

University of Groningen

Genetical genomics with Affymetrix gene expression arrays

Alberts, Rudi

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version

Publisher's PDF, also known as Version of record

Publication date:

2007

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Alberts, R. (2007). *Genetical genomics with Affymetrix gene expression arrays*. s.n.

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

7.2 Dutch summary

Complexe eigenschappen zijn eigenschappen die onder invloed van meerdere factoren tot stand komen. Dit kunnen zowel genetische als omgevingsfactoren zijn. Deze factoren werken vaak samen en/of hebben onderlinge interacties. Dit maakt het moeilijk om deze verschillende factoren te ontrafelen. Deze 'ontrafeling' is het werkgebied van de 'complex trait analysis'. Veel ziekten zijn voorbeelden van complexe eigenschappen, bv. kanker, obesitas en hypertensie.

Twee methodes die vaak worden gebruikt bij het ontrafelen van complexe eigenschappen zijn 'klassieke' QTL mapping en genetical genomics (Jansen and Nap 2001). Een QTL (quantitative trait locus) is een locatie op een chromosoom waar met grote waarschijnlijkheid een of meerdere genen liggen die een bepaalde kwantitatieve eigenschap (bv. lengte) reguleren. Door metingen van de kwantitatieve eigenschap voor meerdere individuen te combineren met zgn. moleculaire merker scores op het genoom (alle chromosomen samen), worden QTL's bepaald. Deze moleculaire merkers zijn plaatsen op het genoom waarin het DNA van de individuen verschilt en voor elk van de individuen de vorm van het DNA kan worden bepaald. Een veel voorkomende moleculaire merker is de SNP (single nucleotide polymorphism). Dit is een plaats op het DNA waar individuen slechts in 1 'letter' verschillen, dus waar bv. sommige individuen een 'A' hebben en andere een 'T'.

Met behulp van genexpressie microarrays kan de activiteit van genen op grote schaal worden gemeten. Het idee achter genetical genomics is dat genactiviteit gebruikt kan worden als kwantitatieve eigenschap in een QTL analyse. In dit geval kunnen per gen waarvan de activiteit is gemeten één of meerdere plaatsen (QTL's) op het genoom worden geïdentificeerd. Deze plaatsen bevatten waarschijnlijk genen die de expressie reguleren van het gen waarvan de activiteit was gemeten. Met behulp van genetical genomics kunnen zodoende regulerende links tussen genen worden gevonden.

'Klassieke' QTL mapping en genetical genomics kunnen op de volgende manier worden gecombineerd: i) bepaal met QTL mapping de QTL's voor een 'klassieke' kwantitatieve eigenschap (elke kwantitatieve eigenschap behalve genexpressie, bv. lengte, kanker vatbaarheid); ii) bepaal met genetical genomics of er genen zijn die zelf in de QTL regio's liggen en ook naar deze regio's mappen (hier QTL's hebben). Het feit dat ze in dit gebied mappen betekent dat ze verschil in expressie hebben. Dit verschil in expressie zou goed de 'klassieke' kwantitatieve eigenschap kunnen reguleren. iii) bekijk met genetical genomics of er genen zijn die elders liggen en naar deze regio's mappen. Co-localisatie van de QTL's is evidentie dat die genen en de 'klassieke' eigenschap iets met elkaar te maken hebben.

Affymetrix is een microarray technologie waarmee genexpressie kan worden geme-

ten. Dit proefschrift richt zich met name op deze technologie. De werking ervan wordt beschreven in Hoofdstuk 1. Hoofdstuk 2 introduceert het concept genetical genomics en laat zien dat het uitvoeren van genetical genomics experimenten op verschillende typen microarrays niet slechts neerkomt op het toepassen van standaard procedures voor microarrays, maar dat betere resultaten worden behaald met aangepaste experimentele ontwerpen en analyses.

De Affymetrix technologie gebruikt 11 tot 20 zgn. probes per gen om de genexpressie te meten. Elk van de probes levert een signaal. Om één expressie waarde per gen te krijgen wordt doorgaans een gemiddelde over de probe signalen genomen. Hoofdstuk 3 laat zien dat dit niet altijd gewenst is en kan leiden tot het verlies van biologisch relevante informatie. Verder wordt een oplossing geboden voor het 'batch' probleem. Vaak wordt de uitvoering van microarray experimenten opgesplitst in verschillende delen ('batches'). Hierdoor kunnen batch effecten in de uiteindelijke data worden geïntroduceerd. Deze kunnen leiden tot ongewenste resultaten. Hoofdstuk 3 geeft hiervoor een oplossing

Affymetrix ontwerpt haar microarrays op grond van gen sequentie informatie die op een bepaald moment beschikbaar is. Echter, doordat deze sequentie informatie in de loop van de tijd evolueert, kan achteraf blijken dat een deel van de probes op een microarray niet meer tot het gewenste gen behoort. Hierom hebben wij in Hoofdstuk 4 een protocol ontwikkeld waarin alle probe sequenties van Affymetrix arrays worden gecontroleerd tegen verschillende databases met gen sequenties en dat de definities van de microarrays bijwerkt. We hebben laten zien dat de detectie van genen die verschil in expressie vertonen, wat vaak met microarrays wordt gedaan, significant verbetert met onze nieuwe definities.

Hoofdstuk 5 behandelt een artefact bij het gebruik van Affymetrix arrays in genetical genomics experimenten. In deze experimenten willen we verschillen in genexpressie detecteren voor verschillende individuen. Het komt echter voor dat deze individuen verschillen in het DNA in de probe regio's hebben, waardoor de samples van de individuen waarvan het DNA gelijk is aan de probe goed hybridizeren en samples van de individuen waarvan het DNA verschilt van de probe minder goed. In dit geval is er (mogelijk) geen verschil in genexpressie maar slechts een verschil in hybridizatie. In Hoofdstuk 5 laten we zien dat dit op grote schaal voorkomt in genetical genomics experimenten en bieden wij een statistische procedure waarmee de probes waarin dit optreedt worden geëlimineerd, zodat QTL analyse kan volgen op de resterende probes.