



## Samenvatting proefschrift Kevin Gerritsen

“RNA analysis of HLA alleles: tidings from the messenger”

**Promotie: 27 mei 2016**  
**Universiteit Maastricht**

**Promotor:**  
Prof. dr. M.G.J. Tilanus

**Co-promotor:**  
Dr. C.E.M. Voorter  
Dr. L. Wieten

De HLA eiwitten zijn de humane variant van het Major Histocompatibiliteits Complex (MHC) en sinds hun initiële ontdekking in de vroege jaren 1900 is gebleken dat ze gecodeerd worden door polymorfe genen in het menselijke genoom. De HLA moleculen zijn geclusterd in klasse I en klasse II eiwitten op basis van hun functie en eiwitstructuur. De HLA klasse I moleculen zijn aanwezig op de meeste kernhoudende cellen en spelen een belangrijke rol bij de immuunreactie, terwijl de HLA klasse II moleculen alleen tot expressie gebracht worden op bepaalde subsets van cellen namelijk de antigeen presenterende cellen. Binnen de HLA klasse I en II clusters van genen zijn drie verschillende klassieke HLA klasse I loci, HLA-A, -B en -C, en in klasse II de loci HLA-DR, -DP en -DQ, geïdentificeerd.

De hoge mate van variatie in de HLA moleculen maakt de presentatie van een grote verscheidenheid aan peptiden aan het immuunsysteem mogelijk. Tegelijkertijd zorgt deze diversiteit voor problemen in transplantatie omdat peptiden van gemismatchte HLA donoren gepresenteerd worden op cellen van de patiënt en andersom, hetgeen complicaties kan veroorzaken bij een transplantatie. Het beste resultaat van een transplantatie is aangetoond in patiënt-donor paren die gematched zijn voor hun HLA-A, -B, -C, -DRB1 en -DQB1 allelen. Wanneer een mismatch in deze loci niet te vermijden is, blijken bepaalde HLA mismatches niet immunogeen en worden getolereerd. Naast HLA allelen die tot expressie komen op het celoppervlak zijn er HLA allelen met een afwezige, lage of nog onbekende oppervlakte expressie die mogelijk het zoeken naar een geschikte donor compliceren. Deze factoren, gecombineerd met de grote diversiteit van de HLA allelen, bemoeilijken de zoektocht naar een geschikt donor-patiënt paar en vereisen een grondige overweging in de HLA matching procedures.

Er zijn verschillende technieken beschikbaar voor HLA typering die met behulp van serologische of moleculaire methoden de HLA allelen van de patiënt en de donor identificeren. Moleculaire technieken zijn afhankelijk van speciaal ontworpen allel- of locus-specifieke PCR primers en probes die hechten aan specifieke polymorfe posities



kenmerkend voor één HLA allel of een groep van allelen. De zeer hoge diversiteit aan HLA allelen vormt een uitdaging voor de moleculaire typeer technieken aangezien zij

het ontwerp van een zeer groot aantal primers en probes vereisen en dan nog meestal beperkt zijn tot het identificeren van HLA polymorfismen van enkele exonen van de HLA genen. Wanneer een HLA typering uitsluitend op DNA is gebaseerd, worden de RNA expressie niveaus en eventueel aanwezige alternatief gespliceerde HLA allelen niet geïdentificeerd.

**In hoofdstuk 2** is een benadering met behulp van RNA Sanger sequencing beschreven, die ontwikkeld en gevalideerd is voor de HLA klasse I allelen. Dit werd bereikt met behulp van een enkele locus-specifieke of twee overlappende groep-specifieke PCR amplificaties en resulteerde in de volledige lengte van de coderende nucleotide sequenties. Om deze RNA benadering te valideren werd een referentie panel geselecteerd met samples die tot de belangrijkste antigen-groepen behoorden. Analyse van deze nieuwe RNA benadering resulteerde in identieke resultaten vergeleken met DNA Sanger typeren. In deze studie zijn van 13 allelen met onbekende exon nucleotide sequenties in de IMGT/HLA database, en 3 allelen met afwijkende expressie, met succes de nucleotide sequenties bepaald met behulp van de groep-specifieke RNA Sanger sequencing methodologie. De RNA Sanger sequencing methodologie is opgezet voor het bepalen van de volledige coderende sequentie van HLA klasse I allelen. Met deze benadering zijn ook alternatief gespliceerde varianten te definiëren.

**In hoofdstuk 3** is de RNA Sanger sequencing methodologie beschreven die ontwikkeld is voor het hemizygoot sequencing van het HLA-DRB1 locus. Meer dan 90% van de HLA-DRB1 allelen hebben onvolledige allel sequenties die de studie van de functionele relevantie van polymorfismen buiten exon 2 niet mogelijk maken. De RNA Sanger sequencing benadering werd succesvol gevalideerd op een panel van DRB1 allelen, waarin alle serologische equivalenten vertegenwoordigd zijn die volledig bekende coderende sequenties in de IMGT/HLA database hadden. Vervolgens werd de methode toegepast op een panel van 52 samples waarin 54 allelen vertegenwoordigd zijn die een onvolledige allel nucleotide sequenties hebben. Van alle 54 allelen werd een volledig coderende sequentie bepaald en toegevoegd aan de HLA allel database (HLA-IMGT) waarbij één nieuw en één gecorrigeerd allel gerapporteerd werden. Deze studie toont de universele toepasbaarheid aan van de op RNA sequencing gebaseerde benadering voor het bepalen van de volledige coderende sequentie. Hierdoor zal het volledige polymorfisme van HLA-DRB1 allelen gedefinieerd kunnen worden. Daarnaast zal het verder ontrafelen van de onbekende nucleotide sequenties van de HLA-DRB1 exonen, studies naar de functie van polymorfismen buiten de peptide bindingsgroeve mogelijk maken.

**In hoofdstuk 4** werd de huidige kennis en de invloed van alternatieve splicing op HLA expressie bestudeerd en besproken. Alternatieve splicing kan worden beïnvloed door één enkel nucleotide polymorfisme (SNP) wanneer het enerzijds de affiniteit verlaagd van een reguliere splice site voor het spliceosome, die verantwoordelijk is voor het katalyseren van alle splicing reacties, of anderzijds als het een secundaire splice site



introduceert. De herkenning van een nucleotide sequentie als mogelijke splice site is gebaseerd op het aantal nucleotiden dat deze sequentie gemeen heeft met een hoog

geconserveerde nucleotide sequentie die geldt voor alle splice sites, met een GT voor de 5' splice sites en voor de 3' splice sites een AG in het upstream intron. Echter, bij

het analyseren van alle HLA klasse I 5' en 3' splice sites werden er aanvullende geconserveerde nucleotiden geïdentificeerd rondom deze GT en AG nucleotiden. Wanneer het aantal nucleotiden van een bepaalde sequentie complementair aan deze consensus sequentie verandert door SNP's, kan het effect sterk genoeg zijn om de plek te laten functioneren als een cryptische splice site en dus te laten concurreren met de reguliere splice site. Deze selectie van splice site bepaalt het splicing patroon, regulier of alternatief, van het HLA mRNA en beïnvloedt de resulterende oppervlakte expressie. Een interessante bevinding was dat niet alle HLA splice sites voldoen aan deze consensus sequentie, maar desondanks nog steeds worden geselecteerd door het spliceosome. Dit suggereert dat er meerdere regulatoire mechanismen actief HLA splicing kunnen beïnvloeden.

**In hoofdstuk 5** werd RNA sequencing toegepast om alternatief gespliceerde mRNA transcripten te bestuderen van het HLA-A\*23:19N allel dat eerder geclassificeerd als een Q allel, HLA-A\*23:19Q, met een nog niet opgelost mRNA en eiwit expressie profiel. De identificatie van HLA allelen met een afwijkend expressie profiel, zoals niet tot expressie komende allelen, zogenaamde nul allelen, zijn klinisch relevant bij een SCT, met name voor donor selectie. Het is echter onduidelijk hoe de allelen met een onbekende alternatieve expressie beschouwd moeten worden in de HLA matching criteria. Het HLA-A\*23:19Q allel werd aangetroffen in een sample uit Guadeloupe: GD23Q, HLA-A\*23:19Q,29:02:01. De volledige DNA sequentie van HLA-A\*23:19Q bleek één enkel polymorfisme op positie 619 (G>A) te hebben ten opzichte van HLA-A\*23:01:01. Serologische typering toonde slechts de aanwezigheid van HLA-A29 op het cel oppervlak, daarentegen werd HLA-A23 niet gedetecteerd. De afwezigheid van een HLA-A23 oppervlakte expressie werd bevestigd door flowcytometrie met gebruik van een rechtstreeks gemerkt monokonaal antilichaam en een panel van vijf indirect gemerkte polyklonale antilichamen gericht tegen alle HLA-A23 (HLA-A9) moleculen. Vervolgens identificeerde een allel specifieke amplificatie de afwezigheid van volle lengte mRNA en de aanwezigheid van twee alternatief gespliceerde mRNA's. Gebaseerd op het gebrek aan HLA-A23 oppervlakte expressie en de aanwezigheid van enkel afwijkende mRNA transcripten, toonde deze studie aan dat HLA-A\*23:19Q een niet tot expressie komend HLA allel is en op grond van deze studie geherclassificeerd is als HLA-A\*23:19N.

In hoofdstuk 6 is de aanwezigheid van een tot co-expressie komende HLA-C mRNA variant beschreven waarin de exon 5 nucleotide sequentie ontbreekt. Deze alternatief gespliceerde mRNA isovorm zou kunnen coderen voor een niet membraan gebonden maar in de circulatie voorkomend HLA-C molecuul. De exacte rol van niet membraan gebonden HLA moleculen in de immunologische processen is grotendeels niet bestudeerd. Maar het is verleidelijk om te speculeren dat deze moleculen zouden kunnen interfereren met de regulatie van het immuunsysteem door het activeren of remmen van de immuun cellen. De in deze studie ontdekte mRNA variant komt tot co-



expressie naast het normale volledige mRNA transcript en de concentratie en aanwezigheid blijkt afhankelijk te zijn van de HLA-C allel groep, met enkele groepen die een ontbrekende, sterke of een zwakke aanwezigheid van de isovorm hebben.

Uitzondering hierop is de HLA-C\*03 allel groep die een variërend expressie patroon van de verschillende allelen groepen vertoont. Aangezien dit mRNA transcript codeert voor een HLA-C molecuul dat zijn transmembraan gedeelte mist, is het zeer waarschijnlijk dat het tot expressie komt als een niet membraan gebonden maar in de

circulatie voorkomend HLA molecuul en dat het mogelijk ook in sommige individuen een rol speelt bij de productie van een normaal aanwezige concentratie oplosbaar HLA-C.

Het introduceren van RNA sequencing maakt de visualisatie van mRNA isovormen mogelijk die geproduceerd worden door alternatieve splicing en die mogelijk anders coderen voor transcripten met een afwijkende oppervlakte of niet membraan gebonden expressie. RNA sequencing zal ook leiden tot het oplossen van de volledige coderende nucleotide sequenties van allelen en daarin nog onbekende HLA polymorfismen aantonen. Dit zal studies mogelijk maken naar het functionele effect van polymorfismen buiten de exonen 2 en 3 van HLA klasse I, en exon 2 van HLA klasse II, coderend voor de peptide bindings groeve