

Whole genome sequencing voor screening op resistentie *Mycobacterium tuberculosis* complex

Rina de Zwaan, Miranda Kamst, Han de Neeling, Tridia van der Laan, Arnout Mulder, Richard Anthony, Dick van Soolingen

Samenvatting

Whole genome sequencing (WGS) van *Mycobacterium tuberculosis* complexisolaten wordt bij het Nationaal Referentielaboratorium Tuberculose van het RIVM gebruikt voor (sub)speciesidentificatie, voor screening op mogelijke resistentie tegen de eerstelijnsantibiotica en voor de epidemiologische typering. De laatste toepassing werd ingevoerd in 2019, de andere twee in 2020. Dit artikel betreft de evaluatie van het screenen op mogelijke resistentie met WGS in het eerste jaar.

Van de 441 isolaten van tuberculosepatiënten hadden er 373 (85 procent) geen resistentiemutaties, waardoor er geen aanvullende fenotypische gevoeligheidsbepalingen hoefden te worden uitgevoerd. De definitieve uitslag was in deze gevallen geheel gebaseerd op de WGS-data. De overige 68 (15 procent) isolaten vertoonden ten minste één mutatie die mogelijk geassocieerd was met resistentie. Dit betrof 93 mutaties, waarvan 44 met een hoge voorspellende waarde voor resistentie en hierbij kwam de WGS-analyse overeen met de fenotypische gevoeligheidsbepaling. In totaal kon bij 92 procent van de cases een betrouwbare uitslag van de gevoeligheid worden afgegeven uitsluitend op basis van de WGS-data. Voor de overige isolaten waren aanvullende fenotypisch gevoeligheidsbepalingen noodzakelijk, vanwege nog ontbrekende kennis over de betekenis van de gevonden mutaties.

Door de toenemende kennis over de correlatie tussen genotypische en fenotypische data is de WGS een betrouwbare methode geworden om te screenen op resistentie bij *M. tuberculosis* complexisolaten.

Abstract

Whole genome sequencing (WGS) of *Mycobacterium tuberculosis* complex isolates is used at the RIVM tuberculosis reference laboratory for (sub)species identification, screening on possible resistance against first line antibiotics and epidemiological typing. The

latter application was implemented in 2019, the other two in 2020. This article is an evaluation of the first year of screening for possible resistance with WGS.

In 373 (85 per cent) isolates from 441 tuberculosis patients no resistance mutation was found, therefore no additional phenotypical testing was performed. In these cases the final report was solely based on WGS data. For the remaining 68 (15 per cent) isolates at least one mutation possibly associated with resistance was found. This concerned 93 mutations, of which 44 with a high predictive value for resistance and in these cases the WGS prediction was in concordance with the results of phenotypical testing. For in total 92 per cent of the cases a reliable susceptibility profile could be generated solely based on WGS data. For the remaining isolates additional phenotypical testing was needed because of incomplete knowledge on the significance of detected mutations.

Because of an expanding knowledge on the correlation between genotypic and phenotypic data, WGS has become a reliable tool in the screening on resistance in *M. tuberculosis* complex isolates.

Introductie

Sinds in 1998 de eerste publicatie over whole genome sequencing (WGS) van de *Mycobacterium tuberculosis*-referentiestam H37Rv verscheen, is de WGS-techniek snel verder ontwikkeld [1].

Rijksinstituut voor Volksgezondheid en Milieu (RIVM),
Centrum Infectieziekteonderzoek, Diagnostiek en
Laboratorium Surveillance, Nationaal
Referentielaboratorium Tuberculose, R.C.C. de Zwaan,
laboratoriumanalist, M.M. Kamst, hoofdanalist,
dr. A.J. de Neeling, bioinformaticus,
T.G. van der Laan, analist, A.A. Mulder, analist,
dr. R.M. Anthony, senior onderzoeker,
prof. (em.) dr. D. van Soolingen, Hoofd Nationaal
Referentielaboratorium Tuberculose.
Correspondentieadres: Rina de Zwaan
(rina.de.zwaan@rivm.nl).

In 2018 publiceerde het internationale CRyPTIC-consortium een uitgebreide studie, met deelname vanuit 16 verschillende landen verdeeld over zes continenten, met meer dan tienduizend vergelijkingen tussen de genotypische en fenotypische gevoeligheidsbepalingen van de eerstelijnsantibiotica: isoniazide (INH), rifampicine (RIF), ethambutol (EMB) en pyrazinamide (PZA). In deze studie werd een dermate hoge negatieve voorspellende waarde gevonden voor resistentie tegen deze antibiotica dat werd geconcludeerd dat met WGS voldoende zekerheid te verkrijgen is om deze ook voor klinisch gebruik in te zetten [2].

Vanaf 2008 wordt bij het Nationaal Referentielaboratorium Tuberculose van het RIVM gewerkt aan WGS van *M. tuberculosis* [3]. Vanaf 2016 zijn alle eerste *M. tuberculosis* complexisolaten van tbc-patiënten aan WGS onderworpen en is de moleculaire analyse vergeleken met de gebruikte routine-methoden: identificatie op basis van GenoType MTBC (Hain Lifescience, GmbH, Nehren, Duitsland), de gevoeligheidsbepaling met behulp van de BACTEC MGIT 960 (Becton Dickinson, New Jersey, Verenigde Staten) en epidemiologische clustering op basis van de Variable Number of Tandem Repeat (VNTR)-typering. Het eerste doel betrof het inzetten van WGS voor de optimalisatie van de epidemiologische typering om het percentage 'valse' clustering bij het epidemiologische onderzoek te verminderen. Naar aanleiding van het clusteronderzoek van de GGD'en bleek de clustering op basis van de WGS voor het aanduiden van tbc-transmissie veel betrouwbaarder dan die op grond van de VNTR-typering [4]. In 2019 werd daarom de overstap gemaakt van VNTR-typering naar WGS.

Na aanvullende WGS-studies in Nederland voor zowel identificatie als gevoeligheidsbepalingen, waarbij onder andere voor de screening op resistentiemutaties de negatieve voorspellende waarden voor de eerstelijnsantibiotica in Nederland hoger bleken te liggen dan bij de internationale CRyPTIC-studie, was er in 2020 voor beide bepalingen voldoende vertrouwen verkregen om de overstap te maken naar WGS [5,6].

In deze studie worden de resultaten van een jaar lang screenen van *M. tuberculosis* complexisolaten op resistentie tegen de eerstelijnsantibiotica met WGS beschreven.

Materiaal en methoden

Isolaten

Van alle kweek-positieve tuberculosepatiënten in Nederland wordt een *M. tuberculosis* complexkweek naar het RIVM gestuurd voor de epidemiologische typering. Tevens wordt hierbij een (sub)species-identificatie uitgevoerd en wordt er gescreend op mogelijke resistentie tegen de eerstelijnsantibiotica. Het RIVM analyseerde de genomische data van 441 eerste positieve kweken van tuberculosepatiënten die in 2020 in Nederland waren geïsoleerd.

DNA-isolatie

De ontvangen kweken werden overgeënt in een MGIT-mediumbuis, waarna van de gegroeide kweek het DNA werd geïsoleerd met behulp van de QIAamp DNA mini kit (Qiagen, GmbH, Hilden, Duitsland).

WGS-sequencing

Illumina sequencing werd uitgevoerd door de firma Baseclear (Leiden). De sequenties werden geanalyseerd in de door het RIVM ontwikkelde bio-informatica pipeline [7]. De pipeline maakt gebruik van Bowtie2 voor de mapping van de paired-end reads tegen het H37Rv-referentiegenoom (GeneBanknr. AL123456.3) en Breseq versie 0.28.1 voor detectie van single nucleotide positions (SNP's). Toekenning van SNP's wordt gedaan op basis van een allelfrequentie van 80 procent of hoger. De WGS-gegevens werden alleen gebruikt bij een minimale gemiddelde coverage depth van 50. Voor screening op contaminatie met non-MTBC-DNA werden de algemene bacteriële identificatiegenen onderzocht (*rrs* & *rrl* minder dan drie SNP). Gemengde infecties en minderheidspopulaties werden gedetecteerd door middel van analyse van de 62 Coll-identificatie-SNP's [8] en 15 INH- en RIF-resistentie-SNP's.

Speciesidentificatie

Speciesidentificatie werd bepaald met behulp van de Coll-methode [8], in combinatie met de SNP-IT tool [5].

Screening op mutaties in resistentiegenen

De WGS-gegevens werden, met behulp van de pipeline, gescreend met een lijst van genoomposities op mutaties die mogelijk met resistentie geassocieerd waren. Voor het detecteren van 'nieuwe' SNP's, substituties (SUB) en inserties/deleties (INDEL's) werd

een lijst met hotspotgebieden in de bekendste resistentiegenen gebruikt.

De screeningslijsten werden samengesteld met als basis de ReSeqTB-databank van 2019 [9], aangevuld met informatie uit onder andere het WHO-rapport van 2019 [10], de CRyPTIC 10K-studie [2], de eigen RIVM-databank (ongepubliceerde data), en verder gepubliceerde data [6]. Deze screeningslijsten bestaan onder andere uit de negen genen zoals beschreven in de internationale CRyPTIC-studie en de aanvullende studie van het RIVM [2,6]: *rpoB*, *katG*, *inhA*, *fabG1*, *ahpC*, *embB*, *embA*, *pncA* en *ppsA*. Daarnaast werd de screeningslijst verder uitgebreid met extra genen, op basis van onder andere de ReSeqTB-databank [9], om de kennis over fenotypische resistentie te verbreden en om minder vaak gevonden mutaties niet te missen; *furA*, *inbR*, *kasA*, *mmaA3*, *mshA*, *mshB*, *mshC*, *mymA*, *nat*, *ndh*, *nudC*, *sigI*, *ponA1*, *rpoA*, *rpoC*, *aftA*, *embC*, *embR*, *ubiA*, *clpC1*, *gpsI*, *mas*, *panD*, *ppsA*, *ppsC*, *ppsD*, *proZ*, *Rv0191*, *Rv1667c*, *Rv2731* en *Rv3008*.

De voorspellende waarde van resistentiemutaties werd gebaseerd op het aantal, het percentage positieve associaties van de gevonden waarnemingen en op het voorkomen van de mutaties in diverse bronnen [2,9,10]. Zo moet een mutatie voor een hoge voorspellende waarde in verschillende bronnen zijn vermeld en dient in meer dan 75 procent van de gevallen de resistentie te zijn bepaald. De meeste mutaties in de lijst hebben echter een lage betrouwbaarheid omdat deze maar weinig zijn gevonden. De screeningslijst werd en wordt steeds aangepast door toevoegingen van nieuwe informatie (literatuur) en door verwijderingen van mutaties

wanneer een vermoedelijke resistentiemutatie minimaal vijf keer werd gevonden, maar alle fenotypische gevoeligheidsbepalingen als gevoelig werden beoordeeld.

Wanneer voor één antibioticum verschillende mutaties werden aangetoond die mogelijk geassocieerd waren met resistentie tegen een bepaald antibioticum, werden deze als één mutatie geteld; bijvoorbeeld een INH-mutatie *katG_S315T* samen met INH-mutatie *fabG1_c-15t* werd geteld als één mutatie, waarbij de hoogste betrouwbaarheid van de twee mutaties werd aangehouden. Mutaties coderend voor verschillende antibiotica werden als verschillende mutaties beschouwd.

Fenotypische gevoeligheidsbepaling

De fenotypische gevoeligheidsbepalingen werden uitgevoerd met de BACTEC MGIT 960. Wanneer voor een antibioticum een resistentiemutatie werd gevonden zonder hoge voorspellende waarde, werd voor dat betreffende antibioticum een fenotypische gevoeligheidsbepaling ingezet. Wanneer er sprake was van een hoge voorspellende waarde of als er mutaties werden gevonden tegen verschillende antibiotica, werden voor alle vier de eerstelijns-antibiotica fenotypische bepalingen ingezet. Dit omdat bij de aanwezigheid van resistentie tegen één antibioticum de kans op andere resistenties toeneemt. Voor *M. bovis* BCG-isolaten die behoorden tot clusters die geassocieerd zijn met intermediaire INH-resistentie werd ook de fenotypische gevoeligheidsbepaling voor INH ingezet, ondanks de afwezigheid van bekende resistentiemutaties. Daarnaast werd voor isolaten van

Tabel 1. Overzicht van de resultaten van fenotypische gevoeligheidsbepalingen per isolaat.

	Aantal	Gevoelig		Verhoogde MIC	
		Aantal	Percentage	Aantal	Percentage
Isolaten met mono fenotypische gevoeligheidsbepaling	28	23	82%	5	18%
Isolaten met complete eerstelijns fenotypische gevoeligheidsbepaling	40	4	10%	36	90%
Totaal isolaten met een fenotypische gevoeligheidsbepaling	68	27	40%	41	60%

Tabel 2. Overzicht van de gevonden mutaties en hun betrouwbaarheid per antibioticumbepaling.

Betrouwbaarheid op basis van WGS	INH			RIF			EMB			PZA			Alle		
	N	I/R		N	R		N	R		N	R		N	R	
Hoogbetrouwbare mutaties	33	33	100%	9	9	100%	0	0	0%	1*	1	100%	43	43	100%
Mediumbetrouwbare / nieuwe mutaties	9	2	22%	4	2	50%	7	3	43%	3	3	100%	23	10	43%
Laagbetrouwbare mutaties	9	0	0%	3	0	0%	0	0	0%	15	4	27%	27	4	15%
Totaal gedetecteerde mutaties	51	35	69%	16	11	69%	7	3	43%	19	8	42%	93	57	61%
Geen relevante mutaties	390	ND	ND	425	ND	ND	434	ND	ND	422	ND	ND			

N = aantal; I/R = aantal intermediair/resistent; *geëxcludeerd; *pncA*_H57D (*M. bovis* (BCG)) en *pncA*_A46A & *panD*_A13A & M117T (*M. canettii*).

M. bovis en *M. canettii*-species, die intrinsiek resistent zijn tegen PZA, geen fenotypische gevoeligheidsbepaling ingezet, tenzij er andere mogelijke resistentiemutaties werden gevonden.

Resultaten

Het RIVM ontving 441 eerste *M. tuberculosis* complexisolaten die geschikt waren voor WGS-analyse. Deze werden gedetermineerd als 415 *M. tuberculosis*, vier *M. africanum*, één *M. canettii*, zes *M. bovis* en 15 *M. bovis* BCG. Onder de *M. bovis* BCG-isolaten was er één geassocieerd met mogelijke intermediaire resistentie tegen INH op basis van de clustering en hiervan werd een fenotypische INH-bepaling uitgevoerd.

Van de 441 isolaten waren er 373 (85 procent) zonder mutaties die mogelijk geassocieerd waren met resistentie tegen de eerstelijns antibiotica. Voor deze isolaten werd geen gevoeligheidsbepaling uitgevoerd. Bij de overige 68 isolaten, met in totaal 93 mutaties, werd ten minste één fenotypische bepaling ingezet. Dit is weergegeven in *tabel 1* (pagina 72). Bij 60 procent van de geteste isolaten werd een verhoogde minimaal remmende concentratie (MIC, minimal inhibitory concentration) gevonden.

Van de 68 isolaten werden in totaal 188 fenotypische

gevoeligheidsbepalingen uitgevoerd (*tabel 2*); 93 bepalingen vanwege de aanwezigheid van een mogelijke resistentiemutatie en 95 vanwege de aanwezigheid van een mutatie met hoge voorspellende waarde voor resistentie tegen een ander antibioticum.

Van de 93 mogelijke resistentiemutaties waren er 43 met een hoge betrouwbaarheid, waarvan de fenotypische gevoeligheidsbepaling voor 100 procent overeenkwam met WGS. Bij de overige 50 fenotypische gevoeligheidsbepalingen met nieuwe mutaties en mutaties met een lagere betrouwbaarheid (laag en medium) werd bij nog eens 14 gevoeligheidsbepalingen een verhoogde MIC gevonden. In *tabel 2* wordt dit per antibioticum weergegeven.

De andere 95 fenotypische gevoeligheidsbepalingen werden extra uitgevoerd vanwege de aanwezigheid van een mutatie met hoge voorspellende waarde voor resistentie tegen een ander antibioticum. Hierbij werd één uitzondering gevonden bij een PZA-bepaling van een isolaat behorend tot lineage 1. Bij nader onderzoek naar de genen die mogelijk geassocieerd zijn met PZA resistentie werden er twee *clpC1*-mutaties (V63A en E810G) aangetoond; maar deze SNP's worden niet geassocieerd met resistentie volgens de ReSeqTB-databank en de WHO-catalogus [9-11].

Discussie

WGS is een zeer effectieve methode gebleken om te screenen op mogelijke resistentie tegen de vier eerstelijnsantibiotica die gebruikt worden in de behandeling van tuberculose. In deze studie konden 373 van de 441 (85 procent) *M. tuberculosis* complexisolaten direct als gevoelig gekenmerkt worden op basis van WGS. Bij de overige 68 isolaten werden mutaties gevonden die mogelijk geassocieerd waren met resistentie en werd aanvullend fenotypisch resistentieonderzoek uitgevoerd. Hierbij waren 43 mutaties met een hoge voorspellende waarde, waarbij geen twijfel over de uitslag was, hetgeen werd bevestigd.

De enige fenotypische gevoeligheidsbepaling die niet overeen kwam met de verwachte voorspelling op grond van WGS betrof een isolaat met PZA-resistentie, behorend tot lineage 1 (Indo_Oceanic). Deze vertoonde 'niet' met resistentie geassocieerde mutaties (*clpC1_V63A* & *E810G*), maar een mutatie die geassocieerd is met lineage 1 (*clpC1_V63A*). In een studie van Modlin *et al.* werd echter aangetoond dat Indo_Oceanic-isolaten oververtegenwoordigd zijn bij mono-PZA-resistente isolaten ten opzicht van de multidrugresistente isolaten met PZA-resistentie. Een voorgestelde theorie is dat er sprake is van een lineage-specifieke, lage fenotypische PZA-resistentie, vergelijkbaar met de PZA-resistentie van *M. bovis* en *M. canettii* [12]. In het nieuwe WHO-rapport (2021) over de voorspellende waarde van resistentiemutaties worden deze *clpC1*-mutaties echter niet geassocieerd met resistentie [11]. In de studie van Modlin *et al.* werden bij de herhaling van de 19 'mono-PZA-resistente' isolaten drie als PZA-gevoelig bepaald; dit waren lineage 1-isolaten met de *clpC1_V63A*-mutatie. De fenotypische bepaling lijkt bij deze stammen geen eenduidig beeld te geven en er is meer onderzoek nodig om dit op te helderen.

Door allerlei factoren kennen fenotypische bepalingen diverse beperkingen, waardoor kleine variaties kunnen ontstaan [13,14]. Deze variaties kunnen weer voor verschillende resultaten en interpretaties zorgen, zeker bij verminderde gevoeligheid of een laag resistentieniveau waarbij de MIC dicht bij het breekpunt ligt. Ook breekpunten die niet voldoende nauwkeurig zijn bepaald zorgen voor variaties in de interpretaties van fenotypische bepalingen en daardoor voor inconsequente resultaten [13,14]. Bij rifampicine is er

veel discussie over dit onderwerp, waarbij wisselende fenotypische resultaten werden vergeleken met de genotypische resultaten en behandeluitkomsten [15]. Uiteindelijk heeft de WHO het breekpunt voor rifampicine in 2021 verlaagd van 1 mg/l naar 0,5 mg/l [16]. Maar ook bij een hoog resistentieniveau tegen rifampicine, zoals bij de meest gangbare *rpoB_S450L*-mutatie, is in het WHO-rapport van 2021 maar 99 procent van de stammen fenotypisch als resistent bepaald [11]. Het gebrek aan reproduceerbaarheid van fenotypische bepalingen, maar ook de gekozen kritische breekpunten (CC) en de epidemiologische cut-offwaarden (ECOFF) blijven een punt van discussie en maken dat de fenotypische bepaling niet altijd de meest betrouwbare resultaten levert [13-15]. De fenotypische bepaling is echter wel van belang voor de uitbreiding van de kennis over de voorspellende waarde van (nieuwe) mutaties, zeker in het geval van tweedelijns- en nieuwe antibiotica. Wereldwijde internationale samenwerking, zoals voor de WHO-catalogus van 2021 met informatie over de voorspellende waarde van grote aantallen getoetste mutaties, is in dit opzicht zeer belangrijk. Dit soort catalogi dient echter dynamisch te zijn en regelmatig bijgewerkt te worden. Idealiter dienen deze databanken te worden uitgebreid met onder andere het niveau van gevoeligheid (MIC) en de behandeluitkomsten. Moleculaire bepalingen zullen in de toekomst steeds meer mogelijkheden bieden, zoals het direct sequencen van *M. tuberculosis* in klinisch materiaal, al is wat dit betreft nog een lange weg te gaan. Het blijft van belang om steeds de meest recente kennis van mogelijke resistentiemutaties toe te passen en hiervoor blijft een grote mate van specifieke deskundigheid vereist.

Referenties

1. Cole ST, Brosch R, Parkhill J, et al. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature*. 1998;393:537-44.
2. Consortium CR, Allix-Béguec C, Arandjelovic I, et al. Prediction of Susceptibility to First-Line Tuberculosis Drugs by DNA Sequencing. *N Engl J Med*. 2018;379:1403-15.
3. Schurch AC, Kremer K, Davieni O, et al. High-resolution typing by integration of genome sequencing data in a large tuberculosis cluster. *J Clin Microbiol*. 2010;48:3403-6.
4. Jajou R, de Neeling A, van Hunen R, et al. Epidemiological links between tuberculosis cases identified twice as efficiently by whole genome sequencing than conventional molecular typing: A population-based study. *PLoS One*, 2018;13:e0195413.
5. Lipworth S, Jajou R, de Neeling A, et al. SNP-IT Tool for Identifying

Subspecies and Associated Lineages of *Mycobacterium tuberculosis* Complex. *Emerg Infect Dis*. 2019;25:482-8.

6. Jajou R, van der Laan T, de Zwaan R, et al. WGS more accurately predicts susceptibility of *Mycobacterium tuberculosis* to first-line drugs than phenotypic testing. *J Antimicrob Chemother*. 2019;74:2605-16.

7. Jajou R, Kohl TA, Walker T, et al. Towards standardisation: comparison of five whole genome sequencing (WGS) analysis pipelines for detection of epidemiologically linked tuberculosis cases. *Euro Surveill*. 2019;24(50).

8. Coll F, McNeerney R, Guerra-Assunção JA, et al. A robust SNP barcode for typing *Mycobacterium tuberculosis* complex strains. *Nat Commun*. 2014;5:4812.

9. Starks AM, Avilés E, Cirillo DM, et al. Collaborative Effort for a Centralized Worldwide Tuberculosis Relational Sequencing Data Platform. *Clin Infect Dis*. 2015;61Suppl 3:S141-6.

10. WHO, The use of next-generation sequencing technologies for the detection of mutations associated with drug resistance in *Mycobacterium tuberculosis* complex: technical guide. , W.H. Organization, Editor. 2018: Geneva.

11. WHO, Catalogue of mutations in *Mycobacterium tuberculosis* complex and their association with drug resistance: supplementary document.

supplementary document. 2021, World Health Organization: Geneva.

12. Modlin SJ, Marbach T, Werngren J, Mansjö M, Hoffner SE, Valafar F. Atypical Genetic Basis of Pyrazinamide Resistance in Monoresistant *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother*. 2021;65(6).

13. Zignol M, Cabibbe AM, Dean AS, et al. Genetic sequencing for surveillance of drug resistance in tuberculosis in highly endemic countries: a multi-country population-based surveillance study. *Lancet Infect Dis*. 2018;18:675-83.

14. Heyckendorf J, Andres S, Köser CU, et al. What Is Resistance? Impact of Phenotypic versus Molecular Drug Resistance Testing on Therapy for Multi- and Extensively Drug-Resistant Tuberculosis. *Antimicrob Agents Chemother*. 2018;62(2).

15. Van Deun A, Aung KJM, Bola V, et al. Rifampin drug resistance tests for tuberculosis: challenging the gold standard. *J Clin Microbiol*. 2013;51:2633-40.

16. WHO, Technical report on critical concentrations for drug susceptibility testing of isoniazid and the rifamycins (rifampicin, rifabutin and rifapentine). 2021, World Health Organization: Geneva.