

‘BiG Grid and beyond’ Een dagje rekenen voor een DNA-structuur

Vijftien jaar duurde het om het volledige menselijk genoom te ontrafelen. Met de geavanceerde technologieën van nu, zou dat huzarenstukje slechts twee weken duren. DNA-onderzoek is in een stroomversnelling geraakt dankzij e-Science. ‘In 2008 leverde één sequence run een halve GigaByte aan data, met de apparatuur van nu is dat al 70 GigaByte’, vertelt Barbera van Schaik, bioinformaticus in het Amsterdam Medisch Centrum. Haar taak is om analyses van DNA-experimenten te optimaliseren: ‘e-Science en e-Infrastructuren zijn daarbij onmisbaar.’

Doorbraak

‘Er is een snelle ontwikkeling geweest in DNA sequencing, de techniek voor het bepalen van de exacte lettercode van een stukje DNA. Dat heeft dit jaar tot een doorbraak geleid’, vertelt Van Schaik. Met haar collega bioinformatici en het e-bioscience team van het Bioinformatica Laboratorium kijkt zij naar genen die betrokken zijn bij zeldzame ziekten.

‘Als het gaat om virusinfecties, werken we samen met de Virus Discovery Unit. Bij onderzoek naar de rol van T- en B-cellen bij reuma werken we samen met de afdeling immunologie. Die cellen zijn belangrijk voor het afweersysteem, maar kunnen overactief zijn bij reuma.

In samenwerking met de Universiteit Leuven hebben we het gen geïdentificeerd dat verantwoordelijk is voor het Nicolaïdes–Baraitser-syndroom, een zeldzame aandoening die leidt tot een verstandelijke handicap, epilepsie en taalachterstand.

Met nieuwe DNA sequenceapparatuur hebben we het DNA bepaald van tien patiënten met dit syndroom. Dit hebben we vervolgens vergeleken met een referentiegenoom en zo konden we de verschillen zien. Door de patiënten ook onderling te vergelijken, zie je of ze dezelfde genmutatie hebben. Dat bleek inderdaad het geval te zijn, waardoor het gen kon worden geïdentificeerd. Vervolgens hebben we in een grote controlegroep van

het Regenboogproject ‘Het Genoom van Nederland’ opgezocht of het inderdaad om een nieuwe mutatie ging.’

Rekentijd

‘Voor het laatste hadden we maar twee weken om die analyse te doen. Dat kostte tot voor kort een maand rekentijd. Maar door de e-Infrastructuur van BiG Grid hebben we de uiteindelijke rekentijd teruggebracht naar anderhalve dag.

Data afkomstig van de DNA sequencers groeien exponentieel. De studies worden steeds omvangrijker, waardoor de kwaliteit van de e-Infrastructuur steeds belangrijker wordt. Het gaat hierbij niet alleen om rekenkracht, maar ook om het delen van kennis met andere onderzoekers en de beschikbaarheid van expertise die het project BiG Grid levert. BiG Grid is hiervoor een betrouwbare partner en dat is belangrijk als je met DNA data werkt.’

DNA-experimenten

De DNA-experimenten worden in het laboratorium uitgevoerd door middel van sequencing. De vrijgekomen data gaan voor analyse naar bioinformatici. Van Schaik: ‘In de jaren zeventig werd één sequentie tegelijkertijd geanalyseerd. De huidige apparatuur kan tot 200 miljoen sequenties tegelijkertijd sequencen. In plaats van naar één gen of een handjevol genen te kijken, kunnen we nu het hele DNA van een patiënt in kaart brengen. Je hebt dus een groter gebied om te zoeken waarin een patiënt afwijkend is ten opzichte van een gezond persoon.’

