

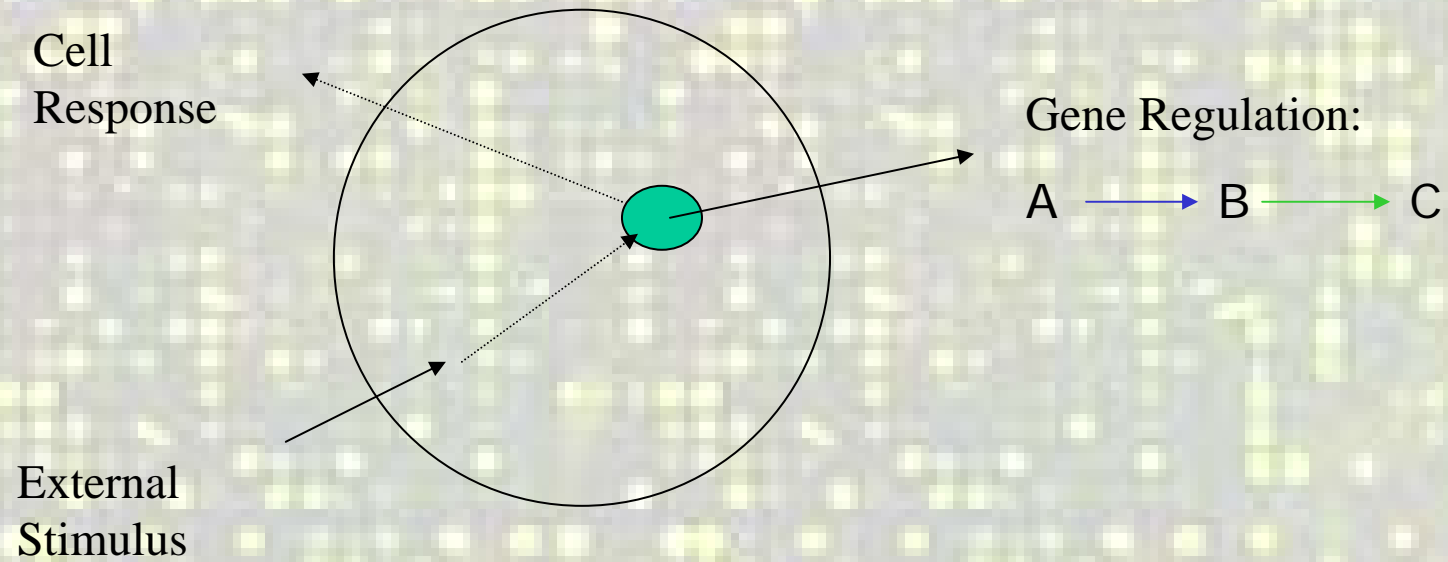
# Inference of Gene Relations from Microarray Data by Abduction

Irene Papatheodorou & Marek Sergot  
Imperial College

# Outline

- Gene Regulation & Microarrays
- Abductive Reasoning
- Model of Gene Interactions
- Applications & Tests
- Evaluation & Further Work

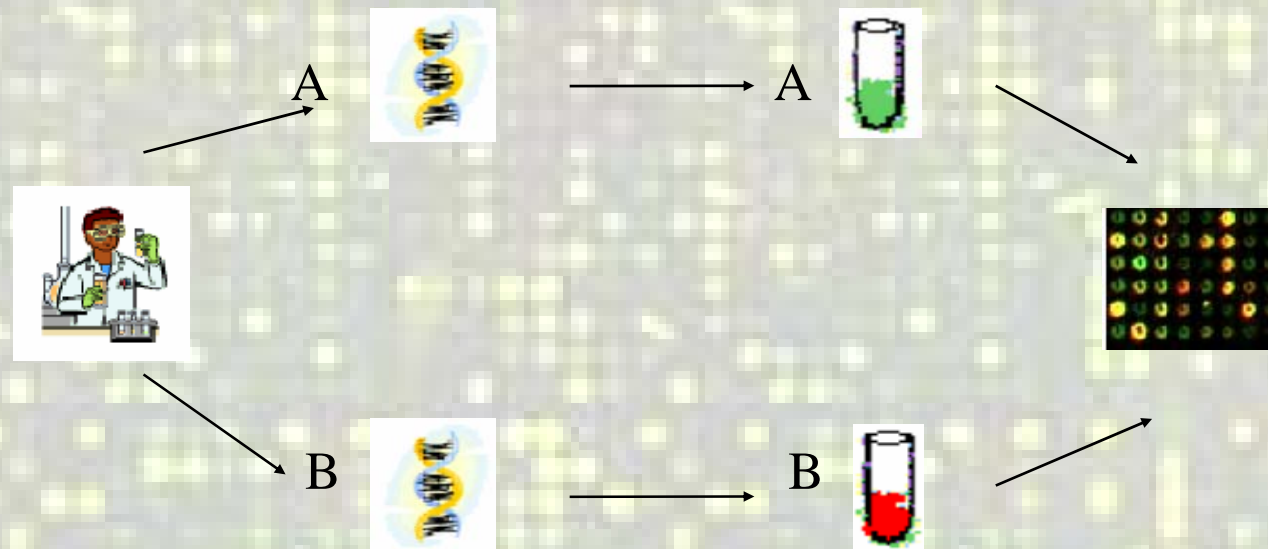
# Gene Regulation



Microarrays measure gene expression

# Microarray Experiment

Measures levels of gene expression



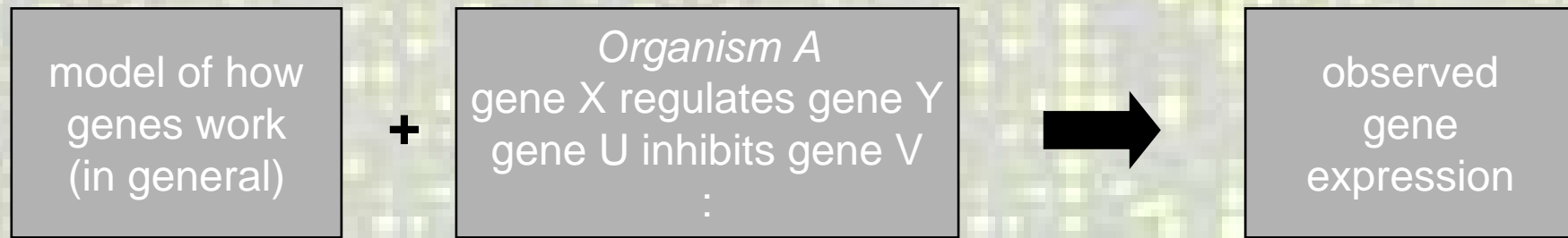
Experiment: gene mutation/environmental stress



# Expression Data to Gene Relations

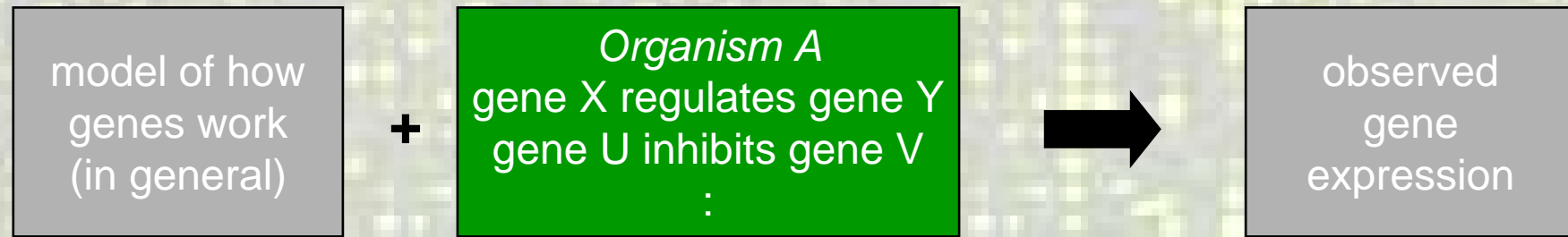
- *Mycobacterium tuberculosis* experiments from CMMI
- Genomic Information-Background Knowledge
- Gene Relations- inhibition/activation
- Inference Method: Abduction

# Deduction



Infer the effect from rules

# Abduction



Inference from effect to cause



# Abductive Inference

- Theory represented by  $(P, A, IC)$ 
  - $P$  is a logic program
  - $A$  is a set of abducible predicates
  - $IC$  Integrity Constraints, logic rules
- Abductive Procedure: Kakas-Mancarella (General Purpose)
- Implementation: Alpha (R. Craven)

# Gene Interaction Model

- Rules & Integrity Constraints of Gene Interactions

- Observables:

*increases\_expression(Expt, Gene)*

*reduces\_expression(Expt, Gene)*

- Abducibles:

*induces(GeneA, GeneB)*

*inhibits(GeneA, GeneB)*

# The Rules (Summary)



\*Unless GENE 2 affected by another gene or  
GENE 2 affected by environmental stress

Recursive rules

2 Parameters

# The Model

- Concept of gene interaction

*increases\_expression(Expt, X) ←*  
*knocks\_out(Expt, G),*  
*inhibits(G, X).*

# The Model: Exceptions

- Top-level: Base case rule

*increases\_expression(Expt, X) ←*  
*knocks\_out(Expt, G),*  
*inhibits(G, X),*  
*not\_affected\_by\_other\_gene(Expt, G, X),*  
*not\_affected\_by\_EnvFactor(Expt, X).*

# Rules of Gene Interaction

- Top-level recursive rule:  
 $\text{increases\_expression}(Expt, X) \leftarrow$   
 $\text{knocks\_out}(Expt, G),$   
 $\text{candidate\_gene}(Expt, G1, G),$   
 $\text{reduces\_expression}(Expt, G1),$   
 $\text{inhibits}(G1, X),$   
 $\text{not\_affected\_by\_EnvFactor}(Expt, X).$

**Parameter: candidate\_gene/3**

# Rules of Gene Interaction

*affected\_by\_other\_gene(Expt, G, X) ←*  
*increases\_expression(Expt, Gx),*  
*Gx ≠ X, Gx ≠ G,*  
*related\_genes(Gx, G),*  
*induces(Gx, X).*

**Parameter: related\_genes/2**

# The Parameters

- “Related Genes” & “Intermediate Genes”
- Focus search on different sets of genes
- Transcription factors
- Similar Function



# Integrity Constraints

- Self-consistency:  
*False: induces(G1,G2), inhibits(G1,G2).*
- Consistency with prior knowledge:  
*False: induces(a,G).*  
*False: induces(G1,X), induces(G2,X),*  
*same\_operon(G1,G2).*
- Experimental Consistency:  
*False: candidate\_gene(E,G1,G2),*  
*mutates(E,G2), not\_affects(E,G1).*

# M.tuberculosis: 1 Observation

- Observation:

*increases\_expression(hspR, 'Rv0350')*

- Hypothesis:

*Hyp = [inhibits('Rv0353', 'Rv0350')]*

- 'Rv0353' is mutated in hspR
- 'Rv0350' is not affected by Environmental Factor
- 'Rv0350' is not affected by other gene

# M.tuberculosis: 2 Hypotheses

- Observation:

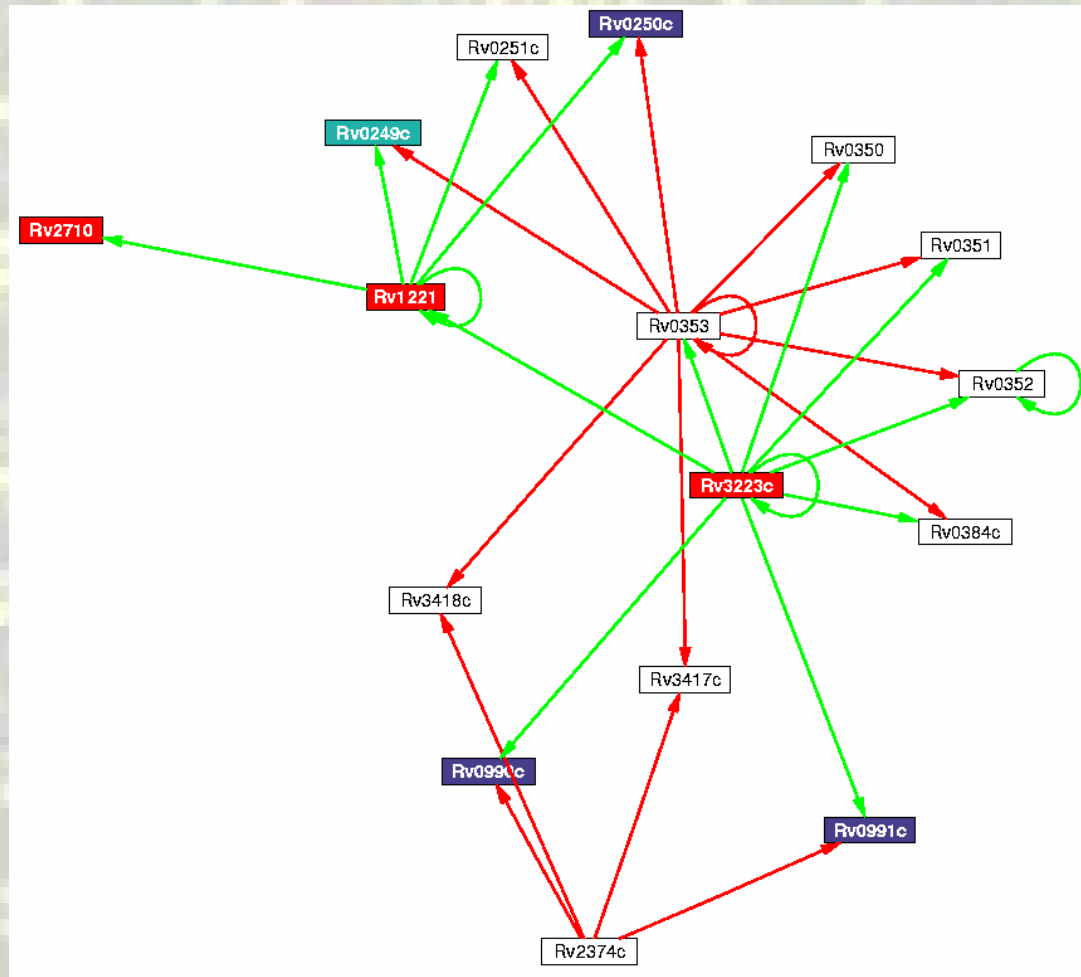
*reduces\_expression(sigH, 'Rv2710')*

- Hypotheses:

*Hyp = [induces('Rv3223c', 'Rv2710')]*

*Hyp = [induces('Rv3223c', 'Rv1221'),  
induces('Rv1221', 'Rv2710')]*

# M.tuberculosis: Regulators



# Evaluation

- General Method for Microarray Analysis
- Simple and Flexible Model
- Enables comparison of experiments
- Reduces Time of Analysis

# Future Work Directions

- Integrate output with pathway information
- Investigate different methods of formulating the problem
- Improve Performance of Abductive Interpreters

# Summary

- Gene Regulation & Microarrays
- Visualising Experiments
- Abductive Model for Gene Interactions
- Applications
- Future Work

# Questions

Irene Papatheodorou

[ivp@doc.ic.ac.uk](mailto:ivp@doc.ic.ac.uk)