Systematic review

Whole genome sequencing of Mycobacterium tuberculosis for detection of drug resistance: a systematic review

D. Papaventsis, N. Casali, I. Kontsevaya, F. Drobniewski, D.M. Cirillo, V. Nikolayevskyy

1) National Reference Laboratory for Mycobacteria, Sotiria Chest Diseases Hospital, Athens, Greece
2) Department of Medicine, Imperial College London, London, UK
3) Division of Immunology, Transplantation and Infectious Diseases, IRCCS San Raffaele Scientific Institute, Milan, Italy
4) PHE National Mycobacterium Reference Laboratory, London, UK

ABSTRACT

Objectives: We conducted a systematic review to determine the diagnostic accuracy of whole genome sequencing (WGS) of Mycobacterium tuberculosis for the detection of resistance to first- and second-line anti-tuberculosis (TB) drugs.

Methods: The study was conducted according to the criteria of the Preferred Reporting Items for Systematic Reviews group. A total of 20 publications were included. The sensitivity, specificity, positive-predictive value and negative-predictive value of WGS using phenotypic drug susceptibility testing methods as a reference standard were determined.

Results: Anti-TB agents tested included all first-line drugs, a variety of reserve drugs, as well as new drugs. Polymorphisms in a total of 53 genes were tested for associations with drug resistance. Pooled sensitivity and specificity values for detection of resistance to selected first-line drugs were 0.98 (95% CI 0.93-0.98) and 0.98 (95% CI 0.98-1.00) for rifampicin and 0.97 (95% CI 0.94-0.99) and 0.93 (95% CI 0.91-0.96) for isoniazid, respectively. Due to high heterogeneity in study designs, lack of data, knowledge of resistance mechanisms and clarity on exclusion of phylogenetic markers, there was a significant variation in analytical performance of WGS for the remaining first-line, reserved drugs and new drugs.

Conclusions: Whole genome sequencing could be considered a promising alternative to existing phenotypic and molecular drug susceptibility testing methods for rifampicin and isoniazid pending standardization of analytical pipelines. To ensure clinical relevance of WGS for detection of M. tuberculosis complex drug resistance, future studies should include information on clinical outcomes.

D. Papaventsis, CMI 2017;23:61

Introduction

Drug-resistant tuberculosis (DR-TB) remains a major challenge to global health [1]. Multidrug-resistant TB (MDR-TB), namely TB that is resistant to both isoniazid and rifampicin, constitutes a major threat. In 2014, the WHO estimated 3.3% of new cases and 20% of previously treated cases to be MDR-TB with an estimated worldwide morbidity and mortality from MDR-TB in 2014 of 480 000 and 190 000 people, respectively [1]. Extensively drug-resistant TB (XDR-TB) is defined as MDR-TB with additional resistance to any fluoroquinolone and one of the second-line injectable anti-TB drugs, amikacin, capreomycin or kanamycin. In 2014, XDR-TB cases, approximately one in every ten MDR-TB cases, were reported by more than 100 countries. Both MDR-TB and XDR-TB require extensive treatment with multiple, potentially toxic drugs and outcomes remain poor. Estimates show that 50% of patients with MDR-TB and <30% of patients with XDR-TB survive [1,2].

In 2015, the WHO proposed expanding rapid testing and detection of cases as one of the five high-priority actions to tackle the global DR-TB crisis [1]. The introduction of new diagnostic and treatment tools for the management of DR-TB enables earlier
diagnosis of MDR-TB, prevents the spread of drug-resistant bacilli and offers more effective treatment in cases where therapeutic options are very limited [3]. Current diagnostic methods such as microbiological drug susceptibility testing and targeted rapid tests, particularly the WHO endorsed Cepheid Xpert MTB/RIF (Mycobacterium tuberculosis/rifampicin) assay (Cepheid, Sunnyvale, CA, USA) and Hain line-probe assays (Hain Lifescience, Nehren, Germany), enable rapid identification of Mycobacterium tuberculosis complex (MTBC) and detection of mutations conferring MTBC resistance to rifampicin, isoniazid and selected reserve anti-TB drugs. However, these methods do not solve many of the major challenges that continue to make the programmatic management of MDR-TB a highly complex public health intervention [3]. More comprehensive diagnostics are needed to ensure effective treatment initiation, improve treatment outcomes and reduce disease transmission.

Whole genome sequencing (WGS) offers new opportunities in the clinical management of DR-TB cases, providing considerable information with unprecedented accuracy when compared with current routine methods [4]. Findings from retrospective [4–10] and prospective [11] studies indicate the potential for WGS to identify various genetic polymorphisms, including single nucleotide polymorphism (SNPs) and small insertions and deletions (indels) that are potentially important for the reliable prediction of drug susceptibility phenotype within a clinically relevant timeframe and a comparable cost range [4,11,12]. Implementation of end-to-end WGS-based diagnostic systems for TB, using decentralized sequencing, centralized analysis models and semi-automated bioinformatics pipelines, have already provided evidence of the capacity of WGS to inform and affect clinical decision-making and to a certain extent replace traditional diagnostic procedures for TB in high-income settings in Europe and North America [12,13]. Current WGS testing of most anti-TB drugs involves culture of the bacteria; however, direct sequencing of MTBC from sputum indicates that culture-free approaches might have a role in the future management of DR-TB [6,14].

Although WGS could transform drug-susceptibility testing not only in high-income settings but also in low-income settings, the growing body of knowledge on mutations conferring drug resistance or consistent with susceptibility has led to the suggestion that identification of true MTBC drug resistance may be more complex than previously thought (low-level resistance, compensatory mutations etc.) [4,15–18]. Systematic reviews with regards to the role of WGS of MTBC for the prediction of drug resistance and drug susceptibility are lacking. The current review attempts to assess the diagnostic value of WGS for the detection of resistance to first- and second-line anti-TB drugs using phenotypic methods as a reference standard, and provide considerations for standardization and future research.

Methods

Criteria for considering studies for inclusion

This systematic review was conducted according to the criteria of the Preferred Reporting Items for Systematic Reviews group [19]. Before conducting the review, a protocol was written and registered with the International Prospective Register for Systematic Reviews (PROSPERO). Cross-sectional and cohort studies conducted on individuals with culture-positive TB disease and/or DNA specimens extracted from MTBC cultures isolated from individuals were considered for inclusion. Studies published between 1 January 2005 and 30 October 2015 with available data on WGS and phenotypic drug susceptibility test (DST) results of MTBC isolates were included; no geographical or language restrictions were applied.

Outcome measures and search methods

A compound search strategy was developed to identify all relevant studies regardless of language or publication status in three online databases. Search methods and outcome measures are detailed in the Supplementary material (Appendix S1).

Data collection and analyses

Selection of studies

All references identified by the compound search strategy were imported into EndNote v.X7 (www.endnote.com, Thomson Reuters). Duplicates were removed. Titles and abstracts were examined in duplicate by two reviewers and studies not meeting the inclusion criteria were excluded. Full text of all potentially relevant studies was obtained and inclusion criteria were applied using a standardized eligibility form. Final agreement on study inclusion was determined through consensus between the two reviewers.

Data extraction and management

Data extraction was performed independently, in duplicate, using three standardized data extraction forms (Excel spreadsheets) as follows:

- General characteristics of studies: year, setting (country, city, low/high TB and DR-TB rates), number of isolates included, risk group/population status, reference standard phenotypic DST method(s) used (including concentrations), data on new/retreatment cases and genetic lineages as well as information on whether study data set was original or downloaded from a public domain.
- Principal findings: number of isolates found to be phenotypically resistant with and without mutations in relevant genes; number of isolates found to be phenotypically sensitive with and without mutations in relevant genes; information on whether phylogenetic markers have been excluded; information on whether WGS results have been confirmed with Sanger sequencing; sensitivity, specificity, positive-predictive value and negative-predictive value;
- Technical details: WGS platform and technology used, genome coverage and sequencing depth, quality control criteria (minimum depth for valid call; minimum consensus rate threshold for valid call; exclusion of PE/PPE/PGRS and repetitive sequences).

Analysis

A Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) diagram was constructed (Fig. 1). The sensitivity, specificity, positive-predictive value and negative-predictive value of WGS compared with phenotypic DST was determined. Forest plots for WGS sensitivities and specificities for selected drugs (rifampicin and isoniazid) were constructed and pooled sensitivities were calculated using MS Excel plugin for forest plots based on a random effects model [20]. For purposes of consistency, authors of the current review manually excluded two recognized lineage-defining SNPs (R463L in katG gene and S95T in gyrA gene, known to define Euro-American lineage and not conferring drug resistance) from the calculations of performance characteristics for the studies where polymorphisms have not been excluded by the respective authors.
Assessment of individual studies’ methodological quality

We assessed the quality of studies using QUADAS-2, a validated tool for diagnostic studies [21]. The QUADAS-2 protocol is available in the online Supplementary material (Appendix S2). Reported technical quality control criteria and cut-off values were derived, analysed and summary recommendations were included in the conclusions.

Results

Study selection

Initially, a total of 2492 publications were identified. Among the 2492 articles, 935 duplicates were excluded, leaving 1557 publications selected for abstract analysis. Following an abstract analysis performed by two reviewers independently, a consensus-based decision was made to exclude a further 1490 articles as not being relevant (not a WGS, not MTBC, and not relevant for DST) (Fig. 1). This left 67 publications for a full-text analysis. Upon completion of full-text analysis, four publications where full texts were not available and 43 more publications, not satisfying eligibility criteria (no direct comparison of WGS with phenotypic DST, no data available to calculate performance characteristics, review papers with no new data, data from one isolate only, data from laboratory-generated mutants) were excluded; 20 publications were finally selected for the current systematic review [4–7,12,14,22–35].

Methodological quality of studies

The overall methodological quality of the included studies is summarized in Fig. 2. Additional information on the quality assessment for each separate study is given in the Supplementary material. The majority of studies had a high risk of bias regarding the selection of samples/patients (see Supplementary material, Table S6) and risk for bias of the index test was unclear in most
studies [14]. We judged applicability for the reference standard to be of ‘low concern’ for all studies.

General characteristics of included studies and data available

General characteristics of included studies and WGS technical data, quality control and polymorphism calling criteria are presented in the Supplementary material (Appendix S3; summarized in Table S3 and S4). Chemistry platforms and hardware for sample analysis used are depicted in the Supplementary material (Fig. S1).

Whole genome data were available for 6906 MTBC isolates from 20 studies with a number of isolates in individual studies varying from 2 to 3651. The full list of drugs and corresponding genes and other genomic regions associated with the resistance to specific drugs, as well as WGS-calculated performance characteristics are presented in Table 1. Drug concentrations used for phenotypic testing were reported in all studies; predominantly standard WHO-recommended methods [36–39] were in use in all laboratories.

Drugs and genes/regions associated with drug resistance

*Mycobacterium tuberculosis* strains included in the reports were tested for resistance for a total of 22 anti-TB agents (Table 1, and see Supplementary material, Table S5); not all drugs were tested in all studies. Polymorphisms in a total of 53 genes were tested for associations with drug resistance; regulatory regions (promoters) of three genes (*inhA*, *pncA*, *ethA*) have also been included into the analysis. Selection of genes harbouring mutations tentatively associated with resistance to specific drugs in individual studies was largely based on the evidence available in the literature as well as in online databases TBDReamDB (https://tbdreamdb.ki.se) [40] and MUBII-TB-DB ([http://omictools.com/mubii-tb-db-tool] [41]. The number of genes tentatively associated with drug resistance ranged from one (rifabutin, minocycline, trimethoprim/sulfamethoxazole, bedaquiline, clofazimine) to 19 (isoniazid) per individual drug. Notably, in only a small number of studies [30,34] were polymorphisms in genes associated with resistance to linezolid and para-aminosalicylic acid included in the analysis; in others no associations were reported because of a lack of available phenotypic data. For the same reason, it was not possible to analyse performance characteristics of WGS for detection of resistance to bedaquiline, clofazimine and trimethoprim/sulfamethoxazole [6].

**Synonymous mutations, phylogenetic markers and other challenges**

Many studies included in the current systematic review lacked clarity as to whether polymorphisms detected in the specific genes have been checked for their associations with drug resistance and whether phylogenetic SNPs had been filtered out (see Supplementary material, Table S5); the same applies to the exclusion of synonymous SNPs. Details on polymorphisms excluded from the analysis and reasons for that have not been systematically reported. In four reports only [4–6,28] both synonymous polymorphisms and phylogenetic markers have been systematically excluded from analysis; in three papers [7,22,29] synonymous SNPs and some phylogenetic markers have been filtered out, and in two more studies [26,30] authors stated that analysis was confined to mutations known to be associated with drug resistance. For consistency, the authors of the current review manually excluded two major recognized lineage-defining SNPs (R463L in *katG* gene and S95T in *gyrA* gene) from the calculations of performance characteristics for the remaining studies as these polymorphisms are known to define Euro-American lineage and do not confer drug resistance [42].

**Performance characteristics**

Performance characteristics of WGS for detection of drug resistance are summarized in Table 1 and Figs 3 and 4, and see Supplementary material (Table S5). The sensitivity for the detection of resistance to rifampicin and isoniazid was high and varied from 0.89 to 1.00 and from 0.9 to 1.00, with pooled estimates of 0.98 (95% CI 0.93–0.98) and 0.97 (95% CI 0.94–0.99), respectively. The specificities for rifampicin and isoniazid varied from 0.67 to 1.00, and 0.83 to 1.00 with pooled estimates of 0.98 (95% CI 0.98–1.00) and 0.93 (95% CI 0.91–0.96), respectively. Sensitivity for isoniazid resistance detection did not vary significantly depending on the number of genes included in the analysis; similarly, there were no differences in analytical performance of WGS for rifampicin and isoniazid resistance depending on the media and methodology of phenotypic tests.
Due to high heterogeneity between studies and/or a very small number of reports (e.g. single reports for para-aminosalicylic acid, minocycline and rifabutin), pooled sensitivities and specificities were not calculated for other drugs. For the remaining first-line drugs, individual sensitivities varied from 0.71 to 1.00; 0.43 to 1.00, and 0.57 to 1.00 for ethambutol, pyrazinamide and streptomycin, respectively; variability in specificity characteristics was even greater (0.15 to 95.8; 0.67 to 1.00; and 0.4 to 1.00, respectively).

There was also a significant variation in analytical performance of WGS for injectable drugs (amikacin, capreomycin and kanamycin) and fluoroquinolones (ciprofloxacin, ofloxacin, moxifloxacin, and levofloxacin).

<table>
<thead>
<tr>
<th>Drug</th>
<th>No of studies</th>
<th>No of strains</th>
<th>Genes and other relevant Mycobacterium tuberculosis genome regions</th>
<th>Sensitivity, % (range)</th>
<th>Specificity, % (range)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rifampicin</td>
<td>19</td>
<td>6286</td>
<td>rpoB, rpoA, rpoC</td>
<td>89.2 100.0</td>
<td>66.7 100.0</td>
</tr>
<tr>
<td>Isoniazid</td>
<td>19</td>
<td>5800</td>
<td>katG, inhA, oxyR-ahpc, fabC, Rv1592C, Rv1772, Rv2242, fabD, fabG1, kasA, accD, oxyR, ndh, fabD24, nat, kasA, mabA, p_inhA, accD6, fabA</td>
<td>90.0 100.0</td>
<td>83.3 100.0</td>
</tr>
<tr>
<td>Ethambutol</td>
<td>17</td>
<td>6059</td>
<td>embA, embB, embC, embR, inhA, inhB, inic, Rv3124, manB, PPE40, mntD, manB</td>
<td>71.4 100.0</td>
<td>15.4 95.8</td>
</tr>
<tr>
<td>Pyrazinamide</td>
<td>13</td>
<td>6130</td>
<td>pncA, p_pncA, rpsA, panD</td>
<td>43.2 100.0</td>
<td>66.7 100.0</td>
</tr>
<tr>
<td>Streptomycin</td>
<td>16</td>
<td>3953</td>
<td>rpsL, rrs, gidB</td>
<td>57.1 100.0</td>
<td>40.0 100.0</td>
</tr>
<tr>
<td>Amikacin b</td>
<td>8</td>
<td>1471</td>
<td>rrs, es, gidB, thyA</td>
<td>80.0 100.0</td>
<td>50.0 100.0</td>
</tr>
<tr>
<td>Capreomycin b</td>
<td>8</td>
<td>1553</td>
<td>rrs, es, gidB, thyA</td>
<td>60.7 100.0</td>
<td>13.7 100.0</td>
</tr>
<tr>
<td>Kanamycin b</td>
<td>7</td>
<td>1289</td>
<td>rrs, es, thyA</td>
<td>75.0 100.0</td>
<td>0.0 100.0</td>
</tr>
<tr>
<td>Injectable drugs c</td>
<td>4</td>
<td>518</td>
<td>rrs, es, gidB, thyA</td>
<td>37.0 100.0</td>
<td>50.0 100.0</td>
</tr>
<tr>
<td>Ciprofloxacin</td>
<td>1</td>
<td>300</td>
<td>gyrA, gyrB</td>
<td>100.0 100.0</td>
<td>98.9 98.9</td>
</tr>
<tr>
<td>Ofloxacin</td>
<td>6</td>
<td>1564</td>
<td>gyrA, gyrB</td>
<td>80.0 100.0</td>
<td>80.0 100.0</td>
</tr>
<tr>
<td>Moxifloxacin                 c</td>
<td>3</td>
<td>1318</td>
<td>gyrA, gyrB</td>
<td>60.0 90.9</td>
<td>68.7 100.0</td>
</tr>
<tr>
<td>Levofoxacin</td>
<td></td>
<td></td>
<td>gyrA, gyrB</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Gatifloxacin                 c</td>
<td></td>
<td></td>
<td>gyrA, gyrB</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Fluoroquinolones c</td>
<td>9</td>
<td>504</td>
<td>gyrA, gyrB</td>
<td>89.2 100.0</td>
<td>100.0 100.0</td>
</tr>
<tr>
<td>Ethionamide</td>
<td>8</td>
<td>867</td>
<td>ethA, ethB, p_inhA, inhA, fabG1</td>
<td>16.7 100.0</td>
<td>50.0 100.0</td>
</tr>
<tr>
<td>Prothionamide</td>
<td>3</td>
<td>502</td>
<td>p_ethA, ethA</td>
<td>40.0 100.0</td>
<td>29.4 80.0</td>
</tr>
<tr>
<td>Rifabutin</td>
<td>1</td>
<td>2</td>
<td>rpoC</td>
<td>100.0 100.0</td>
<td>---</td>
</tr>
<tr>
<td>Para-aminosalicylic acid</td>
<td>1</td>
<td>11</td>
<td>thyA, folC, ribB</td>
<td>75.0 75.0</td>
<td>100.0 100.0</td>
</tr>
<tr>
<td>Trimethoprim/sulfamethoxazole</td>
<td>1</td>
<td>2</td>
<td>dfrA</td>
<td>---</td>
<td>100.0 100.0</td>
</tr>
<tr>
<td>Minocycline</td>
<td>1</td>
<td>2</td>
<td>whiB7</td>
<td>---</td>
<td>100.0 100.0</td>
</tr>
<tr>
<td>Linezolid</td>
<td>2</td>
<td>5</td>
<td>Rrl, rplC</td>
<td>100.0 100.0</td>
<td>---</td>
</tr>
<tr>
<td>Bedaquiline</td>
<td></td>
<td></td>
<td>---</td>
<td>100.0 100.0</td>
<td>---</td>
</tr>
<tr>
<td>Clofazimine</td>
<td></td>
<td></td>
<td>---</td>
<td>100.0 100.0</td>
<td>---</td>
</tr>
<tr>
<td>Amikacin b</td>
<td></td>
<td>ME 718</td>
<td>rr73, eis, gidB</td>
<td>80.0 100.0</td>
<td>75.0 100.0</td>
</tr>
<tr>
<td>Capreomycin b</td>
<td></td>
<td>ME 1553</td>
<td>rr73, eis, gidB, thyA</td>
<td>60.0 100.0</td>
<td>13.7 100.0</td>
</tr>
<tr>
<td>Kanamycin b</td>
<td></td>
<td>ME 1289</td>
<td>rr73, eis, thyA</td>
<td>75.0 100.0</td>
<td>0.0 100.0</td>
</tr>
<tr>
<td>Injectable drugs d</td>
<td></td>
<td>518</td>
<td>rr73, eis, gidB</td>
<td>37.0 100.0</td>
<td>50.0 100.0</td>
</tr>
<tr>
<td>Ciprofloxacin</td>
<td></td>
<td>300</td>
<td>gyrA, gyrB</td>
<td>100.0 100.0</td>
<td>98.9 98.9</td>
</tr>
<tr>
<td>Ofloxacin</td>
<td></td>
<td>1564</td>
<td>gyrA, gyrB</td>
<td>80.0 100.0</td>
<td>80.0 100.0</td>
</tr>
<tr>
<td>Moxifloxacin                 c</td>
<td></td>
<td>1318</td>
<td>gyrA, gyrB</td>
<td>60.0 90.9</td>
<td>68.7 100.0</td>
</tr>
<tr>
<td>Fluoroquinolones c</td>
<td></td>
<td>504</td>
<td>gyrA, gyrB</td>
<td>89.2 100.0</td>
<td>100.0 100.0</td>
</tr>
<tr>
<td>Ethionamide</td>
<td></td>
<td>867</td>
<td>ethA, ethB, p_inhA, inhA, fabG1</td>
<td>16.7 100.0</td>
<td>50.0 100.0</td>
</tr>
<tr>
<td>Prothionamide</td>
<td></td>
<td>502</td>
<td>p_ethA, ethA</td>
<td>40.0 100.0</td>
<td>29.4 80.0</td>
</tr>
<tr>
<td>Rifabutin</td>
<td></td>
<td>2</td>
<td>rpoC</td>
<td>100.0 100.0</td>
<td>---</td>
</tr>
<tr>
<td>Para-aminosalicylic acid</td>
<td></td>
<td>11</td>
<td>thyA, folC, ribB</td>
<td>75.0 75.0</td>
<td>100.0 100.0</td>
</tr>
<tr>
<td>Trimethoprim/sulfamethoxazole</td>
<td></td>
<td>2</td>
<td>dfrA</td>
<td>---</td>
<td>100.0 100.0</td>
</tr>
<tr>
<td>Minocycline</td>
<td></td>
<td>2</td>
<td>whiB7</td>
<td>---</td>
<td>100.0 100.0</td>
</tr>
<tr>
<td>Linezolid</td>
<td></td>
<td>5</td>
<td>Rrl, rplC</td>
<td>100.0 100.0</td>
<td>---</td>
</tr>
<tr>
<td>Bedaquiline</td>
<td></td>
<td></td>
<td>---</td>
<td>100.0 100.0</td>
<td>---</td>
</tr>
<tr>
<td>Clofazimine</td>
<td></td>
<td></td>
<td>---</td>
<td>100.0 100.0</td>
<td>---</td>
</tr>
</tbody>
</table>

*Not all genes were tested in all studies.*

*Calculated only using studies where performances for different injectable drugs (amikacin, capreomycin, and kanamycin) were analysed separately; otherwise please see performance characteristics for injectable drugs.*

*Calculated only using studies where performances for different fluoroquinolones were analysed separately; otherwise please see performance characteristics for fluoroquinolone.*

*Calculated only using studies where combined performance characteristics for injectable drugs were analysed.*

*Calculated only using studies where combined performance characteristics for fluoroquinolones were analysed.*

Due to high heterogeneity between studies and/or a very small number of reports (e.g. single reports for para-aminosalicylic acid, minocycline and rifabutin), pooled sensitivities and specificities were not calculated for other drugs. For the remaining first-line drugs, individual sensitivities varied from 0.71 to 1.00; 0.43 to 1.00, and 0.57 to 1.00 for ethambutol, pyrazinamide and streptomycin, respectively; variability in specificity characteristics was even greater (0.15 to 95.8; 0.67 to 1.00; and 0.4 to 1.00, respectively).

There was also a significant variation in analytical performance of WGS for injectable drugs (amikacin, capreomycin and kanamycin) and fluoroquinolones (ciprofloxacin, ofloxacin, moxifloxacin, and levofloxacin).

![Fig. 3. Forest plot of studies evaluating (a) sensitivity and (b) specificity of whole genome sequencing for detection of resistance to rifampicin.](image-url)
moxifloxacin, gatifloxacin). Overall sensitivity and specificity characteristics were better for fluoroquinolones compared with those for aminoglycosides and cyclic peptides (injectable drugs), e.g. sensitivities 89.2% versus 60.7%, respectively. Both sensitivity and specificity for thioamides were generally suboptimal varying from 0.17 to 1.00 and from 0.29 to 1.00 for ethionamide and prothionamide, respectively.

In studies where M. tuberculosis genetic lineages were reported and strains belonging to more than one lineage were analysed, results have not been stratified by lineages by the original study authors and therefore comparison of analytic performance of WGS depending on genetic group was not possible. Due to the lack of data it was also not possible to assess performance of WGS in new versus retreatment cases and specific patient risk groups.

Discussion

Data derived from the 20 selected publications satisfying selection criteria for the current systematic review on the role and added value of WGS of MTBC for the detection of drug resistance appears to be heterogeneous. Available evidence indicates that analytical performance characteristics of WGS for the detection of resistance to the two most important first-line drugs is high with pooled sensitivity and specificity values of 0.98 (95% CI 0.93–0.98) and 0.98 (95% CI 0.98–1.00) for rifampicin and 0.97 (95% CI 0.94–0.99) and 0.93 (95% CI 0.91–0.96) for isoniazid, respectively. This suggests that WGS could be considered a promising alternative to existing phenotypic and molecular DST methods for rifampicin and isoniazid pending standardization of analytical pipelines, software parameters, and databases and other resources specifying the role of specific genes and mutations in the development of drug resistance.

Performance characteristics for other first-line and reserved drugs varied considerably, preventing authors from drawing any conclusions on the potential role of WGS in routine DST for the majority of anti-TB drugs at present. For the remaining first-line drugs, sensitivities and specificities varied significantly, possibly due to the apparent lack of data on molecular mechanisms and genes involved in the development of resistance to streptomycin, ethambutol and pyrazinamide, as well as lack of standardization in methodologies and drug concentrations for phenotypic tests for pyrazinamide [15,43,44]. For the reserve drugs, better analytical performance of WGS for fluoroquinolones compared with those for aminoglycosides and cyclic peptides (injectable drugs), could be explained by a better understanding of mechanisms and a smaller number of genes associated with resistance to fluoroquinolones (gyrA and gyrB) compared with injectable drugs (rrs, eis, gidB and tlyA).

This systematic review has several strengths and limitations. The review used an extensive search strategy applying international criteria of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses group. Due to heterogeneity of data and lack of internationally agreed quality criteria for WGS studies in the context of MTBC molecular DST, quality assessment was challenging. Limitations of the current review included a relatively small number of studies satisfying the selection criteria, the retrospective nature of the study design in most selected studies, small sample sizes in some studies, lack of standardization of analytical pipelines, software parameters and databases, as well as lack of data related to specific MTBC lineages and high-risk population groups. A significant proportion of studies reviewed here (17/20), exclusively or partially included strains collected in high TB/DR-TB settings characterized by a dominance of specific lineages and higher prevalence of certain resistance mutations; therefore, extrapolation of findings to other TB settings should be made with caution. We have concluded that no evidence currently exists on the variability of WGS performance on strains derived from new versus retreatment patients or strains belonging to specific genetic lineages due to the lack of systematic data. The same applies to cost-effectiveness as few economic data could be extracted from the reports.

The high variations in WGS analytical performance for MTBC molecular DST may not reflect limitations of the method as such, but rather highlight certain limitations of some of the original studies and the importance of the development and standardization of post-processing procedures and algorithms, including, but not limited to (a) the development and extensive validation of databases and other resources containing information on mutations associated with drug resistance (including information on high- and low-level resistance conferring mutations); and (b) the development of standardized analytical pipelines for the differentiation of mutations, associated with drug resistance from the phylogenetic markers and synonymous mutations. Both tasks will most likely require further research into the molecular mechanisms.
of drug-resistance development conducted on large population-based samples of *M. tuberculosis* cultures collected in different geographical areas and epidemiological settings, as some genetic lineages are known to be associated with a greater level of drug resistance and bear specific polymorphisms (e.g. Beijing).

In the context of interpretation of WGS data related to detection of resistance to specific drugs, it is very important to distinguish between at least two major types of polymorphisms found in genes associated with drug resistance [6,45]: (a) polymorphisms actually conferring drug resistance (and therefore being under selective pressure); and (b) other polymorphisms, including lineage-defining and other phylogenetically important variations not associated with drug resistance. These could be further subdivided into synonymous and non-synonymous SNPs, with each group having a specific impact on the functionality and regulation of the MTBC genome. Those SNPs that belong to the latter group are often misclassified as drug resistance markers; obviously inclusion of these polymorphisms in the calculations of performance characteristics of WGS would affect the results significantly. However, lack of standardized analytic pipelines and absence of recognized/widely available sources of such information make the differentiation between these two types of polymorphisms and subsequent calculations of analytical performance of WGS a challenging task.

Another issue adding to the complexity of the analysis of the WGS data in the context of the diagnosis of drug resistance is low-level phenotypic resistance. Some mutations confer ‘low-level’ resistance in MTBC, resulting in ambiguous results from phenotypic tests and discrepancies between the results of tests on liquid (e.g. BACTEC MGIT 960) and solid media and WGS. Currently endorsed methods of phenotypic DST essentially provide only qualitative results (sensitive or resistant with few exceptions, e.g. resistance ratio on solid media), which may result in disagreements between the results of molecular and phenotypic tests, especially where mutations conferring low-level drug resistance are involved. Validation and implementation of microtitre-plate-based methods (e.g. Sensititre MycoTB plates; Trek Diagnostics, Cleveland, OH, USA) and other methodologies based in the determination of MIC could be considered a promising alternative [46–48].

Whole genome sequencing is still a relatively novel approach for TB strain analysis. No commercial kits incorporating all stages of analysis are available so far and both ‘wet’ laboratory technology and data-processing pipelines have not been fully standardized and so require further optimization. Technical data on the sample processing and bioinformatics analysis provided in the reports may help in assessing the quality of the studies and, ideally, developing quality control criteria and parameters that should be applied for the analysis and reported to ensure comparability of studies and reproducibility of results.

Based on the data drawn from the reviewed reports, we would like to propose a minimum set of WGS technical parameters, as well as types of outcomes and performance characteristics that should be included in future reports to ensure comparability of studies and improve the quality of evidence in future (see Supplementary material, Appendix S4).

In conclusion, the current study is the first attempt to summarize and review available evidence on the role and added value of WGS for the detection of resistance to anti-TB drugs in a systematic manner. Heterogeneity in study design and methodologies used for phenotypic testing, lack of certain types of data and insufficient knowledge of mechanisms of the development of resistance to specific drugs, relatively small sample sizes, and, importantly, lack of clarity on principles of exclusion of phylogenetic markers and synonymous SNPs in some studies, affected calculations of WGS analytical performance characteristics and overall conclusions. Future studies will ideally include information on the clinical outcomes to ensure the clinical relevance of results obtained in *vitro* in diagnostic laboratories. Current developments in phenotypic DST (e.g. determination of MIC using microtitre plates) may help in more accurately assessing the performance of WGS for the detection of drug resistance and the role of this method in TB laboratory diagnosis.

**Transparency Declaration**

Authors declare no conflicts of interests.

**Acknowledgements and Role of the funding source**

This study has received funding from the European Centre for Disease Control (ECDC) under the Grants 2009/004 and 2014/001. The funding body had no role in the study design or data analysis.

**Authors’ contributions**

VN, DP ad FD designed the study. VN, DP and IK performed the literature search, selection of publications, data extraction, analysis and drafted the manuscript. All co-authors significantly contributed to the data analysis, writing and critically revised the manuscript. All co-authors read and approved a final version of the manuscript prior to its submission.

**Appendix A. Supplementary data**

Additional Supporting Information may be found in the online version of this article at http://dx.doi.org/10.1016/j.cmi.2016.09.008.

**References**


