

The molecular-matryoshka phenomenon

Citation for published version (APA):

Jamin, C. (2023). *The molecular-matryoshka phenomenon: peeling the layers of spread of antimicrobial resistance genes*. [Doctoral Thesis, Maastricht University]. Maastricht University. <https://doi.org/10.26481/dis.20230329cj>

Document status and date:

Published: 01/01/2023

DOI:

[10.26481/dis.20230329cj](https://doi.org/10.26481/dis.20230329cj)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

10

NEDERLANDSE SAMENVATTING



Het topje van de ijsberg? Implementatie van WGS voor stamtypering en uitbraaksurveillance

Bacteriën worden steeds resistenter tegen antibiotica als gevolg van verkeerd gebruik en overmatig gebruik van deze middelen. Als gevolg daarvan kunnen gewone, momenteel behandelbare infecties onbehandelbaar worden, wat kan leiden tot ernstige complicaties en zelfs overlijden. Daarom is het van belang dat een verdere toename van antibiotica resistentie wordt voorkomen. Een mogelijkheid hierbij is het tegengaan van verspreiding van resistente bacteriën. Voor infectiebestrijding ten behoeve van de volksgezondheid en infectiepreventie in de ziekenhuisomgeving is tijdige detectie van de overdracht van resistente micro-organismen hierbij van vitaal belang. Bacteriële typering kan bepalen of isolaten al dan niet identiek zijn, wat het mogelijk maakt de verspreiding van deze specifieke isolaten tussen patiënten en mogelijk de bron van deze verspreiding te identificeren. De opkomst van "next generation sequencing" (NGS) technologieën, zoals "whole genome sequencing" (WGS), maakt een grondige analyse van de DNA-volgorde van de bacterie mogelijk. Hierdoor is het veld van de moleculaire epidemiologie verschoven naar het gebruik van WGS om de fylogenie of verwantschap tussen bacteriële isolaten op basis van hun genoom te bepalen en zo vast te stellen of er transmissie heeft plaatsgevonden. Met de introductie van deze geavanceerdere technologieën is validatie en evaluatie van de reproduceerbaarheid echter van cruciaal belang. Dit geldt ook voor WGS, waarbij naast het "wet-lab" (bv. kweken, DNA-isolatie en preparatie van DNA voor sequencering) ook het "dry-lab" (dataverwerking en -interpretatie) nauwkeurig moet worden geëvalueerd.

Een van de sterke punten van uitbraaksurveillance is de mogelijkheid om te weten welke isolaten in het ziekenhuis, de regio of zelfs wereldwijd circuleren door middel van snelle gegevensuitwisseling. Een uitstekend voorbeeld van het succesvol delen van data ten behoeve van besluitvorming (op het gebied van de volksgezondheid) is het GISAID-platform (Global Initiative on Sharing Avian Influenza Data, <https://www.gisaid.org/>). GISAID begon met het delen van data over de verspreiding van het influenzavirus wereldwijd. Toch kreeg het platform wereldwijde erkenning en lof voor zijn nut voor het verzamelen van SARS-CoV-2-genoomsequenties tijdens de SARS-CoV-2-pandemie. Hoewel er soortgelijke platformen voor de typering van bacteriële pathogenen bestaan, zoals PATRIC (<https://www.patricbrc.org>) en Pathogenwatch (<https://www.pathogen.watch>), hebben deze platformen nog niet hetzelfde niveau van implementatie bereikt als GISAID.

Voor een succesvolle uitwisseling van genetische informatie is het van belang te weten of de bioinformatische dataverwerking van ruwe sequentiedata van invloed kan zijn op de berekende fylogenie van de bacteriële isolaten. Aangezien laboratoria voor de analyse van WGS-gegevens ofwel eigen analyse pijplijnen gebruiken, dan wel commerciële software implementeren, zijn de verschillende werkwijzen misschien niet uitwisselbaar voor het doel van genotypering en uitbraak-surveillance. De reproduceerbaarheid en uitwisselbaarheid van de "dry-lab" kant van WGS voor uitbraaktypering is tot nu toe niet grondig geëvalueerd. In **hoofdstuk 2** hebben we onderzocht of gedecentraliseerde (lokaal verkregen) WGS data vergelijkbaar zijn tussen centra en welke parameters hierbij van belang zijn, wanneer een gestandaardiseerde procedure voor "wet-lab" en "dry-lab" werd gebruikt. De technische variatie in genetische afstand gemeten tussen deze replica's was klein voor cgMLST, wgMLST en SNP gebaseerde methoden en deze variatie in genetische afstand lag binnen marges die normaal gebruikt worden voor WGS gebaseerde genotypering. Deze geringe technische variatie in genetische afstand tussen de replica's is geruststellend, aangezien dit wijst op de haalbaarheid van gedecentraliseerde WGS-typering met geharmoniseerde protocollen. Hoewel alleen nosocomiale *Enterobacteriaceae* en *Enterococcus*-soorten werden getest, hebben anderen soortgelijke resultaten aangetoond voor *Staphylococcus aureus*, *Listeria monocytogenes*, *Campylobacter jejuni* en *Salmonella enterica*. Harmonisatie van wet-lab protocollen en bioinformatische werkwijzen voor de analyse van WGS-gegevens lijken hiermee toereikend genoeg te zijn voor succesvolle gedecentraliseerde stamgenotypering datavergelijking.

Door de overvloed aan verschillende genoom-assembler algoritmes is het onduidelijk of deze tools onderling uitwisselbaar zijn. In **hoofdstuk 2** hebben we het gebrek aan uitwisselbaarheid aangetoond van bioinformatische werkwijzen voor het genereren van genoom-assemblies voor surveillance van uitbraken. WGS-datasets werden onderworpen aan verschillende *de novo* assemblers om te beoordelen of er genetische dissimilariteit tussen deze replica's was geïntroduceerd. Het aantal SNP's tussen deze replica's overschreed de afkapwaarden die worden gebruikt voor uitbraaktypering, wat aantoont dat het ongeschikt is om verschillende bioinformatische werkwijzen door elkaar te gebruiken voor fylogenetische analyses. Als een platform als GISAID gebruikt zou worden voor het traceren van de verspreiding van klonen over het hele land of de hele wereld, is het essentieel dat WGS-gegevensverwerking, zoals *de novo* assembly, op een uniforme manier wordt uitgevoerd vanwege het gebrek aan uitwisselbaarheid van *de novo* assemblers. Daarnaast hebben bioinformatische werkwijzen ook invloed op het resultaat van stamgenotypering door WGS, zoals aangetoond in **hoofdstuk 3**. In dit hoofdstuk hebben we laten zien dat Nederlandse laboratoria bij analyse van exact dezelfde datasets verschillende genotypering resultaten kregen als gevolg van het gebruik van verschillende bioinformati-

sche werkwijzen. Ondanks verschillen in gebruikte genetische afkapwaarden tussen centra bleken onvoldoende om verschillen tussen typeringsresultaten te verklaren. Door gebruik te maken van één enkele analyse op de genoom-assemblies van de deelnemende centra, hebben we aangetoond dat het gebruiken van uniforme klonale afkapwaarden zinloos is, aangezien de typeringsresultaten tussen centra blijven verschillen, wat eens te meer wijst op de grote bias die genoom assemblers in WGS-gegevens introduceren. Deze resultaten tonen het risico aan van het gebruik van in de literatuur gerapporteerde afkapwaarden. Alleen als exact dezelfde werkwijze wordt gebruikt, kunnen eerder gerapporteerde afkapwaarden voor uitbraak-surveillance worden gebruikt. Als gedecentraliseerde surveillance van (resistente) bacteriestammen door WGS toegankelijk wordt voor een wijder publiek, is het van het grootste belang dat naast de gestandaardiseerde wet-lab procedures ook de bioinformatische werkwijzen worden geharmoniseerd en de meest nauwkeurige bioinformatische werkwijze wordt bepaald. Zonder een grondige harmonisatie zal het nooit mogelijk zijn een brede kennis op te bouwen over de verspreidingsroutes en -mechanismen van resistente micro-organismen en zal men helemaal niet in staat zijn uitbraken buiten de lokale omgeving te identificeren en te voorspellen.

Diep in de ijsberg: Identificatie van antibioticaresistentie (ABR) genen en inzicht in de rol van MGE's

WGS voor de genotypering van pathogenen is in razendsnel tempo dominant geworden in de microbiologie. Daarmee is er een vergelijkbare toename geweest van databases met genetische markers van belang, zoals resistentiegenen, virulentiegenen en mobiele genetische elementen zoals insertiesequenties, transposons en integreerbare & conjugeerbare elementen (ICE). Vaak zijn nieuwe ABR-genen, al over de hele wereld verspreid, voordat ze geïdentificeerd en in databases gecatalogiseerd zijn, zoals het geval was met *mcr-1*. Hoewel de typering van (resistente) bacteriestammen wereldwijd steeds vaker wordt uitgevoerd, wordt vaak over het hoofd gezien hoe deze (nieuwe) resistentiegenen zich tussen bacteriën verspreiden. Daarom is het van cruciaal belang dat i) resistentiegenen betrouwbaar kunnen worden geïdentificeerd, ii) het effect van nieuwe resistentiegenen grondig wordt geëvalueerd en iii) we de rol van mobiele genetische elementen bij de verspreiding van antibioticaresistentie begrijpen.

Identificatie van (nieuwe) ABR-genen

Het is niet eenvoudig om te bepalen welk bacterieel isolaat van de microflora een resistentiegen bevat. De commensale darmmicrobiota vormt een natuurlijk reservoir van resistentiegenen, resistentie genoemd, en kan resistentiegenen bevatten in commensale bacteriën die via horizontale genoverdracht op belangrijke pathogenen kunnen worden overgedragen. Het veelvoorkomende vancomycine resistentiegen

vanC en *vanD* in het humane darmmicrobioom maakt bijvoorbeeld de uitwisseling van deze genen naar *Enterococcus faecium* mogelijk. Daarom is screening op deze genen door PCR om *vanC* en *vanD* positieve *Enterococcus faecium* op te sporen zinloos, omdat deze genen meestal door de commensale bacteriën worden gedragen. Parallel aan dit voorbeeld onderzochten we in **hoofdstuk 5** het belang van de aanwezigheid van *mcr-9*, een nieuw mobiel colistine resistentiegen. Hoewel de prevalentie van deze *mcr-9* in *Enterobacter* soorten alarmerend hoog was, en dragerschap van *mcr-9* zeker niet verwaarloosbaar was, waren we niet in staat om dit gen te koppelen aan daadwerkelijke fenotypische colistine resistentie. Hoewel dit specifieke mechanisme van colistine resistentie niet van belang leek voor de patiëntenzorg, blijft het een belangrijke taak voor de medische microbiologie om de relevantie van nieuwe resistentiegenen en de mate waarin deze verspreid zijn grondig te evalueren. Doen wij dit niet, dan lopen we het risico dat we niet in staat zijn de verspreiding van nieuwe resistentiegenen te voorkomen wanneer ze voor het eerst opduiken, zoals we bijvoorbeeld hebben gezien bij de identificatie van het reeds wereldwijd verspreide *mcr-1*.

Betrouwbare identificatie van ABR-genen door WGS is eveneens belangrijk voor genotyperen van bacteriën. Hoewel momenteel nog weinig gebruik wordt gemaakt van vertaling van genotype naar fenotype voor klinische besluitvorming, is de verwachting dat dergelijke praktijken op grotere schaal zullen worden toegepast wanneer bacterietypering of klinische metagenomics (identificatie van infectieuze agentia rechtstreeks uit klinische monsters met behulp van sequencing technologieën) meer toegankelijk wordt. Een grondige evaluatie van de bioinformatische werkwijze voor de identificatie van ABR-genen is nodig en er zijn geschikte WGS-benchmark datasets samengesteld, zoals beschreven in **hoofdstuk 4**. In **hoofdstuk 3** hebben we problemen geïdentificeerd voor ABR-gen identificatie uit WGS-data als gevolg van bioinformatische werkwijzen. Sommige deelnemers aan een ringonderzoek met WGS-data slaagden er niet in ESBL-genen te identificeren in sommige bacteriële isolaten. Herbeoordeling van de genoom-assemblies van deze centra bevestigde de afwezigheid van deze ABR-genen. Al deze centra hadden één stap in de gegevensverwerking gemeen. Voor de genoom-assemblies werden alle contigs kleiner dan 1000bp uit de assemblies gelaten, wat een veel gebruikte stap is om kleine contaminaties in de dataset te verwijderen. Momenteel worden bij gebruik van WGS in een klinische omgeving de bacteriële isolaten gesequeneerd nadat zij in het lab zijn gekweekt en de antimicrobiële gevoeligheidstests reeds zijn uitgevoerd. Daarom weten de klinici al of bepaalde isolaten bijvoorbeeld een ESBL- of carbapenemase-producerend fenotype hebben. Wanneer NGS-technologieën worden gebruikt voor klinische metagenomics, zou het missen van deze belangrijke ABR-genen ernstige gevolgen kunnen hebben, en kan behandeling, met antibiotica waartegen de ziekteverwekker resistent is, falen.

De rol van MGE's

Ten slotte mag de rol van MGE's voor de verspreiding van ABR-genen niet worden onderschat. Zowel plasmiden als transposons zijn in verband gebracht met uitbraken van ABR-genen, maar dergelijke uitbraken zijn veel moeilijker te identificeren. In de praktijk zijn er geen typeringsmethoden, met uitzondering van "long-read" WGS, in staat om de verborgen aard van mobiele genetische elementen die ABR-genen dragen, goed te identificeren. Sommige ABR-plasmiden kunnen terechtkomen in verschillende bacteriële gastheren, verspreid over verschillende genera, afkomstig uit zowel mens, dier of de omgeving. In **hoofdstuk 6** werd dit gedrag van *mcr-1* plasmiden waargenomen. Hier vonden we een beperkt aantal verschillende plasmide "backbones" die codeerden voor *mcr-1*, maar de *E. coli* die deze plasmiden droegen waren zeer divers. Hoewel deze isolaten afkomstig waren van kippenvlees uit de supermarkt, werden sterk vergelijkbare plasmiden ook aangetroffen in klinische humane isolaten, wat wijst op de wijdverspreide aard van deze plasmiden. Plasmiden die coderen voor het carbapenemase *bla*_{OXA-48} zijn een ander voorbeeld van een zeer succesvol plasmide dat in een groot aantal verschillende bacteriële gastheren voorkomt. Van sommige bacteriestammen is bekend dat zij geassocieerd zijn met specifieke omgevingen; een bekend voorbeeld is *E. coli* sequentie type 131 die in de mens voorkomt. Deze bacteriestammen zijn goed aangepast door de verwerving van specifieke genen, waardoor zij geschikter zijn voor deze omgevingen. Plasmiden kunnen echter tussen verschillende bacteriële gastheren verspreiden middels conjugatie. Alleen de compatibiliteit met andere plasmiden of restrictie-/modificatiesystemen en CRISPR/Cas bepalen of een plasmide kan gedijen in een nieuwe bacteriële gastheer. In **hoofdstuk 7** is het *mcr-4* plasmide in dieren en mensen in Nederland uitgezocht. Hier toonde we aan dat plasmiden zich ook tussen domeinen waarschijnlijk kunnen verspreiden. Daarom is het aan te bevelen dat de surveillance van resistente bacteriële isolaten wordt uitgebreid met de resistentie plasmiden aangezien deze ook een potentieel belangrijke bijdrage leveren aan de wereldwijde verspreiding van antibioticaresistentie.

Ten slotte mag ook de rol van transposons voor de verspreiding van antibioticaresistentie niet worden gemist. De relatieve bijdrage van verspreiding van ABR van deze MGE's is nog weinig onderzocht. Het genestelde poppengedrag zoals bij Matroesjka poppen van deze elementen, waarbij ze van en naar verschillende plasmiden en naar het bacteriële chromosoom bewegen, maakt duidelijk hoe moeilijk het is om deze elementen naar behoren te identificeren en te traceren. Zoals beschreven in **hoofdstuk 8** werden niet-verwante microben met vergelijkbare fenotypische resistentieprofielen gedetecteerd. Vanwege het gelijke fenotype werd de mogelijkheid van eenzelfde ABR-plasmide overwogen, en onderzocht met "long read" WGS. Een groot (46 kilobase) genetisch-eiland, coderend voor 14 ABR-genen, werd geïdentificeerd in verschillende bacteriële soorten, gecodeerd op plasmiden

maar ook in het chromosoom. Veel van deze ABR-genen werden geflankeerd door IS26 insertiesequenties. Deze IS26 worden vaak geassocieerd met ABR-genen en zij voegen zich bij voorkeur in andere IS26-rijke regio's in het genoom, waardoor gen-cassettes worden gegenereerd die bestaan uit ABR-genen die door IS26 worden geflankeerd. De overdracht van dit gen-eiland kan gedeeltelijk worden verklaard door de uitwisseling van genetisch materiaal via conjugatieve plasmiden naar andere leden in deze taxonomische orde. Ten slotte werden 1806 WGS-datasets van ESBL-producerende *Enterobacterales* van gehospitaliseerde patiënten in Nederland 2011 tot 2014 gescreend op dit resistentie-gen-eiland, om na te gaan of dit gen-eiland al vóór 2019 in Nederland aanwezig was. Er werden 50 isolaten geïdentificeerd die coderen voor dit gen-eiland, waarvan er 37 afkomstig waren uit een van de ziekenhuizen uit de studie van **hoofdstuk 8**, wat impliceert dat dit resistentie-gen-eiland al minstens 7 jaar eerder aanwezig was in dit ziekenhuis. De verspreiding van transposons of deze gen-eilanden wordt vaak over het hoofd gezien omdat ze moeilijk te identificeren zijn zonder volledige genomen te analyseren. Alleen wanneer zeldzame en alarmerende resistentiemechanismen (bijvoorbeeld carbapenemases) in het spel zijn, zijn deze elementen geïdentificeerd als de veroorzakers van hun verspreiding. De langdurige aanwezigheid van dit resistentie-gen-eiland wijst op de noodzaak van toezicht op en meer begrip over deze mobiele genetische elementen om de verspreiding van ABR naar nieuwe bacteriestammen in een nosocomiale omgeving te beperken.

Ten slotte heeft WGS voor stamtypering zich bewezen als een betrouwbaar instrument voor infectiepreventie en surveillance van de verspreiding van nosocomiale bacteriën. Toch wordt WGS gebruikt tijdens of na uitbraken en worden mobiele genetische elementen bijna nooit in overweging genomen. WGS zal een belangrijke schakel worden en tot meer inzicht leiden in de rol die MGE's spelen bij verspreiding van antibiotica resistentie. Deze kennis zal vervolgens een bijdrage leveren aan de ontwikkeling van infectiepreventieve interventies die verdere verspreiding van resistente micro-organismen.