Clinical Microbiology and Infection 22 (2016) 1002.e1–1002.e8

Contents lists available at ScienceDirect

Clinical Microbiology and Infection

journal homepage: www.clinicalmicrobiologyandinfection.com

Original article

Using whole genome sequencing to identify resistance determinants and predict antimicrobial resistance phenotypes for year 2015 invasive pneumococcal disease isolates recovered in the United States☆

B.J. Metcalf [1], S. Chochua [1], R.E. Gertz Jr. [1], Z. Li [1], H. Walker [1], T. Tran [1], P.A. Hawkins [1], A. Glennen [2], R. Lynfield [2], Y. Li [1], L. McGee [1], B. Beall [1, *], Active Bacterial Core surveillance team

[1] *National Center for Immunization and Respiratory Diseases, Centers for Disease Control and Prevention, Atlanta, GA, USA*
[2] *Minnesota Department of Health, St Paul, MN, USA*

ARTICLE INFO

Article history:
Received 10 June 2016
Received in revised form
22 July 2016
Accepted 3 August 2016
Available online 17 August 2016

Editor: G. Lina

Keywords:
Accessory genome
Core genome
Minimum inhibitory concentrations
Pneumococcal
Whole genome

ABSTRACT

Our whole genome sequence (WGS) pipeline was assessed for accurate prediction of antimicrobial phenotypes. For 2316 invasive pneumococcal isolates recovered during 2015 we compared WGS pipeline data to broth dilution testing (BDT) for 18 antimicrobials. For 11 antimicrobials categorical discrepancies were assigned when WGS-predicted MICs and BDT MICs predicted different categorizations for susceptibility, intermediate resistance or resistance, ranging from 0.9% (tetracycline) to 2.9% (amoxicillin). For β-lactam antibiotics, the occurrence of at least four-fold differences in MIC ranged from 0.2% (meropenem) to 1.0% (penicillin), although phenotypic retesting resolved 25%–78% of these discrepancies. Non-susceptibility to penicillin, predicted by penicillin-binding protein types, was 2.7% (non-meningitis criteria) and 23.8% (meningitis criteria). Other common resistance determinants included *mef* (475 isolates), *ermB* (191 isolates), *ermB* + *mef* (48 isolates), *tetM* (261 isolates) and *cat* (51 isolates). Additional accessory resistance genes (*tetS*, *tet32*, *aphA-3*, *sat4*) were rarely detected (one to three isolates). Rare core genome mutations conferring erythromycin-resistance included a two-codon *rplD* insertion (*rplD*69-KG-70) and the 23S rRNA A2061G substitution (six isolates). Intermediate cotrimoxazole-resistance was associated with one or two codon insertions within *folP* (238 isolates) or the *folA* I100L substitution (38 isolates), whereas full cotrimoxazole-resistance was attributed to alterations in both genes (172 isolates). The two levofloxacin-resistant isolates contained *parC* and/or *gyrA* mutations. Of 11 remaining isolates with moderately elevated MICs to both ciprofloxacin and levofloxacin, seven contained *parC* or *gyrA* mutations. The two rifampin-resistant isolates contained *rpoB* mutations. WGS-based antimicrobial phenotype prediction was an informative alternative to BDT for invasive pneumococci.
B.J. Metcalf, CMI 2016;22:1002.e1–1002.e8

Introduction

For pneumococcal disease surveillance the two most important features derived from causal strain characterization are presently the capsular serotype and the antimicrobial susceptibility profile [1,2]. The distributions of both pneumococcal serotypes and susceptibility profiles are subject to rapid fluctuations in response to the current highly effective conjugate vaccines that target strains of specific serotypes. Capsular serotyping is of critical importance for evaluating the effectiveness of current vaccines and for generating serotype distribution data that could prove useful for formulating next-generation vaccines. Accurate determination of antimicrobial susceptibility is important both for immediate treatment of patients and for formulating general treatment guidelines for pneumococcal disease [2]. An automated whole genome sequence

☆ The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

* Corresponding author. B. Beall, National Center for Immunization and Respiratory Diseases, Centers for Disease Control and Prevention, Atlanta, GA, USA.
E-mail address: bbeall@cdc.gov (B. Beall).

http://dx.doi.org/10.1016/j.cmi.2016.08.001

(WGS) bioinformatics pipeline extends these features to the detailed identification of genotypes and characteristics such as surface protein antigens [1]. Theoretically, all strain phenotypes can be deduced from a single permanent WGS record. Obtaining key pneumococcal parameters through an automated bioinformatics pipeline ensures that all deduced features are linked to the same pneumococcal culture. Automated extraction of these features from WGS obviates much of the specialized technical skills and labour required for obtaining these phenotypes conventionally. Here we show results of a WGS bioinformatics pipeline predicting antimicrobial phenotypes, and compare these predictions with phenotypic testing of 2316 year 2015 invasive pneumococcal disease (IPD) isolates.

Materials and methods

Isolates

A convenience sampling of 2316 consecutive IPD isolates collected through Active Bacterial Core surveillance (ABCs), and representing approximately 82% of projected year 2015 cases, were subjected to WGS. The surveillance population is described and detailed depictions of IPD disease rates in the USA during the period 1998–2015 are provided at http://www.cdc.gov/pneumococcal/surveillance.html. The year 2015 IPD rates shown at this site include cases represented by the causal isolates that we describe in this study. Additionally, 738 of the year 2015 isolates used for our study presented here were among 2528 isolates previously employed for correlating β-lactam antibiotic MICs to penicillin-binding protein (PBP) types [3].

Whole genome sequencing

Streptococcus pneumoniae strains were cultured on Trypticase soy agar supplemented with 5% sheep blood and incubated overnight at 37°C in 5% CO_2. Genomic DNA for short-read WGS was extracted manually using a modified QIAamp DNA mini kit protocol (Qiagen, Inc., Valencia, CA, USA). Nucleic acid concentration was quantified by an Invitrogen Qubit assay (Thermo Fisher Scientific Inc., Waltham, MA, USA) and samples were sheared using a Covaris M220 ultrasonicator (Covaris, Inc., Woburn, MA, USA) programmed to generate 500-bp fragments. Libraries were constructed on the SciCloneG3 (PerkinElmer Inc., Waltham, MA, USA) using a TruSeq DNA PCR-Free HT library preparation kit with 96 dual indices (Illumina Inc., San Diego, CA, USA) and quantified by a KAPA qPCR library quantification method (Kapa Biosystems Inc., Wilmington, MA, USA). WGS was generated employing two MiSeq instruments and the MiSeq v2 500 cycle kit (Illumina Inc).

Conventional MIC determinations

All 2316 isolates were subjected to broth dilution testing (BDT) for determination of MICs and compared with WGS-based predictions. Wells containing both erythromycin and clindamycin detected inducible clindamycin-resistance as previously described [4]. Discordant results where WGS-based predictions differed from BDT by two or more dilutions (at least four-fold MIC differences) were retested using Etests as described by the manufacturer (Biomérieux, Marcy l'Etoile, France) or D testing as previously described [5]. True discrepancies were those results where both agar diffusion-based retesting (Etests or D tests) and the original BDT results were in agreement and disagreed with WGS-based predictions. Categorical discrepancies were defined as BDT-generated MICs predictive of susceptibility, intermediate resistance or resistance according to established MIC cut-offs [5] that differed from WGS-based predictions for these three categories.

β-lactam antibiotic MICs

PBP types, based upon PBP1a, PBP2b and PBP2x transpeptidase amino acid sequences, were generated for each isolate as previously described [1]. The most frequently encountered conventional MICs for the six antibiotics previously associated with a PBP type, even if previously encountered only once, were assigned for each PBP type. PBP types (including individual protein sequences) and MIC correlates for each PBP type are periodically updated at http://www.cdc.gov/streplab/mic-tables.html. At this website we also record the frequencies of each PBP type and associated MICs for β-lactams. Resistance, intermediate resistance and susceptibility (both non-meningitis and meningitis criteria) employed previously established MIC cut-offs [5]. Categorical discrepancies employed both non-meningitis and meningitis criteria.

Non-β-lactam-resistance determinants

The sequence queries used and bioinformatics pipeline for detection of resistance determinants are provided in a previous study [1], with modifications periodically added at https://github.com/BenJamesMetcalf/Spn_Scripts_Reference. Additionally, the ARG-ANNOT database was employed [6]. A commonly encountered query for identification of a putative inactivating missense mutation within *ermB* was added to the pipeline (TTGGAACAGGTAAAGAGCATTTAACGACGA), where the underlined base indicates a G to A change at base 134 resulting in a glycine to glutamate substitution.

Clonal complex determination

As previously described [1], clonal complexes were loosely defined as sharing four or more alleles with a given reference multilocus sequence type (ST). Reference STs were chosen according to abundance within a given serotype and/or their presence within multiple serotypes of the study, and differed by at least three alleles from other clonal complex reference STs.

Fastq file accession

Accession information for the 2316 fastq files used in this work are provided in the Supplementary material (Table S1).

Results

PBP types among year 2015 IPD isolates

Of the 2316 isolates for which PBP-type-based MIC predictions [1,3] and conventional MICs were compared, 104 (4.5%) were not used for β-lactam antibiotic MIC predictions because their PBP types were newly encountered and not previously associated with conventional MIC testing results. The WGS data from seven additional isolates had assembly errors preventing the generation of a PBP type.

PBP-based predictions of wild-type susceptibility to β-lactam antibiotics

Among the 1680 isolates with WGS-based penicillin MIC of ≤0.03, there were 12 discrepancies (0.7%) between WGS-based predictions and BDT testing (Table 1). For the purpose of this analysis, eight were considered potential categorical discrepancies,

Table 1
Comparison of penicillin MICs predicted by whole genome sequence pipeline penicillin-binding protein types to broth dilution testing MICs

Predicted penicillin MIC (*n*)	No. BDT discrepancies (MIC)[a,b]	CDs	No. with ≥2 dilution differences	Retests of ≥2 dilution differences.		No of PBP types	No of clonal complexes
				Agree with WGS prediction	True discrepancy[a]		
≤0.03 (1680)	4 (0.06) **3 (0.12)** **3 (0.25)** **2 (2.0)**	8	8	8	0	89	123
= 0.06 (44)	12 (0.03) **5 (0.12)**	5	0	NA	NA	8	9
=0.12 (35)	**9 (0.06)** 1 (0.25)	10	0	NA	NA	7	6
=0.25 (201)	**3 (0.03)** **2 (0.06)** 12 (0.12) 16 (0.5)	5	5	5	0	13	6
=0.5 (33)	3 (0.25) 2 (1.0)	0	0	NA	NA	5	3
= 1.0 (33)	14 (0.5) 2 (2.0)	0	0	NA	NA	6	7
= 2.0 (116)	19 (1.0) **8 (4.0)** **2 (8.0)** 2 (0.5) 2 (0.25)	10	6	4	2 (0.25,0.5)	7	5
= 4.0 (55)	**8 (2.0)** **5 (8.0)** **2 (0.5)** **1 (1.0)** **1 (0.25)**	17	4	0	4 (0.5, 1.0, 0.25)	7	7
= 8 (8)	**4 (4.0)**	4	0	NA	NA	3	2
Totals (%, 95% CI 2205 (100%, NA)	147 (6.7%, 5.7%–7.8%)	59 (2.7%, 2.0%–3.4%)	23 (1.0%, 0.66%–1.6%)	17	6 (0.27%, 0.10%–0.59%)	145	

BDT, broth dilution testing; PBP, penicillin-binding protein; WGS, whole genome sequence.
[a] The discordant BDT result(s) in parentheses.
[b] Categorical discrepant BDT results are in bold. These apply for both meningitis and non-meningitis guidelines [3]. Instances of at least four-fold MIC differences (two dilutions) between BDT result and WGS prediction are underlined.

because they had MICs above the cut-offs for meningitis cases, even though all eight of these isolates were from non-meningitis cases (data not shown). Since all eight of these categorical discrepancies had actual BDT MICs of at least two dilutions more than the predicted MIC of ≤0.03 mg/L, they were retested using conventional agar diffusion methodology (Etests). The eight isolates were in agreement with the WGS prediction, since they yielded MICs in the range of 0.012–0.032 mg/L (data not shown). Similar results were found for the other five β-lactam antibiotics (see Supplementary material, Tables S2 to S6) in that when PBP profiles were indicative of little or no decreased susceptibility, no true discrepancies were observed. The lowest BDT MIC was obtained for 75.7%–89.5% of the 2205 isolates for the six different β-lactam antibiotics (Table 1, and see Supplementary material, Tables S2 to S6). There is a lower number of clonal complexes within resistant PBP types than within PBP types that predict basal susceptibility (Table 1, and see Supplementary material, Tables S2 to S6). reflecting the preponderance and greater genetic diversity of basally susceptible pneumococcal strains (data not shown). The majority of true discrepancies for the six β-lactam antibiotics reflected isolates with lower MICs than those predicted by PBP type correlates (Table 1, and see Supplementary material, Tables S2 to S6).

WGS-based predictions for erythromycin, clindamycin and quinupristin-dalfopristin

Almost all macrolide resistance was detected through a query that broadly detects *mef* genes [1,7,8] and *ermB* [1,9,10]. BDT testing results indicated that in about 5% of *ermB*-positive isolates, clindamycin-resistance was inducibly expressed rather than constitutively (data not shown). Also included in the pipeline is a putative inactivating (null) missense mutation within *ermB* (provided in methods) that we have identified fairly frequently (representative isolate 20155282). We correctly predicted complete erythromycin and clindamycin susceptibility for 17 such isolates based upon this specific pipeline query sequence. There were two true discrepancies for *ermB*-positive isolates. Both strains contained full-length *ermB* structural gene alleles (isolates 20156290 and 20161788) that we have previously observed within isolates resistant to erythromycin and clindamycin (data not shown).

Three true discrepancies were observed among the 48 isolates positive for both *ermB* and *mef* (Table 2). Two of these isolates were also unexpectedly non-susceptible to quinupristin-dalfopristin (BDT MIC of >1 mg/L verified with subsequent Etest results of 2–3 mg/L (isolates 20154278 and 20155427). Another isolate positive for the pipeline *ermB* query was clindamycin-susceptible and was retrospectively found to contain a truncated *ermB* gene due to a premature stop codon at base 642 of the 735-bp structural gene (isolate 20155328). The 23S rRNA gene A2061G substitution predicted constitutive resistance to erythromycin and clindamycin (clindamycin MIC of 1 mg/L) in four isolates, as previously observed [11,12]; however, one isolate was only intermediately clindamycin-resistant. Single occurrences of two isolates with 23S rRNA gene substitutions (C2049T or A2069G) other than the resistance-conferring A2061G substitution shown (Table 2) were susceptible to erythromycin and other ribosome-targeting antibiotics.

One isolate with a previously observed two codon insertion within *rplD* (isolate 20153241) (Table 2), a mutation previously

Table 2
Macrolide-resistance determinants and corresponding phenotype predictions compared with BDT testing results of 2316 IPD isolates

WGS result (no. isolates)	interpretation [c]	BDT discrepant result (*n*)	Re-test result	
			Agreed with WGS	True discrepancies
ermB (191) [a]	eryR cliR QP-S linS	eryS cliS (3)	2	1
		eryR cliS (2)	2	0
		eryI cliS (1)	0	1
ermB + *mef* (48)	eryR cliR QP-S	eryR cliR QP-NS (2)	0	2
		eryR cliS (1)	0	1
mef (475)[b]	eryR cliS	eryS (3)	3	0
		eryR cliR (2)	2	0
rRNA A2061G (5)	eryR cliR QP-S	eryR cliR QP-NS (1)	1	0
		eryR cliI (1)[b]	ND	ND
rplD 2 codon insertion (KG) between codons 69 and 70 (1)	eryI QP-R	eryR QP-S (1)	0	1
Negative (1597)	eryS cliS	eryR (3)	2	1[d]
		eryI (3)	3	0
		eryR cliI (2)	1	1 [d]
Totals (%, 95% CI) (2316, 100%, NA)		25 (1.1%, 0.70%–1.6%)	15 (0.7%, 0.36%–1.1%)	9 (0.4%, 0.18%–0.74%)

BDT, broth dilution testing; IPD, invasive pneumococcal disease; WGS, whole genome sequence.
[a] Seventeen of these isolates were eryS cliS apparently due to a missense substitution (G41E) within the *ermB* structural gene that is detected by the WGS pipeline. These isolates are not listed in the interpretation of discrepant results.
[b] The single discrepant intermediately clindamycin-resistant result differed from the other four by a single dilution (clindamycin MIC = 0.5 rather than 1.0).
[c] Abbreviations: eryS and R, erythromycin-susceptible and resistant; QP-S and R, quinupristin-susceptible and resistant; cliR and S.
[d] Isolate was found to be inducibly clindamycin-resistant using the double disc diffusion test.

associated with intermediate erythromycin-resistance and resistance to quinupristin-dalfopristin [12], was subsequently Etested as erythromycin-resistant (MIC of 16 mg/L) and quinupristin-dalfopristin-susceptible (MIC of 1 mg/L). Six isolates with the ribosomal protein L4 (*rplD*) E30K substitution (including isolates 20151636 and 20153922), an alteration previously associated with macrolide-resistance [8] or combined resistance to macrolides and quinupristin-dalfopristin [12], were susceptible to both antimicrobials (data not shown). A single instance of the *rplD* Q67K substitution was found in an isolate that was susceptible to all antibiotics on the BDT panel, although a strain with the two substitutions Q67R and R72G has been associated with resistance to quinupristin-dalfopristin and linezolid [12]. We found no instances of linezolid-resistance. Of 1597 isolates that were not predicted by the WGS pipeline to be resistant to erythromycin, clindamycin and quinupristin-dalfopristin, we found two isolates that were erythromycin non-susceptible and inducibly clindamycin-resistant (isolates 20154490, 20154507), suggesting a possible undetected determinant of the *erm* (23S rRNA methylase) class [10–13].

Cotrimoxazole (trimethoprim/sulfamethoxazole)-resistance

The *folA* I100L substitution [14] or one or two codon insertions within *folP* [15], were highly associated with intermediate resistance to trimethoprim/sulfamethoxazole (MIC of 1/19 to 2/38 mg/L) as determined by BDT (221 of 276 isolates, 80.1%) (Table 3). The majority of the discrepancies (29/45, 64%) where single mutants in *folA* or *folP* were not intermediately resistant was due to the BDT MIC result of 0.5/9.5 mg/L, which is the highest MIC categorized as cotrimoxazole-susceptible [5]. Of the 2004 year 2015 ABCs isolates that were determined as cotrimoxazole-susceptible through BDT, only 109 (5.4%) had an MIC of 0.5/9.5 mg/L. In this study we determined that of 101 isolates with cotrimoxazole MICs of 0.5/9.5 mg/L, 16 (15.8%) contained *folP* insertions and 13 (12.8%) contained the *folA* I100L substitution. Double mutants with the *folA* I100L substitution and a *folP* one or two codon insertion accounted for 172 isolates, the majority (163/172, 94.8%) of which were fully cotrimoxazole-resistant (Table 3). We found nine different one or two codon insertions between codons 59 and 69 of the 351 codon *folP* gene (Fig. 1), each identified by different pipeline designations based upon the location of the insertion within *folP* and also influenced by localized sequence divergence of immediate flanking region (not shown) due to the apparent mosaicism we observed within many of the mutant *folP* alleles (data not shown). With the exception of *folP* insertions 188 and 176, these one or two amino acid insertions have been previously identified in pneumococci resistant to sulfonamides or cotrimoxazole [15–17].

Tetracycline resistance

Eight of the 261 *tetM*-positive isolates were found to be tetracycline-susceptible, potentially indicating a defective *tetM* gene (Table 4). Two *tetM*-positive isolates were additionally *tetS*-positive [18,19] (isolates 20152656 and 20156655) and one isolate negative for tetracycline-resistance according to our original pipeline [1] contained *tet32* [20] alone (isolate 20153228). The *tetS* and *tet32* determinants were detected solely through the ARG-ANNOT database [6] that we have added to our pipeline. Both *tetS* and *tetM* were identified through the ARG-ANNOT database for isolates 20152656 and 20156655, whereas only *tetM* was detected for these two isolates through our 100-bp pipeline query. This indicated that the positive ARG-ANNOT pipeline results could have been due to the presence of two different tetracycline-resistance genes that share ≥70% identity over a ≥70% structural gene overlap [6], or could have been due to the presence of a single somewhat divergent *tetM* allele that has localized high sequence identity to our *tetM* pipeline query. We found that the latter circumstance was true, with the two full-length structural genes from the two isolates sharing 96.0%–96.1% and 78.2%–78.3% sequence identity to the ARG-ANNOT *tetM* (1920 bp) and *tetS* (1926 bp) database sequences, respectively.

One tetracycline-resistant isolate (isolate 20152918) tested negative for resistance genes in our pipeline and in the ARG-ANNOT database.

Chloramphenicol resistance

All *cat*-positive [21] isolates were chloramphenicol-resistant with no associated BDT discrepancies (see Supplementary material, Table S7). Of the 2265 *cat*-negative isolates, there were

Table 3
Determinants of cotrimoxazole resistance and corresponding phenotype predictions compared to BDT testing results of 2316 IPD isolates

Cotrimoxazole resistance determinants (no. of occurrences)	Prediction[a]	BDT discrepant result (*n*)	Retest result for >2 dilution discrepancy	
			Agrees with WGS	True discrepancy
folP insertion (238)	Cot-I (MIC 1.0–2.0 mg/L)	0.5 mg/L (16)	NA	NA
		>4 mg/L (5)	3	2[a]
		=4 mg/L (6)	NA	NA
folA I100L (38)	Cot-I (MIC 1.0–2.0 mg/L)	0.5 mg/L (13)	NA	NA
		0.25 mg/L (1)	0	1
		4 mg/L (1)	NA	NA
		>4 mg/L (3)	2	1 [b]
folP insert + *folA* I100L (172)	Cot-R (MIC ≥4.0 mg/L)	1–2 mg/L (8)	NA	NA
		1 mg/L (1)	0	1[c]
Negative (1868)	Cot-S (MIC ≤0.5 mg/L)	1 mg/L (7)	7	0
		2 mg/L (3)	3	0
		4 mg/L (2)	2	0
Totals (%, 95% CI) (2316, 100%, NA)		66 (2.8%, 2.2%–3.6%)	17	5 (0.22%, 0.07%–0.5%)

BDT, broth dilution testing; IPD, invasive pneumococcal disease; WGS, whole genome sequence.
[a] MICs shown correspond to cotrimoxazole (trimethoprim/sulfamethoxazole) MICs of ≤0.5/9.5 to >4/76 mg/L.
[b] Etest MIC of 3 mg/L for two of these three isolates. For one of the two *folP* insertion isolates the Etest MIC was >32 mg/L.
[c] Etest MIC of 1.5 mg/L for this isolate.

folP designation	Amino acid sequence	No. of isolates
Wild-type	56 STRPGSSYVEIE 67	1906[a]
169,173,177	56 STRP**RP**GSSYVEIE 69	153
176	56 STRPG**RG**SSYVEIE 69	5
178	56 STRPG**R**SSYVEIE 68	80
179	56 STRPGSS**S**YVEIE 69	1
182	56 STRPGSSY**SY**VGIE 69	4
185	56 STRPGSSY**Y**VEIE 68	1
188	56 STRPGSSYV**V**EIE 68	1
189,186	56 STRPGSSYVE**VE**IE 69	94
192,195	56 STRPGSSYVEI**EI**E 69	71

Fig. 1. One and two codon insertions within localized region of *folP* between codons 59 and 69, their designations, and number of isolates detected with insertion. [a]Contain no insertions within indicated region; however, sequence not examined.

23 BDT discrepancies, 12 of which were subsequently determined to be chloramphenicol-susceptible by Etests (MICs of ≤6 mg/L). More than 50% of ABCs isolates determined to be chloramphenicol-susceptible through BDT during 2015 had a chloramphenicol BDT MIC of 4.0 mg/L, differing by a single dilution from the criteria for resistance [4].

Fluoroquinolone resistance

In pneumococci and other species, fluoroquinolones are most active in targeting localized regions within the ParC subunit of DNA topoisomerase IV and the GyrA subunit of DNA gyrase [22]. Both levofloxacin-resistant isolates identified by BDT (isolates 20160273 and 20153527) also had a ciprofloxacin MIC of 4 mg/L and were associated with *parC* substitutions (see Supplementary material, Table S8). One of these two strains additionally contained the *gyrA* S81Y substitution and was the only double mutant identified. Although there are no CLSI guidelines for pneumococcal ciprofloxacin MICs [5], we considered MICs of ≥4 mg/L found among 12 isolates (see Supplementary material, Table S8) to be phenotypically significant and potentially associated with causal genomic alterations. Although levofloxacin-resistance is considered to be ≥4 mg/L, we also considered the levofloxacin MIC of 2 mg/L as moderately elevated, since only 38 of the 2316 isolates had levofloxacin MICs of ≥2 mg/L. Of the nine levofloxacin-susceptible isolates with elevated MICs to both fluoroquinolones (see Supplementary material, Table S8), five contained substitutions within ParC or GyrA, all of which have been associated with elevated MICs to these antibiotics [23]. Although our pipeline successfully detected the two levofloxacin-non-susceptible isolates in our isolate set, there were six instances of elevated ciprofloxacin MICs (five with MIC of 4 mg/L and one with MIC >4 mg/L) that were not detected (isolates 20154278, 20154285, 20153966, 20155459, 20160446, 20161660).

Rifampin resistance

Rifampin binds to the β-subunit of bacterial RNA polymerase to prevent its essential function [24]. Two rifampin-resistant isolates (isolates 20160910 and 20153860) were flagged by the pipeline, each showing a single substitution (H499N or D489V) within *rpoB* previously associated with rifampin-resistance [25,26]. Retesting (Etests) of these isolates revealed a rifampin MIC of 8.0 mg/L for the H499N substitution, and an MIC of >32 mg/L for the D489V substitution (see Supplementary material, Fig. S1). Ten additional substitutions within this region of RpoB were flagged in one or two isolates, but were associated with rifampin-susceptibility (BDT MIC of ≤2 mg/L and Etest results of ≤0.05 mg/L), including S481L, M488I, M488V, D489N, H491P, H491R, H491Y, N492T, S495C and H499L. Of the 12 substitutions, all present as single RpoB substitutions, only the D489N and H499L substitutions were previously associated with pneumococcal rifampin-resistance [25,26].

Vancomycin

We did not detect any isolates that were positive for the vancomycin-resistance determinant queries in our pipeline and in

Table 4
Determinants of tetracycline resistance and corresponding phenotype predictions compared with BDT testing results of 2316 IPD isolates

WGS data (no. occurrences)	Prediction	BDT discrepant results (No)	Re-test result for ≥2-fold dilution discrepancy	
			Agrees with WGS	Agrees with BDT
tetM (259)	Tet-R	≤0.5 (10)	2	8
	(MIC ≥4)	2.0[b] (4)	ND	ND
(*tetM* + *tetS*)[a] (2)	Tet-R	None		
tet32 (1)	Tet-R	None		
Negative (2054)	Tet-S	≥4 (7)	5	2
Totals (%, 95% CI) (2316, 100%, NA)		21 (0.91%, 0.56%–1.4%)	8 (0.35%, 0.15%–0.68%)	9 (0.4%, 0.18%–0.74%)

BDT, broth dilution testing; IPD, invasive pneumococcal disease; WGS, whole genome sequence.
[a] Single tetracycline-resistance determinant within both isolates detected solely as *tetM* through CDC pipeline [1] and as both *tetM* and *tetS* through ARG-ANNOT database [6].
[b] 2.0 mg/L is intermediate resistance.

ARG-ANNOT [6]. Four of the 2316 isolates were initially assigned BDT vancomycin MICs of >1 mg/L, all of which were retested and found to have MICs of ≤0.5 mg/L.

Aminoglycosides and streptothricins

The *aphA-3* and *sat4* genes, associated with resistance to kanamycin and streptothricins, respectively, and often co-linked to each other or to other resistance genes on insertion elements in Gram-positive species [9,27], were detected from the ARG-ANNOT database [6] among four isolates. Two unrelated isolates (20156295 and 20153959) were positive for both determinants, along with *ermB* and *tetM*. The third isolate (20154707) was positive only for *aphA-3* and *ermB*. The fourth isolate (20152957) was positive for only *sat4* and *ermB*. We did not perform phenotypic testing for kanamycin and streptothricins.

Discussion

A significant finding from this work was the absence of discrepant phenotypic testing results for PBP types predictive of wild-type susceptibility to β-lactam antibiotics. These 'sensitive' PBP alleles are all highly conserved with wild-type unaltered pneumococcal PBP genes, consistent with strong associations of altered PBPs with β-lactam non-susceptibility [1,3]. The high proportion of clonal complexes relative to PBP types among β-lactam-susceptible isolates reflects the broad clonal distribution of individual 'sensitive' PBP alleles [1]. In contrast, the genetic diversity of isolates with reduced susceptibility or resistance to β-lactams is relatively constrained in the USA, especially now after sequential introductions of two highly effective conjugate vaccines within the past 17 years [1,2].

Most instances where at least four-fold differences between BDT MIC and WGS-predicted MICs could not be resolved involved decreased BDT MICs compared with pipeline-based predictions. This is a much better scenario than, for example, underestimated β-lactam MICs. For antibiotics such as β-lactams, this potentially reflects compensatory changes that alleviate fitness costs that occur within PBP domains not included within the PBP type, or other enzymes affecting peptidoglycan structure such as MurM that in some strains are known to influence β-lactam MICs [28,29]. There was a small number of isolates with unexplained non-susceptibility or elevated MICs to other antibiotics, including erythromycin, clindamycin, tetracycline and quinupristin-dalfopristin. For these isolates, deeper searches for resistance determinants that include expansion of the original gene query targets and identification of potentially new accessory genes or chromosomal mutations are warranted and are ongoing. Conventional testing of isolates that are flagged due to detection of rare alterations within highly conserved determinants, such as those that are involved in translation (e.g. *rplD*, *rplV*, 23S rRNA genes), transcription (e.g. *rpoB*) and DNA replication (e.g. *parC*, *gyrA*), is warranted, not only due to the lack of abundant pre-existing phenotypic data, but because of the likelihood that resistance may often involve additional undetected determinants that could affect resistance levels and additional antimicrobial susceptibility phenotypes. It will be necessary to continue BDT MIC determinations for a carefully selected significant minority of surveillance isolates to be vigilant for the successful emergence of new resistance mechanisms. Where isolates are actually predicted to be susceptible due to identified mutations within known accessory element resistance genes (for example the *ermB* missense substitution associated with erythromycin-susceptible isolates described in this work), these isolates should still be considered potentially reversibly resistant. For other antimicrobials, WGS-based predictions might influence current phenotypic guidelines. For example, 58% of the most recently tested 2611 ABCs isolates tested as phenotypically chloramphenicol-susceptible had the MIC of 4 mg/L, suggesting intrinsic reduced susceptibility that might not be compatible with assigning the MIC of 8 mg/L as fully chloramphenicol-resistant.

The *folA* I100L substitution and insertions within a small region of *folP* confer resistance to trimethoprim and sulfamethoxazole, respectively [14–16]. Here we have shown that individually the *folA* I100L and *folA* insertions are usually associated with intermediate resistance to cotrimoxazole (trimethoprim-sulfamethoxazole) while mutations in both genes are nearly always indicative of full cotrimoxazole resistance. The rarity of cotrimoxazole MIC discrepancies (five instances within 2316 isolates) is probably dictated by the essential roles of mutant *folA* and *folP* alleles in the folic acid biosynthetic pathway and the lack of other major genetic influences on cotrimoxazole-resistance.

The ARG-ANNOT database [6] provides a useful expansion of our original resistance pipeline [1] for detecting accessory genome resistance determinants. For these 2316 isolates it identified all of the accessory element genes among these isolates that were detected by our pipeline (*mef*, *ermB*, *tetM*, *cat*). ARG-ANNOT also detected a distinct tetracycline-resistance gene (*tet-32*) within one isolate and potential resistance to aminoglycosides and/or streptothricin antibiotics in four isolates. Although for this isolate set ARG-ANNOT was slightly more encompassing than our own specialized pipeline for detection of accessory element resistance features, the potential exists for our smaller, broadly encompassing gene queries to detect genes that are not detected through ARG-ANNOT. For example, in groups A and B streptococci, our broadly encompassing *ermTR* and *mef* queries (also included in our pneumococcal pipeline) appear to quantitatively detect these determinants; however, these two genes were not detected through ARG-ANNOT (data not shown). We have observed that ARG-ANNOT has invariably detected the macrolide-resistance determinant *msrD* [30] among the *mef*+ streptococcal isolates that we have

characterized through our pipelines (data not shown). Although ARG-ANNOT was not designed to recognize signatures of resistances within core genomic determinants (e.g. PBP types, *folA*, *folP*, 23S rRNA genes, *rpoB*) it is useful for recognizing additional accessory determinants among streptococci not detected through our species-specific pipelines (manuscripts in preparation). We conclude that our streptococcal accessory resistance pipelines and the ARG-ANNOT database are complementary resources for detecting accessory resistance determinants.

Here we demonstrated the capability of our WGS pipeline to accurately and reliably predict antimicrobial testing results for currently recovered IPD isolates to six different β-lactam antibiotics, erythromycin, clindamycin, cotrimoxazole, tetracycline and chloramphenicol. In addition, the pipeline detected rare instances of fluoroquinolone-resistance and rifampin-resistance. This sampling of recent national surveillance IPD isolates gives an accurate assessment of the pipeline's reliability for detecting and measuring individual resistance phenotypes, and for determining the frequencies of different resistance mechanisms. WGS-based methodology is an effective substitute for BDT that also serves for concurrent determination of other critical pneumococcal strain features. As technical advancements continue, WGS technology may be employed in future clinical settings for deducing key pathogen parameters relevant to patient care. The high concordance of WGS-based predictions with pneumococcal antimicrobial phenotypes is compatible with this goal.

Acknowledgements

Major funding for this work was provided through support from the Advanced Molecular Detection (AMD) initiative at CDC. We are grateful to the Minnesota Department of Public Health laboratory for susceptibility testing of all isolates recovered in Minnesota. We thank the following individuals for their contributions to the establishment and maintenance of the ABCs system. California Emerging Infections Program: A. Reingold, S. Brooks and H. Randel. Colorado Emerging Infections Program: L. Miller, B. White, D. Aragon, M. Barnes and J. Sadlowski. Connecticut Emerging Infections Program: S. Petit, M. Cartter, C. Marquez and M. Wilson. Georgia Emerging Infections Program: M. Farley, S. Thomas, A. Tunali and W. Baughman. Maryland Emerging Infections Program: L. Harrison, J. Benton, T. Carter, R. Hollick, K. Holmes and A. Riner. Minnesota Emerging Infections Program: C. Holtzman, R. Danila and K. MacInnes. New Mexico Emerging Infections Program: K. Scherzinger, K. Angeles, J. Bareta, L. Butler, S. Khanlian, R. Mansmann and M. Nichols. New York Emerging Infections Program: N. Bennett, S. Zansky, S. Currenti and S. McGuire. Oregon Emerging Infections Program: A. Thomas, M. Schmidt, J. Thompson and T. Poissant. Tennessee Emerging Infections Program: W. Schaffner, B. Barnes, K. Leib, K. Dyer and L. McKnight. CDC: R. Gierke, O. Almendares, J. Hudson, L. McGlone and G. Langley.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.cmi.2016.08.001.

Transparency declaration

The authors are aware of no relationships/conditions/circumstances that present a potential conflict of interest.

References

[1] Metcalf BJ, Gertz Jr RE, Gladstone RA, Walker H, Sherwood LK, Jackson D, et al. Strain features and distributions in pneumococci from children with invasive disease before and after 13-valent conjugate vaccine implementation in the USA. Clin Microbiol Infect 2016;22:60.e9–60.e29.
[2] Kim L, McGee L, Tomczyk S, Beall B. Biological and epidemiological features of antibiotic-resistant *Streptococcus pneumoniae* in PRE- AND POST-CONJUGATE VACCINE ERAS: a United States perspective. Clin Microbiol Rev 2016;29:525–52.
[3] Li Y, Metcalf BJ, Chochua S, Li Z, Gertz Jr RE, Walker H, et al. Penicillin-binding protein transpeptidase signatures for tracking and predicting β-lactam resistance levels in *Streptococcus pneumoniae*. Mbio 2016:e00756–16.
[4] Jorgensen JH, McElmeel ML, Fulcher LC, McGee L, Glennen A. Evaluation of disk approximation and single-well broth tests for detection of inducible clindamycin resistance in *Streptococcus pneumoniae*. J Clin Microbiol 2011;49:3332–3.
[5] Clinical and Laboratory Standards Institute (CLSI). Performance standards for antimicrobial susceptibility testing; Twenty-third informational supplement. CLSI document M100–S22.CLSI. Wayne, PA: CLSI; 2013.
[6] Gupta SK, Padmanabhan BR, Diene SM, Lopez-Rojas R, Kempf M, Landraud L, et al. ARG-ANNOT, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. Antimicrob Agents Chemother 2014;58:212–20.
[7] Clancy J, Petitpas J, Did-Hajj F, Yuan W, Cronan M, Kamath AV, et al. Molecular cloning and functional analysis of a novel macrolide-resistance determinant, *mefA*, from Streptococcus pyogenes. Mol Microbiol 1996;22:867–79.
[8] Tait-Kamradt A, Clancy J, Cronan M, Did-Hajj F, Wondrack L, Yuan W, et al. *mefE* is necessary for the erythromycin-resistance M phenotype in *Streptococcus pneumoniae*. Antimicrob Agents Chemother 1997;41:2251–5.
[9] Courvalin P, Carlier C. Transposable multiple antibiotic resistance in *Streptococcus pneumoniae*. Mol Gen Genet 1986;205:291–7.
[10] Trieu-Cuot P, Poyart-Salmeron C, Carlier C, Courvalin P. Nucleotide sequence of the erythromycin resistance gene of the conjugative transposon Tn1545. Nucleic Acids Res 1990;18:3660.
[11] Tait-Kamradt A, Davies T, Appelbaum PC, Depardieu F, Courvalin P, Petitpas J, et al. Two new mechanisms of macrolide resistance in clinical strains of *Streptococcus pneumoniae* from Eastern Europe and North America. Antimicrob Agents Chemother 2000;44:3395–401.
[12] Hawkins PA, Chochua S, Jackson D, Beall B, McGee L. Mobile elements and chromosomal changes associated with MLS resistance phenotypes of invasive pneumococci recovered in the United States. Microb Drug Resist 2015;21:121–9.
[13] Wierzbowski AK, Nichol K, Laing N, Hisanaga T, Nikulin A, Karlowsky JA, et al. Macrolide resistance mechanisms among *Streptococcus pneumoniae* isolated over 6 years of Canadian Respiratory Organism Susceptibility Study (CROSS) (1998–2004). J Antimicrob Chemother 2007;60:733–40.
[14] Pikis A, Donkershoot JA, Rodriguez WJ, Keith JM. A conservative amino acid mutation in the chromosome-encoded dihydrofolate reductase confers trimethoprim resistance in *Streptococcus pneumoniae*. J Infect Dis 1998;178:700–6.
[15] Haasum Y, Strom K, Wehelie R, Luna V, Roberts MC, Maskell JP, et al. Amino acid repetitions in the dihydropteroate synthase of *Streptococcus pneumoniae* lead to sulfonamide resistance with limited effects on substrate Km. Antimicrob Agents Chemother 2001;45:805–9.
[16] Schmitz FJ, Perdikouli M, Beeck A, Verhoef J, Fluit AC, European SENTRY participants. Resistance to trimethoprim-sulfamethoxazole and modifications in genes coding for dihydrofolate reductase and dihydropteroate synthase in European *Streptococcus pneumoniae* isolates. J Antimicrob Chemother 2001;48:935–6.
[17] Cornick JE, Harris SR, Parry CM, Moore MJ, Jassi C, Kamng'ona A, et al. Genomic identification of a novel co-trimoxazole resistance genotype and its prevalence amongst *Streptococcus pneumoniae* in Malawi. J Antimicrob Chemother 2014;69:368–74.
[18] Burdett V, Inamine J, Rajagopalan. Heterogeneity of tetracycline resistance determinants in Streptococcus. J Bacteriol 1982:995–1004.
[19] Charpentier E, Gerbaud G, Courvalin P. Characterization of a new class of tetracycline-resistance gene tet(S) in *Listeria monocytogenes* BM4210. Gene 1993;131:27–34.
[20] Melville CM, Scott KP, Mercer DK, Flint HJ. Novel tetracycline resistance gene, tet(32), in the Clostridium-related human colonic anaerobe K10 and its transmission in vitro to the rumen anaerobe *Butyrivibrio fibrisolvens*. Antimicrob Agents Chemother 2001;45:3246–9.
[21] Horinouchi S, Weisblum B. Nucleotide sequence and functional map of pC194, a plasmid that specifies inducible chloramphenicol resistance. J Bacteriol 1982;150:815–25.
[22] Tankovic J, Perichon B, Duval J, Courvalin P. Contribution of mutations in *gyrA* and *parC* genes to fluoroquinolone resistance of mutants obtained in vivo and in vitro. Antimicrob Agents Chemother 1996;40:2505–10.
[23] Brueggemann AB, Coffman SL, Rhomberg P, Huynh H, Almer L, Nilius A, et al. Fluoroquinolone resistance in *Streptococcus pneumoniae* in United States since 1994–1995. Antimicrob Agents Chemother 2002;46:680–8.
[24] Campbell EA, Korzheva N, Mustaev A, Murakami K, Nair S, Goldfarb A, et al. Structural mechanism for rifampicin inhibition of bacterial RNA polymerase. Cell 2001;104:901–12.

[25] Ferrándiz MJ, Ardanuy C, Liñares J, García-Arenzana JM, Cercenado E, Fleites A, et al. New mutations and horizontal transfer of *rpoB* among rifampin-resistant *Streptococcus pneumoniae* from four Spanish hospitals. Antimicrob Agents Chemother 2005;49:2237–45.

[26] Ferrándiz MJ, Ardanuy C, Liñares J, Balsalobre L, García MT, de la Campa AG. New species genetic approach to identify strains of mitis group streptococci that are donors of rifampin resistance to *Streptococcus pneumoniae*. Antimicrob Agents Chemother 2011;55:368–72.

[27] Werner G, Hildebrandt B, Witte W. Linkage of erm(B) and *aadE-sat4-aphA-3* in multiple-resistant *Enterococcus faecium* isolates of different ecological origins. Microb Drug Resist 2003;9(Suppl. 1):S9–16.

[28] Sauerbier J, Maurer P, Rieger M, Hakenbeck R. *Streptococcus pneumoniae* R6 interspecies transformation: genetic analysis of penicillin resistance determinants and genome-wide recombination events. Mol Microbiol 2012;86:692–706.

[29] Hakenbeck R, Brückner R, Denapaite D, Maurer P. Molecular mechanisms of β-lactam resistance in *Streptococcus pneumoniae*. Future Microbiol 2012;7:395–410.

[30] Daly MM, Doktor S, Flamm R, Shortridge D. Characterization and prevalence of MefA, MefE, and the associated *msr*(D) gene in *Streptococcus pneumoniae* clinical isolates. J Clin Microbiol 2004;42:3570–4.