IncFIA/IncFII

**CAVp50 (K. pneumoniae)**

- pKPC_CAvp50 12,072 bp NT
- 54,722 bp IncN

**CAVp50 (K. pneumoniae)**

- pKPC_CAv50 12,072 bp NT
- 54,722 bp IncN

- Mixed plasmid population
Within-patient evolution

• One patient with 4 species carrying Tn4401::Tn5403 over 3 months
Conclusions

• Rapid evolutionary changes mediated by MGE activity in a single-hospital $bla_{KPC}$ outbreak

• Importance of homologous recombination in resistance gene dissemination
  • Tn4401 spread via Tn2-like elements
  • Multiple plasmid acquisition, with recombination between different Tn4401 variants
  • Integration of Tn4401::Tn5403 into Tn5403, with recombination between Tn5403 copies

• Long-read sequencing of epidemiologically related isolates enables detection of specific MGE activity
  • But can be laborious to disentangle due to complex rearrangements
  • Need better tools for high-throughput analysis
Acknowledgements

Oxford
Derrick Crook
Nicole Stoesser
Sarah Walker
Tim Peto
Ali Vaughan
David Eyre
Stephen Bush
Nick Sanderson

UVA
Amy Mathers
Hardik Parikh
Kasi Vegesana
Ian German-Mesner
Katie Barry
Shireen Kotay

Mt Sinai
Robert Sebra

The research was funded by the National Institute for Health Research Health Protection Research Unit (NIHR HPRU) in Healthcare Associated Infections and Antimicrobial Resistance at Oxford University in partnership with Public Health England (PHE) [grant HPRU-2012-10041], and by a BAA subcontract from US Centers for Disease Control and Prevention. The views expressed in this publication are those of the authors and not necessarily those of the funders.