

Samenvatting en algemene conclusie

Dit proefschrift richt zich op het verbeteren van de gepersonaliseerde geneeskunde door middel van technologische innovaties. Daarbij ligt de focus op het optimaliseren van genotyperingsstrategieën in farmacogenetica (PGx) en op het bestuderen van nieuwe informatie uit single cell transcriptomics. Gepersonaliseerde geneeskunde is gericht op het bieden van geïndividualiseerde gezondheidszorg die precies is afgestemd op het genetische profiel van de patiënt. Het doel hiervan is om bij te dragen aan de preventie, de diagnostisering en de behandeling van ziekten. Door het DNA van de patiënt te onderzoeken binnen PGx, kunnen genetische verschillen worden geïdentificeerd die een invloed hebben op de verwerking van geneesmiddelen. Hierdoor is het mogelijk om de dosering van geneesmiddelen proactief aan te passen aan de individuele behoeften van een patiënt. Binnen het luik van single cell transcriptomics laat de identificatie van tumorspecifieke neoantigenen of tumordodende T-cellen de ontwikkeling van interessante, gepersonaliseerde therapieën toe. Ons vermogen om die informatie te verkrijgen, is echter afhankelijk van de beschikbare technologische methodes. In dit proefschrift worden daarom nieuwe technologieën ontwikkeld en wordt de toegevoegde waarde van deze innovaties voor het bevorderen van gepersonaliseerde geneeskunde besproken.

Hoofdstuk 1 fungeert als een overkoepelend, inleidend hoofdstuk en biedt de nodige achtergrondinformatie over de structurele organisatie van de cel. Daarnaast wordt uitgelegd hoe DNA, RNA en eiwitten als essentiële bouwstenen bijdragen aan het leven.

Het eerste deel van dit proefschrift, bestaande uit **Hoofdstuk 2**, **Hoofdstuk 3** en **Hoofdstuk 4** richt zich op PGx. **Hoofdstuk 2** geeft een overzicht van de huidige stand van zaken rond PGx binnen de farmacologie. Het bespreekt de relevantie van PGx, de bestaande therapeutische richtlijnen en aanbevelingen voor het uitvoeren van PGx-tests, en behandelt verschillende genotyperingsstrategieën. Samenvattend is er een groot potentieel voor de klinische implementatie van PGx. Dit wordt onderbouwd door de grote hoeveelheid genetische varianten die in de literatuur zijn geïdentificeerd en hun aangetoonde impact op verschillende farmacokinetische en farmacodynamische parameters. De meeste informatie is beschikbaar over varianten in de genen van de cytochroom P450 isoenzym-familie. Verschillende internationale consortia stellen therapeutische richtlijnen en testadviezen ter beschikking, waarin aanbevelingen worden gedaan over welke genen getest moeten worden en hoe die testresultaten geïnterpreteerd dienen te worden. Hoewel traditionele genotyperingsstrategieën, die gebaseerd zijn op PCR of Sanger sequencing, nog steeds hun plaats hebben binnen de huidige

testmogelijkheden, bieden next-generation sequencing gebaseerde technologieën veelbelovende nieuwe toepassingen omdat alle varianten binnen een gen gedetecteerd kunnen worden. Specifiek de long-read sequencingstrategieën leveren lange reads op, waarmee varianten gefaseerd kunnen worden en grotere structurele varianten in kaart gebracht kunnen worden.

In **Hoofdstuk 3** tonen we aan dat het ‘adaptive sampling’ algoritme van Oxford Nanopore Technologies (ONT) succesvol kan worden ingezet om genetische varianten te identificeren binnen een panel van 1036 genen uit de PharmGKB database. Onder die genetische varianten vallen zowel SNVs/INDELs als complexe, structurele varianten. Onze resultaten laten zien dat het mogelijk is om drie DNA-stalen tegelijkertijd te analyseren in één sequencing run, wat de kosteneffectiviteit aanzienlijk verbetert. Door de varianten die met onze strategie zijn opgespoord te vergelijken met varianten in literatuurreferenties, kunnen we aantonen dat de mate van recall (99.35%) en precisie (99.84%) groot is. Daarnaast hebben we star-allele-bepalingen uitgevoerd om de PGx-diploypes te achterhalen. Hieruit blijkt een sterke correlatie tussen de gouden standaard en onze resultaten. Voor bepaalde genen leidde onze methode zelfs tot verbeterde star-allelbepalingen. We concluderen dat PGx op basis van long-read sequencing een zeer waardevolle tool is om de implementatie van farmacogenetica te bevorderen.

In **Hoofdstuk 4** voerden we een heranalyse uit van de hierboven beschreven data. Hierbij gebruikten we de laatste nieuwe versies van basecalling-, variant calling- en faseringsalgoritmes. Ook implementeerden we een nieuwe star-allelcaller, StarPhase, in onze analysepipeline. De resultaten van de heranalyse tonen aan dat de correcte star-allelen teruggevonden werden, zelfs voor het complexe *CYP2D6*-gen. Bijkomend voerden we een vergelijkende studie uit tussen onze strategie en het Twist Alliance PGx-panel, dat gesequencet werd op het PacBio en/of ONT platform. De resultaten van die studie tonen aan dat de gevonden star-allelen tussen beide methoden zeer sterk gelijklopend waren. Samengevat illustreren onze resultaten het potentieel van doelgerichte long-read sequencing voor PGx-toepassingen.

Het tweede deel van dit proefschrift bestaat uit **Hoofdstuk 5**, **Hoofdstuk 6** en **Hoofdstuk 7** en richt zich op het veld van single cell transcriptomics. **Hoofdstuk 5** biedt de nodige achtergrondkennis over de regulatie van genexpressie en RNA-sequencing. Genexpressie in de cel wordt sterk gereguleerd, zowel op (post)transcriptioneel als op translationeel niveau. Bulk RNA-sequencingmethodes laten toe om differentiële genexpressie tussen twee of meer condities te onderzoeken. Ze missen echter de resolutie om subtiele veranderingen in de expressie tussen

individuele cellen te gaan bestuderen. De afgelopen jaren zijn verschillende single cell RNA-sequencingstrategieën (scRNA-seq) ontwikkeld. De keuze voor de meest geschikte strategie hangt af van het aantal cellen dat onderzocht moet worden, de fractie van het transcript dat onderzocht moet worden, en de beschikbare instrumenten.

In **Hoofdstuk 6** introduceren we een nieuwe, plaatgebaseerde scRNA-seq-strategie, genaamd PB10X. Die methode maakt het mogelijk om individuele cellen direct te sorteren in 384-well platen, waardoor het gebruik van het 10X Genomics microfluidics-platform overbodig wordt. Hierdoor is het niet langer vereist om bij de start van het experiment over een groot aantal cellen te beschikken, wat standaard vereist is bij 10X Genomics. De PB10X-strategie ondersteunt zowel de opbouw van sequencing libraries gericht op genexpressie als op de karakterisering van het T-celreceptor (TCR) repertoire. Door PB10X te combineren met TCR-aanrijking, konden we bij 81,71% van de geanalyseerde cellen zowel de alfa- als de bèta-ketens van de TCR reconstrueren. Bij vergelijking van de prestaties van genexpressie-libraries tussen PB10X en andere strategieën, bleek PB10X het laagste aantal unieke genen en unieke UMIs per cel te detecteren. Desondanks biedt PB10X een waardevolle scRNA-seq-strategie, vooral in onderzoekomstandigheden waarin slechts een beperkt aantal gesorteerde cellen beschikbaar is en 10X compatibiliteit nodig is.

Hoofdstuk 7 beschrijft de ontwikkeling van een single-cell long-read isoform sequencing strategie (scLIS-seq). De scRNA-seq-methodes die in Hoofdstuk 5 zijn besproken, zijn vooral geschikt voor short-read sequencing, waardoor ze beperkingen hebben bij het in kaart brengen van alternatieve isoformen. Om de scLIS-seq-methode te valideren, werd deze gebenchmarkt ten opzichte van de short-read Smart-seq3xpress (SS3X)-strategie. De vergelijking werd uitgevoerd voor zowel Jurkat- als voor HEK293T-cellen, waarbij hetzelfde cDNA als uitgangspunt werd gebruikt. De resultaten tonen een sterke correlatie in genexpressie tussen beide strategieën. Daarnaast werd aangetoond dat voldoende sequencing diepte essentieel is om vergelijkbare UMI-counts terug te vinden. Expliciet wordt aangetoond dat scLIS-seq het mogelijk maakt om langere en meer accurate isoforms te identificeren vergeleken met isoform reconstructies op basis van short-read sequencing. Hieruit kan worden geconcludeerd dat scLIS-seq een duidelijke meerwaarde biedt bij het identificeren van isovormen op het niveau van individuele cellen.

Het laatste hoofdstuk, **Hoofdstuk 8**, bespreekt de bredere internationale context, relevantie, en toekomstperspectieven van het uitgevoerde onderzoek. We beschrijven de implementatie van PGx binnen Europa en België, en gaan dieper in op de huidige discussie binnen het veld over

reactief testen versus preemtief testen van genotypes. Daarnaast bespreken we de huidige hindernissen voor de klinische implementatie van PGx, waaronder de kosten, en lichten we toe hoe de door ons ontwikkelde strategie hieraan kan bijdragen. Verder bespreken we de toekomstige stappen die nodig zijn om optimaal gebruik te maken van de informatie uit long-read sequencing data. Voor het luik rond single cell transcriptomics bespreken we hoe scRNA-seq, en de in Hoofdstuk 6 en Hoofdstuk 7 ontwikkelde methodes, zouden kunnen bijdragen aan gepersonaliseerde neoantigen-vaccins en gepersonaliseerde TIL-therapie. We zien dat deze innovatieve therapieën het potentieel hebben om een levensreddende impact te hebben op patiënten.

Het onderzoek dat we in dit proefschrift presenteren, levert waardevolle bouwstenen voor het versnellen van de doorbraak van gepersonaliseerde geneeskunde. De methodologieën die we hebben ontwikkeld, sluiten naadloos aan bij de strategische doelstellingen van 'Vlaanderen in transitie. Prioriteiten voor wetenschap, technologie en innovatie op weg naar 2025'. We zetten expliciet in op gepersonaliseerde zorg en het centraal stellen van de patiënt binnen het ziektemanagement. De daadwerkelijke implementatie van gepersonaliseerde geneeskunde in de klinische praktijk is echter afhankelijk van diverse factoren, waaronder de economische haalbaarheid. Een correcte inschatting van de kosteneffectiviteit vereist dat de onderliggende biologische metingen met de hoogste nauwkeurigheid en betrouwbaarheid kunnen worden uitgevoerd. Dit onderstreept het belang van geavanceerde analysemethoden, zoals die we in dit werk hebben ontwikkeld.