J Antimicrob Chemother 2017; **72**: 2764–2768
doi:10.1093/jac/dkx217 Advance Access publication 19 July 2017

Journal of Antimicrobial Chemotherapy

PointFinder: a novel web tool for WGS-based detection of antimicrobial resistance associated with chromosomal point mutations in bacterial pathogens

Ea Zankari[1,2], Rosa Allesøe[2], Katrine G. Joensen[3], Lina M. Cavaco[1], Ole Lund[2] and Frank M. Aarestrup[1]*

[1]National Food Institute, Technical University of Denmark, 2800 Kgs Lyngby, Denmark; [2]Center for Biological Sequence Analysis, Department of Systems Biology, Technical University of Denmark, 2800 Kgs Lyngby, Denmark; [3]Department of Microbiology and Infection Control, Statens Serum Institut, Copenhagen, Denmark

*Corresponding author. Tel: +45-35886281; E-mail: fmaa@food.dtu.dk

Received 20 March 2017; returned 19 April 2017; revised 2 June 2017; accepted 7 June 2017

Background: Antibiotic resistance is a major health problem, as drugs that were once highly effective no longer cure bacterial infections. WGS has previously been shown to be an alternative method for detecting horizontally acquired antimicrobial resistance genes. However, suitable bioinformatics methods that can provide easily interpretable, accurate and fast results for antimicrobial resistance associated with chromosomal point mutations are still lacking.

Methods: Phenotypic antimicrobial susceptibility tests were performed on 150 isolates covering three different bacterial species: *Salmonella enterica*, *Escherichia coli* and *Campylobacter jejuni*. The web-server ResFinder-2.1 was used to identify acquired antimicrobial resistance genes and two methods, the novel PointFinder (using BLAST) and an in-house method (mapping of raw WGS reads), were used to identify chromosomal point mutations. Results were compared with phenotypic antimicrobial susceptibility testing results.

Results: A total of 685 different phenotypic tests associated with chromosomal resistance to quinolones, polymyxin, rifampicin, macrolides and tetracyclines resulted in 98.4% concordance. Eleven cases of disagreement between tested and predicted susceptibility were observed: two *C. jejuni* isolates with phenotypic fluoroquinolone resistance and two with phenotypic erythromycin resistance and five colistin-susceptible *E. coli* isolates with a detected *pmrB* V161G mutation when assembled with Velvet, but not when using SPAdes or when mapping the reads.

Conclusions: PointFinder proved, with high concordance between phenotypic and predicted antimicrobial susceptibility, to be a user-friendly web tool for detection of chromosomal point mutations associated with antimicrobial resistance.

Introduction

Horizontal gene transfer among bacterial isolates is often considered the main mediator of acquired antimicrobial resistance. However, mutational resistance is another important way to confer resistance.

It has previously been shown that WGS can be an alternative to phenotypic susceptibility testing of bacterial isolates for detection of horizontally acquired resistance.[1,2] Databases for mapping to chromosomal mutations have also been developed for *Mycobacterium tuberculosis*.[3] However, at present there is a lack of suitable bioinformatics methods to provide easily interpretable results for antimicrobial resistance associated with chromosomal point mutations for most bacterial species.

In this study, a novel web tool, PointFinder, was developed for detection of chromosomal point mutations associated with antimicrobial resistance, in bacterial WGS data. PointFinder may be run in parallel and become an extension to the already existing web server tool ResFinder,[1] which detects horizontally acquired resistance genes in WGS data. The performance was compared with that of an in-house mapping method for detecting point mutations and both results were compared with phenotypic antimicrobial susceptibility tests in order to validate the possibilities of using these methods as alternatives to standard phenotypical testing.

Table 1. Overview of chromosomal point mutations for each species included in the database

Species	Gene	Chromosomal mutations	Resistance	Reference(s)
E. coli	*gyrA*	A51V, A67S, G81C, D82G, S83L, S83W, S83A, S83V, S83I, A84P, A84V, D87N, D87G, D87Y, D87H, D87V, Q106H, Q106R, A196E	quinolone	4
	gyrB	R136L, R136C, R136H, R136S, R136G, R136I, R136E, D426N, K447E	quinolone	4,5
	parC	A56T, S57T, F60I, F60L, G78D, G78K, S80R, S80I, S80L, S80Y, S80F, E84G, E84K, E84V, E84A, A108V, A108T	quinolone	4,6,7
	parE	L416F, G423R, P439S, I444F, S458T, E460D, E460K, I464F, I470M, D475E, D476N, I529L	quinolone	4,6,7
	pmrA	S39I, R81S	colistin	8
	pmrB	V161G	colistin	8
	folP	P63R, P64L, P64S, P64A, P64H	sulphonamides	9
	rpoB	V146F, Q513L, Q513P, H526Y, R529C, R529S, S531F, L533P, T563P, P564L, R687H	rifamycin	10
	23S[a]	A2059G	macrolide	11
	16S *rrsB*[a]	A523C, G527T, C528T, G1064T, G1064C, G1064A, C1066T, G1068A	spectinomycin	12–15
	16S *rrsB*[a]	A964G, G1053A, C1054T, A1055G, G1058C	tetracycline	13,16
	16S *rrsB*[a]	T1406A, A1408G	gentamicin	17
	16S *rrsC*[a]	A794G, A794T, G926A, G926T, G926C, A1519G, A1519C, A1519T	kasugamycin	18
	16S *rrsH*[a]	C1192T	spectinomycin	19
S. enterica	*gyrA*	A67P, D72G, V73I, G81C, G81S, G81H, G81D, D82G, D82N, S83Y, S83F, S83A, D87N, D87G, D87Y, D87K, L98V, A119S, A119E, A119V, A131G, E139A	quinolone	4
	gyrB	Y421C, R438L, S464Y, S464F, E466D	quinolone	4,20
	parC	T66I, G78D, S80R, S80I, E84K, E84G	quinolone	4,21
	parE	M438I, E454G, S458P, V461G, H462Y, A499T, V514G, V521F	quinolone	4,20,22
	pmrA	G15R, G53E, G53R, R81C, R81H	colistin	23
	pmrB	L22P, S29R, T92A, P94Q, E121A, S124P, N130Y, T147P, R155P, T156P, T156M, V161M, V161L, V161G, E166K, M186I, G206W, G206R, S305R	colistin	23
	16S *rrsD*[a]	C1065T, C1192T	spectinomycin	24
C. jejuni	*gyrA*	A70T, D85T, T86I, T86A, T86K, T86V, D90A, D90N, D90T, P104S	quinolone	25–28
	23S[a]	A2074G, A2074T, A2074C, A2075G	macrolide	28
	cmeR	A86G	macrolide	29
	rplV	A103C	macrolide	29
	rpsL	K88E, K88R, K88Q	spectinomycin	30

[a]rRNA gene, mutation shown in DNA.

Materials and methods

Chromosomal mutation database

Information regarding mutations in chromosomal genes associated with antimicrobial resistance was collected from published papers (Table 1). The reference sequences were selected from WT *Escherichia coli* strain K-12 (MG1655) for the *E. coli* database, *Salmonella* Typhimurium strain LT2 for the *Salmonella enterica* database and *Campylobacter jejuni* NCTC 11168 for the *C. jejuni* database.

Bacterial isolates and WGS data

In total, 150 isolates covering three species were included in the study: *E. coli* ($n = 50$) and *Salmonella* ($n = 50$) isolates from the in-house strain collection at the National Food Institute and *C. jejuni* ($n = 50$) isolates from the in-house strain collection at Statens Serum Institut. The isolates were selected on the basis of having both WGS data and phenotypes available. The *Salmonella* isolates included strains from 10 different serovars (Tables S1 to S3, available as Supplementary data at *JAC* Online). All bacterial isolates were sequenced using the Miseq platform (Illumina) to obtain paired-end sequences and assembled *de novo* using Velvet (reference software). Bacterial strains were screened for phenotypic resistance using MIC determinations interpreted according to EUCAST (www.eucast.org). Only the susceptibility tests relevant for antimicrobial resistance associated with chromosomal point mutations for each species were analysed (Table 2). As resistance to some of the antimicrobial agents can be caused by either acquired genes or chromosomal point mutations, ResFinder-2.1 (www.genomicepidemiology.org)[31] was used to detect known acquired resistance genes in the WGS data, using a threshold of 98% identity (%ID) and 60% length (minimum percentage length of the resistance gene to be covered). All isolates with disagreement between the phenotypic and predicted susceptibility were re-tested.

PointFinder

PointFinder consists of two databases: a chromosomal gene database, with all reference sequences in fasta format; and a chromosomal mutation database containing information on codon positions and substitutions. PointFinder uses BLASTn for identifying the best match for each gene in the chromosomal gene database, and only hits with an identity of $\geq$80% are further analysed. The program goes through each alignment comparing each position for the query (sequence found in input sequence) with the corresponding position in the subject (database sequence). All mismatches are saved and compared with the chromosomal mutation database. It is

possible for users to select whether they want to see all mismatches or only known mismatches found in positions from the chromosomal database. In this study we have only looked at mismatches found in positions known to confer resistance, and thus specified in the database.

Mapping method

The fastq files corresponding to the paired-end reads were mapped against the chromosomal gene sequence database using the assimpler.py script described in Joensen *et al.*[32] In brief, 17mers from the reads were mapped to the reference sequence and extended to ungapped alignments that were considered significant if they had a score of at least 50, using a match score of 1 and a mismatch score of −3. A base was called if $Z = (X - Y)/\sqrt{(X + Y)}$ was >3.29, where X is the number of observations of the most common nucleotide and Y is the number of other nucleotides at that position. Furthermore, nucleotide calls were considered significant only when the most common nucleotide was at least 10 times more abundant than other nucleotides at the position. All mismatches in positions from the chromosomal mutation database were outputted except silent mutations, which were discarded. In cases with disagreement between PointFinder and mapping, the isolates were re-assembled *de novo* using SPAdes[33] and re-analysed by PointFinder.

Results

MIC and predicted antimicrobial resistance

The 150 isolates were each tested against four to six different antimicrobial agents (Table 2), leading to a total of 684 susceptibility test results (Tables S1–S3). These results were compared with the results from PointFinder, mapping and ResFinder. Resistance to colistin, sulphonamides, tetracycline, erythromycin and spectinomycin can be caused by both chromosomal point mutations and acquired resistance genes; therefore results from both PointFinder and ResFinder were used to explain resistance.

Table 2. Antimicrobial agents used for susceptibility tests for each species

Species	Antimicrobial agents
E. coli	ciprofloxacin, nalidixic acid, colistin, sulphonamide, tetracycline, spectinomycin
Salmonella	ciprofloxacin, nalidixic acid, colistin, spectinomycin
C. jejuni	ciprofloxacin, nalidixic acid, erythromycin, spectinomycin

Acquired resistance genes, chromosomal point mutations or both can cause resistance to antimicrobial agents.

For all *Salmonella* isolates, complete agreement between tested and predicted susceptibility was observed (Tables S1–S3). Disagreements in *E. coli* and *C. jejuni* were observed in five and four cases, respectively (Table 3).

The point mutation *pmrB* V161G was found by PointFinder in five *E. coli* isolates (E30–E34), but all tested phenotypically susceptible to colistin (MIC ≤1 mg/L). In *C. jejuni*, two isolates (C23 and C39) tested phenotypically resistant to ciprofloxacin (MIC 8) and nalidixic acid (MIC >64 mg/L), while two (C8 and C24) tested erythromycin resistant (MIC >128 mg/L), but neither mutations nor acquired genes were found that could explain the resistance.

PointFinder versus mapping

Mapping and PointFinder found the same mutations in all isolates except the five *pmrB* V161G mutations found by PointFinder in *E. coli* strains (E30–E34). The five isolates were re-assembled *de novo* using SPAdes[33] and run through PointFinder, and this time no mutations were found in *pmrB* in any of the isolates. The codon change detected in the Velvet assembly of the five isolates was GTG→GGG, and when looking further into the sequences, the mapping showed that 28%–37% (Table 4) of the reads mapping to *pmrB* contained GGG instead of GTG.

Discussion

This study showed a high agreement between phenotypic susceptibility tests and WGS-predicted resistance, with only 11 (1.6%) mismatches. However, since the number of isolates included in the evaluation was very limited and selected, this has to be further verified in future studies. The six disagreements observed in *C. jejuni* all involved predicted susceptibility, whereas the isolates were phenotypically resistant, which may be due to unknown novel genes or mutations, as neither ResFinder nor PointFinder can detect novel resistance mechanisms.

We found that the BLAST-based method was dependent on the assembly method, which can cause either false-positive or -negative results. As the mapping method does not depend on the assembly this method gives a more precise result, which is consistent with a recent study by Clausen *et al.*[34] Exploring the sequences mapping to *pmrB*, we found that ∼1/3 of the isolate sequences for

Table 3. Disagreements between phenotypic and predicted resistance

Predicted genotype			Conventional test			
PointFinder	mapping	ResFinder	resistant	susceptible	No. of isolates	Isolate ID
pmrB V161G	–	–	–	CIP, NAL, **CST**, SMX, TET	4	E30, E32, E33, E34
pmrB V161G	–	*tet*(B)	TET	CIP, NAL, **CST**, SMX	1	E31
–	–	–	**ERY**	CIP, NAL, SPE	1	C8
–	–	–	**CIP**, **NAL**	ERY, SPE	2	C23, C39
gyrA T86I, *gyrA* P104S	*gyrA* T86I, *gyrA* P104S	–	CIP, NAL, **ERY**	SPE	1	C24

CIP, ciprofloxacin; CST, colistin; ERY, erythromycin; NAL, nalidixic acid; SMX, sulphonamide; TET, tetracycline; SPE, spectinomycin.
Bold, mismatch between predicted and conventional results.

Table 4. Mapped sequences to *pmrB* position 161 (amino acid position)

Isolate ID	Total mapped sequences	No. of sequences mapping to	
		GTG (%)	GGG (%)
E30	69	50 (72)	19 (28)
E31	70	44 (63)	26 (37)
E32	80	54 (68)	26 (33)
E33	70	47 (67)	23 (33)
E34	89	64 (72)	25 (28)

each isolate contained the V161G mutation, indicating that there may be more than one copy of the *pmrB* gene present or that the sequence consisted of more than a single isolate. For many professionals working with WGS data, assemblies are still the preferred format. Due to their smaller size, assemblies are easier to share, upload and manage. Therefore, users working with assembly-based methods should consider that the data quality and method of assembly might influence the output.

Both ARG-ANNOT[35] and CARD[36] have tried to incorporate chromosomal point mutations in their databases. ARG-ANNOT has a database with partial sequences for chromosomal mutational regions of genes associated with mutational resistance, as well as information about position and mutation in the corresponding gene. ARG-ANNOT does not automatically detect these mutations, so the user has to manually browse through the alignment to detect potential mutations. CARD's resistance gene identifier (RGI) protein variant models use curated SNP matrices to detect and report mutations associated with resistance.[37] Unfortunately, neither ARG-ANNOT nor CARD takes the bacterial species into account. This means that both methods also output possible mutations/sequences related to mutational resistance, which is not relevant for the bacteria in question. The user must therefore have prior knowledge of which mutational genes and specific mutations they are looking for in order to use these methods. To cope with some of these problems, we have developed PointFinder, with the purpose of facilitating user-friendly detection of chromosomal point mutations associated with resistance. In addition to being user friendly, the output from the web tool is easily understandable, reporting the detected mutations, nucleotide and amino acid codon changes, predicted resistance and links to papers describing the detected mutations. In the current version this covers mutations conferring resistance to quinolones, macrolides and polymyxin in *E. coli, Salmonella* and *C. jejuni*, but will be developed continuously with additional species.

Conclusions

This study showed a high concordance between phenotypic antimicrobial susceptibility and predicted genotype by ResFinder and PointFinder from WGS data. PointFinder is a user-friendly method for detection of chromosomal point mutations associated with antimicrobial resistance.

Funding

This study was supported by the Center for Genomic Epidemiology (http://www.genomicepidemiology.org/), grant 09–067103/DSF from the Danish Council for Strategic Research and has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 643476.

Transparency declarations

None to declare.

Supplementary data

Tables S1 to S3 are available as Supplementary data at *JAC* Online.

References

1 Zankari E, Hasman H, Kaas RS *et al*. Genotyping using whole-genome sequencing is a realistic alternative to surveillance based on phenotypic antimicrobial susceptibility testing. *J Antimicrob Chemother* 2013; **68**: 771–7.

2 Stoesser N, Batty EM, Eyre DW *et al*. Predicting antimicrobial susceptibilities for *Escherichia coli* and *Klebsiella pneumoniae* isolates using whole genomic sequence data. *J Antimicrob Chemother* 2013; **68**: 2234–44.

3 Bradley P, Gordon NC, Walker TM *et al*. Rapid antibiotic-resistance predictions from genome sequence data for *Staphylococcus aureus* and *Mycobacterium tuberculosis*. *Nat Commun* 2015; **6**: 10063.

4 Hopkins KL, Davies RH, Threlfall EJ. Mechanisms of quinolone resistance in *Escherichia coli* and *Salmonella*: recent developments. *Int J Antimicrob Agents* 2005; **25**: 358–73.

5 Confreres A, Maxwell A. *gyrB* mutations which confer coumarin resistance also affect DNA supercoiling and ATP hydrolysis by *Escherichia coli* DNA gyrase. *Mol Microbiol* 1992; **6**: 1617–24.

6 Cavaco LM, Frimodt-Møller N, Hasman H *et al*. Prevalence of quinolone resistance mechanisms and associations to minimum inhibitory concentrations in quinolone-resistant *Escherichia coli* isolated from humans and swine in Denmark. *Microb Drug Resist* 2008; **14**: 163–9.

7 Nawaz M, Sung K, Kweon O *et al*. Characterisation of novel mutations involved in quinolone resistance in *Escherichia coli* isolated from imported shrimp. *Int J Antimicrob Agents* 2015; **45**: 471–6.

8 Quesada A, Porrero MC, Téllez S *et al*. Polymorphism of genes encoding PmrAB in colistin-resistant strains of *Escherichia coli* and *Salmonella enterica* isolated from poultry and swine. *J Antimicrob Chemother* 2015; **70**: 71–4.

9 Vedantam G, Guay GG, Austria NE *et al*. Characterization of mutations contributing to sulfathiazole resistance in *Escherichia coli*. *Antimicrob Agents Chemother* 1998; **42**: 88–93.

10 Jin DJ, Gross CA. Mapping and sequencing of mutations in the *Escherichia coli rpoB* gene that lead to rifampicin resistance. *J Mol Biol* 1988; **202**: 45–58.

11 Jensen LB, Aarestrup FM. Macrolide resistance in *Campylobacter coli* of animal origin in Denmark. *Antimicrob Agents Chemother* 2001; **45**: 371–2.

12 Melancon P, Lemieux C, Brakier-Gingras L. A mutation in the 530 loop of *Escherichia coli* 16S ribosomal RNA causes resistance to streptomycin. *Nucleic Acids Res* 1988; **16**: 9631–9.

13 Yassin A, Fredrick K, Mankin AS. Deleterious mutations in small subunit ribosomal RNA identify functional sites and potential targets for antibiotics. *Proc Natl Acad Sci USA* 2005; **102**: 16620–5.

14 Makosky PC, Dahlberg AE. Spectinomycin resistance at site 1192 in 16S ribosomal RNA of *E. coli*: an analysis of three mutants. *Biochimie* 1987; **69**: 885–9.

15 Brink MF, Brink G, Verbeet MP *et al*. Spectinomycin interacts specifically with the residues G1064 and C1192 in 16S rRNA, thereby potentially freezing this molecule into an inactive conformation. *Nucleic Acids Res* 1994; **22**: 325–31.

16 Ross JI, Eady EA, Cove JH *et al*. 16S rRNA mutation associated with tetracycline resistance in a Gram-positive bacterium. *Antimicrob Agents Chemother* 1998; **42**: 1702–5.

17 Recht MI, Puglisi JD. Aminoglycoside resistance with homogeneous and heterogeneous populations of antibiotic-resistant ribosomes. *Antimicrob Agents Chemother* 2001; **45**: 2414–9.

18 Vila-Sanjurjo A, Squires CL, Dahlberg AE. Isolation of kasugamycin resistant mutants in the 16 S ribosomal RNA of *Escherichia coli*. *J Mol Biol* 1999; **293**: 1–8.

19 Sigmund CD, Ettayebi M, Morgan EA. Antibiotic resistance mutations in 16S and 23S ribosomal RNA genes of *Escherichia coli*. *Nucleic Acids Res* 1984; **12**: 4653–63.

20 O'Regan E, Quinn T, Pagès J-M *et al*. Multiple regulatory pathways associated with high-level ciprofloxacin and multidrug resistance in *Salmonella enterica* serovar Enteritidis: involvement of RamA and other global regulators. *Antimicrob Agents Chemother* 2009; **53**: 1080–7.

21 Eaves DJ, Randall L, Gray DT *et al*. Prevalence of mutations within the quinolone resistance-determining region of *gyrA*, *gyrB*, *parC*, and *parE* and association with antibiotic resistance in quinolone-resistant *Salmonella enterica*. *Antimicrob Agents Chemother* 2004; **48**: 4012–5.

22 Thong KL, Ngoi ST, Chai LC *et al*. Quinolone resistance mechanisms among *Salmonella enterica* in Malaysia. *Microb Drug Resist* 2016; **22**: 259–72.

23 Sun S, Negrea A, Rhen M *et al*. Genetic analysis of colistin resistance in *Salmonella enterica* serovar Typhimurium. *Antimicrob Agents Chemother* 2009; **53**: 2298–305.

24 O'Connor M, Dahlberg AE. Isolation of spectinomycin resistance mutations in the 16S rRNA of *Salmonella enterica* serovar Typhimurium and expression in *Escherichia coli* and *Salmonella*. *Curr Microbiol* 2002; **45**: 429–33.

25 Wang Y, Huang WM, Taylor DE. Cloning and nucleotide sequence of the *Campylobacter jejuni gyrA* gene and characterization of quinolone resistance mutations. *Antimicrob Agents Chemother* 1993; **37**: 457–63.

26 Alfredson DA, Korolik V. Antibiotic resistance and resistance mechanisms in *Campylobacter jejuni* and *Campylobacter coli*. *FEMS Microbiol Lett* 2007; **277**: 123–32.

27 Engberg J, Aarestrup FM, Taylor DE *et al*. Quinolone and macrolide resistance in *Campylobacter jejuni* and *C. coli*: resistance mechanisms and trends in human isolates. *Emerg Infect Dis* 2001; **7**: 24–34.

28 Payot S, Bolla J-M, Corcoran D *et al*. Mechanisms of fluoroquinolone and macrolide resistance in *Campylobacter* spp. *Microbes Infect* 2006; **8**: 1967–71.

29 Martinez A, Lin J. Effect of an efflux pump inhibitor on the function of the multidrug efflux pump CmeABC and antimicrobial resistance in *Campylobacter*. *Foodborne Pathog Dis* 2006; **3**: 393–402.

30 Olkkola S, Juntunen P, Heiska H *et al*. Mutations in the *rpsL* gene are involved in streptomycin resistance in *Campylobacter coli*. *Microb Drug Resist* 2010; **16**: 105–10.

31 Zankari E, Hasman H, Cosentino S *et al*. Identification of acquired antimicrobial resistance genes. *J Antimicrob Chemother* 2012; **67**: 2640–4.

32 Joensen KG, Scheutz F, Lund O *et al*. Real-time whole-genome sequencing for routine typing, surveillance, and outbreak detection of verotoxigenic *Escherichia coli*. *J Clin Microbiol* 2014; **52**: 1501–10.

33 Bankevich A, Nurk S, Antipov D *et al*. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 2012; **19**: 455–77.

34 Clausen PT, Zankari E, Aarestrup FM *et al*. Benchmarking of methods for identification of antimicrobial resistance genes in bacterial whole genome data. *J Antimicrob Chemother* 2016; **71**: 2484–8.

35 Gupta SK, Padmanabhan BR, Diene SM *et al*. ARG-ANNOT, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. *Antimicrob Agents Chemother* 2014; **58**: 212–20.

36 McArthur AG, Waglechner N, Nizam F *et al*. The comprehensive antibiotic resistance database. *Antimicrob Agents Chemother* 2013; **57**: 3348–57.

37 McArthur AG, Wright GD. Bioinformatics of antimicrobial resistance in the age of molecular epidemiology. *Curr Opin Microbiol* 2015; **27**: 45–50.