Toward a Whole-Cell Model of *Mycoplasma genitalium*
Markus Covert
Stanford Bioengineering

The ultimate test of understanding a simple cell, more than being able to build one, would be to build a computer model of the cell, because that really requires understanding at a deeper level.
— Clyde Hutchison
The New York Times, 1999

With a whole cell model, we could…

**Interpret** large-scale datasets
**Predict** complex behaviors
**Design** novel organisms rationally
**Reduce** development time and cost

**E. coli** (FBA): 600 metabolic genes
**E. coli**: 127 genes
**M. pneumoniae**

<table>
<thead>
<tr>
<th>Year</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>1999</td>
<td>E. coli (FBA) 600 metabolic genes</td>
</tr>
<tr>
<td>2001</td>
<td>Expanded “minimal cell” model</td>
</tr>
<tr>
<td>2003</td>
<td>M. pneumoniae “tour de force”</td>
</tr>
<tr>
<td>2007</td>
<td>BioCyc: 160 genomes</td>
</tr>
<tr>
<td>2009</td>
<td>H. salinarium regulatory model</td>
</tr>
</tbody>
</table>
Question 1: Can we build a whole-cell model today?

Lessons of the past decade

Whole-cell modeling requires a **variety** of approaches

Integration of diverse approaches will be essential
Implementation

Simulation Integration Simulation

Metabolite concentrations Protein monomers Complexes
Polymerase positions Ribosome positions RNAs
Metabolic fluxes Cell mass...

Mycoplasma genitalium

Small number of genes
Annotated genome sequence
Homology to model bacteria
Free-living for culture
**Project status (8/2010)**

**M. genitalium**: 525 genes

- **Implemented**: 358 (89%)
- **Remaining**: 43 (11%)
- **Annotated**: 401 (76%)
- **Unannotated**: 124 (24%)

**Model goal**: 401 genes

**Metabolism**: 147
**Translation**: 70
**tRNA aminoacylation**: 60
**RNA modification**: 13
**DNA replication**: 12
**Protein decay**: 9
**Protein translocation**: 9
**Attachment**: 9
**Transcription**: 8
**Protein folding**: 6
**Ribosome assembly**: 4
**Protein processing**: 4
**Protein modification**: 3
**DNA repair**: 2
**DNA damage/repair**: 18
**Cytokinesis**: 8
**RNA processing**: 6
**Transcription regulation**: 6
**DNA supercoiling**: 5

**92.4% of KO simulations qualitatively correct**

**Question 2**: Can we predict the outcome of perturbations?
DNA primase-related protein
FAD-dependent glycerol-3-phosphate dehydrogenase
Transketolase
Ribulose-phosphate 3-epimerase
Dihydroxyacetone dehydrogenase
Chromosome segregation protein
DNA polymerase III, delta subunit, putative
ATP synthase F1, epsilon subunit
Uracil phosphoribosyltransferase
Glycerol kinase
Purine nucleoside phosphorylase
Deoxyribose-phosphate aldolase
ABC transporter subunit (ribose?)
ABC transporter subunit (ribose?)
Modification methylase, HemK family
Potassium uptake protein, TrkA family
isoleucyl-tRNA synthetase
DAK2 phosphatase domain protein
Thiamine biosynthesis tRNA modification protein Thil

MG_030
MG_038
MG_049
MG_050
MG_119
MG_120
MG_259
MG_323
MG_345
MG_369
MG_372

MG_471 knockout simulation

Total protein production in two single-cell M. genitalium simulations

WT
MG_471 KO
Question 3: Can we interpret complex phenotypes?

M. genitalium
M. capricolum

Flux-balance analysis
Whole-cell model

Cai et al., Nature 2008
Can we build a whole cell model today?

Can we predict the outcome of perturbations?

Can we interpret complex phenotypes?

Acknowledgements

Covert Lab
Whole-cell modeling
Jonathan Karr
Jayodita Sanghvi

NF-κB single cell dynamics
Timothy Lee
Jake Hughey

Virus-pathogen interactions
Nate Maynard
Elsa Birch

Support
NIH Director’s Pioneer Award
NIH Pathway to Independence Award

covertlab.stanford.edu