

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/138641> holds various files of this Leiden University dissertation.

**Author:** Du, C.

**Title:** Multi-omics studies of the control of growth and antibiotic production of streptomyces

**Issue Date:** 2020-12-09

## Nederlandse samenvatting

Actinobacteriën staan bekend om hun complexe levenscyclus en het vermogen om bioactieve natuurlijke producten (NP) te produceren (Hopwood, 2007, Fedorova *et al.*, 2012). Door de uitgebreide screening van deze bacteriën in de 20e eeuw is de frequentie van het vinden van echt nieuwe antibiotica dramatisch afgenomen. Nieuwe systematische benaderingen zouden daarom nieuwe geneesmiddelen moeten opleveren die nodig zijn om de problemen van de opkomende antibioticaresistentie tegen te gaan. Aangezien conventionele NP-screeningsmethoden voornamelijk die stoffen identificeren die onder laboratoriumomstandigheden in hogere concentraties worden geproduceerd, zijn nieuwe methoden nodig om de bioactieve stoffen te ontdekken die worden gespecificeerd door cryptische biosynthetische genclusters (BGCs) (Baltz, 2007). Vanuit metabolomics-perspectief betekent dit dat we de efficiëntie van dereplicatie moeten verbeteren (Wu *et al.*, 2015b, Krug & Muller, 2014). Om ons in staat te stellen de NPs te identificeren die zijn geproduceerd uit BGCs die onder laboratoriumomstandigheden cryptisch zijn, moeten we manieren vinden om ze te activeren. Mutagenese, veranderende groeiomstandigheden en het nabootsen van omgevingsfactoren zijn drie belangrijke methoden om algemene activering van BGCs te bereiken. De mutagenesemethode wordt geïllustreerd in **Hoofdstuk 3**, waar streptomycine-resistente mutanten werden gegenereerd en stammen met verhoogde of verminderde antibiotica-activiteit werden geselecteerd voor verdere analyse. Het verantwoordelijke BGC werd vervolgens geïdentificeerd door de productie van de bioactieve metaboliet met LC-MS te correleren aan de expressie van BGCs welke door proteoom-analyse is verkregen. Een alternatief om de antibioticaproductie te activeren is het nabootsen van omgevingsfactoren door meerdere kleine moleculen, zoals aangetoond voor de veelbelovende NP-producent *S. roseifaciens* in **Hoofdstuk 4**. Deze studie bracht strategieën aan het licht die op een systematische manier kunnen worden toegepast.

Systeembioologie is een essentieel onderdeel van NP-onderzoek, omdat het ons in staat stelt het biologische systeem als geheel te begrijpen. Dit onderzoeksveld profiteert nu van de enorme technologische vooruitgang die in het verleden ontbrak (Butcher *et al.*, 2004). De recente vooruitgang in omics-technologieën, met name transcriptomics, metabolomics en proteomics, heeft dit type onderzoek beter haalbaar gemaakt. Onder deze technologieën is proteomics belangrijk voor de studie van NP-biosynthese, omdat het wetenschappers in staat stelt de expressieniveaus van de metabole enzymen te bepalen (Du & van Wezel, 2018). Zoals aangetoond in **Hoofdstuk 4** bleek uit statistische analyse dat jasmonzuur, N-acetylglucosamine, chitosan en benzeencarbonzuur een verrassend vergelijkbaar effect op proteoom-niveau in *S. roseifaciens* veroorzaakten. Dit leidde echter niet tot een vergelijkbaar effect in de antimicrobiële activiteit, wat kan betekenen dat dit mogelijk verband houdt met de veranderingen in enkele belangrijke enzymen,

kleine verschillen in de groeiomstandigheden tussen experimenten of de effecten van elicitors op groeisnelheid of morfologie. Ten tweede brengt verrijkingsanalyse met genontologie de resolutie terug naar het tussenliggende niveau. De resultaten toonden aan dat hydroxycoumarine, dat een goede antimicrobiële activiteit opwekt, waarschijnlijk het energiemetabolisme van *S. roseifaciens* beïnvloedt. Door het gebruik van hoge-resolutie proteomics voor het onderzoeken van de kortetermijnreactie op jasmonzuur werd ten slotte een gencluster (*jar*) geïdentificeerd. Dit gencluster is gerelateerd aan de antimicrobiële activiteit van *S. roseifaciens*, maar correleert niet direct aan de biosynthese van de biologische activiteit.

Het tot uiting brengen van cryptische BGCs om meer inzicht te krijgen in de metabolieten die ze specificeren, is niet altijd ongecompliceerd, vooral niet voor die van niet-kweekbare micro-organismen of voor die geïdentificeerd zijn in metagenoom studies. Om toegang te krijgen tot dit deel van de NP-repository is heterologe expressie van BGCs in een geoptimaliseerde host een belangrijke methode (Nepal & Wang, 2019, Gomez-Escribano & Bibb, 2011). Deze methode is effectief gebleken bij het vinden van nieuwe antibiotica (Du *et al.*, 2013), het bepalen van de grenzen van BGCs (Komatsu *et al.*, 2013, Liu *et al.*, 2018), het creëren van kunstmatige natuurlijke producten door verschillende BGCs te combineren (Park *et al.*, 2011) en voor het karakteriseren van functies van individuele genen (Waldman *et al.*, 2015). Het blijft echter moeilijk om een geschikte gastheer te kiezen voor specifieke BGCs. Dit feit, samen met andere technische problemen, zoals het overdragen van grote genfragmenten en de vereiste van biosynthetische repertoire van de natuurlijke producent dat mogelijk ontbreekt in de nieuwe gastheer, vertraagt de toepassing van heterologe productieplatforms. Wat is bovendien de basis om te beslissen welke BGC prioriteit te geven, als we niet weten wat de verwante NP is? Gensynthese is nog steeds relatief duur, en daarom is grootschalige synthese van genen niet gebruikelijk. Om deze problemen op te lossen werd een geoptimaliseerde stam, *S. coelicolor* M1152, gemaakt om de productie te maximaliseren en tegelijkertijd de complexiteit van het maken van extracten te verminderen (Gomez-Escribano & Bibb, 2011). Het optimaliseren van deze stam is echter voornamelijk gebaseerd op ervaring, en het was onbekend wat het effect van het optimalisatieproces op het systeemniveau is. Daarom hebben we een multi-omics studie uitgevoerd over een tijdreeks in een volledig gedefinieerd fermentatiesysteem en vervolgens de monsters geanalyseerd via proteomics met een hoge resolutie (**Hoofdstuk 5**).

Significante verschillen werden waargenomen in de globale proteïenprofielen tussen de stam met gereduceerde achtergrond en zijn wildtype ouder, waaronder veel leden van de regulon van de globale transcriptiefactoren PhoP, GlnR en ScbR. We vonden ook positieve regulatie van biosynthetische enzymen van ectoïne, welke een verhoogde osmotische stress in *S. coelicolor* M1152 kunnen aangeven.

Deze bevindingen bevorderen een verder rationeel ontwerp van deze stam of een vormen een leidraad voor de optimalisatie van andere veelbelovende gastheerstammen.

Het eerste volledig gesequentieerde modelorganisme in het phylum van meest belangrijke antibioticaproductanten, Actinobacteriën, is *S. coelicolor* (Bentley *et al.*, 2002). Het vermogen om veel verschillende antibiotica te produceren, in het bijzonder PKS en NRPS, suggereert dat het een logische keuze is als gastheer voor de heterologe expressie van BGCs. *S. coelicolor* is sinds 1960s het modelorganisme voor onderzoek naar streptomyceten (Rigali *et al.*, 2018, Urem *et al.*, 2016). Dit regelgevende netwerk hangt ook nauw samen met de ontwikkeling van *Streptomyces* (van Wezel & McDowall, 2011).

Een interessant proteïne is SCO1839, dat lid is van een nog niet bestudeerde proteïnefamilie die uniek is voor Actinobacteriën. Eerdere studies naar genen betrokken bij de ontwikkeling van *S. coelicolor* hebben voorgesteld dat SCO1839 mogelijk deelneemt aan het reguleren van sporulatie (Kim *et al.*, 2015). **Hoofdstuk 6** geeft een gedetailleerde analyse van de DNA-bindingseigenschappen van SCO1839. We ontdekten dat SCO1839 een waarschijnlijk nucleoïde-geassocieerd proteïne (NAP) is dat zich bindt aan een zeer kleine DNA-sequentie rond het palindroom GATC. Zoals verwacht van het korte motief, is de bindingslocatie van SCO1839 wijdverspreid op het chromosoom. Naast SCO1839 worden veel andere kleine DNA-bindende proteïnen gecodeerd door het *S. coelicolor*-genoom dat de ontwikkeling controleert, zoals geïllustreerd door BldC (Hunt *et al.*, 2005). De wijze van DNA-binding lijkt te verschillen van die van andere regulatoren, welke functioneren via asymmetrische kop-staart-oligomerisatie op directe DNA-herhalingssequenties, wat resulteert in dramatische DNA-vertanding (Schumacher *et al.*, 2018, Bush *et al.*, 2019). In tegenstelling tot SCO1839 is er tot nu toe geen significant bindingsmotief voorgesteld voor BldC. Aangezien de bindingsplaatsen op het DNA van BldC echter aanzienlijk in lengte variëren, is er ook variatie in de lengtes van de SCO1839-bindingsequenties, zoals blijkt uit de ChIP-Seq-gegevens. Het is dus nodig om het DNA-binding model voor SCO1839 te ophelderen en te vergelijken met dat van BldC en andere actinobacterie regulatoren.

*Streptomyces* hebben veel verschillende NAPs waarvan de functies nog niet volledig zijn ontcijferd. Deze proteïnen geven niet alleen structuur aan de nucleoïde, maar spelen ook een belangrijke rol bij de regulering van verschillende cellulaire processen. De studie van SCO1839 en zijn specifieke bindingsmotief heeft een nieuwe route opgeleverd om de functie van NAPs in *Streptomyces* te decoderen. Het benadrukte ook het belang van NAPs in deelnemende ontwikkelingsprocessen van *Streptomyces*, die belangrijke regulerende factoren zijn in het secundaire metabolisme.

Concluderend, dit proefschrift beschrijft de toepassing van geavanceerde proteomics en andere 'omics-technologieën in NP-onderzoek in *Streptomyces*, in het bijzonder als onderdeel van systeembrede studies. Dit omvat de proteomining-technologie om BGCs te verbinden met een bioactiviteit die van belang is, studies van de gevolgen op proteoom-niveau van *Streptomyces* op kleine moleculen die omgevingsignalen nabootsen, en analyse van de gevolgen van stam ontwerpbenaderingen voor de optimaliseren van *S. coelicolor* als heterologe gastheer voor productie van de NP. Bovendien werd een nieuwe familie van kleine DNA-bindende proteïnen beschreven die zouden kunnen functioneren in de epigenetische regulatie van de antibioticaproductie, waarbij de analyse van moleculaire biologie en bio-informatica worden gecombineerd. Tegenwoordig genereren nieuwe high-throughput-methoden, en in het bijzonder 'omics-technieken, veel meer gegevens dan ooit tevoren. De uitdaging ligt daarbij in het zo combineren van multidimensionale datasets dat we grote biologische vragen kunnen beantwoorden. Ik hoop dat dit proefschrift een bijdrage heeft geleverd aan het vervullen van deze lastige taak. Systeembrede benaderingen zijn essentieel voor het verkrijgen van een meer uitgebreid begrip van de regulerende netwerken die NP-biosynthese en de complexe biologie van *Streptomyces* controleren.