

Samenvatting

In het wetenschappelijk onderzoek naar de werking van alle genen (genomica) is de laatste jaren een enorme vooruitgang geboekt door nieuwe technologische ontwikkelingen waarmee de expressie van genen (het transcriptoom) op grote schaal in kaart gebracht kan worden. De toepassing van deze nieuwe technologieën heeft duidelijk gemaakt dat de transcriptomen van zowel prokaryote als eukaryote organismen veel complexer zijn dan voorheen werd gedacht. Bovendien heeft het gebruik van deze technologieën voor transcriptoomanalyse nieuwe inzichten verschaft in de moleculaire basis van ontwikkelings- en ziekteprocessen. De technologieën voor grootschalige transcriptoomanalyse berusten op twee benaderingen, microarrayanalyse en sequentieanalyse. Microarrayanalyse is een veelgebruikte techniek waarmee de expressieniveaus van duizenden transcripten gelijktijdig bepaald kunnen worden. Deze techniek heeft bijgedragen aan de oplossing van een groot aantal biologische vraagstellingen. Echter, microarrayanalyse heeft verschillende beperkingen die inherent zijn aan alle analysemethoden die gebaseerd zijn op hybridisatie van DNA of RNA. Bovendien is onderzoek waarbij microarrays gebruikt worden bevooroordeeld in die zin dat de samenstelling van de genen op de array voorafgaand aan het experiment is bepaald. Met de opkomst van de nieuwe generatie van technologieën voor sequentieanalyse is het mogelijk geworden om miljoenen transcripten tegelijk te detecteren en zijn de beperkingen van microarrayanalyse overwonnen. Transcriptoomanalyse is onmisbaar om de regulering van genexpressie aan het licht te brengen. Genexpressie wordt gereguleerd op verschillende niveaus, van transcriptie tot post-translationele modificatie van het gecodeerde eiwit. Recentelijk is het duidelijk geworden dat kleine, niet-coderende RNA-moleculen, zoals microRNA's (miRNA's) belangrijke componenten van het reguleringsmechanisme van genexpressie zijn. MiRNA's werken als negatieve regulatoren van mRNA-translatie of mRNA-stabiliteit en spelen een rol bij een groot aantal biologische processen, bijvoorbeeld de ontwikkeling en de functie van het immuunsysteem. Afwijkende expressie van miRNA's is waargenomen bij verschillende ziekten, waaronder infectieziekten en kanker. Het in kaart brengen van de verschillen in miRNA-expressie tijdens ziekteprocessen en het ontrafelen van de biologische functies van miRNA's is daarom van groot belang voor het begrip van de moleculaire mechanismen van ziekten. Om te onderzoeken hoe ziekteprocessen de expressiepatronen van mRNA's en miRNA's beïnvloeden, worden zowel *in vitro*-systemen (celculturen) als *in vivo*-systemen (proefdiermodellen) toegepast. De zebrafis is een veelgebruikt modelorganisme voor ontwikkelingsbiologie en recentelijk ook voor immunologisch onderzoek, vanwege de opvallende overeenkomsten tussen het immuunsysteem van de zebrafis en dat van de mens. In de laatste jaren hebben zebrafismodellen voor verschillende infectieziekten en vormen van kanker zich bewezen als nuttige aanvullingen op zoogdiermodellen.

Het doel van het in dit proefschrift beschreven onderzoek was om inzicht te krij-

gen in de complexiteit van het transcriptoom tijdens infectieziekten en kanker en in de regulerende functies van miRNAs bij deze ziekteprocessen, door gebruik te maken van de nieuwe technologische mogelijkheden voor transcriptoomanalyse en de mogelijkheden van de zebra-vis als modelorganisme.

In **Hoofdstuk 2** rapporteren wij voor het allereerst over de toepassing van groot-schalige sequentieanalyse voor onderzoek naar de reactie van het transcriptoom van een gastheerorganisme op een bacteriële infectieziekte. Hierbij hebben wij gebruik gemaakt van het Solexa/Illumina-systeem voor digitale genexpressie (DGE of Tag-Seq), een methode die is gebaseerd op het detecteren van korte transcript-specifieke sequenties (“tags”). In deze studie onderzochten wij de gastheerreactie van de zebra-vis op een infectie met *Mycobacterium marinum*, een model voor humane tuberculose. Door twee Tag-Seq-banken te maken van controlevissen en twee Tag-Seq-banken van mycobacterium-geïnfecteerde vissen konden wij de technische reproduceerbaarheid van de methode bevestigen. Met deze benadering konden wij 5-8 miljoen sequenties per bank detecteren en de kartering daarvan toonde aan deze representatief waren voor meer dan 70% van alle genen die aanwezig zijn in de publieke transcript-databanken. De Tag-Seq-sequenties die een significant verschillende expressie vertoonden na mycobacterium-infectie omvatten ongeveer 2% van de transcripten in de databases. Behalve kwantitatieve informatie over genexpressie, kon uit de Tag-Seq-analyse ook worden afgeleid of de transcriptsequenties van de coderende (“sense”) of van de niet-coderende (“antisense”) streng van het DNA afkomstig waren. Er is toenemend bewijs voor het bestaan van antisense-transcripten, hoewel de biologische relevantie hiervan nog veel vragen oproept. In onze studie vonden wij ongeveer 40% antisense-sequenties, in overeenstemming met andere recente grootschalige sequentiestudies in de muis en in menselijke cellen. Echter, het grootste deel van de sequenties die significant verschillend waren tussen geïnfecteerde vissen en de controles betrof sense-transcripten. Dit geeft aan dat de transcriptionele regulering van de door mycobacterium-geïnduceerde immunreactie vooral op het niveau van de sense-transcripten plaatsvindt. Daarom hebben wij onze verdere studie gericht op de analyse van de sense-transcripten, hoewel de regulering van antisense-expressie ook zeer interessant is en verder onderzoek behoeft.

Om onze resultaten te valideren hebben wij onze Tag-Seq-data vergeleken met een eerder samengestelde referentieset van mycobacterium-gereguleerde genen van de zebra-vis, die bepaald was door middel van een multiplatform-microarrayanalyse. Deze vergelijking liet een sterke correlatie tussen de twee methoden zien. Daarnaast hebben wij een nieuwe microarrayanalyse uitgevoerd met een zelfontworpen Agilent-platform, waaruit ook een aanzienlijke overlap met de Tag-Seq-data bleek. Het feit dat met Agilent-microarrayanalyse en met Tag-Seq-analyse dezelfde functionele groepen van genen gedetecteerd werden, toonde aan dat deze twee benaderingen voor transcriptoomanalyse goed vergelijkbaar zijn. Aan de andere kant bracht de vergelijking ook verschillen aan het licht die verklaard kunnen worden door de beperkingen van microarrayanalyse (zoals beperkte gevoeligheid, hybridisatie-artefac-

ten en onvolkomenheden in het array-ontwerp) of van de Tag-Seq-methode (waarbij sommige transcripten gemist worden omdat ze geen unieke sequentie bevatten of een knipplaats voor het enzym dat gebruikt wordt voor de vervaardiging van Tag-Seq-banken). Daarom kunnen microarrayanalyse en Tag-Seq-analyse beschouwd worden als complementaire technieken. Naast de vergelijking met microarrayanalyse, hebben wij ook enkele van de differentieel tot expressie komende transcriptsequenties geverifieerd met behulp van kwantitatieve PCR-analyse. Dit ondersteunde de Tag-Seq-resultaten zelfs in gevallen waar er een discrepantie was tussen de Tag-Seq- en microarray-resultaten. Vervolgens hebben wij aan de hand van verschillende voorbeelden laten zien hoe transcriptoomanalyse met behulp van Tag-Seq additieve waardevolle informatie over genexpressie kan opleveren. Op basis van de Tag-Seq-data konden wij concluderen dat er groepen zijn van gecoördineerd-gereguleerde genen op bepaalde chromosomen. De Tag-Seq-data waren tevens bruikbaar om voorspellingen van genmodellen (intron-exon-structuren) te verifiëren. Bovendien toonden de Tag-Seq-data aan dat infectie kan leiden tot expressie van andere transcript-isovormen en konden wij nieuwe mycobacterium-gereguleerde transcripten in kaart brengen die ontbraken in de beschikbare transcript-databanken. Bij elkaar genomen hebben wij met deze resultaten aangetoond dat het Solexa/Illumina-systeem voor digitale genexpressie een uiterst waardevolle technologie is om expressie te kwantificeren en om de complexiteit van het transcriptoom te ontrafelen.

In **Hoofdstuk 3** hebben wij sequentieanalyse gebruikt om meer kennis te vergaren over het transcriptoom van zebrafisembryo's tijdens de reactie van het aangeboren ("innate") immuunsysteem op een infectie met *Salmonella enterica* serovar Typhimurium (*S. typhimurium*). Hierbij hebben wij twee sequentiemethoden toegepast, de Tag-Seq-methode, die is gebaseerd op het bepalen van korte transcript-specifieke sequenties verkregen door enzymatische digestie, en de RNA-Seq-methode, waarbij de gehele transcripten worden gefragmenteerd en de sequentie bepaald. Wij analyseerden 1 dag oude zebrafisembryo's op 8 uur na infectie met *S. typhimurium*, welk tijdstip was gebaseerd op eerder microarraydata waarbij een sterke inflammatoire reactie was waargenomen. Onze analyse resulteerde in een set van 165 transcripten die een overeenkomstige reactie vertoonden in zowel de Tag-Seq-analyse als in de microarrayanalyse. Bij de vergelijking van Tag-Seq met RNA-Seq vonden wij een overlap van 241 transcripten die differentieel tot expressie kwamen, waarmee wij hebben laten zien dat deze twee sequentiebepalingsmethoden een vergelijkbare prestatie leveren bij het kwantificeren van de transcriptoomreactie op infectie. De gecombineerde transcriptoomdata van de sequentie- en microarrayanalyses hebben wij vervolgens gebruikt om een geannoteerde referentieset samen te stellen van infectie-responsieve genen in zebrafisembryo's. De genen in deze referentieset coderen voor transcriptiefactoren, signaaltransductie-eiwitten, cytokinen en chemokinen, complementfactoren, eiwitten betrokken bij apoptose en proteolyse alsmede eiwitten met anti-microbiële activiteiten. Daarnaast omvatte de referentieset genen voor zowel bekende als nieuwe eiwitten die niet eerder met de immuunreactie in verband

waren gebracht. Wij hebben de Tag-Seq-data van de *S. typhimurium*-infectie van zebravisembryo's ook vergeleken met de Tag-Seq-data van *M. marinum*-infectie in volwassen zebravissen uit Hoofdstuk 2. Het eindstadium van de *M. marinum*-infectie in volwassen vissen gaat gepaard met een sterke inflammatoire reactie, vergelijkbaar met de reactie van zebravisembryo's op *S. typhimurium*-infectie. De vergelijking resulteerde in een gemeenschappelijke set van 206 genen die een verhoogd expressieniveau vertoonden in beide infectiemodellen. Deze set omvatte transcriptiefactorgenen en signaaltransductiegenen betrokken bij de aangeboren immuunreactie, evenals genen waarvoor een rol bij de immuunreactie niet eerder bekend was. Deze transcriptoomdata vormen een waardevolle referentie voor toekomstig onderzoek naar gastheer-pathogeen-interacties met gebruik van zebravis-infectiemodellen.

Het onderzoek beschreven in **Hoofdstuk 4** was gericht op functies van miRNA's bij het immuunsysteem van gewervelde dieren (vertebraten) tijdens infectieziekten. Net als in de voorgaande hoofdstukken, gebruikten wij als ziektemodellen zowel volwassen zebravissen als zebravisembryo's die geïnjecteerd waren met de bacteriële pathogenen *M. marinum* and *S. typhimurium*. Door embryonale en volwassen zebravissen in onze experimenten te gebruiken konden wij onderscheid maken in miRNA-expressiepatronen die kenmerkend waren voor het aangeboren ("innate") immuunsysteem en voor het verworven ("adaptive") immuunsysteem, aangezien in embryo's alleen nog het aangeboren systeem is ontwikkeld. Door gebruik te maken van een nieuw Agilent-microarrayplatform voor miRNA-detectie konden we een set van miRNA's identificeren die algemeen reageren op verschillende typen infecties, waaronder leden van de miR-21-, miR-146- en miR-181-families. De resultaten waren in overeenstemming met eerdere studies die een verband hadden aangetoond van deze miRNA's met het immuunsysteem en met kanker in humane patiëntmateriaal en celcultuursystemen en in proefdiermodellen van zoogdieren. Echter, onze studie is de eerste die deze miRNA's in verband brengt met mycobacterium- en salmonella-infecties. Bovendien ondersteunen onze resultaten eerdere hypothesen dat regulering van de immuunreactie en kankerprocessen afhankelijk is van gemeenschappelijke miRNA's.

De miR-146-familie is eerder in verband gebracht met de regulering van signaalroutes van het aangeboren immuunsysteem in de mens en in diermodellen. Onze data zijn hiermee in overeenstemming, aangezien miR-146a en miR-146b niet alleen geïnduceerd werden in volwassen zebravissen, maar ook in embryo's die volledig afhankelijk zijn van het aangeboren immuunsysteem. Er wordt gedacht dat miR-146a en -b een rol spelen bij de fijnafstelling van de Toll-receptor- en interleukine-1-receptor-signalroutes (TLR- en IL1R-signalroutes) die een sleutelrol spelen bij de herkenning van pathogenen en bij de activering van de aangeboren immuunreactie. Om te onderzoeken of de functie van miR-146 geconserveerd is in de zebravis hebben wij een voorspelling gemaakt van de genen die door deze miRNA's gereguleerd zouden kunnen worden (doelwitgenen of "target genes"). Wij vonden herkenningsplaatsen voor miR-146-a en -b van de zebravis in *irak1* en *traf6*, twee genen die een rol spelen

in de TLR- en IL1R-sigtaalroutes en die beide herkend ook worden door de miR-146-familie van zoogdieren. Daarnaast vonden wij dat het transcript dat codeert voor MyD88, een essentieel adaptereiwit in de TLR- en IL1R-sigtaalroutes, ook mogelijke herkenningsplaatsen voor miR-146a en -b van de zebravis bevat. De inductie van miR-146 door bacteriële pathogenen is in overeenstemming met de genexpressiedata uit hoofdstukken 2 en 3, die laten zien dat tijdens infecties in de zebravis er verschillende genen van de TLR-sigtaalroute geïnduceerd worden, evenals doelwitgenen en negatieve regulatoren van deze route. Daarom ondersteunen onze data de hypothese dat miRNA-regulering een belangrijke rol speelt als een negatief terugkoppelingsmechanisme om de immuunreactie op de juiste manier onder controle te houden. Om andere mogelijke doelwitgenen van de miR-146-familie te identificeren hebben wij gebruik gemaakt van softwareprogramma's waarmee herkenningsplaatsen voor miRNA's voorspeld kunnen worden. De meeste van de voorspelde doelwitgenen van miR-146 die geconserveerd waren tussen de zebravis en de mens, waren genen die eerder in verband gebracht waren met infecties, tumorprogressie en processen die onderdeel uitmaken van de immuunreactie, zoals TLR-sigtaaltransductie, regulering van de activiteit van de centrale transcriptiefactor NF- κ B, apoptose, hemostase en T-cell-ontwikkeling en -functie. Onze voorspelling van geconserveerde miR-146-doelwitgenen ondersteunt daarom de veronderstelde functie van miR-146 bij de regulering van de immuunreactie. Met deze studie hebben wij laten zien dat de zebravis een bruikbaar model is om miRNA-functies bij het immuunsysteem van vertebraten te onderzoeken, aangezien de miRNA's die tijdens bacteriële infecties tot expressie komen sterk geconserveerd zijn en er een aanzienlijk overlap is in de voorspelde doelwitgenen van miRNA's in de mens en in de zebravis.

Om meer inzicht te krijgen in de rol van miRNA's in mechanismen van kanker hebben wij in **Hoofdstuk 5** miRNA-expressiepatronen in levertumoren van de zebravis onderzocht. Eerder was aangetoond dat de zebravis een waardevol ziektemodel is voor leverkanker, aangezien er opvallende moleculaire overeenkomsten zijn met humane hepatocellulaire carcinoma (HCC), zoals geconserveerde patronen van genexpressie en transcriptionele regulering van verschillende bij kanker betrokken gengroepen. Op te bepalen of er ook conservering is op het niveau van miRNA-expressie tussen levertumoren van de zebravis en van de mens, hebben wij een miRNA-transcriptoomanalyse uitgevoerd van carcinogeen-geïnduceerde levertumoren van zebravissen. Wij hebben miRNA's die een significant verschillende expressie vertoonden tussen levertumoren en gezonde lever van de zebravis vergeleken met de miRNA's die eerder in verband waren gebracht met humane HCC. Deze vergelijking resulteerde in een set van miRNA's waarvan de expressie ontregeld was in levertumoren van zowel de zebravis als de mens. De meest opvallende overeenkomsten waren de inductie van miR-21, miR-23a, miR-146a/b, miR-221 en miR-222, en de repressie van miR-1 en miR-122. Bovendien konden wij diverse andere miRNA's aantonen in levertumoren van de zebravis die niet eerder in verband gebracht waren met humane HCC, maar die wel eerder geassocieerd waren met verschillende andere typen van

kanker. Voor deze miRNA's zijn functies gepostuleerd bij het ontstaan van tumoren, bij de suppressie van tumoren, bij apoptose, proliferatie, differentiatie, invasiviteit en metastase. Naast de ontregeling van primaire miRNA's, vonden wij in verschillende gevallen ook ontregeling van de zogenoemde ster-sequenties. Dit zijn miRNA's die afkomstig zijn van de tegenovergestelde arm van de haarspeldstructuur van precursor-miRNA's. Bovendien konden wij een differentiële expressie waarnemen van een grote groep van sequenties die overeenkomen met voorspelde haarspeldstructuren in het zebrawisgenoom. Van deze structuren is het nog niet bekend of het mogelijk ook miRNA-genen zijn of dat er andere typen niet-coderende RNA's van worden afgeschreven, die nog verdere studie behoeven. Samen met de eerder gepubliceerde genexpressiestudies tonen onze miRNA-expressiedata aan dat zebrawis-levertumoren een geschikt model zijn voor het bestuderen van humane levertumoren, gezien de aanzienlijke moleculaire overeenkomsten. Tevens tonen onze resultaten aan dat de Agilent-microarraytechnology zeer bruikbaar is om nieuw voorspelde miRNA's of andere niet-coderende, regulerende RNA-moleculen te detecteren.

Vervolgens hebben wij de set van levertumor-gerelateerde miRNA's vergeleken met de set van miRNA's die geïnduceerd waren door mycobacterium-infectie, zoals bepaald in hoofdstuk 4. Deze vergelijking resulteerde in een gemeenschappelijke groep van tumor- en immuun-gerelateerde miRNA's, waaronder leden van de miR-15-, miR-16-, miR-21-, miR-34- en miR-146-families. Dit ondersteunt de opvatting dat veel miRNA's die betrokken zijn bij regulering van de immuunreactie, ook een rol spelen bij processen van kanker. Tenslotte hebben wij onderzocht of de doelwitgenen van kanker-gerelateerde miRNA's mogelijk ook geconserveerd zijn. Functionele annotatie van experimenteel-gevalideerde doelwitgenen van miRNA's die betrokken zijn bij humane leverkanker, toont aan dat deze genen een rol spelen in diverse kanker-gerelateerde processen en signaalroutes. Voorspelling van de miRNA-herkenningsplaatsen in de overeenkomstige genen van de zebrawis liet zien dat verschillende doelwitgenen van de miR-146-familie (*IRAK1*, *TRAF6*, *IRF5*), de miR-221/miR222-familie (*CDKN1B*, *CDKN1C*, *KIT*) en de miR-1-familie (*GJA1*, *PDCD4*, *TPM4*, *LASP1*, *TMSB4X*, *RABL2A*, *RABL2B*) mogelijk geconserveerd zijn tussen de mens en de zebrawis. Deze genen zijn betrokken bij het immuunsysteem zowel als bij kanker-gerelateerde processen, zoals regulering van de celcyclus, proliferatie, apoptose, celadhesie, celmotiliteit en organisatie van het cytoskelet. Bovendien functioneren deze genen in kanker- en immuun-gerelateerde signaalroutes, zoals MAPK-, NF- κ B-, p53- en ErbB-siginaaltransductie, alsmede siginaaltransductie door kleine G-eiwitten. Door deze resultaten wordt de overeenkomst tussen leverkanker van de zebrawis en de mens verder versterkt, evenals de hypothese dat miRNA's een belangrijke rol spelen bij deze kanker-gerelateerde processen in beide soorten. Onze miRNA-resultaten hebben een nieuw niveau van conservering tussen levertumoren van de zebrawis en de mens aan het licht gebracht naast de al aangetoonde sterke conservering van genexpressiepatronen en regulering door transcriptiefactoren. Daarom kan de zebrawis als leverkankermodel bijdragen om miRNA-functies

bij kanker te ontrafelen en om de interactie tussen kanker en het immuunsysteem verder te onderzoeken.

Concluderend heeft het in dit proefschrift beschreven onderzoek ondersteunend bewijs geleverd dat de zebravis een waardevol modelsysteem is voor infectieziekten en kanker, evenals voor het bestuderen van de regulerende functies van miRNAs bij het immuunsysteem en ziekten van vertebraten. Met het gebruik van de nieuwe generatie van technologieën voor sequentieanalyse voor de studie van mycobacterium- en salmonella-infectiemodellen van de zebravis, hebben wij voor het eerst aangetoond dat het Solexa/Illumina-systeem voor digitale genexpressie (DGE of Tag-Seq) een zeer bruikbare technologie is voor transcriptoomanalyse van de gastheerreactie op een infectie. Onze resultaten hebben aangetoond dat deze technologie waardevolle additionele informatie oplevert ten opzichte van microarrayanalyse, zoals inzicht in infectie-geïnduceerde expressie van alternatieve transcript-isovormen. Bovendien hebben wij laten zien dat een andere grootschalige sequentietechnologie, RNA-Seq, evenzeer geschikt is voor het kwantificeren van de transcriptoomreactie op infectie. Aangezien RNA-Seq een nog krachtiger techniek is dan Tag-Seq om transcriptionele landschappen in kaart te brengen, is het te verwachten dat een toekomstige dieptestudie van RNA-Seq-data verdere inzichten zal opleveren in de expressie van alternatieve isovormen tijdens infectie. Door combinatie van de transcriptoomdata van de sequentie- en microarrayanalyses hebben wij een waardevolle referentieset samengesteld van infectie-responsieve genen in zebravis-infectiemodellen. Deze infectie-responsieve genen omvatten niet alleen welbekende immuunreactiegenen van vertebraten, maar ook genen die niet eerder met infectieziekten in verband zijn gebracht en die daarom zeer interessant zijn voor functioneel onderzoek naar hun mogelijke rol bij gastheer-pathogeen-interacties. Daarnaast hebben wij een set van infectie-responsieve en leverkanker-gerelateerde miRNAs in kaart gebracht die geconserveerd zijn tussen de mens en de zebravis en hiermee ondersteunend bewijs geleverd dat gemeenschappelijke miRNAs een regulerende rol spelen bij de immuunreactie en bij processen die ten grondslag liggen aan de ontwikkeling van kanker. Onze gecombineerde mRNA- en miRNA-transcriptoomdata vormen een sterke basis voor de toepassing van de zebravis als infectie- en kankermodel in toekomstig onderzoek.

