

10 tips voor microarrays zonder ruis

Microarrays leveren genomics-onderzoekers gigantisch veel data op, maar die zijn niet allemaal even biologisch relevant. Drie ervaringsdeskundigen geven samen tien tips om de signaal-ruisverhouding van microarrays te optimaliseren.

Door Els van den Brink, gepubliceerd in oktober 2008 in het Laboratorium Magazine

Hoe onderscheid je biologisch relevante signalen van kunstmatige ruis? Dat is een vraag die in principe bij elke labtechniek gesteld kan worden. Bij de microarray-techniek (zie kader voor korte uitleg) speelt deze vraag nog nadrukkelijker. Dat komt vooral door de grote hoeveelheden data die de techniek genereert, de extreem hoge gevoeligheid en de relatief hoge kosten, waardoor onderzoekers experimenten maar in beperkte mate herhalen. De uitslag moet het liefst in één keer goed zijn. Wat vinden en wensen de experts die met deze techniek werken en wat zijn hun tips voor optimalisatie?

Doelstelling

“Wij willen dat de microarray-techniek een meettechniek is, waar we op kunnen vertrouwen”, zegt Frank Holstege, hoogleraar Genomics binnen het UMC Utrecht en hoofd van de Microarray Faciliteit aldaar. Hij legt uit: “Twee aspecten zijn van belang: de betrouwbaarheid en de reproduceerbaarheid. Meestal kijken onderzoekers vooral naar dat laatste. Maar als een meting goed reproduceert, wil dat nog niet zeggen dat het resultaat biologisch relevant is. De belangrijkste vraag is of het resultaat klopt met wat er in de onderzochte cellen aanwezig was.” De doelstelling van een microarray experiment kan verschillende eisen stellen aan de betrouwbaarheid van de data. Holstege legt uit: “Veel microarrays worden gebruikt voor screening. Daarbij is minder belangrijk dat de data betrouwbaar zijn, omdat de geselecteerde genen toch nog geverifieerd worden met een andere techniek.” Zelf gebruikt Holstege microarrays op een heel andere manier. Holstege: “Wij kijken naar genexpressie-patronen, bijvoorbeeld na uitschakeling van een van de genen. Dat patroon kun je beschouwen als een fenotype waaraan je deze cellen kunt herkennen. Hiermee kunnen we uitspraken doen over de samenstelling en interacties van eiwitcomplexen. Dat is een manier van werken die veel verder gaat, waarbij het er meer op aankomt dat de data betrouwbaar zijn.”

Uitdaging

In de praktijk is een analyse pas echt betrouwbaar als een onderzoeker zonder twijfel biologische relevante signalen kan onderscheiden van de ruis die altijd in metingen aanwezig is. Ron Kerkhoven, hoofd van de Centrale Microarray Faciliteit van het NKI in Amsterdam, zegt hierover: “Signalen van de meest interessante genen komen meestal maar net boven het achtergrondniveau uit. We staan voor de uitdaging om signalen op te krikken zonder dat de achtergrondsignalen mee omhoog gaan. We hebben al veel vooruitgang geboekt, want het achtergrondniveau is nu nog maar vijf procent van het oorspronkelijke niveau.” Deze vooruitgang was alleen mogelijk dankzij een zorgvuldige optimalisatie van de microarray experimenten. Holstege, Kerkhoven en Fiona Nielsen, AIO bij het Nijmegen Centre for Molecular Life Science, geven een tiental tips om biologisch relevante signalen van ruis te onderscheiden.

Tip 1: Begin met een goed ontwerp.

“Wat vaak mis gaat, is dat onderzoekers hun vraagstelling niet goed gedefinieerd hebben”, geeft Holstege aan. “Je moet goed weten wat je er precies uit wilt halen, en welke analyses je precies wilt gaan doen. Anders blijkt achteraf vaak dat de juiste controles niet zijn meegenomen of dat er onvoldoende ruimte is voor statistiek.”

Tip 2: Amplificeer vooraf al het RNA.

“Wij amplificeren (vermenigvuldigen, red.) altijd het RNA, zelfs als de hoeveelheid in principe voldoende is”, vertelt Kerkhoven. “De reden is dat het RNA er zuiverder door wordt, waardoor we minder last hebben van achtergrondsignalen. Het is daarbij wel van belang om geamplificeerde monsters niet te vergelijken met niet-geamplificeerde monsters.”

Tip 3: Automatiseer zo veel mogelijk.

“In ons automatische hybridisatiestation worden de microarrays heel gelijkmatig behandeld. Eerder ontstond er variatie in data doordat de hybridisatie aan de randen van microarrays minder goed was verlopen. Daar hebben we nu geen last meer van”, vertelt Kerkhoven. Holstege heeft niet alleen de hybridisatie geautomatiseerd, maar ook de

daaraan voorafgaande voorbehandeling. Een pipetteerrobot is sinds kort in staat om alle honderdtwintig voorbehandelingsstappen achter elkaar uit te voeren. “Groot voordeel is dat de kwaliteit constant is en dat de apparatuur een stuk sneller is. Alleen de isolatie van RNA uit weefsels doen we nog handmatig, omdat we daarbij voor elk weefsel een ander protocol nodig hebben”, vertelt Holstege.

Tip 4: Gebruik kalibratiestandaards.

“Wij gebruiken kalibratiestandaards om de betrouwbaarheid van onze data vast te stellen. Daarin is ons lab echt uniek”, vertelt Holstege trots. In de praktijk komt het erop neer dat hij een bekende hoeveelheid RNA toevoegt aan het RNA dat hij wil analyseren. Op zijn arrays heeft Holstege een aantal oligonucleotides toegevoegd waaraan dit externe controle-RNA kan binden. Zo kan Holstege precies vaststellen hoe de verhouding is van het uiteindelijke signaal ten opzichte van de hoeveelheid RNA die hij heeft toegevoegd. “Dankzij deze standaards hebben wij al onze procedures, vanaf de amplificatie tot aan de data analyse, kunnen optimaliseren voor weergave van het juiste signaal”, vertelt Holstege. Consequentie is wel dat hij meestal gebruik maakt van zelfgeprinte microarrays. Holstege legt uit: “Sommige commerciële arrays hebben wel standaards, maar die werken niet altijd ideaal en er zijn vaak niet genoeg. Er is wel een *External RNA Control Consortium* opgericht waarin alle microarray-leveranciers zitting hebben. Het is hun bedoeling om te komen tot gezamenlijke standaards, maar dat is tot nu toe nog niet gebeurd.”

Tip 5: Stap over op bead-arrays.

“Wij zijn voor genexpressie studies vrijwel helemaal overgestapt op de bead-arrays van Illumina”, vertelt Kerkhoven. De relatief lage prijs (een factor drie lager dan de microarrays van andere leveranciers) maakte het financieel interessant om over te stappen op commerciële arrays. Maar volgens Kerkhoven zijn er ook veel kwaliteitsvoordelen. Op de microarrays van Illumina zitten de oligonucleotides niet vast aan het glazen plaatje, maar aan microscopische bolletjes van 2-3 micrometer. Door de kleine afmetingen van deze bolletjes heeft Illumina de ruimte om voor elk gen 40 bolletjes op te nemen, die verspreid liggen over het hele plaatje. Lokale effecten (bijvoorbeeld door een kras) worden automatisch uitgemiddeld. Bovendien worden de analyses niet verstoord door binding van fluorescente labels (zie tip 6). Volgens Kerkhoven is er sprake van een brede trend richting deze illumina-arrays.

Tip 6: Pas op voor storende fluorescente labels.

Een gebruikelijke microarray-aanpak is om twee monsters te labelen met twee verschillende kleuren (rood en groen), en die te vergelijken op een en dezelfde microarray. “Het probleem is dat deze fluorescente labels zelf ook een zekere affiniteit hebben voor bepaalde oligonucleotides. Die affiniteit is verschillend is voor de rode en groene labels. Vandaar dat we een meting altijd twee keer uitvoerden, waarbij we de tweede keer de labels omwisselen”, vertelt Kerkhoven. Zelf heeft hij het probleem opgelost door zijn overstap naar de Illumina arrays. Nu labelt hij het RNA of DNA alleen nog met een biotine-molecuul en voegt hij pas later, nadat het RNA of DNA gebonden is aan de microarray, een fluorescent label toe, dat via het eiwit streptavidine aan deze biotine kan binden. Holstege heeft nog wel met het label-probleem te maken, en heeft daar onlangs speciale software voor ontwikkeld. Holstege legt uit: “We kunnen precies uitrekenen wat het effect is van de binding van de fluorescente labels en daarmee de data corrigeren. Zo kunnen we veel subtielere effecten meten. Als er eerder bij vijf à tien genen een relevant effect was, sneeuwde dat onder tussen de driehonderd genen waar een effect was door het fluorescente label. Nu kunnen we zulke kleine effecten wel zien.”

Tip 7: Controleer de kwaliteit van de oligonucleotides.

In een ideaal microarray-experiment bindt elk stukje DNA of RNA keurig aan de oligonucleotide die exact overeenkomt met zijn eigen sequentie. De praktijk werkt vaak anders. Sommige oligonucleotides vormen alleen interne verbindingen, anderen blijken niet specifiek genoeg, waardoor ze verschillende stukjes DNA of RNA kunnen binden. Fiona Nielsen onderzocht hoe je de sequenties van de oligonucleotides op de microarrays zou kunnen optimaliseren. Nielsen ontdekte dat vooral de smelttemperatuur, de temperatuur waarbij gebonden DNA weer loslaat, een kritisch punt is. Daarnaast is het van belang dat een sequentie uniek is. Niensens optimalisatie verbeterde de signaal-ruis verhouding met twintig procent, ten opzichte van het standaard systeem van microarray-leverancier Nimblegen. De rekenmethode is in principe ook bruikbaar voor andere systemen. Nimblegen, die zelf ook partner was in het onderzoeksproject, gaat de uitkomsten waarschijnlijk toepassen.

Kerkhoven en Holstege herkennen het probleem, maar hebben weinig ruimte om er wat aan te doen, zelfs niet bij hun zelfgeprinte microarrays. Kerkhoven legt uit: “Het probleem is dat het veel geld kost om een bibliotheek van oligonucleotides aan te schaffen. Dat doe je niet zomaar opnieuw als een paar niet meer goed zijn. Feit is wel dat bij onze huidige bibliotheek een op de vijf oligonucleotides niet meer goed is. Dat komt vooral doordat de kennis over het genoom nog sterk in ontwikkeling is. Elke drie maanden zijn er nieuwe inzichten over de locaties waar genen

beginnen en eindigen. Een oligonucleotide blijkt dan opeens niet meer complementair te zijn aan een stukje van een gen, maar aan een stukje DNA dat daar net buiten ligt.”

Tip 8: Pas op voor storende primersequenties.

Microarray-gebruikers moeten alert zijn op de kleinste details. Zo ontdekte Kerkhoven onlangs dat een klein stukje extra RNA, dat tijdens het amplificatieproces wordt toegevoegd aan alle RNA fragmenten, een belangrijke bron van ruis kan zijn. Het gaat om een zogenaamde primersequentie, die bestaat uit een serie thymine-basen, gevolgd door zes tot negen andere basen. De storende invloed van de serie thymines was al wel bekend. Vandaar dat onderzoekers meestal polyA (een serie Adenine-basen) toevoegen om deze thymines af te dekken en zo te voorkomen dat alle RNA-fragmenten binden aan een oligonucleotide waar toevallig een aantal thymines in voorkomen. Kerkhoven: “De laatste zes tot negen basen kunnen we niet blokkeren. Het is te duur om daar iets voor te kopen en bovendien is het stukje niet altijd hetzelfde. Als je het amplificatie-enzym bij een andere leverancier koopt, is dit stukje RNA ook weer anders. Daardoor kunnen we deze ruis niet voorkomen. We kunnen er bij de analyse wel rekening mee houden.”

Tip 9: Zet microarray-apparatuur in een ozonvrije klimaatkamer.

“Ik wordt in de zomer regelmatig gebeld door mensen bij wie de metingen niet meer goed lopen”, vertelt Holstege. “Meestal ligt de oorzaak bij een afwijkend ozongehalte en luchtvochtigheid. Zelf hebben we daar geen last meer van, dankzij onze speciale klimaatkamers.” In deze ruimtes wordt al het ozon weggezogen, blijft de temperatuur en luchtvochtigheid constant en is alles helemaal stofvrij. Dat is van belang voor het printen van de microarrayplaten, maar ook voor de kwaliteit van de fluorescente labels. Ook Kerkhoven heeft zijn printers in een klimaatkamer staan, en alle scanners in speciale ozonvrije kasten.

Tip 10: Wees voorbereid op de next generation sequencing-technologie.

De enige vraag die overblijft, is hoe zinvol het nog is om grote investeringen te doen op het gebied van microarrays. Kerkhoven, Holstege en Nielsen verwachten namelijk alle drie dat microarray experimenten binnen een paar jaar gedeeltelijk vervangen zullen worden door *next generation sequencing*: een nieuwe technologie waarmee in korte tijd een compleet genoom geanalyseerd kan worden, en die ook mogelijkheden biedt voor de analyse van bijvoorbeeld genexpressie en eiwitbinding.

“De techniek is nu nog erg duur, maar over een jaar of twee is het waarschijnlijk wel betaalbaar en kun je er alles mee vaststellen”, denkt Kerkhoven. Nielsen heeft de focus van haar onderzoek al in deze richting aangepast. Holstege is voorzigtiger: “Meestal duurt het tien tot vijftien jaar voor een techniek echt betrouwbaar is. Het sequencen zelf gaat wel goed, maar de stappen die eraan voorafgaan staan nog in de kinderschoenen.”

Microarrays in het kort

Microarrays zijn te vergelijken met een borstel, waarbij korte stukjes DNA (oligonucleotides genoemd) als borstelharen gebonden zitten aan een plaatje van glas, kunststof of silica. Het te onderzoeken DNA of RNA wordt gefragmenteerd, gelabeld met een fluorescent label om het vervolgens te laten binden (hybridiseren) met de oligonucleotides op de microarray. Een scanner meet vervolgens de fluorescentie en levert zo informatie over de verschillende genen.

De toepassingen liggen vooral in het genomics onderzoek, bijvoorbeeld om te bepalen welke genen tot expressie komen of welke eiwitten binden aan het DNA.

Illumina-microarrays (zie afbeelding) zijn gebaseerd op kleine bolletjes van silica, een soort kralen. Elk bolletje heeft een diameter van drie micrometer en draagt meer dan 100.000 korte stukjes DNA.



©2006, Illumina Inc. All rights reserved.