

TI2800 Contextproject: Samenvatting week 4

Groep 5/E

7 maart 2012

1 Wat is Bioinformatica?

In de *review paper* WHAT IS BIOINFORMATICS? AN INTRODUCTION AND OVERVIEW uit 2001 wordt een overzicht gegeven van de uitdagingen die aanwezig zijn in het opslaan en analyseren van de datastroom die biologen in hun werk genereren. Om de grote datastroom overzichtelijk te kunnen houden zijn verschillende soorten databases ontwikkeld. Van het opslaan van afzonderlijke eiwitten met hun eigenschappen en functie, tot het opslaan van de volgorde van de baseparen in de het hele DNA van verschillende organismen, de zogenaamde *whole-genome*-databases. Door toenemende integratie van deze verschillende databases is de informatie veel beter in zijn context te plaatsen en kunnen waardevoller resultaten worden behaald.

Nu de gegevens goed kunnen worden opgeslagen en geraadpleegd kan er worden nagedacht over wat voor analyses we erop kunnen loslaten. Er wordt bijvoorbeeld gekeken naar eiwitvolgordes, 3D-structuren en uitlijningen, welke delen van het DNA encoderen voor welke eiwitten en het voorspellen van secundaire en tertiäre structuur van eiwitsequenties.

Om al deze zaken mogelijk te maken zijn verschillende informatietechnieken nodig: simulaties en voorspellingen leunen zwaar op beschikbare kennis uit de informatica en de statistiek.

Tot slot worden enkele toepassingen opgesomd: Het zoeken naar vergelijkbare biomoleculen van een molecuul waar nog weinig van bekend is om zo meer te weten te komen over het onbekende molecuul. Het ontwerpen en produceren van medicijnen.

2 Uitdagingen in gepersonaliseerde medicijnen

Zoals in het einde van de vorige *paper* al even langs kwam is een van de belangrijke toepassingen van bioinformatica in de ontwikkeling van medicijnen. In BIOINFORMATICS CHALLENGES FOR PERSONALIZED MEDICINE uit 2011 wordt een overzicht gegeven van de recente ontwikkelingen op dit gebied. Er worden vier methodes beschreven die noodzakelijk zijn om de realisatie van deze medicijnen mogelijk te maken:

1. Verwerking op grote schaal van genoom-data.
2. Interpreten van functieveranderingen van genetische veranderingen.
3. Genetische interactie relateren aan het fenotype (waarneembare eigenschappen) van organismen.
4. Ontdekkingen vertalen naar toepassingen in de medische praktijk.

3 Trends in *computational biology*

Tot slot wordt in TRENDS IN COMPUTATIONAL BIOLOGY – 2010 een aantal trends beschreven in de hedendaagse bio-informatica.

Als eerste wordt de zogenaamde *next-generation sequence analysis* beschreven, waarin *RNA-splicing* beter voorspeld en geanalyseerd kan worden.

Vervolgens wordt onderzoek gedaan in beschikbare dataset om verbanden te zoeken tussen genetische data en fenotypes. Statistische analyse laat nieuwe verbanden zien.

Machine-leren wordt gebruikt om taken waartoe mensen prima in staat zijn te automatiseren waardoor menselijke vooringenomenheid en afleiding door andere substructuren uit de vergelijking gehaald kan worden waardoor beter vergeleken kan worden.

Tenslotte wordt een soort web 2.0 in de bio-informatica gesignaleerd. Een online-spel gebruikt menselijk ruimtelijk denken om computers beter in staat te stellen eiwitstructuren te voorspellen. Een ander bedrijf vraagt via een website aan mensen waarvan het DNA al bekend is informatie over hun eigenschappen om zo nieuwe verbanden te kunnen leggen.