

Protein-DNA interactions underlie practically any biological process, from the development of an embryo to diseases such as cancer. Therefore, understanding protein-DNA interactions is essential not only for fundamental biological research, but also in applied research including drug development. Methods to study protein-DNA interactions comprise genome-wide techniques, such as chromatin immunoprecipitation followed by micro-array hybridization (ChIP-chip) or high throughput sequencing (ChIP-seq), as well as techniques in which binding to a short DNA fragment is studied, as for example the electrophoretic mobility shift assay (EMSA).

In this thesis, I aimed to improve the methodology to study protein-DNA interactions in several ways. First, the 'Densitometric Image Analysis Software' (DIAS) has been developed to determine binding constants and cooperativity parameters from EMSAs with an accuracy of about 20% instead of the usual factor of 2. Second, software has been implemented to identify genomic binding regions in ChIP-chip data, and to predict binding sites within these regions using an energy-based position weight matrix derived from EMSAs analyzed with the DIAS software. Finally, nanopore simulation software has been developed to simulate current-voltage characteristics of nanopores, and a nanopore-based measurement instrument has been designed to detect the short DNA fragments previously used with EMSAs. This may lead to future measurements of protein-DNA binding constants at a significantly increased speed compared to EMSAs.

Proteïne-DNA-interacties liggen aan de basis van praktisch elk biologisch proces, van de ontwikkeling van een embryo tot ziektes zoals kanker. Daarom is het begrijpen van proteïne-DNA-interacties niet alleen cruciaal voor fundamenteel biologisch onderzoek, maar ook voor toegepast onderzoek waaronder de ontwikkeling van geneesmiddelen. Methodes om proteïne-DNA-interacties te bestuderen omvatten genomwijde technieken, zoals chromatine-immunoprecipitatie gevolgd door micro-array-hybridisatie (ChIP-chip) of hoge-doorvoersequencerings (ChIP-seq), en technieken waarin binding op een kort DNA-fragment bestudeerd wordt, zoals bv. de elektroforetische mobiliteitsveranderingsassay (EMSA).

Het doel van deze thesis was de methodologie om proteïne-DNA-interacties te bestuderen op verscheidene manieren te verbeteren. Ten eerste heb ik de 'Densitometric Image Analysis Software' (DIAS) ontwikkeld om bindingsconstanten en coöperativiteitsparameters te bepalen met EMSA's, met een nauwkeurigheid van ongeveer 20% in plaats van de gebruikelijke factor 2. Ten tweede heb ik software geïmplementeerd om genomische bindingsregio's te identificeren uit ChIP-chip data, en om bindingssites te voorspellen binnen deze regio's, gebruik makende van een energie-gebaseerde positie-specifieke gewichtsmatrix bepaald a.d.h.v. EMSA's geanalyseerd met de DIAS-software. Als laatste stap heb ik simulatiesoftware ontwikkeld om stroom-spanningskarakteristieken van nanoporiën te bepalen, en een nanopore-gebaseerd meetinstrument ontworpen om korte DNA-fragmenten te detecteren die voorheen gebruikt werden in EMSA's. Dit kan in de toekomst leiden tot metingen van proteïne-DNA-bindingsconstanten tegen een significant hogere meetnelheid dan verkrijgbaar met EMSA's.

Van Oeffelen Liesbeth