



## UvA-DARE (Digital Academic Repository)

### Genetic regulatory networks inference : modeling, parameters estimation & model validation

Fomekong Nanfack, Y.

**Publication date**  
2010

[Link to publication](#)

#### **Citation for published version (APA):**

Fomekong Nanfack, Y. (2010). *Genetic regulatory networks inference : modeling, parameters estimation & model validation*.

#### **General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

#### **Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

---

## Samenvatting

---

Tijdens de vroege embryonale ontwikkeling van dieren specificeren morfogen-gradiënten vanaf het eencellige stadium verschillende structuren. Morfogenen verschaffen spatiële informatie door middel van de vorming van concentratiegradiënten die het ontwikkelende embryo in aparte domeinen verdeelt. De combinatie van verschillende morfogenen kan lokaal leiden tot de ontwikkeling van verschillende gespecialiseerde celtypes en weefselstructuren. Met behulp van dit algemene principe kunnen organismen zeer diverse en complexe lichaamsbouwplannen en zeer gespecialiseerde weefsels vormen.

Om een beter inzicht te krijgen in de vorming van het lichaamsbouwplan is kennis over de onderliggende biochemische processen essentieel, en meer specifiek op het niveau van genetische regulatie. Gen regulatoire netwerken (GRNs) omvatten een verzameling van genen met onderlinge interacties die op dynamische wijze het expressieniveau van elke gen aanstuurt. Kennis over de betrokken genen en de architectuur van het netwerk zijn vereist om inzicht te krijgen in het mechanisme dat leidt tot specifieke fenomenen zoals patroonvorming.

In sommige gevallen is het direct ophelderen van de aard van de interacties tussen de genen mogelijk met behulp van experimentele methoden. Een andere optie is het gebruik van wiskundige modellen en simulatietechnieken. Kwalitatieve en kwantitatieve modellen kunnen zowel inzicht geven in de aard van de netwerkinteracties als wel inzicht geven in de dynamische eigenschappen van het netwerk.

Het gebruik van kwantitatieve modellen maakt het mogelijk het patroonvormingsmechanisme, aangestuurd door het gen regulatoire netwerk verder te onderzoeken. Over het algemeen bevatten gen regulatoire netwerken die spatio-temporele patroonvorming beschrijven een groot aantal parameters zoals bi-

voorbeeld kinetische coëfficiënten van biochemische reacties, diffusie coëfficiënten, bindingseigenschappen van transcriptiefactoren, de geometrie van het systeem etc. Wanneer de beginwaarden en de parameters bekend zijn (bijvoorbeeld verkregen uit de literatuur, experimentele studies of andere bronnen) dan kunnen de vergelijkingen (partiële differentiaalvergelijken) die de patroonvorming beschrijven geformuleerd worden als een rechtstreeks beginwaardeprobleem. De partiële differentiaalvergelijkingen kunnen met numerieke methoden rechtstreeks opgelost worden. Vervolgens kan bijvoorbeeld met behulp van een gevoeligheidsanalyse de robuustheid van het model ten opzichte van de modelparameters worden onderzocht. In de praktijk is de exacte waarde van de meeste modelparameters jammer genoeg niet te verkrijgen met behulp van een directe meting van de parameters, het is dan nodig deze te schatten aan de hand van andere type experimentele data zoals bijvoorbeeld experimenteel gemeten expressiepatronen. Het wiskundige probleem kan dan geformuleerd worden als een invers beginwaardeprobleem. Bij deze aanpak zijn alleen de dynamische vergelijkingen die de patroonvorming beschrijven en een deel van de parameters bekend.

Inverse beginwaardeproblemen die continue spatio-temporele patroonvorming beschrijven vergen nauwkeurige kwantitatieve data, een voldoende correcte beschrijving van het systeem en een efficiënte methode om de onbekende parameters te schatten. Over het algemeen zijn dit soort inverse beginwaardeproblemen, waarbij de parameters gevonden worden met behulp van optimalisatietechnieken ("fitten") niet goed gedefinieerd. De oorzaak hiervan ligt deels in de kwaliteit van de experimentele data in combinatie met het grote aantal modelparameters. In het geval van een probleem met spatio-temporele data kan de optimalisatie zo rekenintensief zijn dat minder grote numerieke nauwkeurigheid nodig is om binnen redelijke tijd oplossingen te vinden. Het is daarom van cruciaal belang de juiste optimalisatietechniek te kiezen.

Hoewel de parameters die verkregen zijn met de optimalisatietechniek het experimenteel gemeten expressiepatroon kan reproduceren is het niet zeker dat de werkelijke onderliggende netwerkstructuur geïdentificeerd is. In sommige gevallen genereert de optimalisatietechniek een unieke oplossing van het netwerk, echter in de meeste gevallen worden meerdere verschillende oplossingen gevonden. Het is wenselijk op grond van eigenschappen anders dan de kwaliteit van de "fit" om de "beste" oplossingen te selecteren. Een systematische analyse van de gevonden netwerken is nodig om de validiteit verder te onderzoeken.

In dit proefschrift staan de resultaten beschreven van een uitgebreide analyse van gen regulatoire netwerken die in staat zijn spatio-temporele genexpressiepatronen te simuleren. De nadruk ligt op het schatten van modelparameters die geïnfereerd zijn met behulp van optimalisatietechnieken en daaropvolgend een uitgebreide analyse van de verkregen oplossingen. Als concreet biologisch onderzoeksobject is gekozen voor de vroege embryonale

ontwikkeling van *Drosophila melanogaster*.

Als eerste presenteren we een hybride  $(\mu, \lambda)$ -evolutionaire strategie voor het schatten van de modelparameters, dit is een efficiënte methode die gebaseerd is op evolutionaire algoritmes. Vergeleken met andere methodes zoals "simulated annealing" die eerder gebruikt zijn voor dit systeem is dit algoritme veel sneller, waardoor het mogelijk is een veel groter aantal netwerken die het expressiepatroon goed reproduceren te verkrijgen in dezelfde rekentijd. Vervolgens is de gevoeligheid, robuustheid, uniekheid en stabiliteit van alle netwerken systematisch onderzocht en vergeleken. We laten zien dat de variatie in de gesimuleerde expressieniveaus geïdentificeerd kan worden met clustertechnieken en dat deze variatie deels verband houdt met variatie in stabiliteit van de netwerken. De gevoeligheidsanalyse maakt het mogelijk oplossingen te identificeren die relatief gezien robuust zijn voor parametervariatie. Een andere maat voor robuustheid is de gevoeligheid van het netwerk voor fluctuaties in genexpressie door stochasticiteit tijdens de patroonvorming. Uit beide analyses blijkt dat er een duidelijk verband bestaat tussen gevoeligheid voor parametervariatie en gevoeligheid voor stochastische fluctuaties in genexpressieniveaus. De eindconclusie van dit onderzoek is dat bij het infereren van gen regulatoire netwerken met behulp van een inverteerd beginwaardeprobleem niet alleen gefocuseerd zou moeten worden op de kwaliteit van de fit maar ook op andere modeleigenschappen. De resultaten suggereren dat een optimalisatietechniek met meerdere optimalisatiecriteria betere resultaten kan geven.