Stochastic Modeling of Gene Transcription in a Single Cell

SEA 2012

David Knox
Computational Bioscience Program
University of Colorado Anschutz Medical Campus
Your World
My World
Computational Biology

Adapted from Dr. Alan Qi
 Dept of Computer Science and Department of Statistics
 Purdue University
 SEA Conference February 2011
Modeling takes everything we know

- Population Studies
- Whole Cell Simulations
- Heart/Kidney Simulations
- Signaling Pathways
- Protein Folding
- Transcriptional Regulation
- Molecular Dynamics
Stochastic Modeling of Gene Transcription in a Single Cell

- What is Transcription Regulation
- Modeling Paradigm
- Stochastic Rule Builder (SRB)
- Visualization of Simulation Results
Stochastic Modeling of Gene Transcription in a Single Cell

• What is Transcription Regulation
• Modeling Paradigm
• Stochastic Rule Builder (SRB)
• Visualization of Simulation Results
Transcriptional Regulation
Transcriptional Diversity and Regulation
Modeling Transcriptional Regulation: The Players

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Component</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RNA Polymerase II, Transcription Machinery (TM)</td>
</tr>
</tbody>
</table>
Requirements and benefits of modeling methods. Adapted from “Modelling and analysis of gene regulatory networks” Guy Karlebach and Ron Shamir, 2008.

All models are wrong; some are useful.  
-- George Box
Modeling with the Stochastic Simulation Algorithm

• The SSA is a Monte Carlo simulation of the master chemical equation

• Every interaction between components must be explicitly defined

Reactant + Reactant ... → Resultant + ... @ RATE in STEPS
SSA Modeling of Transcription Factors
Activating / Repressing Transcription

\[
\begin{align*}
Tf_a + \text{DNA}[7] & \rightarrow \ TF_a_{-}\text{DNA}[7] @ \text{on\_rate}_a \text{ in } S_a \text{ steps} \\
\text{TF}_a_{-}\text{DNA}[7] & \rightarrow \ TF_a + \text{DNA}[7] @ \text{off\_rate}_a \\
\text{TF}_a_{-}\text{DNA}[7] + \text{TM} & \rightarrow \ \text{TM\_DNA}[7] + \text{TF}_a @ \text{on\_rate}_c \text{ in } S_c \text{ steps} \\
\text{TM} + \text{DNA}[90] & \rightarrow \ \text{TM\_DNA}[90] @ \text{rate} \text{ in } S \text{ steps} \\
\text{TM\_DNA}[90] + \text{DNA}[91] & \rightarrow \ \text{TM\_DNA}[91] + \text{DNA}[90] @ \text{rate} \text{ in } S \text{ steps} \\
\text{TM\_DNA}[90] + \text{DNA}[89] & \rightarrow \ \text{TM\_DNA}[89] + \text{DNA}[90] @ \text{rate} \text{ in } S \text{ steps}
\end{align*}
\]
Stochastic Modeling of Gene Transcription in a Single Cell

- What is Transcription Regulation
- **Modeling Paradigm**
- Stochastic Rule Builder (SRB)
- Visualization of Simulation Results
State Diagram for one DNA Position

- Free<sub>i-2</sub>
- Free<sub>i-1</sub>
- Free<sub>i</sub>
- Free<sub>i+1</sub>
- Free<sub>i+2</sub>

- TF bound
- Nucleosome bound
- TM bound

Bioinformatics Supergroup - February 2012
State Diagram for TM
Advancing one DNA Position
Stochastic Modeling of Gene Transcription in a Single Cell

• What is Transcription Regulation
• Modeling Paradigm
• Stochastic Rule Builder (SRB)
• Visualization of Simulation Results
Model System Overview

- State Diagrams
- DNA sequence
- TF Motif, Strength
- Nucleosome Prob.
- Transcriptional Machinery Rates
  - initiation,
  - elongation,
  - termination
- Component Concentrations

Stochastic Rule Builder (SRB) → Rules & Reactants → Simulator DIZZY

Sim Results → Reactant counts at each iteration

SRB Visualizer

- State Diagrams
- DNA sequence
- TF Motif, Strength
- Nucleosome Prob.
- Transcriptional Machinery Rates
  - initiation,
  - elongation,
  - termination
- Component Concentrations

Bioinformatics Supergroup - February 2012
### Biological Detail vs Computer Resources

**Sizes of SRB Generated Models**

<table>
<thead>
<tr>
<th>Nucleotides per group</th>
<th># Reactants</th>
<th># Rules</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>54366</td>
<td>405326</td>
</tr>
<tr>
<td>2</td>
<td>27361</td>
<td>127106</td>
</tr>
<tr>
<td>3</td>
<td>18361</td>
<td>68205</td>
</tr>
<tr>
<td>6</td>
<td>9361</td>
<td>26206</td>
</tr>
<tr>
<td>10</td>
<td>5758</td>
<td>14127</td>
</tr>
</tbody>
</table>

**IME4 (2700 nt)**

#### Using 30 nucleotides per group
- My manual creation, ~2000 rules, ~1000 reactants
- SRB produces ~4700 rules, ~2150 reactants
Time to Simulate Models using DIZZY

<table>
<thead>
<tr>
<th>Group Size</th>
<th># Iterations</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>6</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td></td>
<td>---</td>
<td>142</td>
<td>91</td>
<td>31</td>
<td>17</td>
</tr>
<tr>
<td>2000</td>
<td></td>
<td>---</td>
<td>227</td>
<td>173</td>
<td>44</td>
<td>35</td>
</tr>
<tr>
<td>4000</td>
<td></td>
<td>---</td>
<td>635</td>
<td>271</td>
<td>103</td>
<td>66</td>
</tr>
<tr>
<td>8000</td>
<td></td>
<td>---</td>
<td>---</td>
<td>725</td>
<td>208</td>
<td>174</td>
</tr>
<tr>
<td>16000</td>
<td></td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>611</td>
<td>484</td>
</tr>
<tr>
<td>32000</td>
<td></td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>798</td>
</tr>
</tbody>
</table>

Dizzy Stochastic Simulator
Stochastic Modeling of Gene Transcription in a Single Cell

• What is Transcription Regulation
• Modeling Paradigm
• Stochastic Rule Builder (SRB)
• Visualization of Simulation Results
Visualizing the Process

<table>
<thead>
<tr>
<th>State</th>
<th>Character</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initiation</td>
<td>0&gt;&gt;,1&gt;&gt;</td>
</tr>
<tr>
<td>Transcribing</td>
<td>==}</td>
</tr>
<tr>
<td>Transcribed</td>
<td>==}</td>
</tr>
</tbody>
</table>

+ time

Position on DNA

State Character
Initiation 0>>,1>>
Transcribing ==}
Transcribed ==}
Visualizing Transcriptional Interference

<table>
<thead>
<tr>
<th>State</th>
<th>Character</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initiation</td>
<td>0&gt;&gt;,1&gt;&gt;</td>
</tr>
<tr>
<td>Transcribing</td>
<td>==}</td>
</tr>
<tr>
<td>Transcribed</td>
<td>==] [==</td>
</tr>
</tbody>
</table>

Position on DNA

Gene

Bioinformatics Supergroup - February 2012
Visualizing Nucleosomes and Transcription Factors

<table>
<thead>
<tr>
<th>State</th>
<th>Character</th>
<th>0&gt;&gt;,1&gt;&gt;</th>
<th>&lt;&lt;0,&lt;&lt;1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initiation</td>
<td>0&gt;&gt;,1&gt;&gt;</td>
<td>&lt;&lt;0,&lt;&lt;1</td>
<td></td>
</tr>
<tr>
<td>Transcribing</td>
<td>==}</td>
<td>{==</td>
<td></td>
</tr>
<tr>
<td>Transcribed</td>
<td>==]</td>
<td>[==</td>
<td></td>
</tr>
<tr>
<td>Stable Nucleosome</td>
<td>(...)n...)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Binding Nucleosome</td>
<td>(---b---)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unbinding Nucleosome</td>
<td>(<em><strong>u</strong></em>)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Position on DNA
Summary

- State Diagrams
- DNA sequence
- TF Motif, Strength
- Nucleosome Prob.
- Transcriptional
  Machinery Rates
  initiation,
elongation,
termination
- Component
  Concentrations

- New Data
- Updated Data
Future Work

- Parameter Search
- More Detailed Components and Complex Interactions
- Parallelize / Distributed
- Visualization via Animation
Acknowledgements

Computational Bioscience Program
University of Colorado Anschutz Medical Campus

Dr. Robin Dowell – MCDB Univ. of Colorado Boulder

- Alex Poole
- Jess Vera
- Dr. Mary Allen
- Tim Read
- Phil Richmond
- JaeAnn Dwulet
- Amber Sorenson
- Cynthia Page
- Joe Rokicki
- Li Wang
GOD... THE HUMAN GENOME CODE'S BEEN UNRAVELLED

DAMN HACKERS!!! NOW, I HAVE TO CHANGE THE PASSWORD