



Universiteit
Leiden
The Netherlands

The mechanical genome : inquiries into the mechanical function of genetic information

Tompitak, M.; Tompitak M.

Citation

Tompitak, M. (2017, October 11). *The mechanical genome : inquiries into the mechanical function of genetic information*. *Casimir PhD Series*. Retrieved from <https://hdl.handle.net/1887/53236>

Version: Not Applicable (or Unknown)

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/53236>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/53236> holds various files of this Leiden University dissertation.

Author: Tompitak, M.

Title: The mechanical genome : inquiries into the mechanical function of genetic information

Issue Date: 2017-10-11

SAMENVATTING

DNA, als de drager van genetische informatie een fundamenteel onderdeel van al het leven op aarde, is een lang, kettingvormig molecuul: een polymeer. Het is echter een speciaal soort polymeer, omdat de monomeren waar DNA uit bestaat (de nucleotiden) niet allemaal identiek zijn. Er zijn vier soorten nucleotiden, meestal aangeduid met A, T, C en G. Het onderscheid tussen deze monomeren is ook wat het coderen van informatie mogelijk maakt: de letters spelen een rol analoog aan de nullen en enen van het binair systeem dat onze digitale apparaten gebruiken.

Het onderscheid is echter niet puur informatie-theoretisch. De vier nucleotiden zijn noodzakelijkerwijs niet-identieke objecten, met verschillende fysische en chemische eigenschappen. Daarom is een verschil in DNA-sequentie niet alleen een verschil in gecodeerde informatie, maar ook een verschil in flexibiliteit, intrinsieke buiging, en andere elastische eigenschappen van het molecuul. Dergelijke verschillen kunnen verstrekkende gevolgen hebben.

DNA moet regelmatig worden vervormd. Het voornaamste voorbeeld is dat DNA sterk moet worden gebogen om in een cel te passen. Het menselijk genoom, bijvoorbeeld, is zo'n twee meter lang, en past dan ook alleen maar in onze kleine cellen omdat het op ingenieuze wijze is opgevouwen. Sommige delen van het DNA zijn makkelijker op te vouwen dan andere, vanwege de variaties in de elastische eigenschappen. Het gevolg is dat de DNA-sequenties het opvouwen van het genoom beïnvloeden. Dit leidt tot een systeem met een rijke fysica, waarbij de informatie die in het DNA ligt opgeslagen nauw is verbonden met hoe het zich gedraagt.

Het compactificeren van DNA is slechts één voorbeeld van zulk samenspel tussen de informatie in en de fysica van DNA (maar waarschijnlijk het meest belangrijke voorbeeld). In bredere context zijn er twee belangrijke vragen om te stellen. Ten eerste: maakt de natuur gebruik van de nauwe verbintenis tussen DNA-sequentie en fysiek gedrag? Het ligt voor de hand dat het evolutieproces DNA-sequenties ook selecteert op gunstige fysische eigenschappen, als daarmee enig voordeel te behalen valt. Ten tweede: hoe kunnen we deze link gebruiken om DNA te manipuleren? Wat voor eigenschappen en gedrag kunnen we een DNA-molecuul meegeven door de sequentie te veranderen?

De vijf projecten beschreven in dit proefschrift zijn stuk voor stuk pogingen om onze grip op deze vraagstukken te versterken. In het eerste project, beschreven in Hoofdstuk 2, zoeken we de grenzen op van de intrinsieke buiging van DNA-moleculen. We ontwerpen DNA-sequenties die leiden tot moleculen die sterk, steeds in dezelfde richting gebogen zijn, zodat deze moleculen van zichzelf superhelische vormen aannemen. Dergelijke superhelische structuren zien eruit als kleine veren, en gedragen zich ook soortgelijk: ze bieden een sterkere weerstand tegen een externe trekkracht dan we zouden verwachten van DNA.

In Hoofdstuk 3 wenden we ons tot het DNA-systeem dat zowel in de rest van dit proefschrift, als in de literatuur, de meeste aandacht krijgt: het nucleosoom. Het opvouwen van het DNA, zodat het in cellen past, gebeurt (in eukaryoten, e.g. dieren, planten en schimmels) met behulp van kleine eiwitcilinders waar het DNA omheen gewikkeld wordt. Het resulterende DNA-eiwitcomplex noemen we een nucleosoom. Omdat DNA sterk moet worden gebogen om dit complex te vormen, hangt de affiniteit van een stuk DNA voor het nucleosoom sterk af van de sequentie. Er is en wordt veel onderzoek verricht naar de sequentie-afhankelijke affiniteit van DNA voor het nucleosoom, hoe deze affiniteit de organisatie van een genoom beïnvloedt en in hoeverre deze invloed van belang is in levende cellen. In Hoofdstuk 3 proberen we dieper te kijken dan een simpele, scalaire eigenschap zoals algehele affiniteit.

Wanneer we aan de uiteinden van een stuk DNA trekken dat om een eiwitcilinder gewonden is, verwachten we uiteraard dat we het DNA los zullen trekken. Echter blijkt dat, vanwege de geometrie van het proces, nucleosomen kinetisch beschermd zijn tegen het geforceerd afwikkelen: er bestaat een energetische barrière, en deze barrière wordt des te hoger, naarmate de kracht toeneemt. Het resultaat is dat het lostrekken van DNA een significante hoeveelheid kracht vereist, hetgeen van pas komt omdat we over het algemeen niet willen dat de nucleosomen in onze cellen uit elkaar vallen zodra er aan het DNA getrokken wordt.

Niet geheel onverwacht is de manier waarop nucleosomen afwikkelen afhankelijk van de DNA-sequentie. We weten dat nucleosomen soms asymmetrisch afwikkelen, waarmee we bedoelen dat één van de uiteinden eerder loskomt dan het andere, doordat de DNA-sequentie aan dat uiteinde minder grote affiniteit voor het nucleosoom heeft.

In Hoofdstuk 3 proberen we dit idee verder te voeren, en ontwerpen we nucleosomen met een gat in de barrière tegen het afwikkelen. Het resultaat is dat we nucleosomen kunnen maken die niet sterk kinetisch beschermd zijn, en dat we ze via specifieke paden kunnen laten afwikkelen.

Het feit dat dit mogelijk is demonstreert dat de term ‘nucleosoom’, net als ‘DNA’, niet refereert aan een enkel systeem, maar aan een hele klasse van systemen, en dat nucleosomen zeer verschillend gedrag kunnen vertonen.

In Hoofdstukken 4 and 5 introduceren we nieuwe methodologie. Veel van het werk in dit proefschrift bouwt voort op de Mutation Monte Carlo (MMC) methode van Behrouz Eslami-Mossallam. Het idee achter deze methode is even eenvoudig als krachtig: neem een Monte Carlo-simulatie van een DNA-systeem, en voeg er mutaties aan toe. Door de simulatie tegelijkertijd zowel de fysische configuraties van het systeem, als de ruimte van mogelijke DNA-sequenties te laten doorzoeken, convergeert hij automatisch naar sequenties die hoge affiniteit voor het systeem hebben, en geeft ons de statistische eigenschappen van die sequenties. In Hoofdstukken 4 and 5 breiden we deze methodologie uit.

In Hoofdstuk 4 laten we eerst zien hoe MMC kan worden gebruikt om een bioinformatisch model te genereren dat de sequentie-afhankelijke affiniteit van DNA benadert voor het systeem waarvoor we een fysisch model simuleren. Dit nieuwe bioinformatische model geeft onze methoden significant meer bereik, omdat het feit dat het een benadering is, wordt gecompenseerd door een grote besparing in computationele complexiteit. Dit stelt ons in staat om vraagstukken onder de loep te nemen die geenszins te behappen zijn met een gedetailleerd biofysisch model zoals het nucleosoommodel uit Hoofdstuk 3. De rest van Hoofdstuk 4 is gewijd aan het benchmarken van het model, en het onderzoeken van wat de voorwaarden zijn voor een zo nauwkeurig mogelijke benadering.

In Hoofdstuk 5 onderzoeken we in meer detail de relatie tussen de nieuwe methode uit Hoofdstuk 4 en de MMC-methode. We zien dat onze nieuwe methode de kloof dicht tussen MMC-simulaties en SELEX-experimenten (een klasse van experimenten waarin sequenties worden geselecteerd op hun affiniteit voor een gegeven systeem, zoals het nucleosoom; de experimentele methodologie is in veel opzichten vergelijkbaar met de computationele MMC-methode). Ook krijgen we dieper inzicht in hoe de MMC-methode werkt, bovenal in de rol van de temperatuur in een MMC-simulatie.

In Hoofdstuk 6 passen we het nieuwe model van Hoofdstuk 4 toe op biologische data. Dankzij de grote winst in computationele efficiëntie die we met dit model boeken, kunnen we gehele genomen analyseren. We richten onze blik op de promotoren van de genen van verschillende organismen. Dit zijn de delen van een genoom, die zich voor de genen bevinden, en die invloed uitoefenen op de mate waarin een gen tot expressie komt. Hierin is een interessante rol weggelegd voor de elastische eigenschappen van de

betreffende DNA-sequenties, en specifiek voor hun affiniteit voor nucleosomen. Dit omdat DNA dat in een nucleosoom is gewikkeld, niet kan worden uitgelezen; het nucleosoom zit andere systemen die aan het DNA willen binden in de weg. Een grote of kleine affiniteit in de regio waar de machinerie die de genen uitleest wil binden heeft zo direct invloed op hoe vaak een gen wordt gelezen.

Het is bekend dat echte genomen in de DNA-sequenties van promotoren voor mechanische signalen coderen. Gist, een simpel ééncellig organisme, heeft in deze regio's bijvoorbeeld DNA-sequenties met een lage affiniteit voor nucleosomen, om het DNA toegankelijk te houden. Het menselijk genoom heeft juist signalen die nucleosomen sterk aantrekken. Men denkt dat dit het genoom in staat stelt in deze regio's nucleosomen te behouden in zaadcellen, waarin de meeste nucleosomen van het genoom worden verwijderd.

Bij het analyseren van deze signalen in de promotoren van zo'n 50 verschillende organismen uit verschillende takken van de fylogenetische stamboom vinden we een opmerkelijk universele overeenkomst: ééncellige organismen lijken allemaal op gist, en hebben signalen die nucleosomen afstoten, terwijl meercellige organismen allemaal signalen bevatten die nucleosomen aantrekken. Daarnaast bestaat er in het geval van meercellige organismen een correlatie tussen hoe sterk deze signalen zijn, en hoe complex het organisme is: zoogdieren hebben DNA-sequenties met zeer hoge affiniteit voor nucleosomen, terwijl simpelere dieren, zoals fruitvliegjes, zwakkere signalen vertonen. Of al deze signalen, en de strakke scheiding die we zien tussen ééncelligen en meercelligen, in alle gevallen dezelfde functies hebben als in gist en in mensen, zal nog verder moeten worden onderzocht.

Het beschreven onderzoek en de bijbehorende bevindingen pogen antwoorden te verschaffen op de twee vragen die we aan het begin van deze samenvatting stelden. We hebben gekeken naar mechanische signalen die in echte genomen te vinden zijn; we hebben de grenzen opgezocht van de eigenschappen die we via de sequentie aan een DNA-systeem kunnen meegeven; en we hebben nieuwe methoden aangedragen die ons in staat stellen om de mogelijkheden die de sequentie-afhankelijke mechanische eigenschappen van DNA ons bieden nog verder te onderzoeken.