Supplementary Online Content

Reuter S, Ellington MJ, Cartwright EJ, et al. Rapid bacterial whole-genome sequencing to enhance diagnostic and public health microbiology. *JAMA Intern Med*. Published online June 24, 2013. doi:10.1001/jamainternmed.2013.7734.

This supplementary material has been provided by the authors to give readers additional information about their work.

Supplementary information for: Rapid whole genome sequencing to improve diagnostic and public health microbiology. Sandra Reuter, et al.

eMethods

Reference laboratory tests

Minimum inhibitory concentrations of carbapenem-resistant isolates were determined by the reference laboratory using doubling-concentration agar plate dilutions.[1] Extended-spectrum beta-lactamase (ESBL) activity was detected phenotypically by potentiation of cefotaxime and/or ceftazidime by clavulanic acid. High-level AmpC beta-lactamase activity was detected phenotypically by potentiation of cefotaxime by cloxacillin. Acquired carbapenemase gene detection was done by multiplex PCR assays to detect genes for class A (KPC, IMI, NMC, SME), class B (IMP, VIM, AIM, GIM, KHM, SIM, SPM until 2011 when NDM was added to the panel tested), and class D (OXA-23-like, OXA-24-like, OXA-48-like, OXA-51-like, OXA-58-like) carbapenemases. For *A. baumannii*, the PCR panel was limited to class D (OXA-23, OXA-40, OXA-48, OXA-51, OXA-58, OXA-143) carbapenemases. Loss of outer-membrane porin expression was inferred by interpreting the PCR results in conjunction with the phenotypic resistance patterns. PFGE for the *E. cloacae* isolates was performed using XbaI as previously described,[2] whereas SmaI was used for the *E. faecium* isolates.

DNA extraction

Isolates were retrieved from –80 °C stocks, from which secondary bacterial stocks were prepared by growing a single colony in brain heart infusion broth (BHI) (Colindale, UK) at 37 °C in air for 18 hours (with the exception of the *N. meningitidis* isolates, which were grown for 48 hours in 5% CO_2). DNA was then extracted using the QIAamp DNA Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions using the Gram-positive or Gram-negative protocol. For the Gram-positive protocol, one change to the standard procedure was made; rather than incubating the sample for 15 min at 95 °C, which resulted in significant DNA degradation, we incubated for 10 min at 70 °C.

Sequencing and bioinformatic analyses

DNA fragment libraries were prepared using the Nextera kit (Illumina Inc., San Diego, USA), and 150bp paired end DNA sequences were determined using the Illumina MiSeq (Illumina Inc.), as previously described.[3] The exception was *A. baumannii* (strain AB223) for which only 58bp reads (associated with 80 fold coverage) were achieved for the reverse read due to a run interruption. Genome sequence data were analysed at the Wellcome Trust Sanger Institute, blinded to all clinical, antimicrobial resistance pattern and epidemiological data. Sequencing reads were mapped to complete reference genome sequences (EMBL accession CP001918 for *E. cloacae*, CP003583 for *E. faecium*, and AE002098.2 for *N. meningitidis*) using SMALT. Single nucleotide polymorphisms (SNPs) were identified using a standard approach[4] by removing SNPs with low quality scores and by filtering for SNPs that were present in at least 75% of the mapped reads. The resulting whole genome alignments were used to estimate maximum likelihood phylogenies using RAxML.[5] For *E. cloacae* and *N. meningitidis* the mobile genetic elements such as phage and transposases were excluded from the alignment. The general time-reversible model with gamma correction was used for among-site variation. For the high resolution phylogeny of the *E. cloacae* outbreak isolates the genome of the earliest isolate (EC1a) was assembled de novo using Velvet[6] and used as the reference for read mapping. Contigs with less than 500bp were excluded. In addition, SNP calls were visually curated to exclude those with low quality, ambiguous mapping or a position at the end of contigs, leaving a reference of 4,886,929bp. The genome sequence data have been deposited in the European Nucleotide Archive (eTable 1).

For detection of antibiotic resistance determinants, a reference of relevant genes was compiled (eTable 3). The genomes of test isolates were assembled *de novo* as described above and reference resistance genes then mapped against test assemblies using SMALT, allowing the same gene to map multiple times to the assembly using a cutoff for detection of 90% DNA sequence identity. Alignments terminating at contig ends were

retained if at least the last 20 bases of the contig mapped with greater than or equal to 90% identity to one end of the reference gene, which allows the identification of potential matches in difficult-to-assemble, repetitive regions. In cases where multiple genes mapped at the same location of the assembly, the best match gene based on percentage identity was retained. To further assess whether the candidate genes identified from the assembly were present, the raw sequencing reads were mapped to each candidate gene, allowing assessment of coverage and SNP/indel variation in the isolate of interest. To investigate potential loss of function in porin genes, the assemblies were searched for the amino acid sequence of the porin gene of the respective species (eTable 4). Hits were confirmed using BLAST searches, checked for length compared to the reference, and respective disruptions noted.

For the analysis of the *N. meningitidis* genomes, single locus gene typing (PorA variable region (VR1 & VR2), *porB* allele, FetA variable region, and *fHbp* allele) was achieved using an automated online tool (PubMLST Neisseria BIGSdb database, http://pubmlst.org/neisseria/).[7] Capsular typing was performed by compiling a pseudomolecule consisting of *csa*, *csb*, *csc*, *csw*, and *csy* genes of respective reference genomes (eTable 5). This pseudomolecule was used to blast against the assemblies as in the approach used for the detection of antibiotic resistance determinants. Results were visualized as intensities and manually inspected to correctly derive the serogroup (eFigure 1).

eTable 1: Bacterial isolates sequenced in this study

Strain name used in study	Study patient Code	Strain name in ENA	ENA accession number	Organism
EC1a	EC1	EC606	ERS184249	*Enterobacter cloacae*
EC2a	EC2	EC875	ERS184250	*Enterobacter cloacae*
EC2b	EC2	EC293	ERS184251	*Enterobacter cloacae*
EC3a	EC3	EC262	ERS184252	*Enterobacter cloacae*
EC4a	EC4	EC543	ERS184245	*Enterobacter cloacae*
EC5a	EC5	EC146	ERS184246	*Enterobacter cloacae*
EC6a	EC6	EC356	ERS184247	*Enterobacter cloacae*
EC7a	EC7	EC507	ERS184248	*Enterobacter cloacae*
EF1a	EF1	EF301	ERS184258	*Enterococcus faecium*
EF2a	EF2	EF509	ERS184259	*Enterococcus faecium*
EF3a	EF3	EF705	ERS184260	*Enterococcus faecium*
EF3b	EF3	EF920	ERS184261	*Enterococcus faecium*
EF3c	EF3	EF752	ERS184262	*Enterococcus faecium*
EF3d	EF3	EF605	ERS184263	*Enterococcus faecium*
EF2b	EF2	EF359	ERS184264	*Enterococcus faecium*
EF4a	EF4	EF360	ERS184265	*Enterococcus faecium*
AB223	n/a[a]	AB223	ERS184244	*Acinetobacter baumannii*
EC302	n/a[a]	EC302	ERS184253	*Enterobacter cloacae*
Eco216	n/a[a]	Eco216	ERS184257	*Escherichia coli*
KP652	n/a[a]	KP652	ERS184255	*Klebsiella pneumoniae*
NM1	NM1	NM125	ERS184237	*Neisseria meningitidis*
NM2	NM2	NM499	ERS184236	*Neisseria meningitidis*
NM3	NM3	NM642	ERS184238	*Neisseria meningitidis*
NM4	NM4	NM456	ERS184239	*Neisseria meningitidis*
NM5	NM5	NM804	ERS184240	*Neisseria meningitidis*
NM6	NM6	NM205	ERS184241	*Neisseria meningitidis*
NM7	NM7	NM052	ERS184242	*Neisseria meningitidis*

[a] Four isolates were used that were not linked to a patient described in the study

eTable 2: Molecular determinants involved in beta-lactam resistance identified in this study

Organism	Isolate number	Source	Acquired beta-lactamase genes detected by WGS				Outer membrane porin (Omp) genes	MIC (mg/L)									
			Carbapenemases	ESBL	AmpC[a]	Others		IPM	MEM	ETP	CTX	CTX/CLA	CAZ	CAZ/CLA	CTX/CLO	FOX	TZP
A. baumannii	AB223	Ascitic fluid	**OXA-23**	-	-	-	Not examined	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
E. cloacae	EC1a	Blood	**IMP-1**	SHV-12	-	TEM-1[b]	OmpC truncated, potentially due to *IS* element and OmpF – premature stop codon (W77STOP)	16	16	>16	128	>32	>256	>32	128	>64	>64
E. cloacae	EC302	Urine	-	-	-	-	**OmpF – premature stop codon (S106STOP)**	0.5	1	8	>256	>32	256	>32	128	>64	>64
K. pneumoniae	KP652	Urine	-	**CTX-M-15**	-	SHV-133[c]	**OmpK36 – premature stop codon (W44STOP)**	2	2	16	>256	0.5	256	1	>256	16	>64
E. coli	Eco216	Urine	-	**CTX-M-15**	-	TEM-1[b]	**OmpF – frameshift mutation at position 252**	0.25	2	>16	>256	8	256	8	>256	>64	>64

Text in bold denotes the dominant mechanism(s) of resistance.
Abbreviations: WGS, whole-genome sequencing; ESBL, extended-spectrum beta-lactamase; MIC, minimum inhibitory concentration; ETP, ertapenem; IPM, imipenem; MEM, meropenem; CTX, cefotaxime; CTX/CLA, cefotaxime and clavulanic acid; CAZ, ceftazidime; CAZ/CLA, ceftazidime and clavulanic acid; CTX/CLO, cefotaxime and cloxacillin; FOX, cefoxitime; TZP, piperacillin and tazobactam. ND, not done.

[a]This does not include chromosomal AmpC beta-lactamases.
[b]TEM-1 is a penicillinase and has no significant ESBL activity.[8]
[c]bla_{SHV} gene identical to SHV-133 (GenBank accession # AB551737). No information available regarding the activity spectrum of this enzyme.

eTable 3: List of antibiotic resistance genes used for analysis of Gram-negative bacilli

Gene	Accession number
blaSCO-1	EF063111
blaTEM	EU977570
blaSHV-5	EF653399
blaVEB-1	AF010416
blaPER-1	ABQ10556
ampC (*Acinetobacter*)	CP001172
blaACC-1	AJ133121
blaFOX-1	X77455
blaCMY-1	X92508
blaDHA-1	AJ237702
blaLAT-1	X78117
blaMIR-1	M37839
blaOXA-1	J02967.2
blaOXA-2	X07260
blaOXA-9	M55547
blaOXA-10	U37105.2
blaCTXM-1	FJ235693
blaCTXM-2	DQ125241
blaCTXM-8	AF189721
blaCTXM-9	AF252621.2
blaCTXM-25	AF518567.2
blaCTXM-64	GQ456156
blaOXA-23	AJ132105
blaOXA-24	AF509241
blaOXA-48	AY236073
blaOXA-50	AY306130
blaOXA-51	AJ309734
blaOXA-55	AY343493
blaOXA-58	AY665723
blaOXA-60	AF525303
NMC-R/NMC-A	Z21956
blaSME-1	JF974075
blaGES	FJ696404
blaKPC-1	AF297554
blaIMI-R/IMI-2	JN412066
blaIMP-1	S71932
blaIMP-2	AJ243491
blaIMP-4	DQ532122
blaIMP-5	AF290912
blaIMP-7	AF416736.2
blaIMP-13	AJ550807
blaIMP-14	FN397627
blaIMP-16	AJ584652.2
blaIMP-18	EF184215
blaIMP-22	FM876313
blaIMP-29	HQ438058
blaVIM-1	EF690696
blaNDM-1	JN872329
blaSPM	AJ492820
blaGIM	JF414726
blaAIM	AM998375
blaDIM	GU323019
blaKHM	AB364006
blaSIM	AY887066

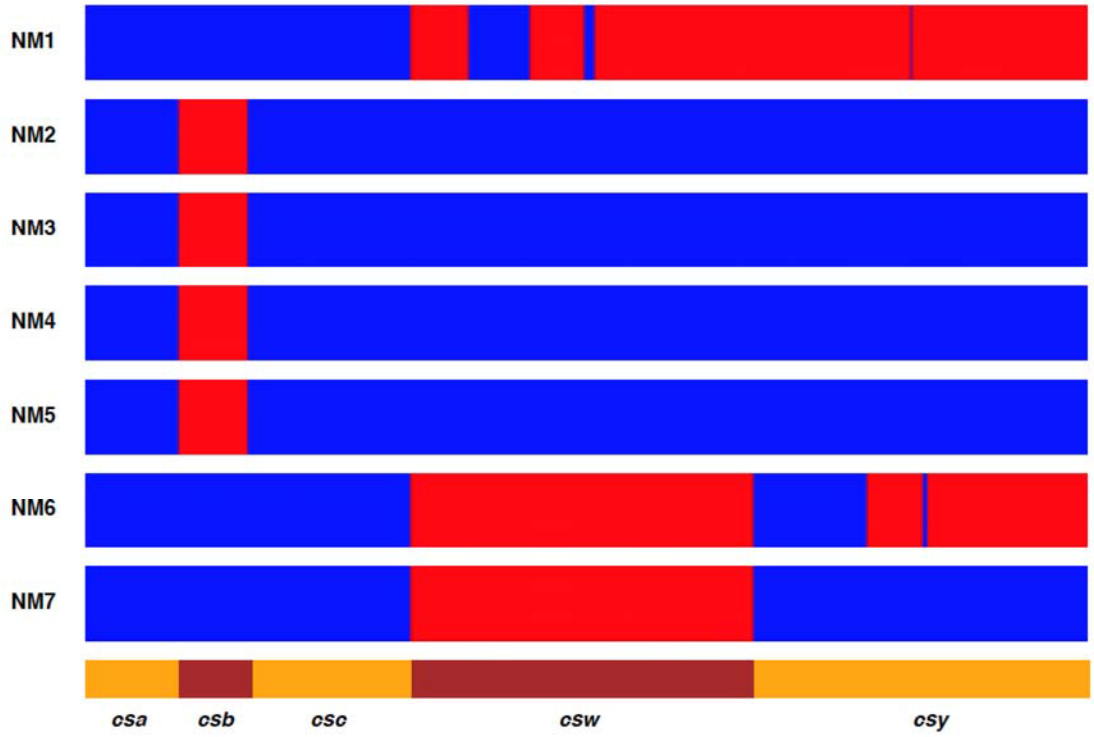
eTable 4: List of porin genes

Gene	Accession number
Escherichia coli ompC	U00096.2
Escherichia coli ompF	U00096.2
Enterobacter cloacae ompC	AJ316539.1
Enterobacter cloacae ompF	AJ316540.1
Klebsiella pneumoniae ompK35	AJ303057.1
Klebsiella pneumoniae ompK36	Z33506.1

eTable 5: List of accession numbers for genes used for serogrouping *N. meningitidis*

Gene	Serogroup	Accession number
Neisseria meningitidis Z2491	A	AL157959.1
Neisseria meningitidis MC58	B	AE002098.2
Neisseria meningitidis FAM18	C	AM421808.1
Neisseria meningitidis	W135	HF562987
Neisseria meningitidis	Y	HF562989

eFigure 1: Heat map showing the presence (red) or absence (blue) of genes encoding capsule serogroups



Note: Some cross-mapping occurs between genes for serogroups W and Y given the similarity of the *csw* and *csy* genes.

eReferences

1. British Society of Antimicrobial Chemotherapy. BSAC methods for antimicrobial susceptibility testing. Version 11.1, May 2012. 2012.
2. Shet V, Gouliouris T, Brown NM, Turton JF, Zhang J, Woodford N. IMP metallo-β-lactamase-producing clinical isolates of *Enterobacter cloacae* in the UK. *J Antimicrob Chemother.* Jun 2011;66(6):1408-1409.
3. Köser CU, Holden MT, Ellington MJ, et al. Rapid whole-genome sequencing for investigation of a neonatal MRSA outbreak. *N Engl J Med.* Jun 14 2012;366(24):2267-2275.
4. Harris SR, Feil EJ, Holden MT, et al. Evolution of MRSA during hospital transmission and intercontinental spread. *Science.* Jan 2010;327(5964):469-474.
5. Stamatakis A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics.* Nov 1 2006;22(21):2688-2690.
6. Zerbino DR, Birney E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* May 2008;18(5):821-829.
7. Jolley KA, Maiden MC. BIGSdb: Scalable analysis of bacterial genome variation at the population level. *BMC Bioinformatics.* 2010;11:595.
8. Bush K, Jacoby GA, Medeiros AA. A functional classification scheme for beta-lactamases and its correlation with molecular structure. *Antimicrob Agents Chemother.* Jun 1995;39(6):1211-1233