
Learning genetic networks from gene expression data

Dr. Dirk Husmeier

Biomathematics & Statistics Scotland (BioSS)
JCMB, The King's Buildings, Edinburgh EH9 3JZ
United Kingdom

<http://www.bioss.ac.uk/~dirk>

Paradigm Shift in Molecular Biology

Paradigm Shift in Molecular Biology

Pre-Genomic

- Reductionist (DNA or RNA or protein)
- Generally qualitative, non-numeric
- Hypothesis driven

Paradigm Shift in Molecular Biology

Pre-Genomic

- Reductionist (DNA or RNA or protein)
- Generally qualitative, non-numeric
- Hypothesis driven

Post-Genomic

- Holistic, systems approach: DNA and RNA and protein
- Quantitative, highly numeric
- Data driven

Paradigm Shift in Molecular Biology

Pre-Genomic

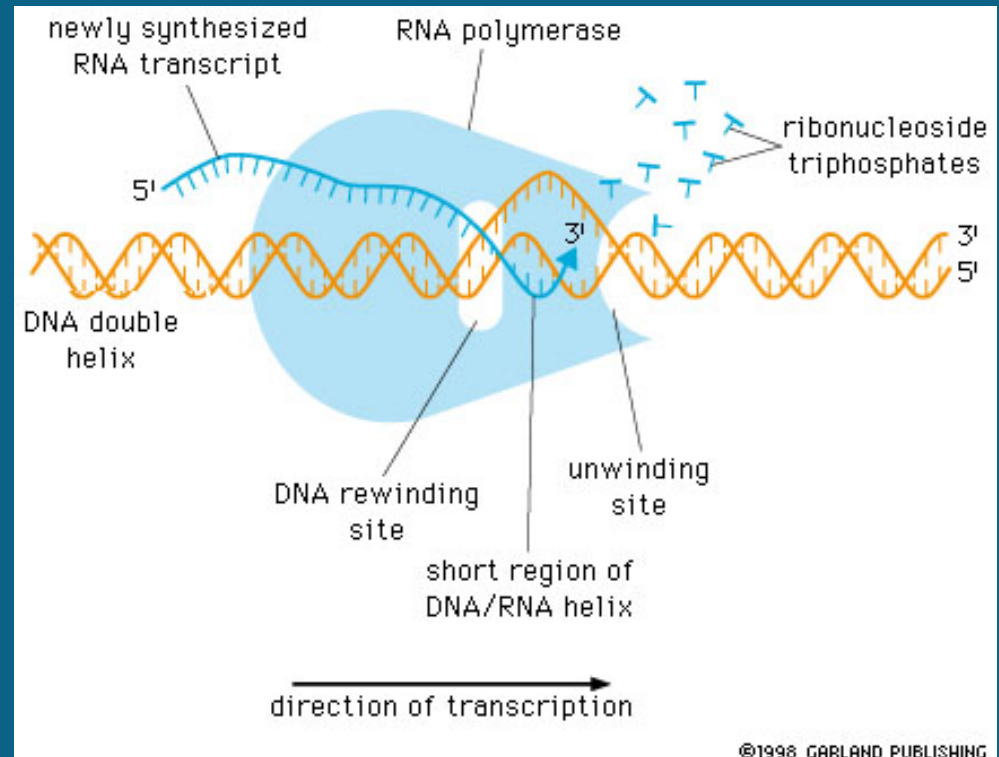
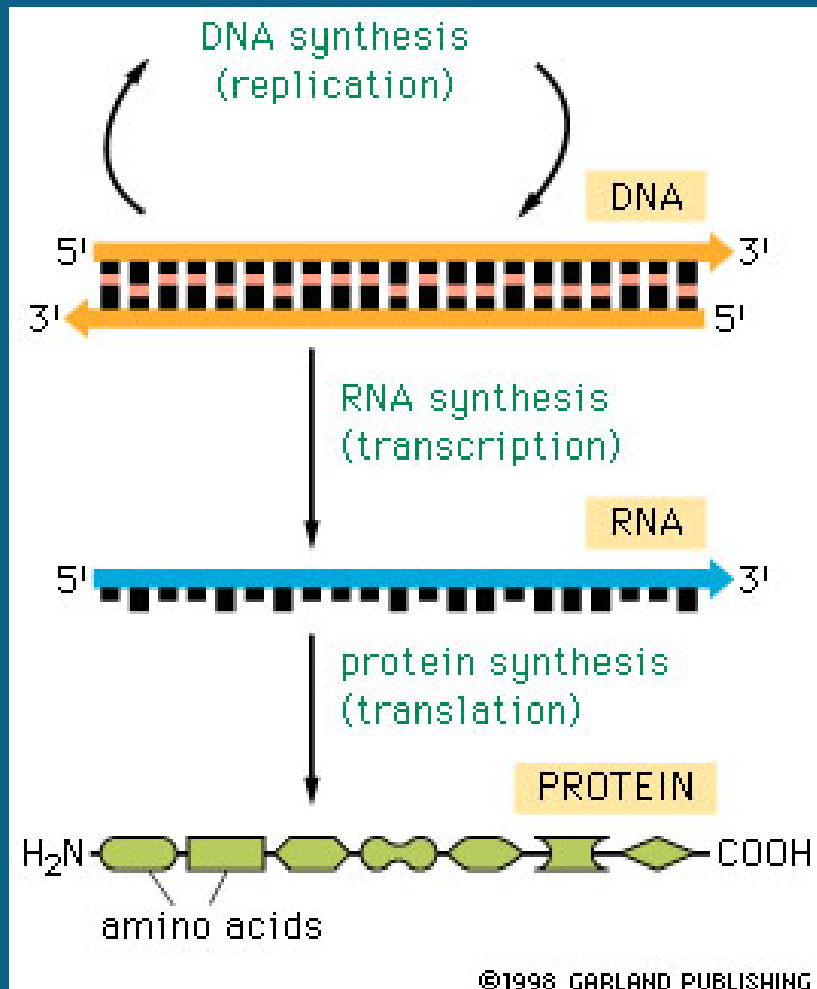
- Reductionist (DNA or RNA or protein)
- Generally qualitative, non-numeric
- Hypothesis driven

Post-Genomic

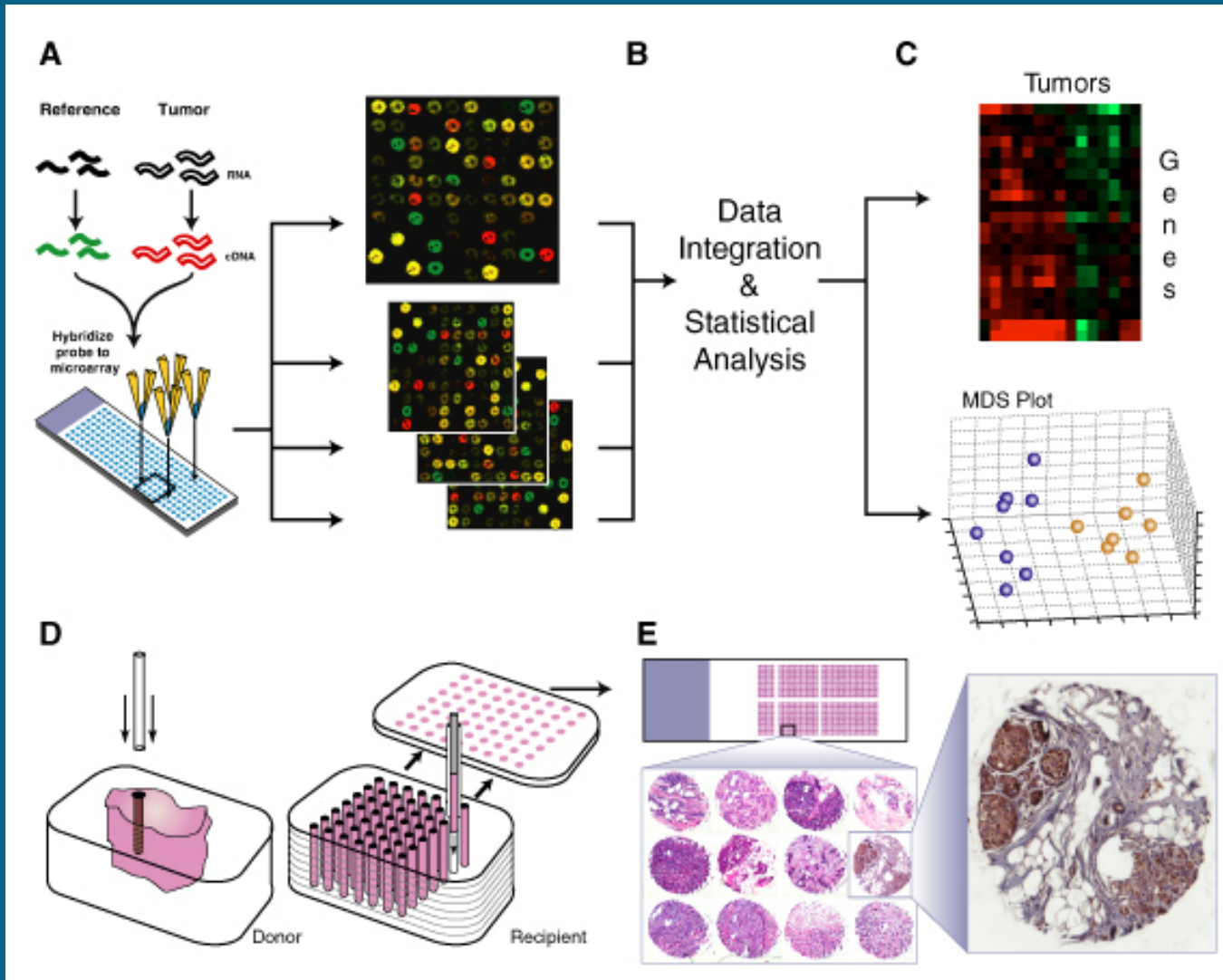
- Holistic, systems approach: DNA and RNA and protein
- Quantitative, highly numeric
- Data driven

⇒ Need for **machine learning** and **statistics**

Inferring genetic networks
from
microarray gene expression data



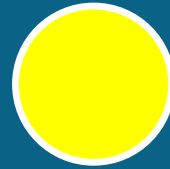
From <http://www.csu.edu.au/faculty/health/biomed/subjects/molbol/images>



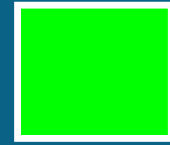
From http://www.nhgri.nih.gov/DIR/Microarray/NEJM_Supplement/



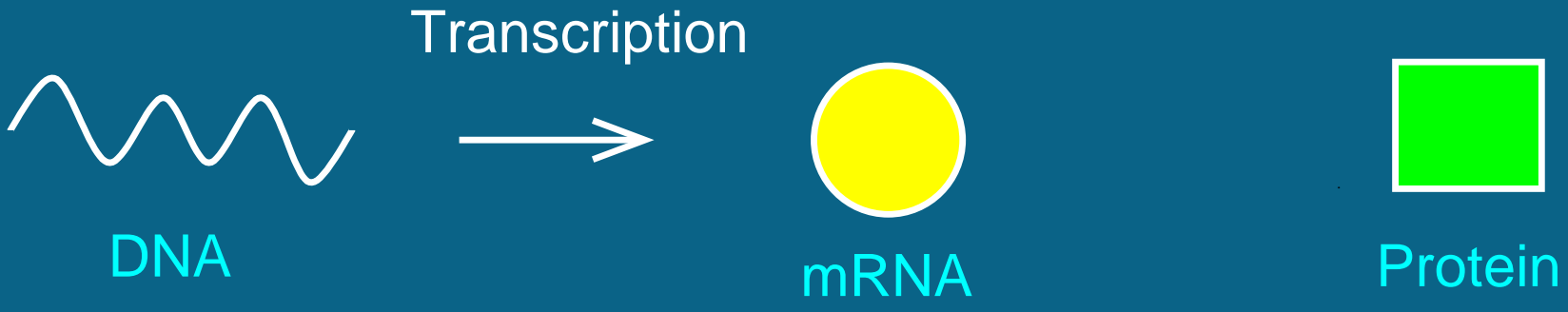
DNA

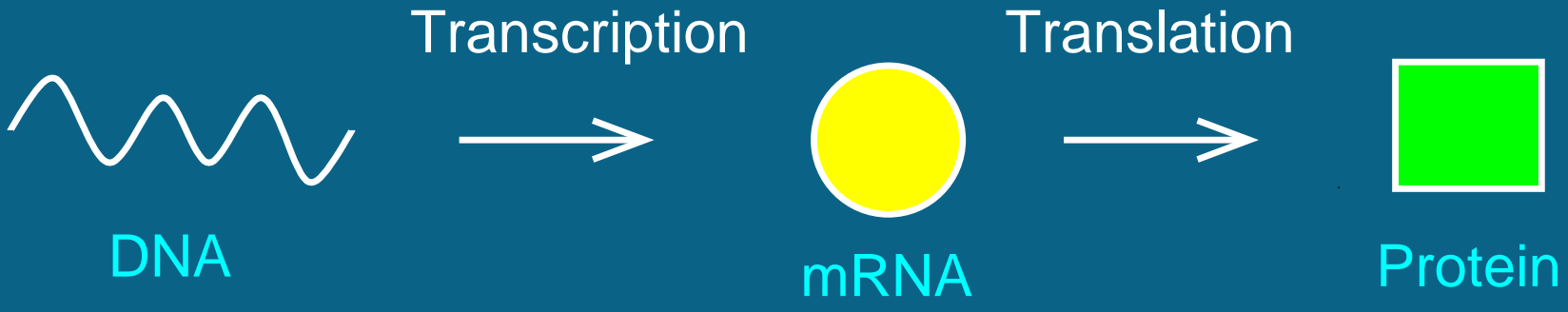


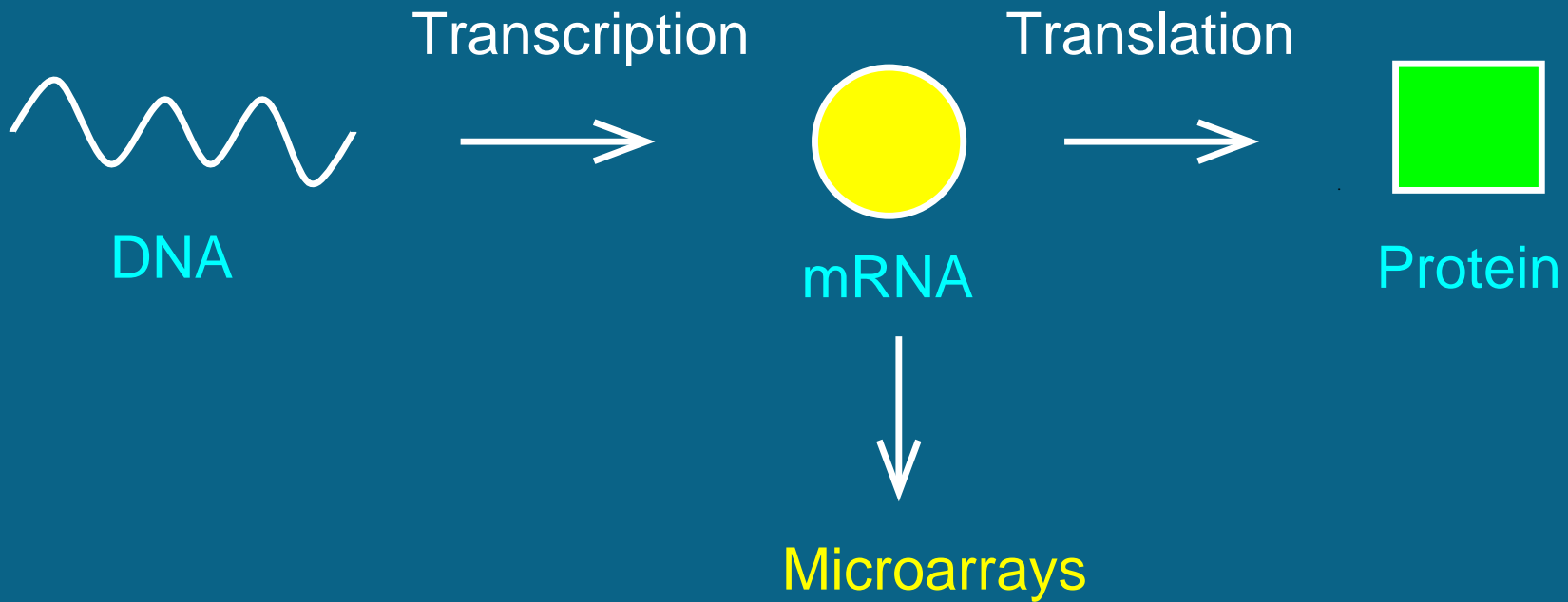
mRNA

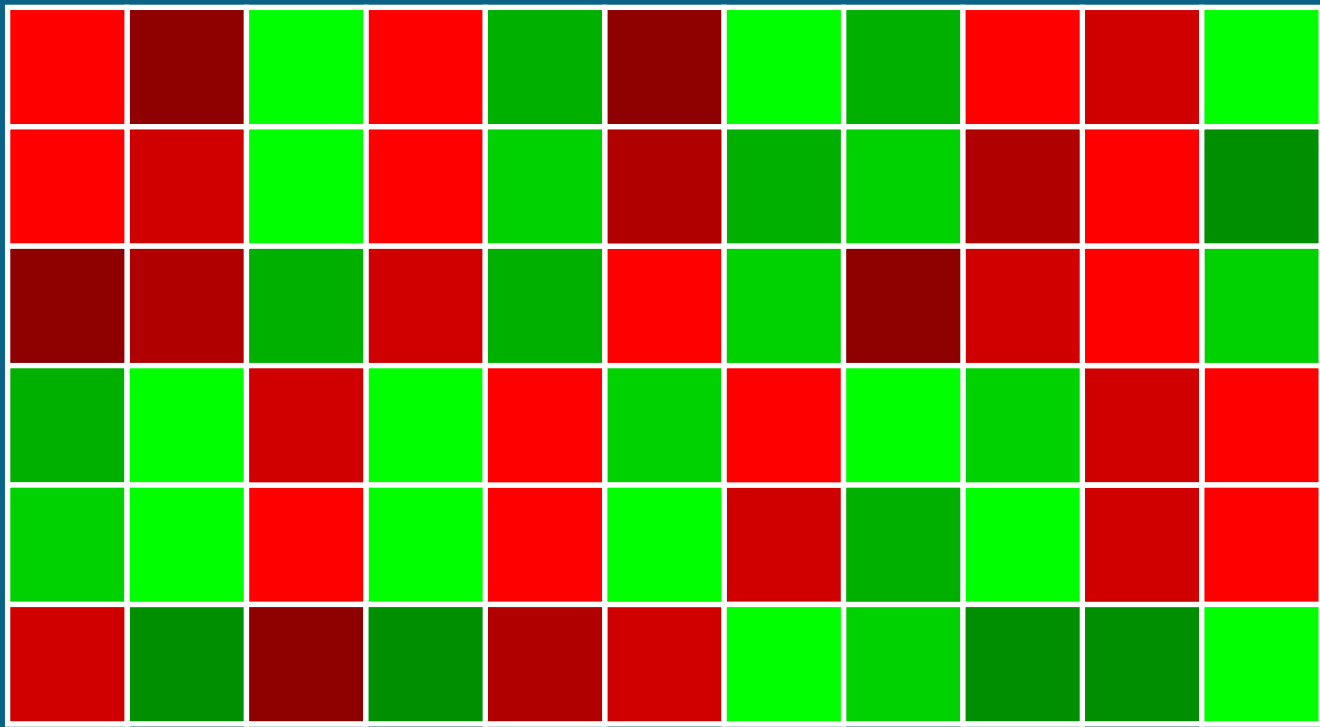


Protein



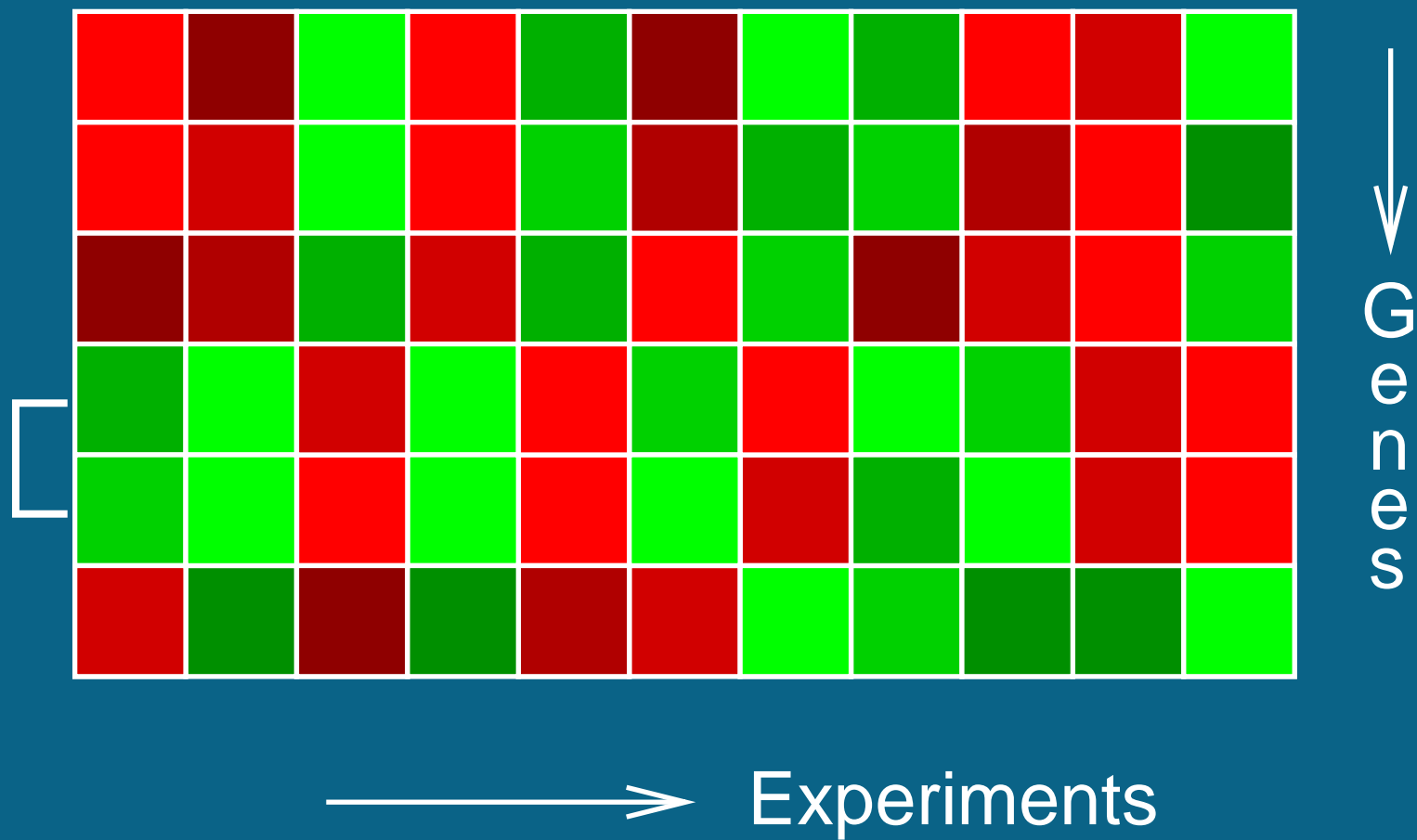


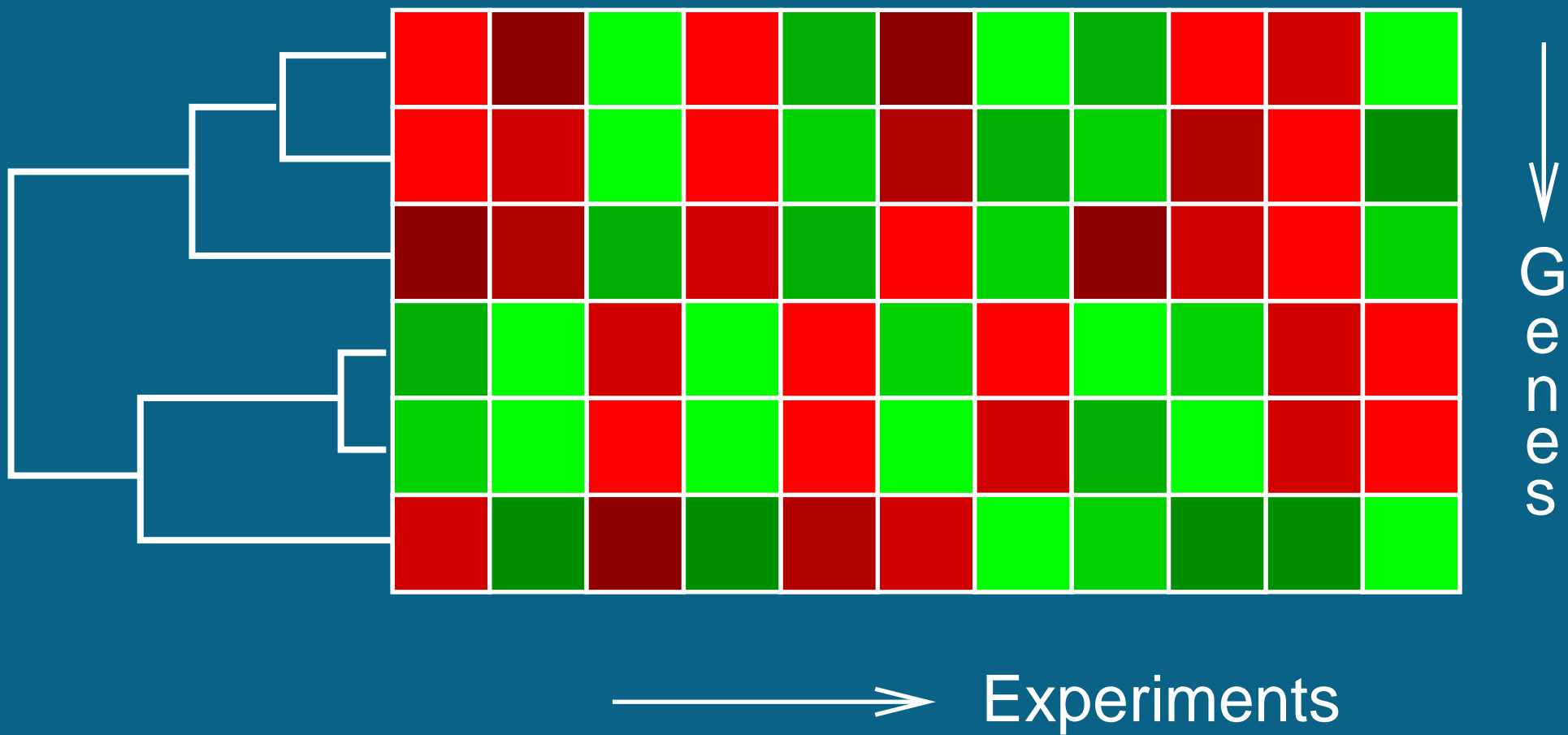


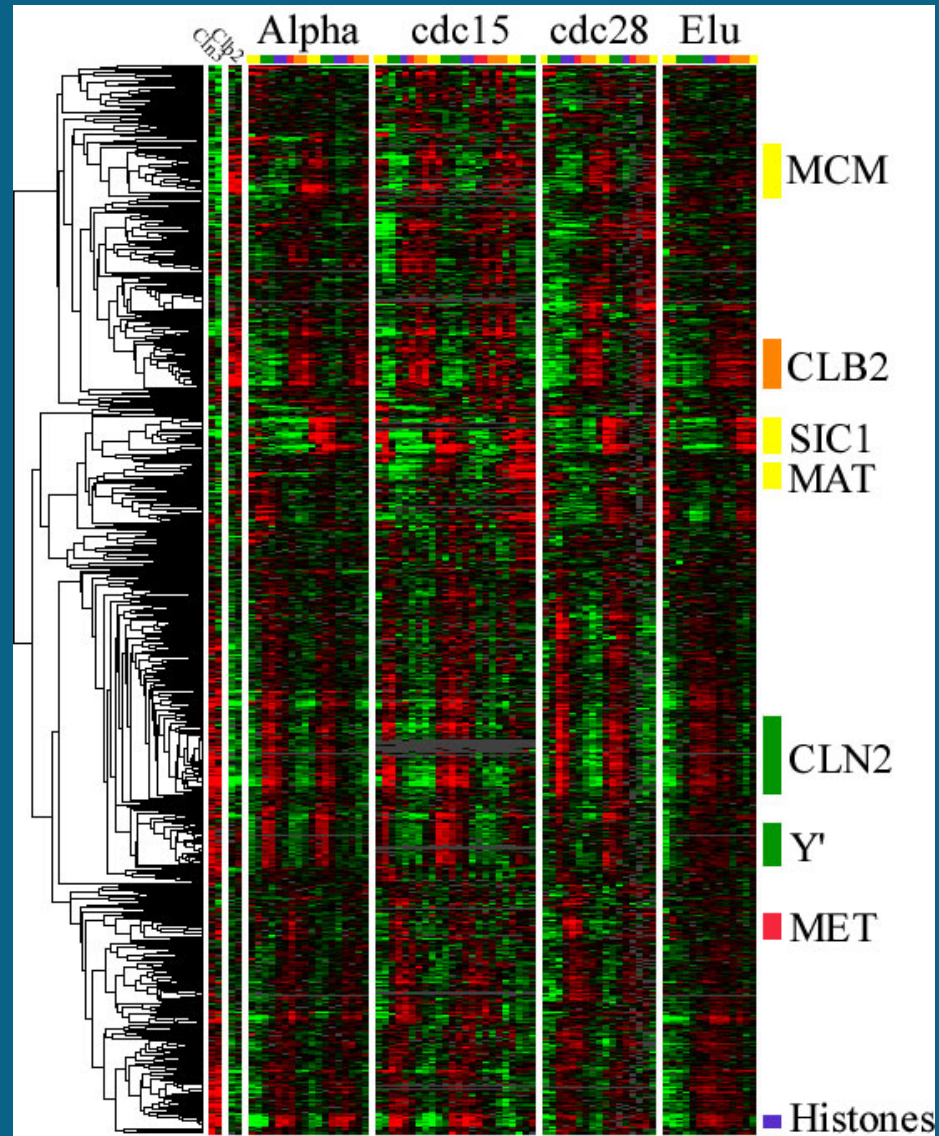


↓
Genes

→ Experiments







From Spellman et al., <http://cellcycle-www.stanford.edu/>

Advantage of clustering

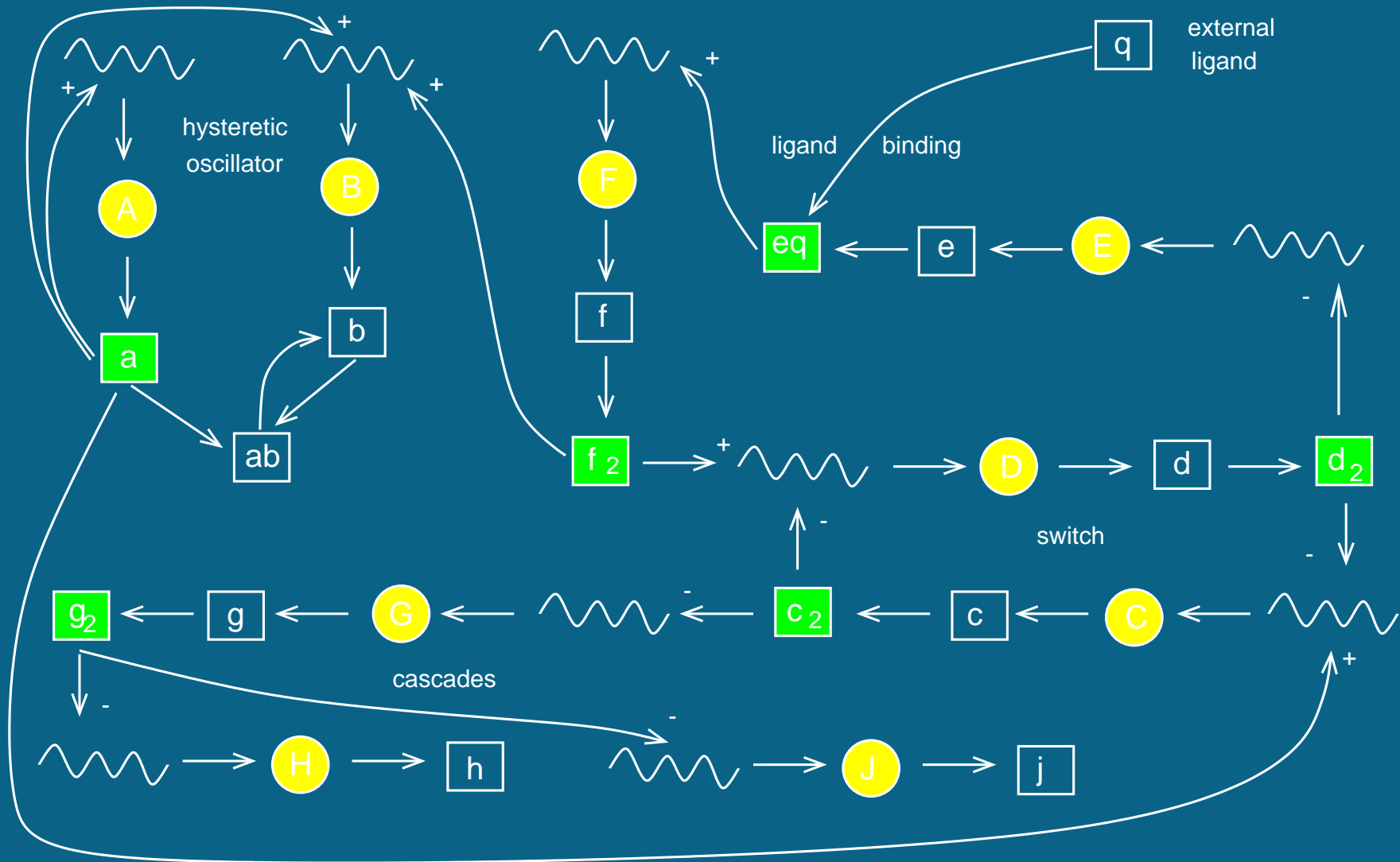
Fast, computationally cheap

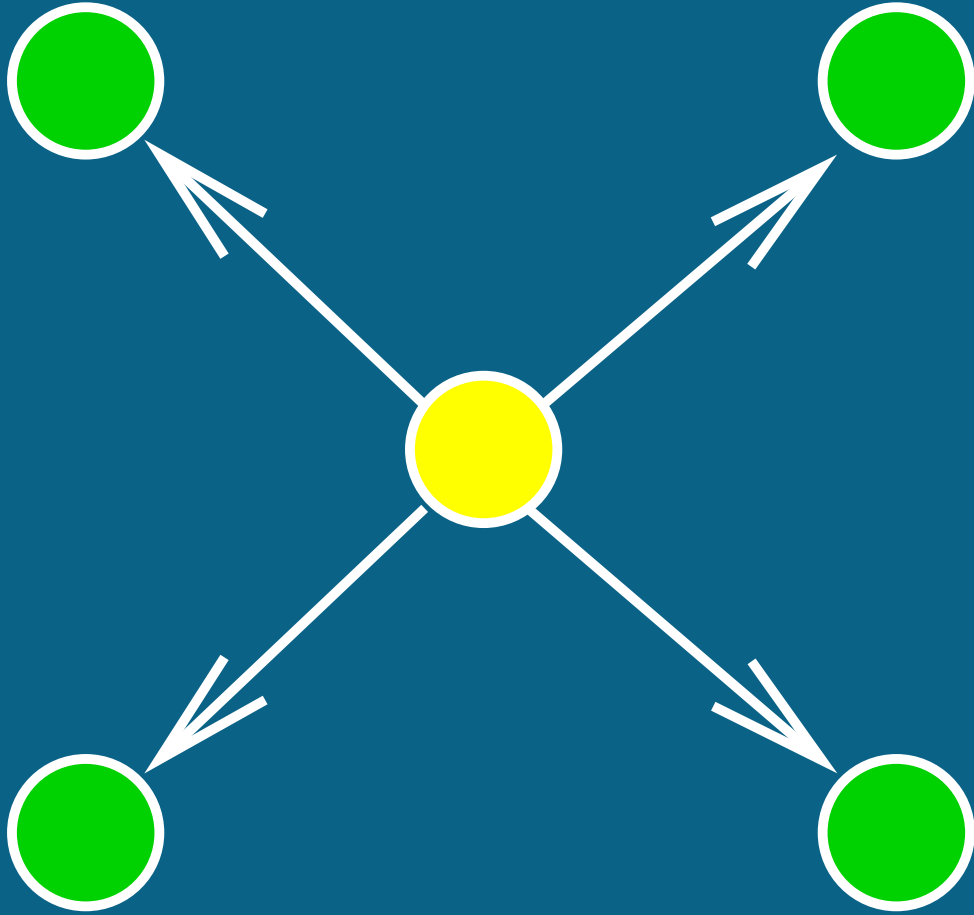
Advantage of clustering

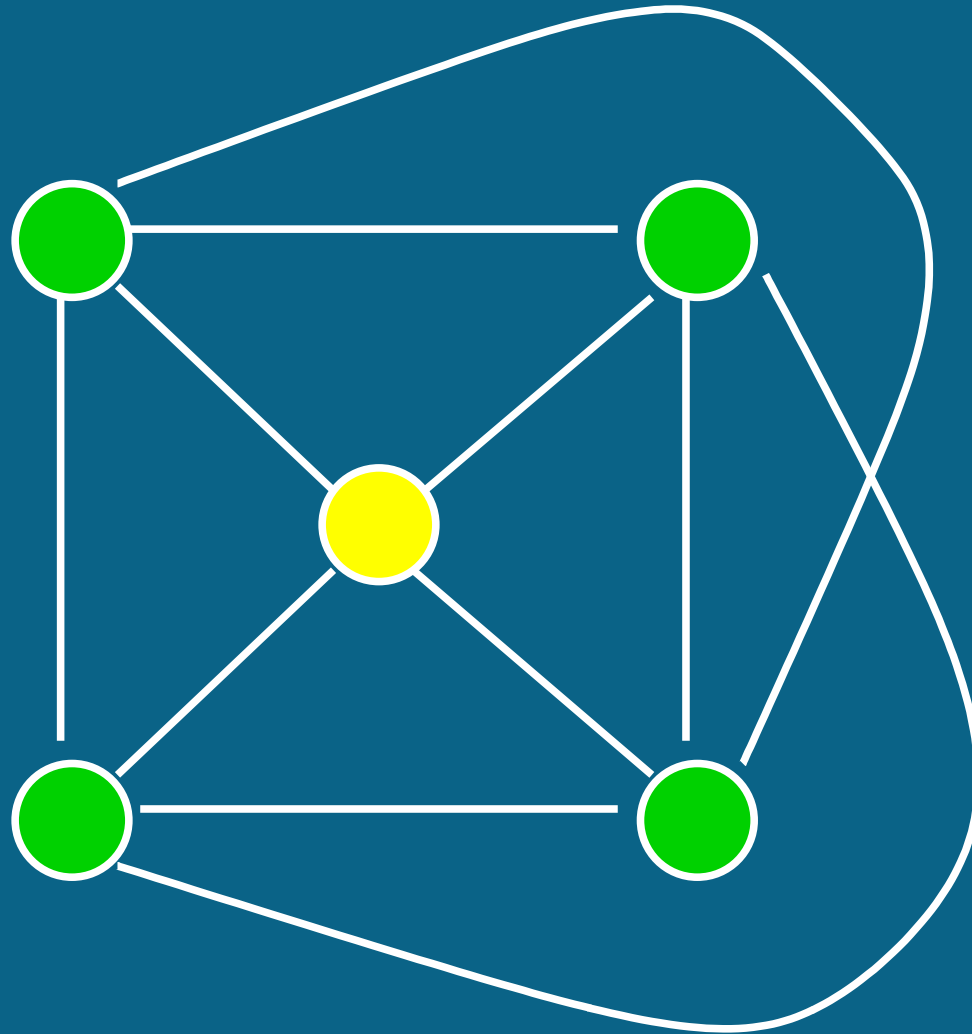
Fast, computationally cheap

Shortcoming of clustering

It is **NOT** reverse engineering







Reverse engineering

Learn the network structure from gene expression data.

Problem: Noise, sparse data

Bayesian networks

Probabilistic framework for
robust inference of interactions
in the presence of noise

Nir Friedman et al. (2000)
Journal of Computational Biology 7: 601-620

Outline of the talk

- Recapitulation: Bayesian networks
- Reverse engineering:
Learning networks from data
- Application to the yeast cell cycle
- Estimating the accuracy of inference

Outline of the talk

- **Recapitulation: Bayesian networks**
- Reverse engineering:
 - Learning networks from data
- Application to the yeast cell cycle
- Estimating the accuracy of inference

A

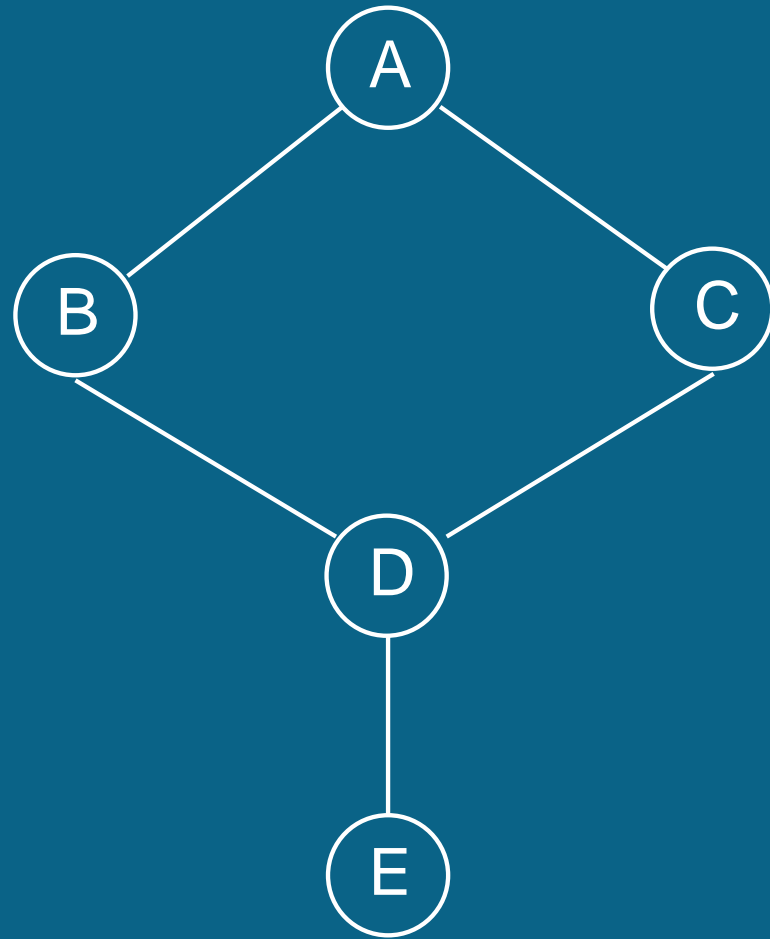
B

C

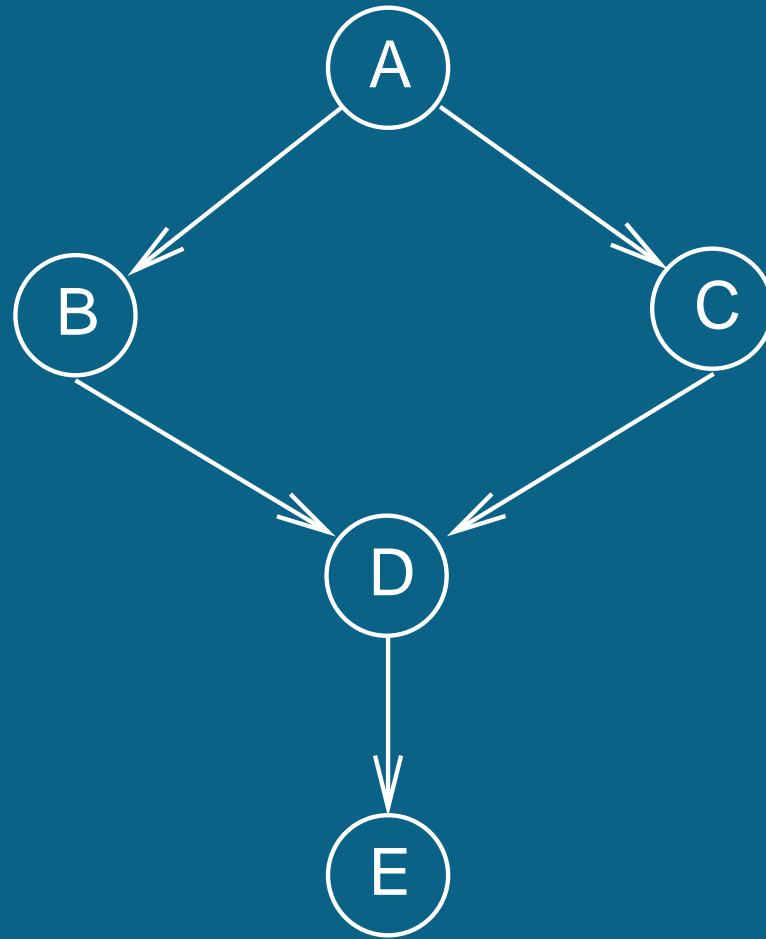
D

E

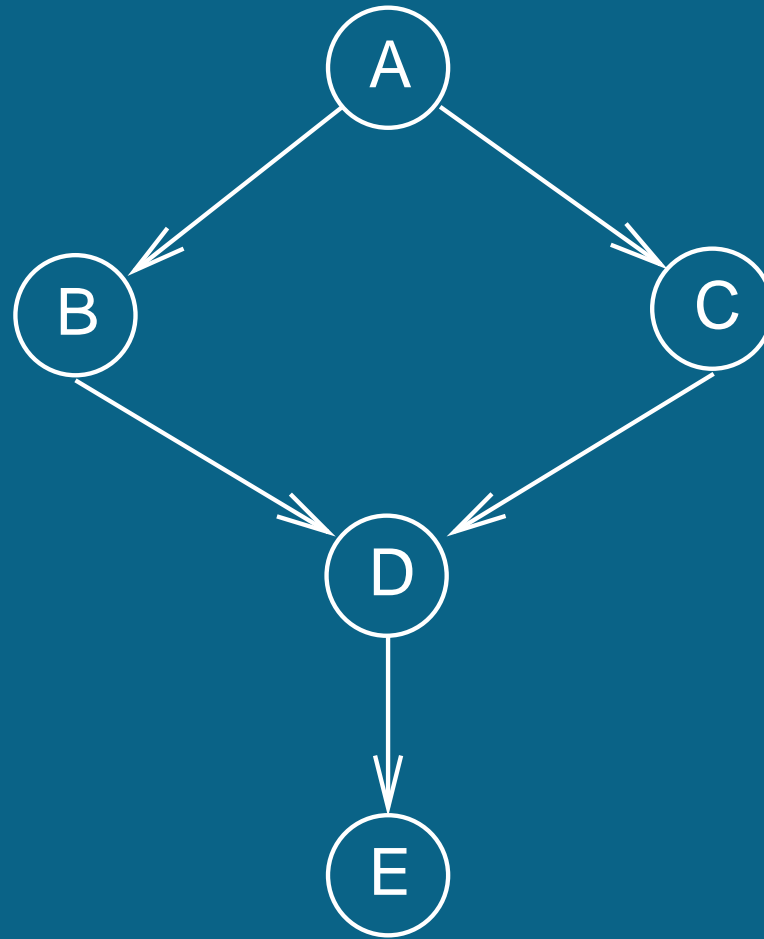
Nodes



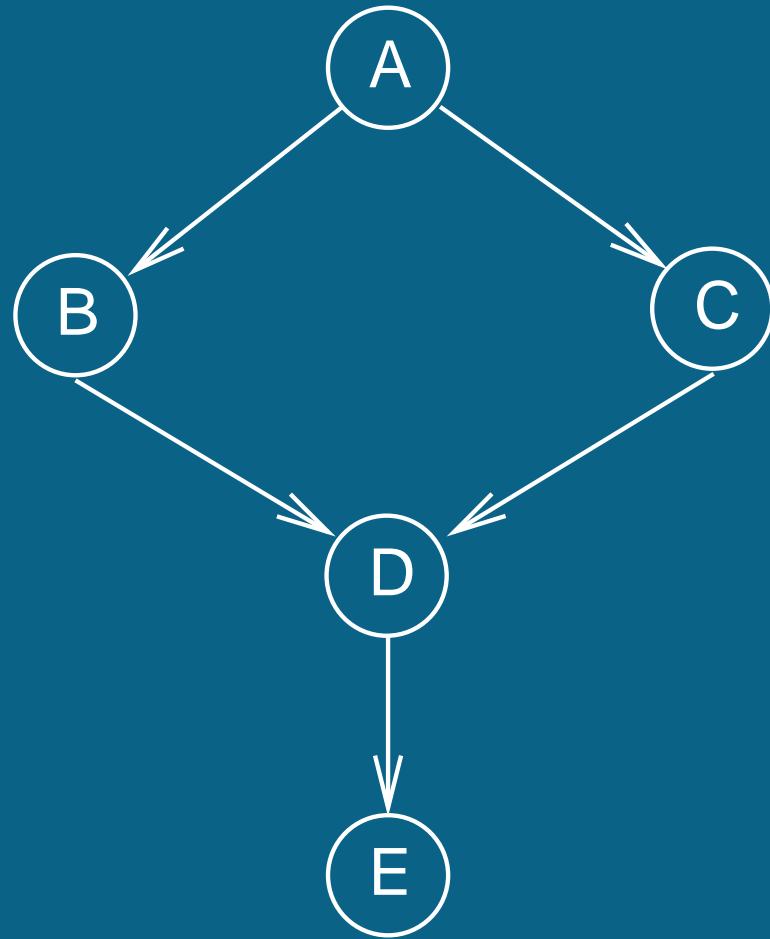
Edges



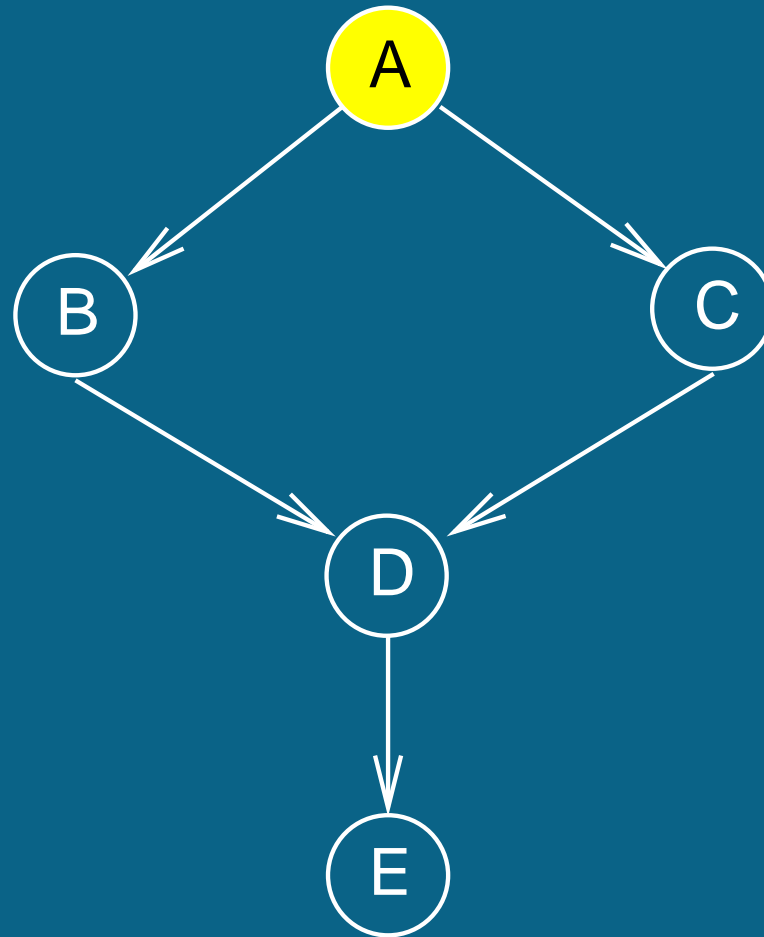
Edges = directed



No directed cycles !

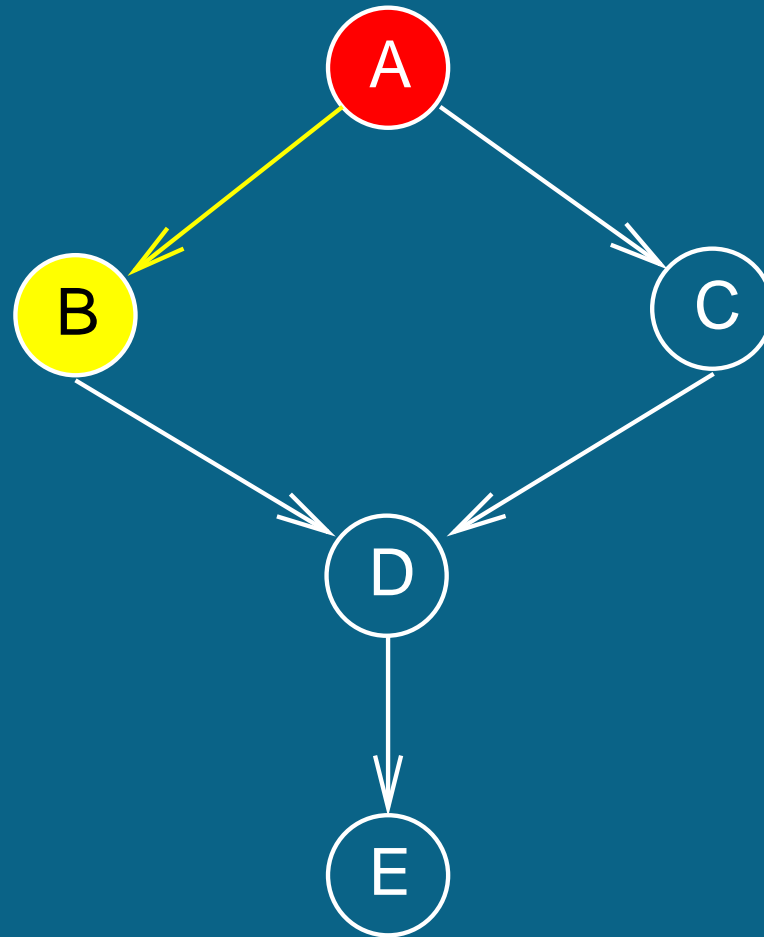


$$P(A, B, C, D, E) = \prod_i P(\text{node}_i | \text{parents}_i)$$

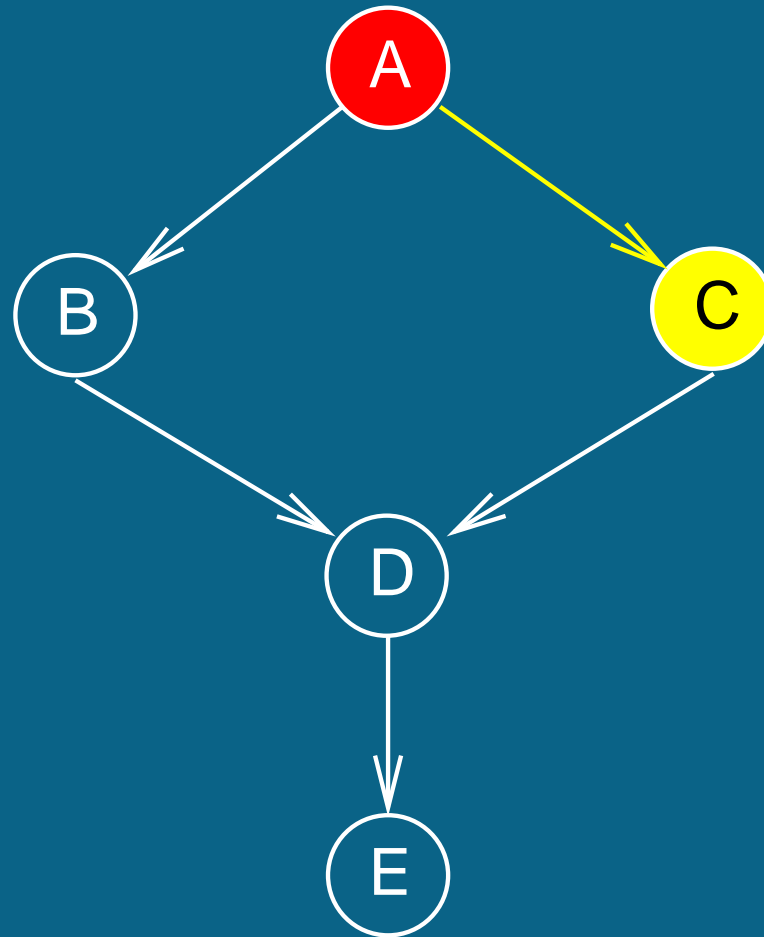


$$P(A, B, C, D, E) =$$

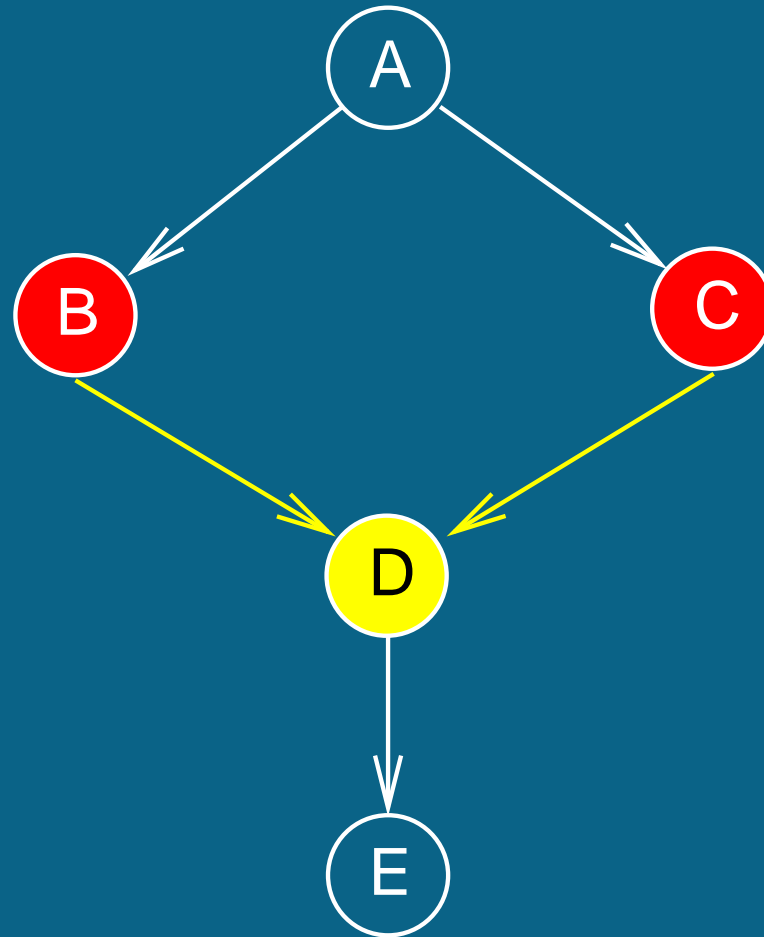
$$P(A)$$



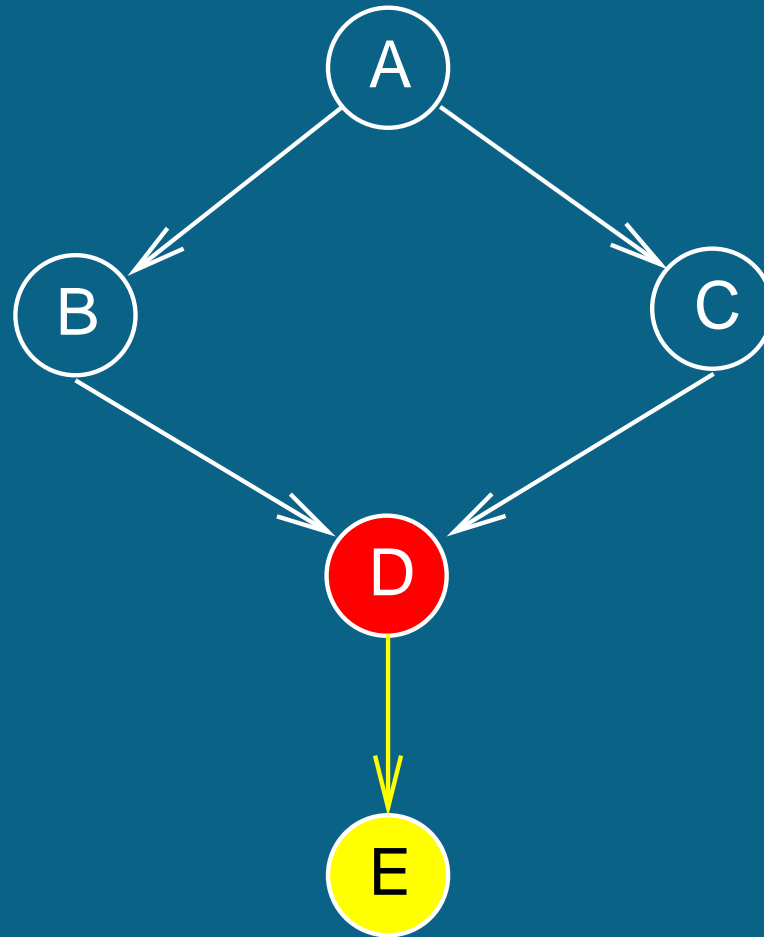
$$P(A, B, C, D, E) = P(A)P(B|A)$$



$$P(A, B, C, D, E) = P(A)P(B|A)P(C|A)$$

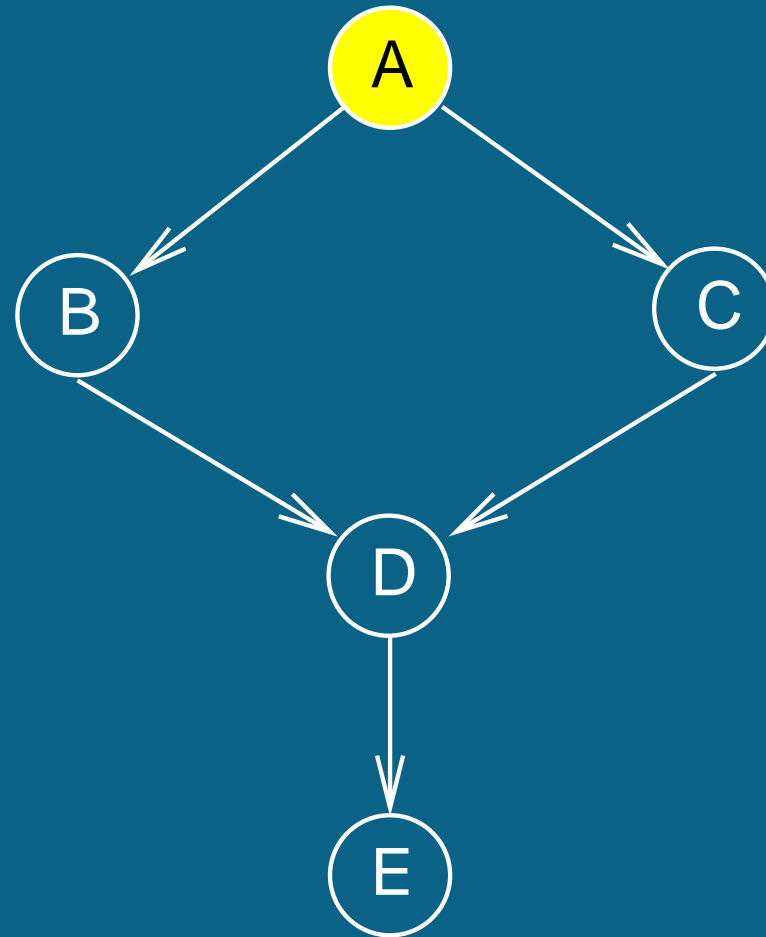


$$P(A, B, C, D, E) = P(A)P(B|A)P(C|A)P(D|B, C)$$

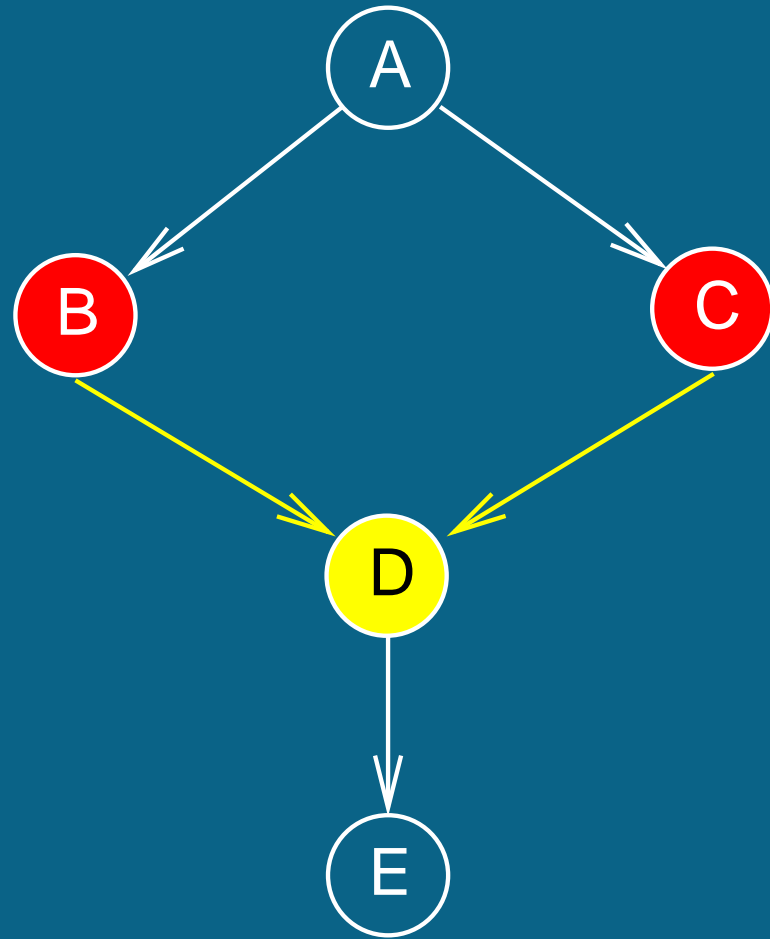


$$P(A, B, C, D, E) = P(A)P(B|A)P(C|A)P(D|B, C)P(E|D)$$

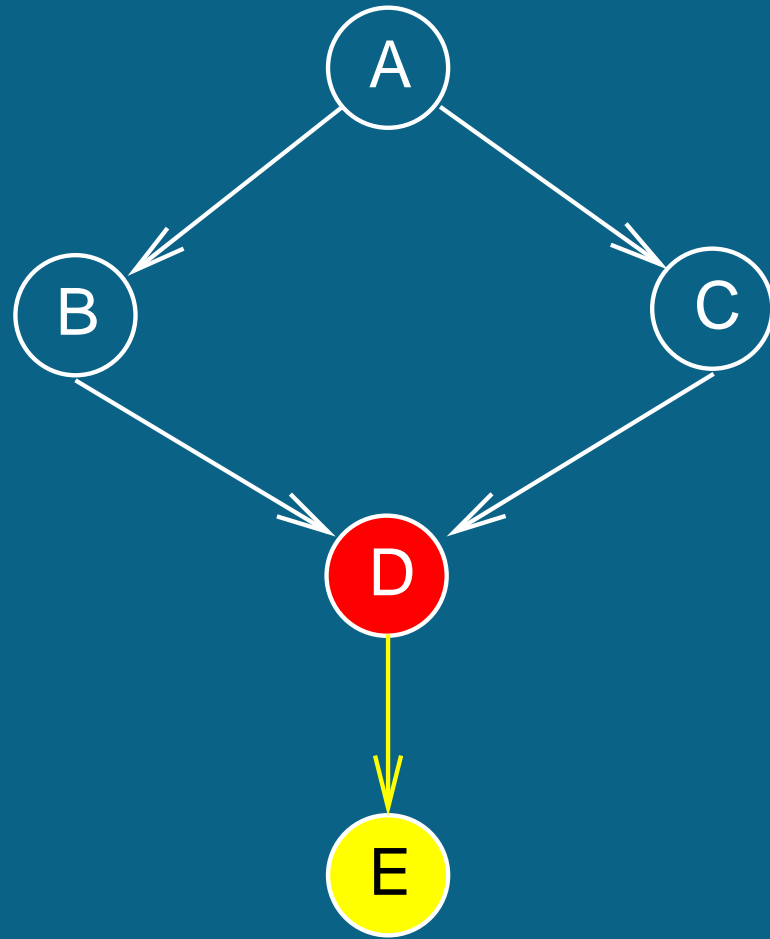
Biological interpretation



Initiation of cell (sub-)cycle

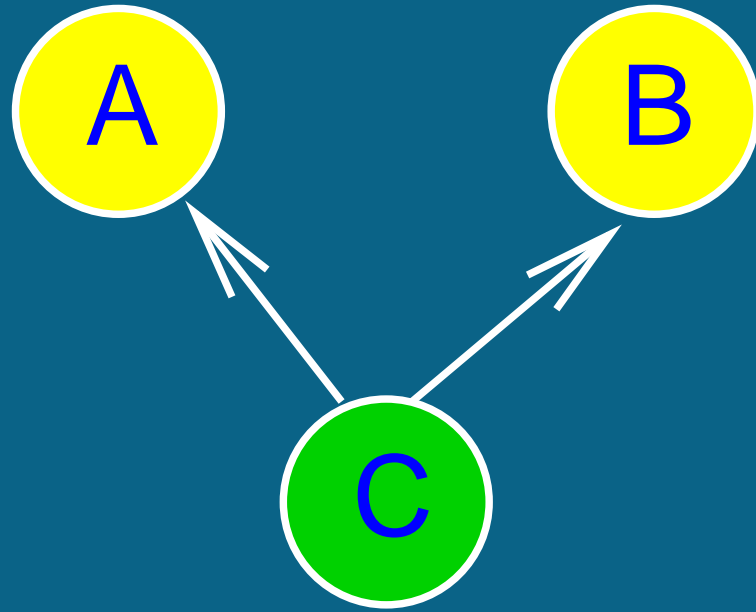


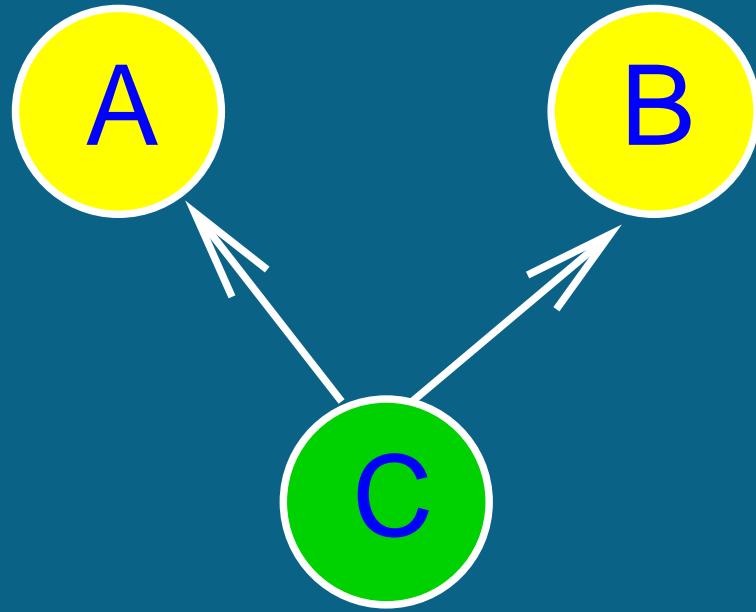
Co-regulation



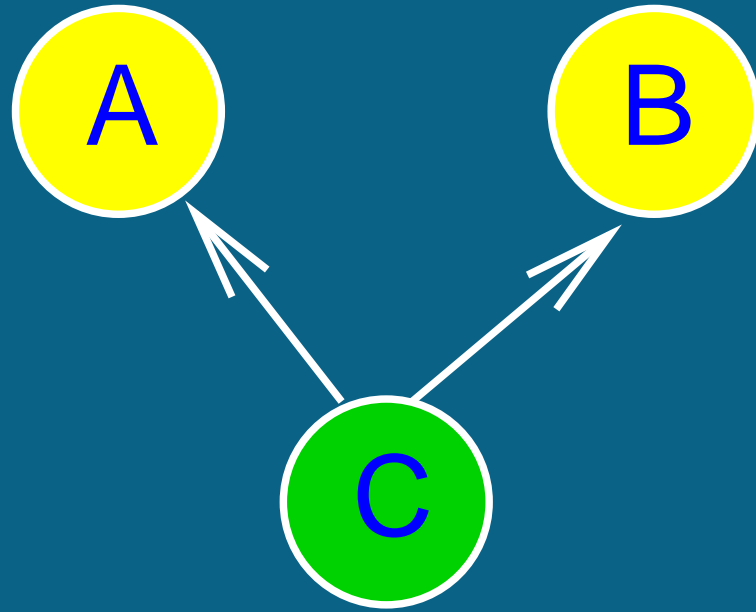
Mediation

Conditional independence relations



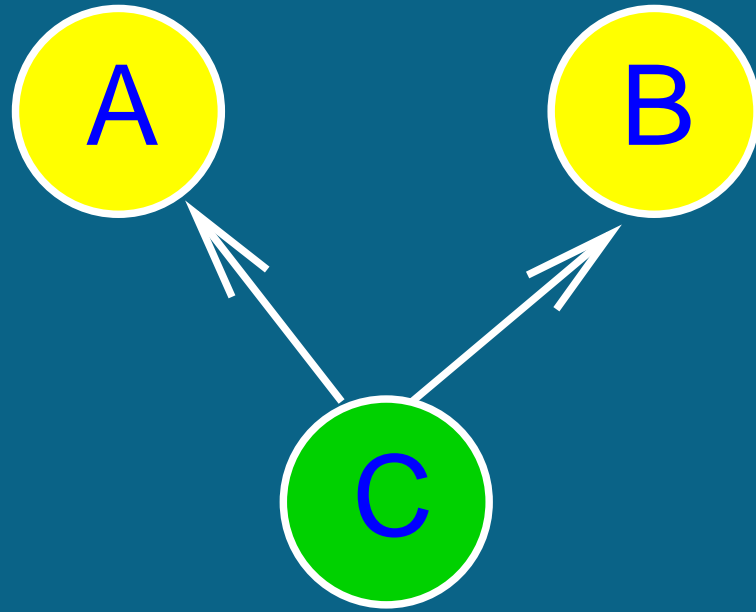


$$P(A, B, C) = P(A|C)P(B|C)P(C)$$



$$P(A, B, C) = P(A|C)P(B|C)P(C)$$

$$P(A, B|C) = \frac{P(A, B, C)}{P(C)} = P(A|C)P(B|C)$$



$$P(A, B, C) = P(A|C)P(B|C)P(C)$$

$$P(A, B|C) = \frac{P(A, B, C)}{P(C)} = P(A|C)P(B|C)$$

But: $P(A, B) \neq P(A)P(B)$

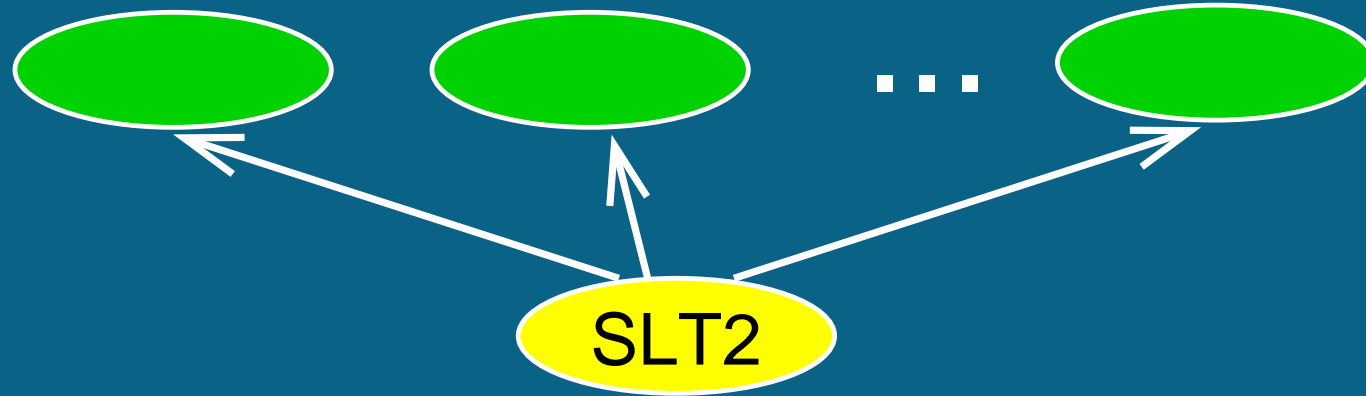
Biological example

Yeast cell cycle

Nir Friedman et al. (2000)

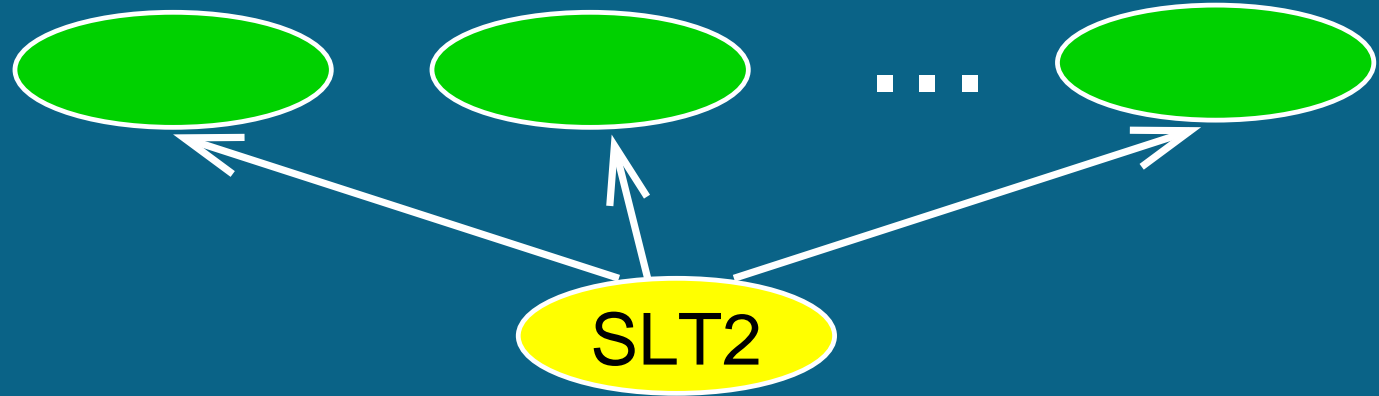
Journal of Computational Biology 7: 601-620

Low osmolarity response genes

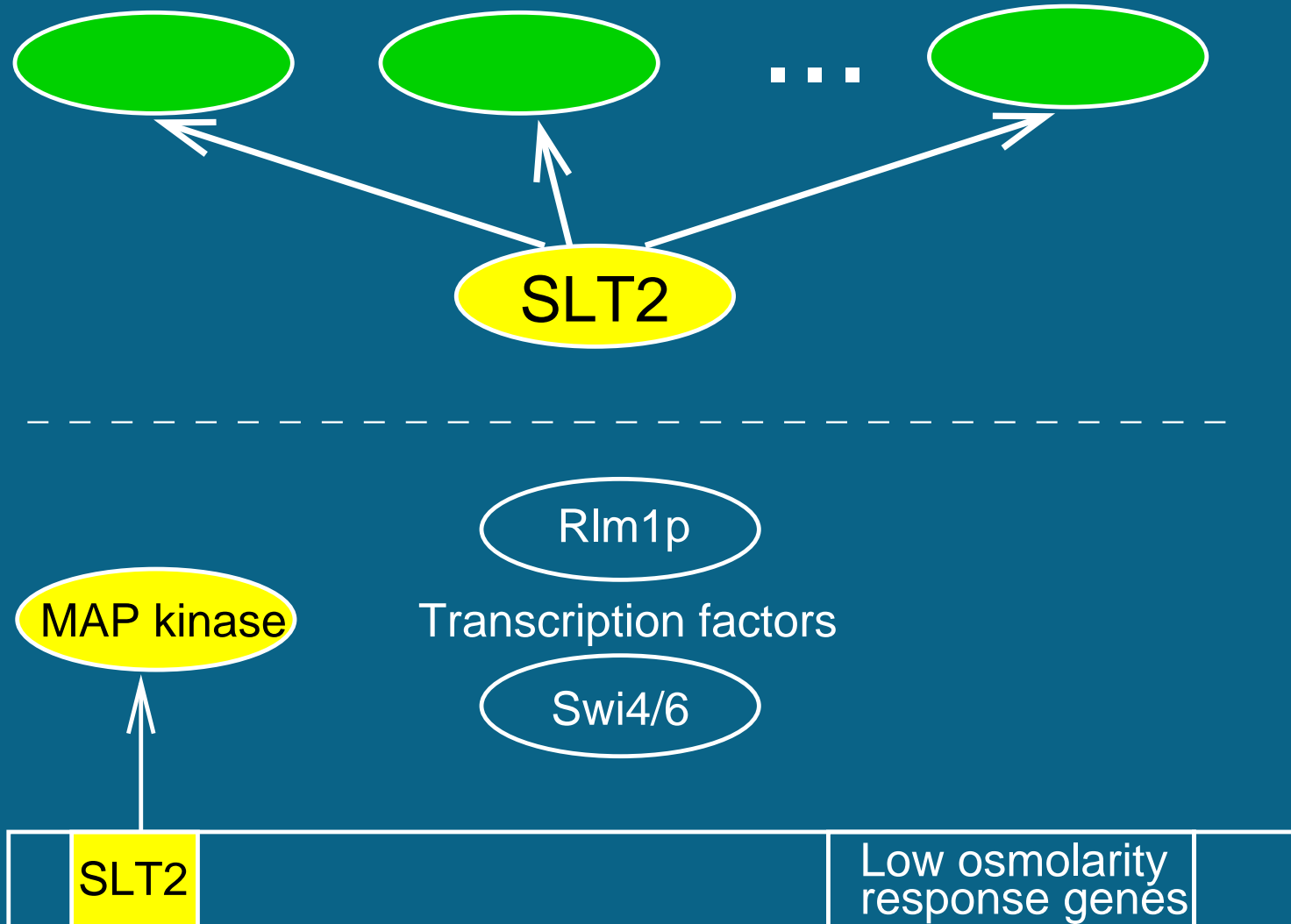


	SLT2		Low osmolarity response genes	
--	------	--	-------------------------------	--

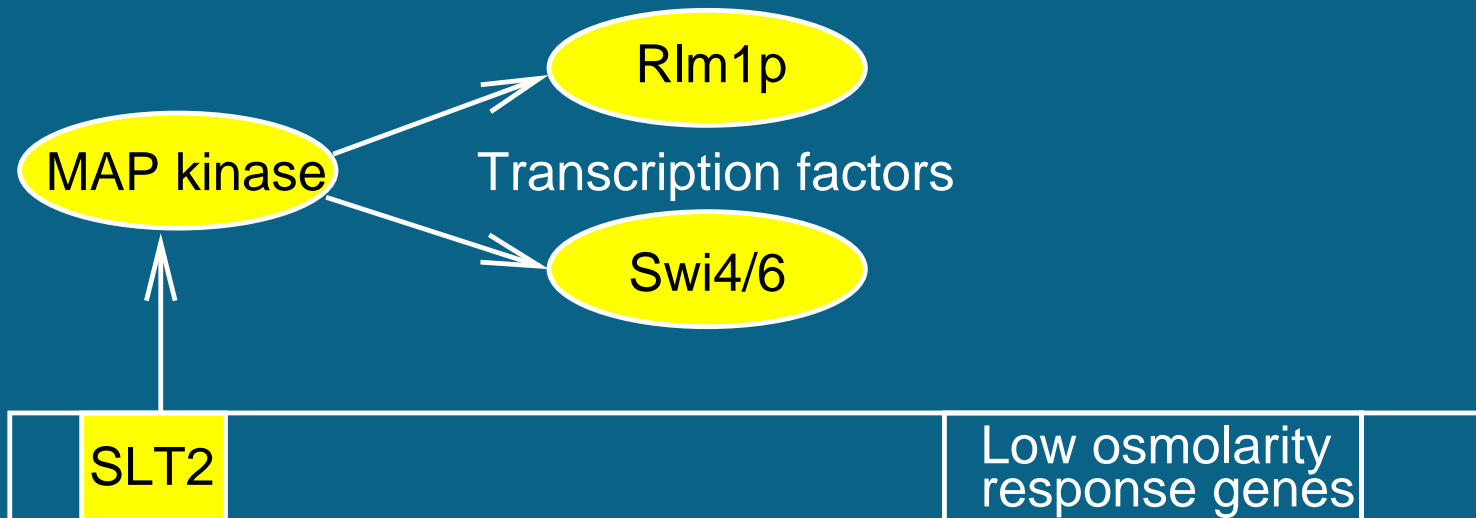
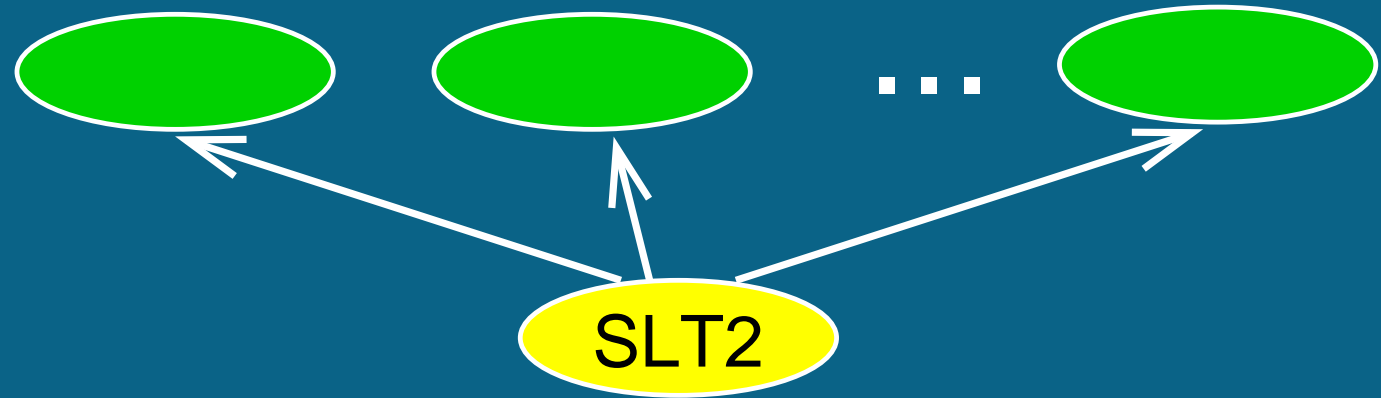
Low osmolarity response genes



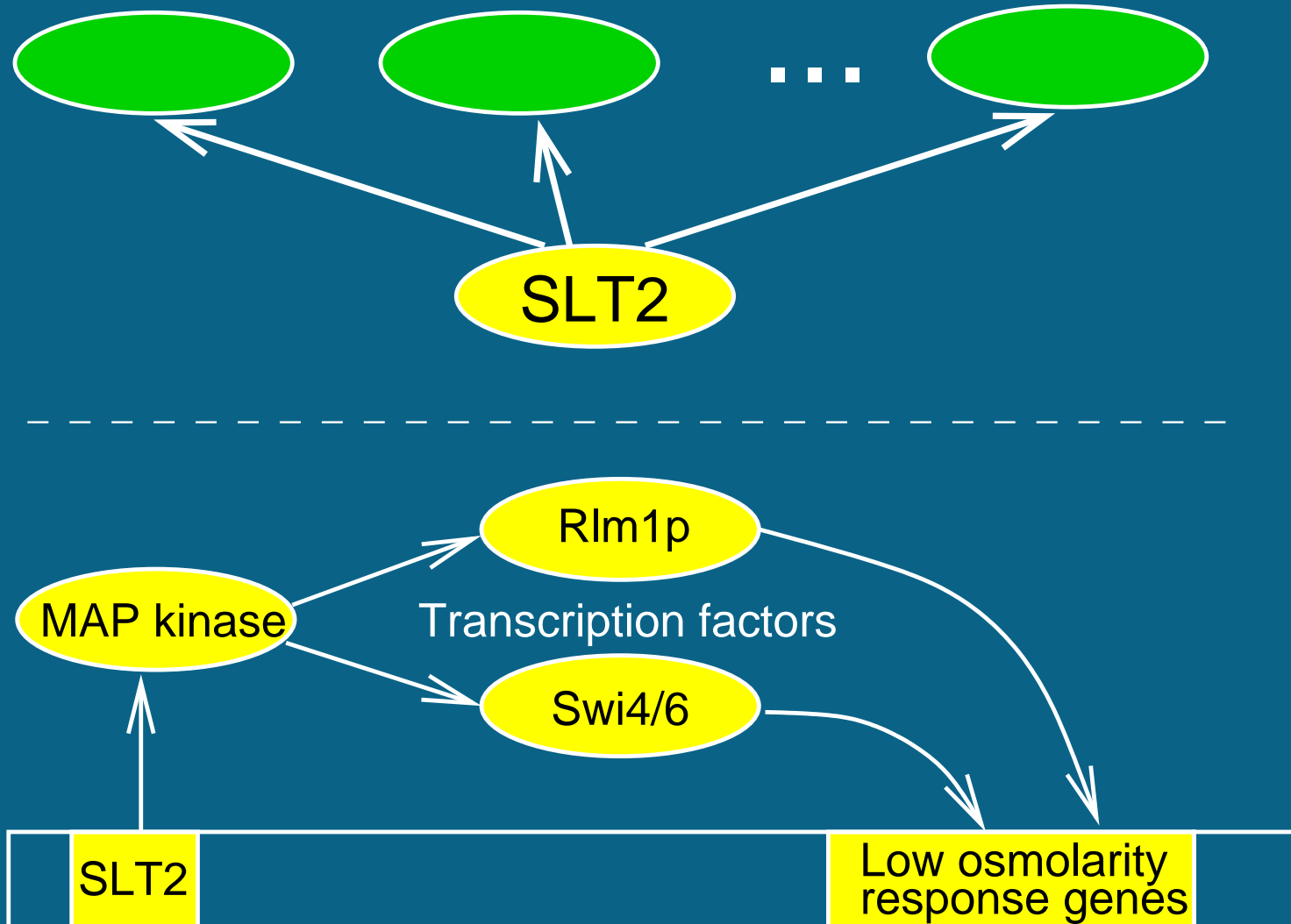
Low osmolarity response genes



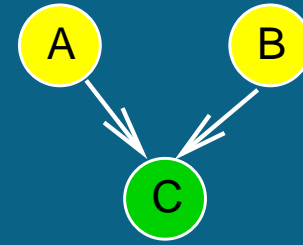
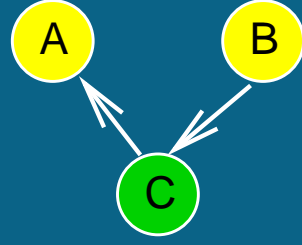
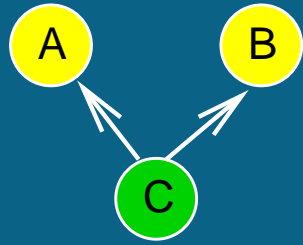
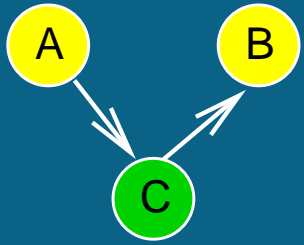
Low osmolarity response genes

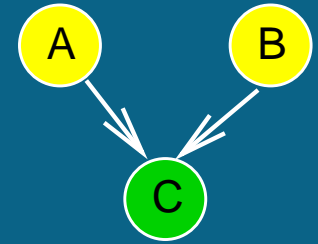
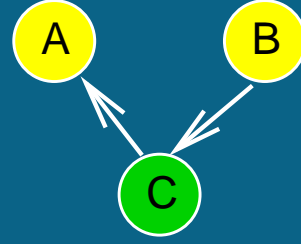
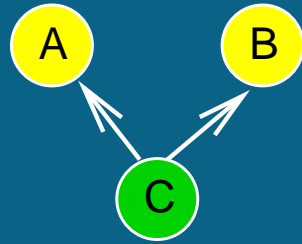
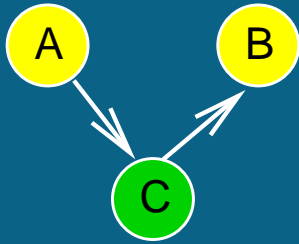


Low osmolarity response genes



Equivalence classes and PDAGs





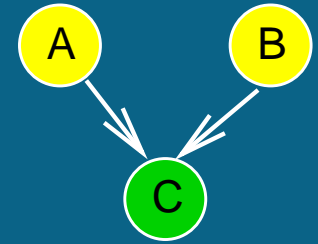
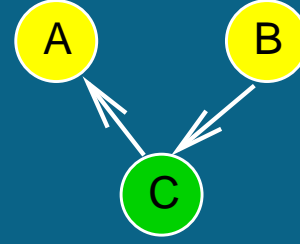
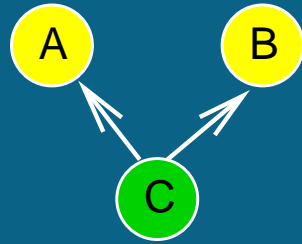
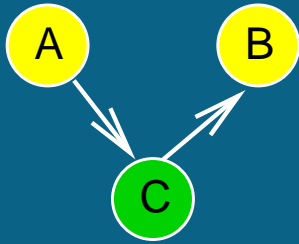
$P(A,B,C) =$

$P(B|C) P(C|A) P(A)$

$P(A|C) P(B|C) P(C)$

$P(A|C) P(C|B) P(B)$

$P(C|A,B) P(A) P(B)$



$P(A,B,C) =$

$$P(B|C) P(C|A) P(A)$$

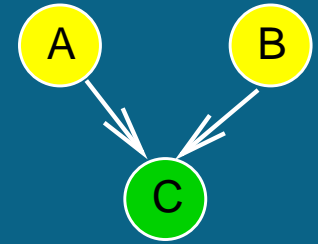
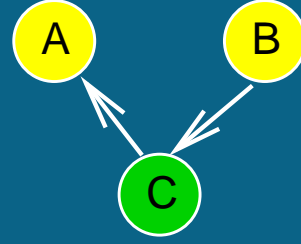
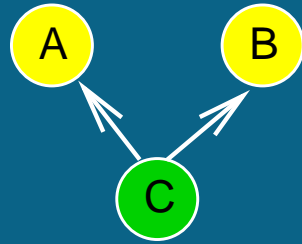
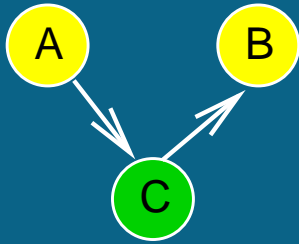
$\underbrace{\hspace{10em}}$
 $P(A|C) P(C)$

$$P(A|C) P(B|C) P(C)$$

$$P(A|C) P(C|B) P(B)$$

$\underbrace{\hspace{10em}}$
 $P(B|C) P(C)$

$$P(C|A,B) P(A) P(B)$$



$P(A,B,C) =$

$P(B|C) P(C|A) P(A)$

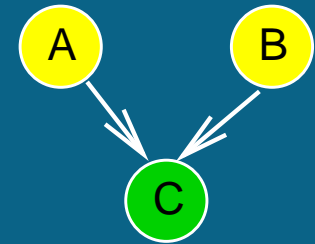
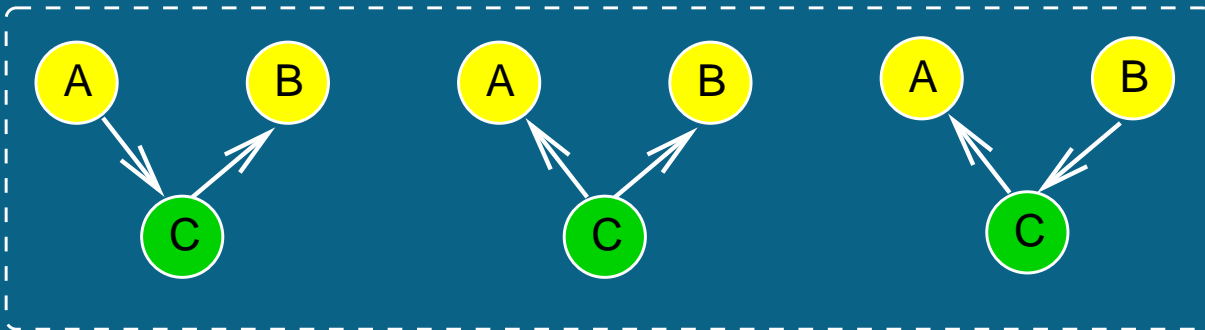
$P(A|C) P(B|C) P(C)$

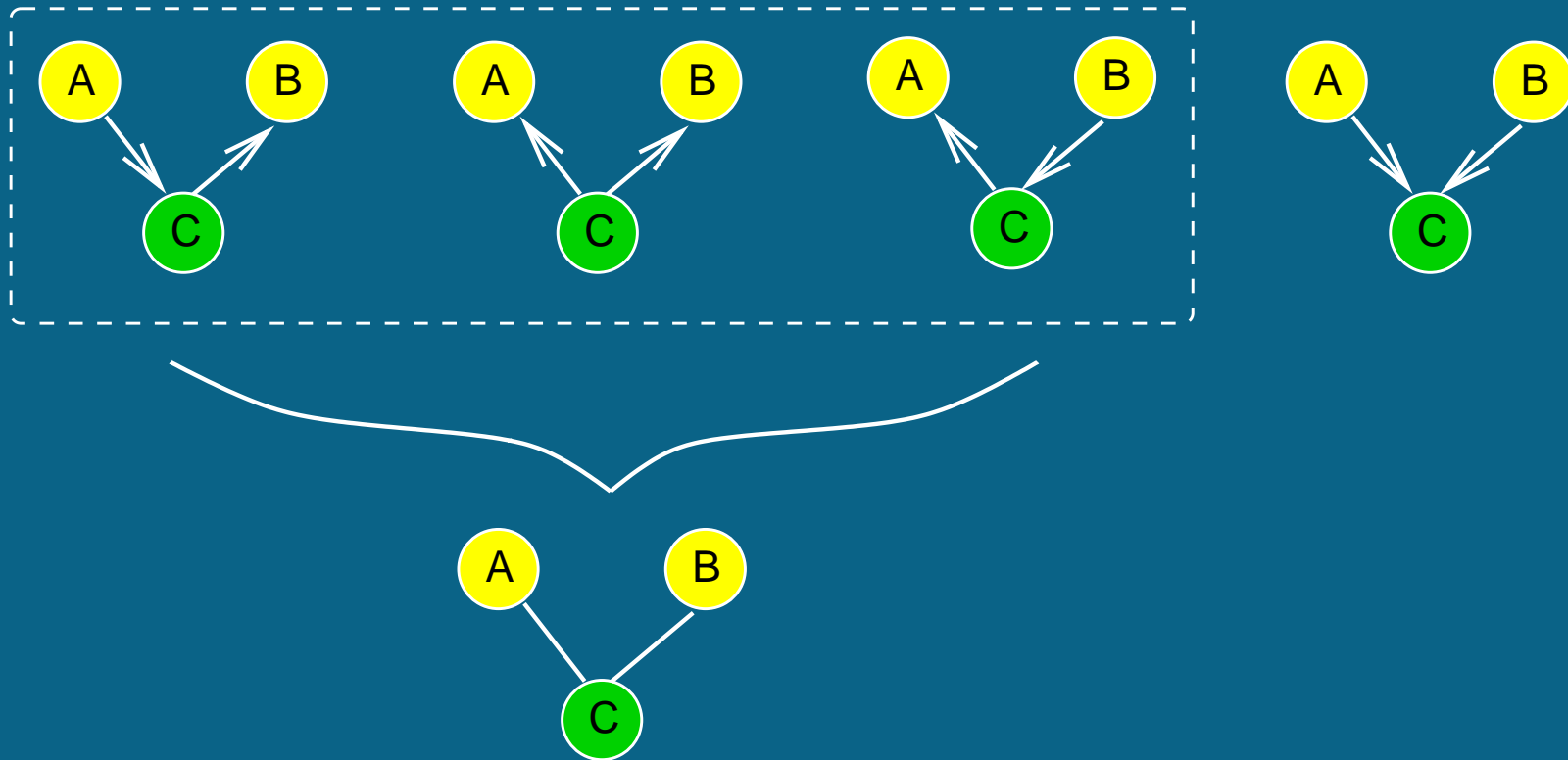
$P(A|C) P(C|B) P(B)$

$P(C|A,B) P(A) P(B)$

$\underbrace{\hspace{10em}}_{P(A|C) P(C)}$

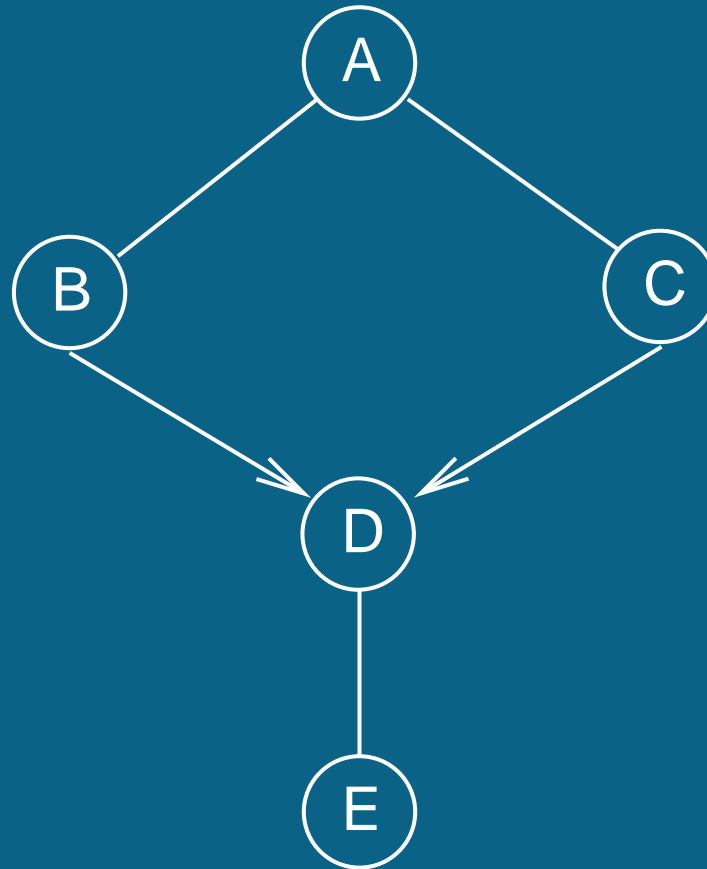
$\underbrace{\hspace{10em}}_{P(B|C) P(C)}$





- Two DAGs are **equivalent** iff they have the same **skeleton** (= the underlying undirected graph) and the same **v-structure**.
- **v-structure**: Converging directed edges into the same node without an edge between the parents.

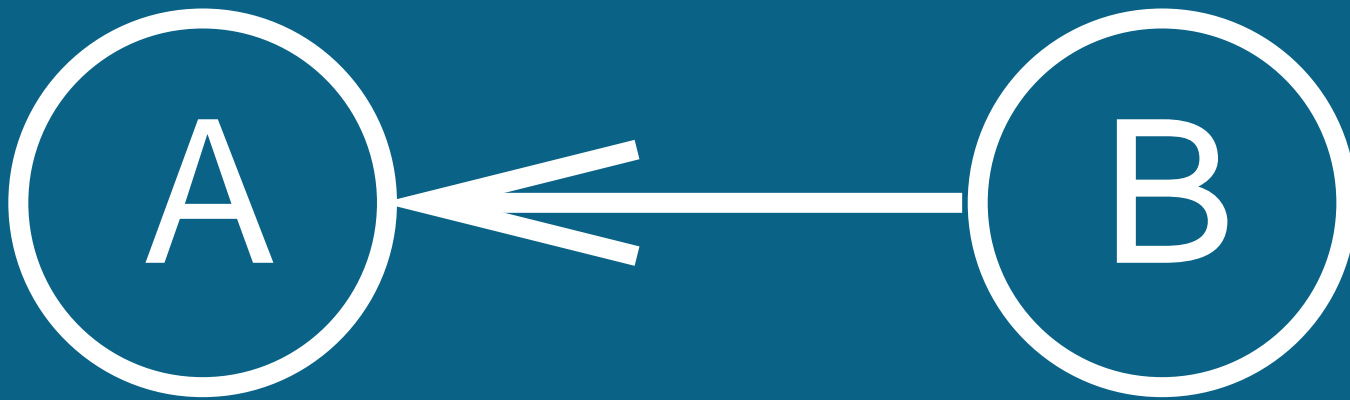
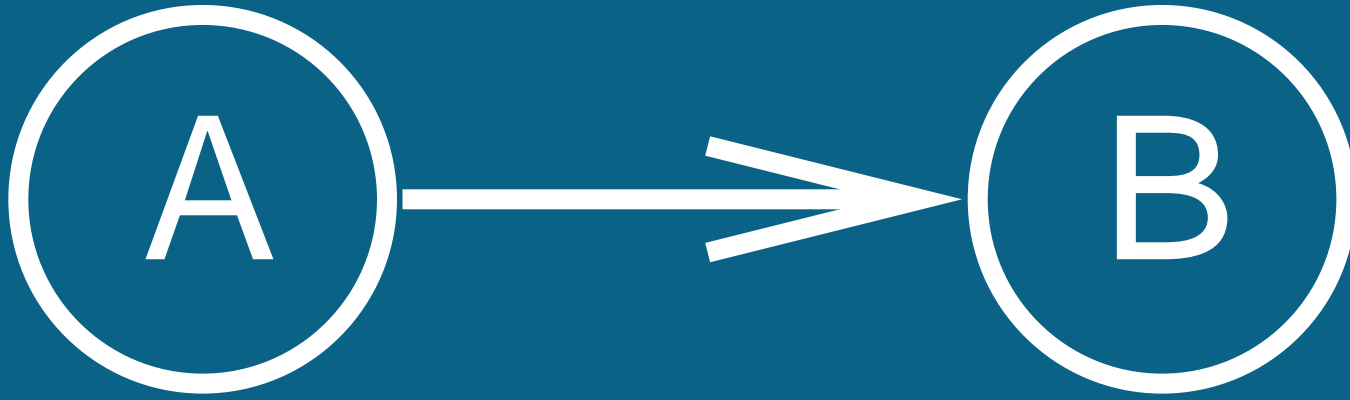
An **equivalence class of DAGs** can be represented by a **PDAG**
(partially directed acyclic graph).

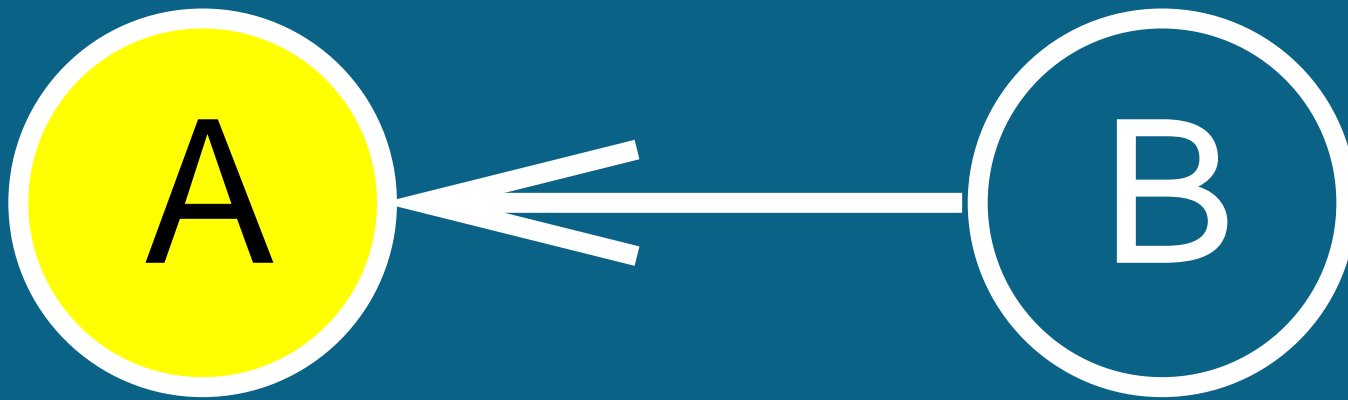
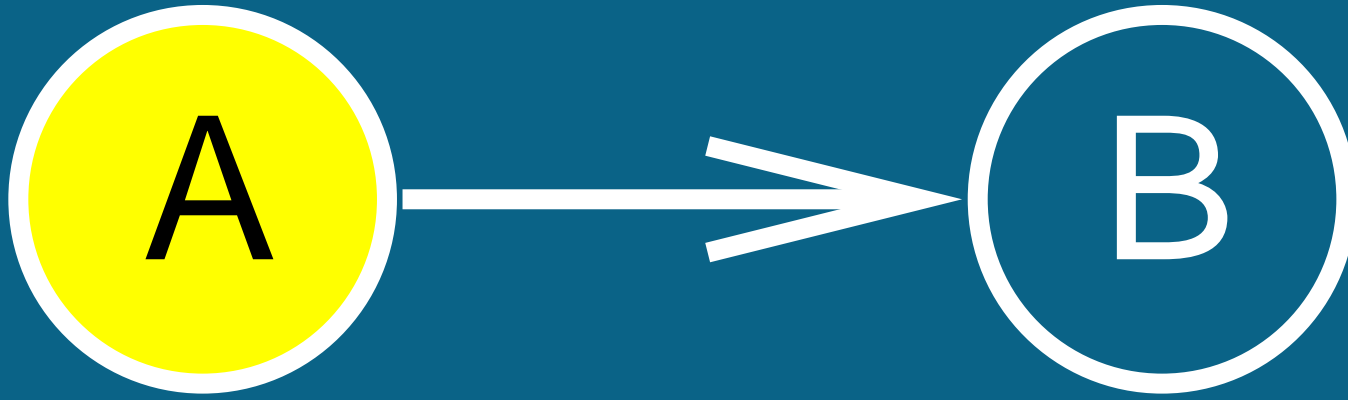


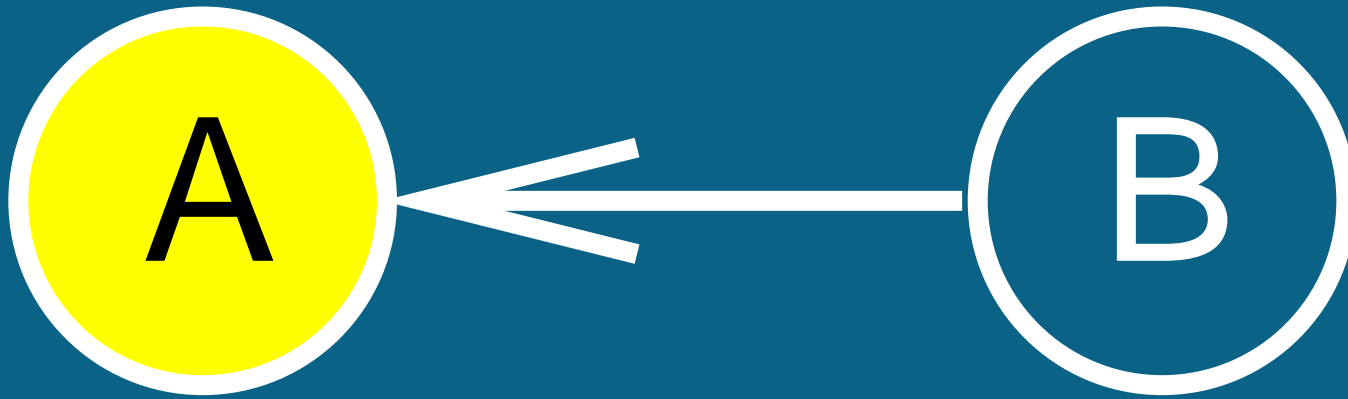
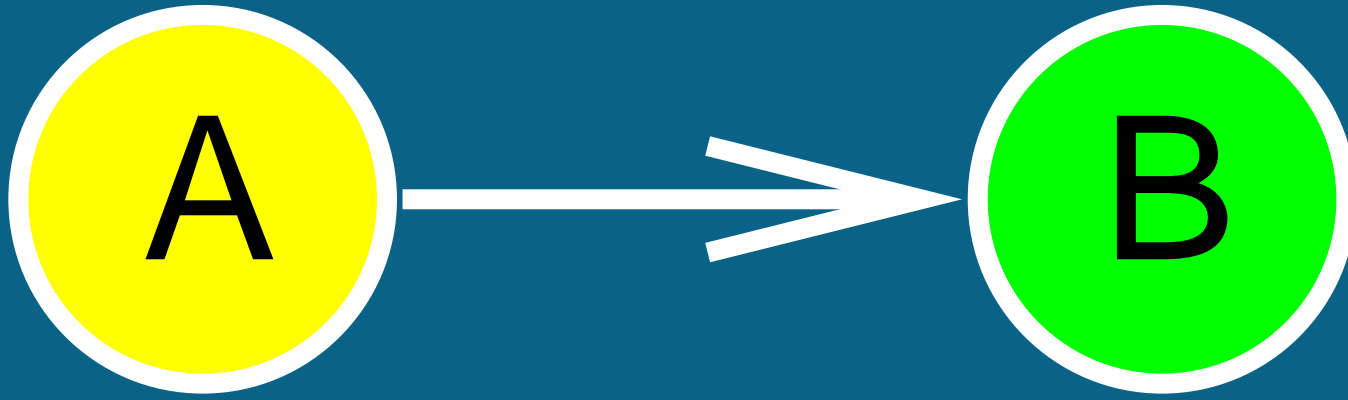
We can only learn PDAGs from the data!

- **Observation:** A passive measurement of the domain of interest.
- **Intervention:** Setting the values of some variables using forces outside the causal model, e.g., **gene knockout** or **over-expression**

- **Observation**: A passive measurement of the domain of interest.
- **Intervention**: Setting the values of some variables using forces outside the causal model, e.g., **gene knockout** or **over-expression**
- **Interventions** can destroy the symmetry within an equivalent class.







Learning with interventions

$$P(M|D) = P(D|M)P(M); \quad P(D|M) = \int P(D|\theta, M)P(\theta|M)d\theta$$

$$P(D|\theta, M) = \prod_i P(X_i|Pa(X_i), \theta)$$

Two models $\mathcal{M}, \mathcal{M}'$ with the same score are **structure equivalent**.

Int: Set of interventions \longrightarrow **Modified score**:

$$P(D|\theta, M) = \prod_{i, X_i \notin Int} P(X_i|Pa(X_i), \theta)$$

This score is no longer **structure equivalent**.

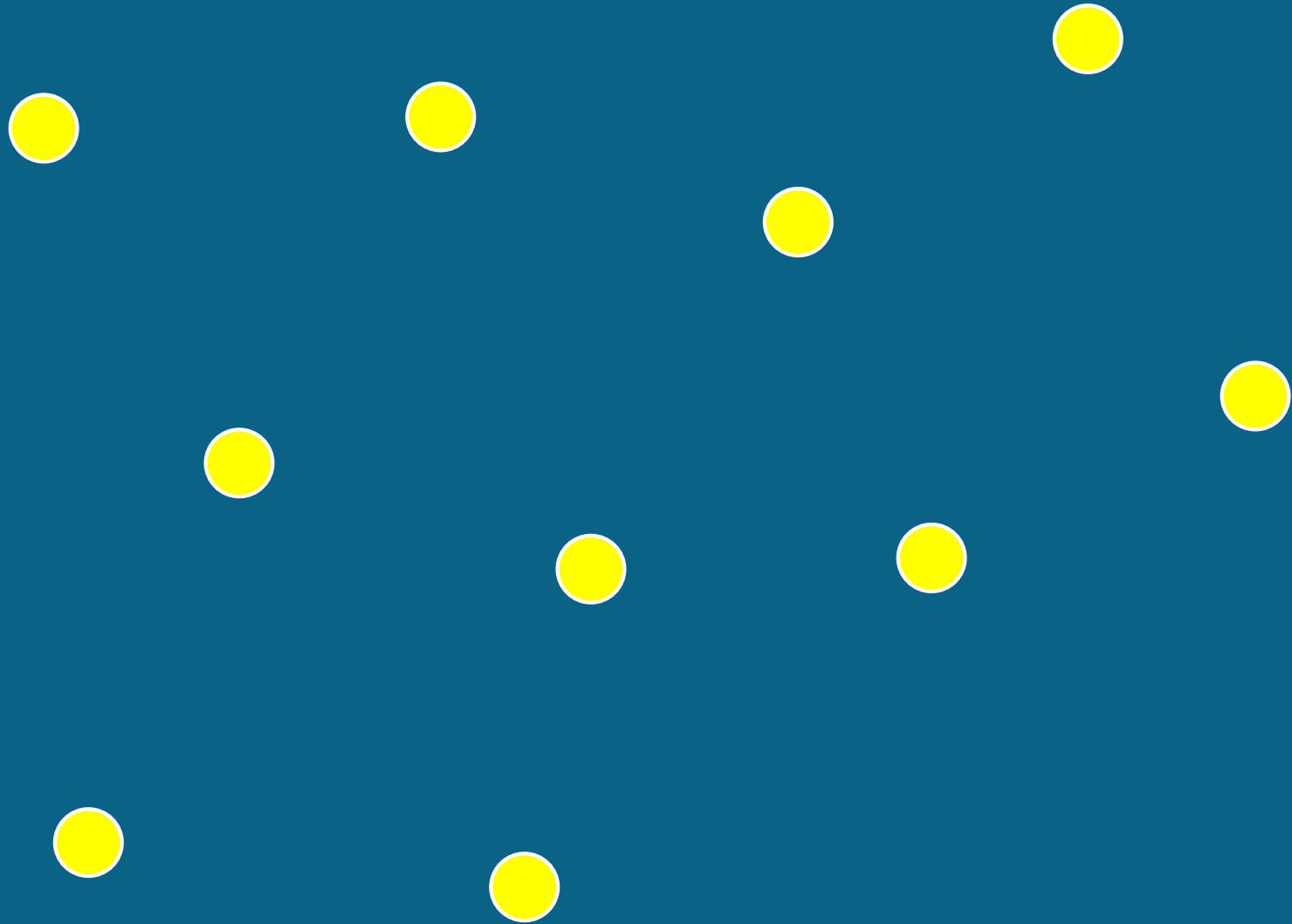
$\mathcal{M}, \mathcal{M}'$ are **intervention equivalent**, if they have the same score given the data and the interventions.

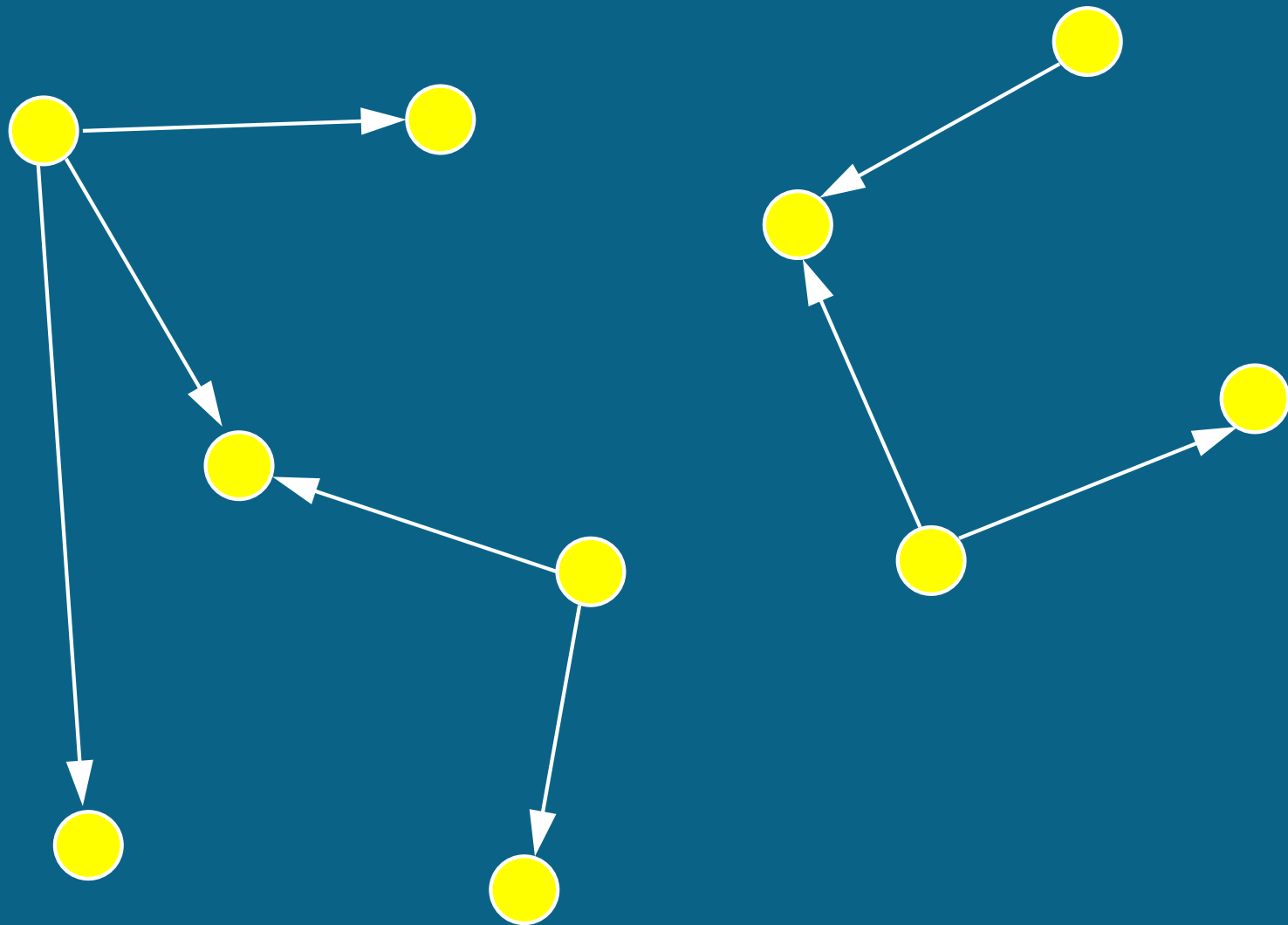
$$\{\mathcal{M}|\text{intervention equivalent}\} \subset \{\mathcal{M}|\text{structure equivalent}\}$$

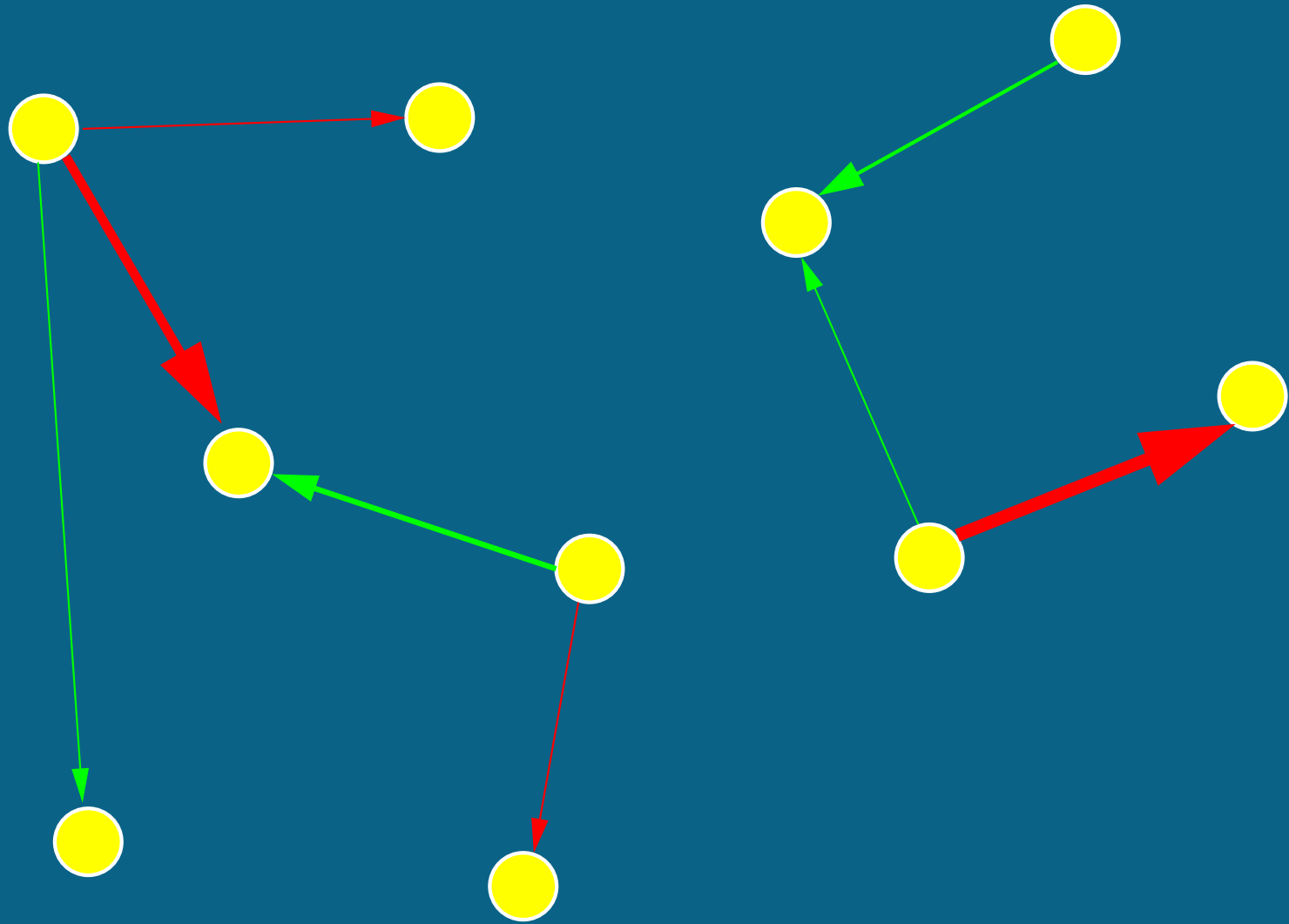
Result of interventions: More edges in the PDAG will be directed, including all edges entering or leaving an intervened variable.

Outline of the talk

- Recapitulation: Bayesian networks
- **Reverse engineering:**
Learning networks from data
- Application to the yeast cell cycle
- Estimating the accuracy of inference







Classical learning paradigm

Classical learning paradigm

Find the best network structure M :

$$M^* = \operatorname{argmax}\{P(M|D)\}$$

Classical learning paradigm

Find the **best network structure** M :

$$M^* = \operatorname{argmax}\{P(M|D)\}$$

Find the **best parameters** θ^*

$$\theta^* = \operatorname{argmax}\{P(\theta|D, M^*)\}$$

Find the best model M , that is, the best network

$$P(M|D) \propto P(D|M)P(M)$$

Find the best model M , that is, the best network

$$P(M|D) \propto P(D|M)P(M)$$

$$P(D|M) = \int P(D|\theta, M)P(\theta|M)d\theta$$

When is the integral **analytically tractable**?

Find the best model M , that is, the **best network**

$$P(M|D) \propto P(D|M)P(M)$$

$$P(D|M) = \int P(D|\theta, M)P(\theta|M)d\theta$$

When is the integral **analytically tractable**?

- Complete observation: **No missing values.**
- $P(D|\theta, M)$ and $P(\theta|M)$ must satisfy certain regularity conditions.
- Examples: **Multimodal** with a Dirichlet prior, **linear Gaussian** with a normal-gamma prior.

Naive approach

- Compute $P(M|D)$ for all possible network structures M .
- Select network structure M^* that maximizes $P(M|D)$

Naive approach

- Compute $P(M|D)$ for all possible network structures M .
- Select network structure M^* that maximizes $P(M|D)$

Problem 1:

Number of different network structures increases super-exponentially with the number of nodes.

N of nodes	2	4	6	8	10
N of structures	3	543	3.7×10^6	7.8×10^{11}	4.2×10^{18}

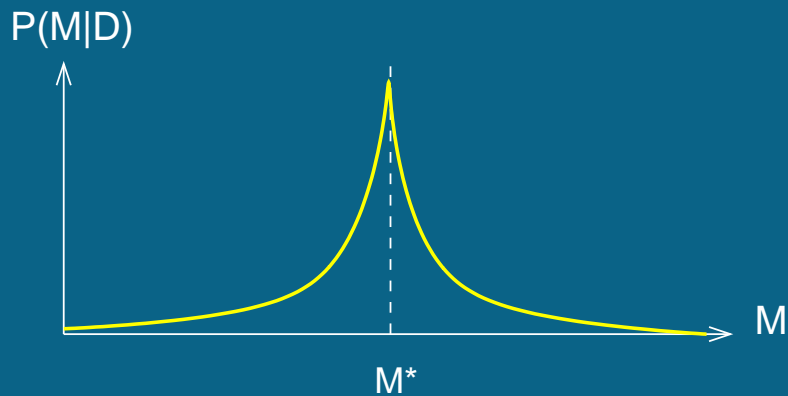
→ Optimization problem intractable for large N of nodes

Naive approach

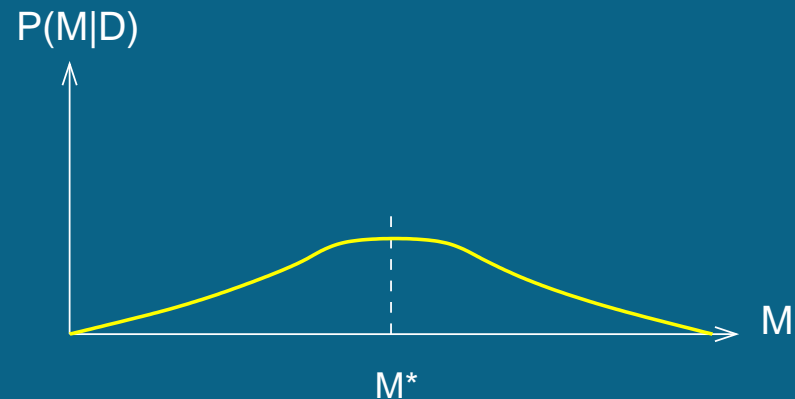
- Compute $P(M|D)$ for all possible network structures M .
- Select network structure M^* that maximizes $P(M|D)$

Problem 2:

Data are sparse \rightarrow Intrinsic uncertainty of inference



Large data set D :
Best network structure M^* well defined



Small data set D :
Intrinsic uncertainty about M^*

Objective: Sample from the posterior distribution

$$P(M_k|D) = \frac{P(D|M_k)P(M_k)}{\sum_i P(D|M_i)P(M_i)}$$

Direct approach intractable due to $\sum_i P(D|M_i)P(M_i)$

Objective: Sample from the posterior distribution

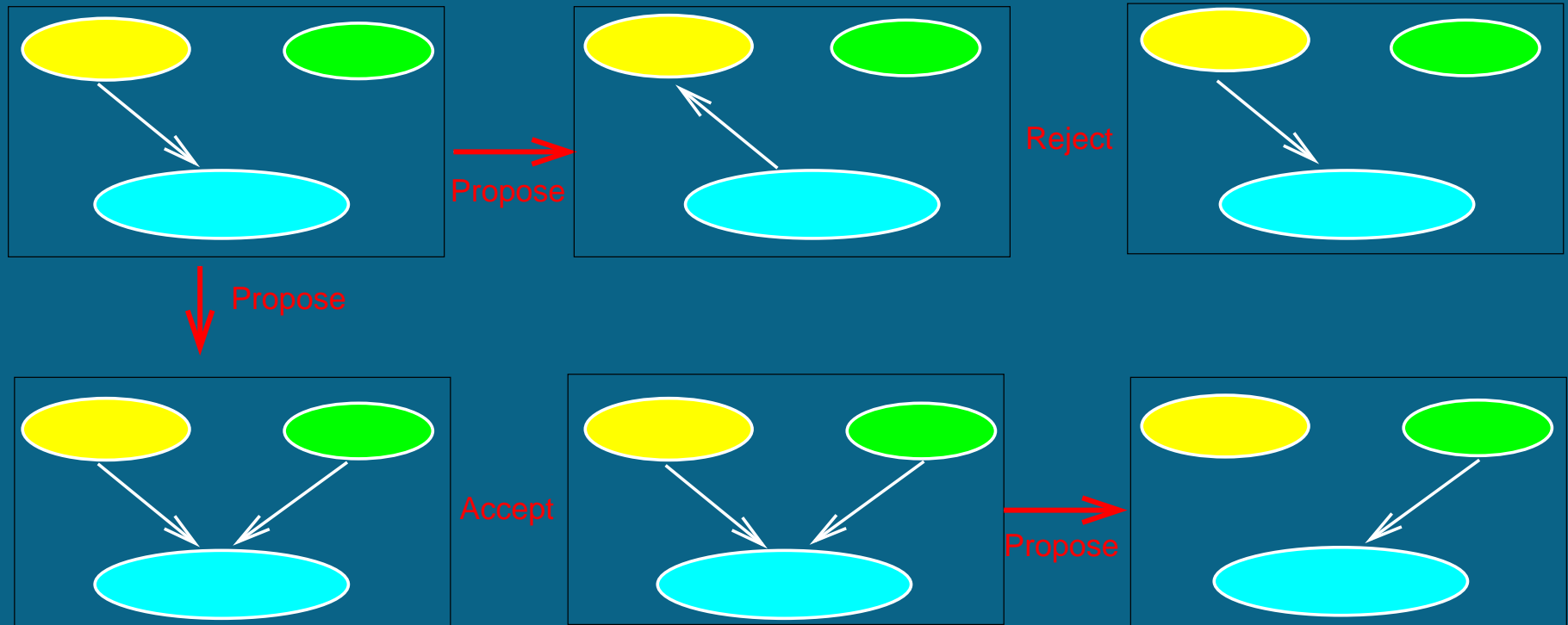
$$P(M_k|D) = \frac{P(D|M_k)P(M_k)}{\sum_i P(D|M_i)P(M_i)}$$

Direct approach intractable due to $\sum_i P(D|M_i)P(M_i)$

Markov chain Monte Carlo (MCMC):

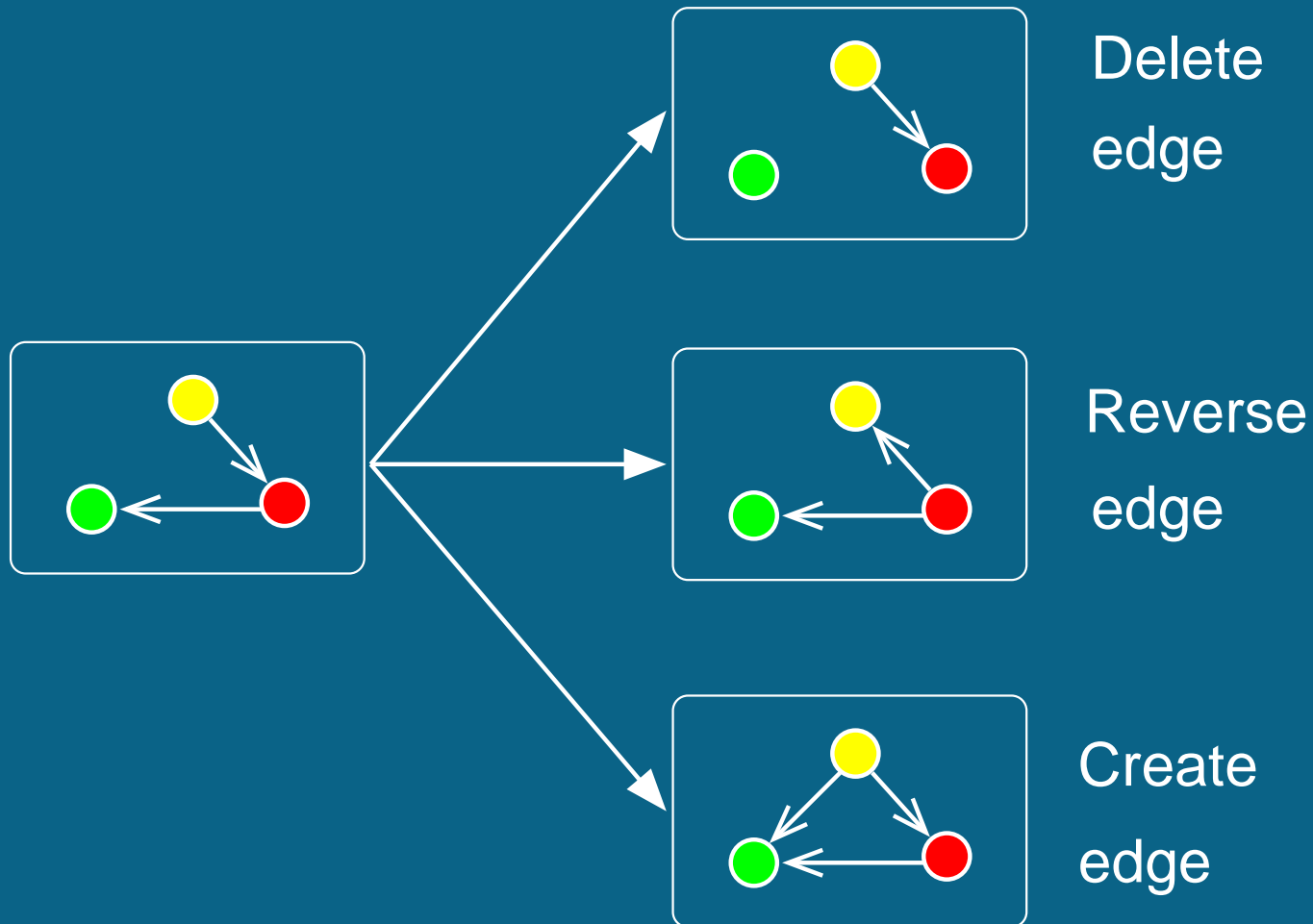
- **Proposal move:** Given network M_{old} , propose a new network M_{new} with probability $Q(M_{new}|M_{old})$.
- **Acceptance/Rejection:** Accept this new network with probability $\min \left\{ 1, \frac{P(D|M_{new})P(M_{new})}{P(D|M_{old})P(M_{old})} \times \frac{Q(M_{old}|M_{new})}{Q(M_{new}|M_{old})} \right\}$

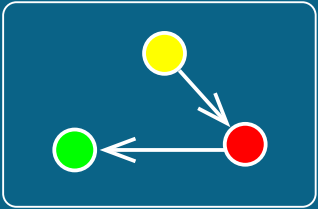
Markov chain Monte Carlo (MCMC)



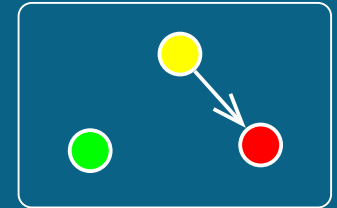
Accept move with probability: $\min \left\{ 1, \frac{P(M_{new}|D)}{P(M_{old}|D)} \times \frac{Q(M_{old}|M_{new})}{Q(M_{new}|M_{old})} \right\}$

MCMC moves

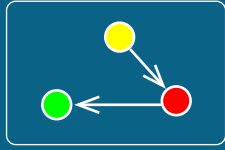




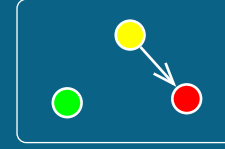
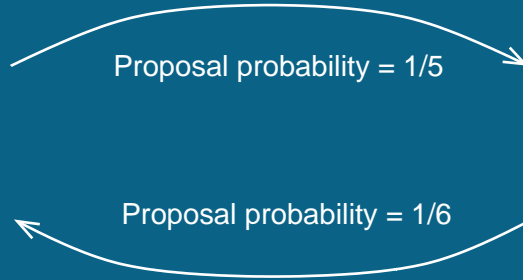
Proposal probability = ?



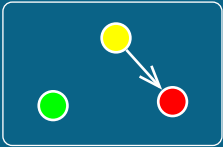
Proposal probability = ?



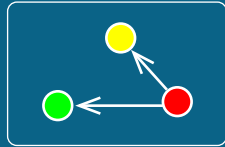
Neighbourhood



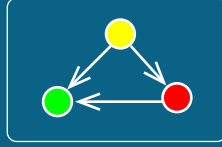
Neighbourhood



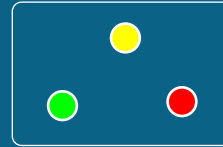
Delete



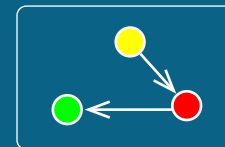
Reverse



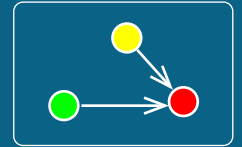
Add



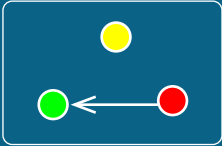
Delete



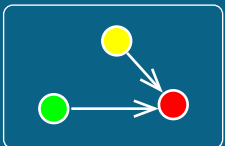
Add



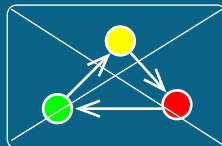
Add



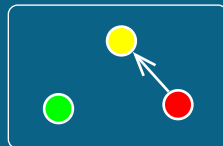
Delete



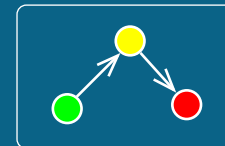
Reverse



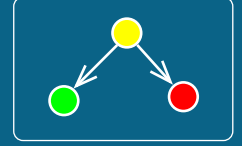
Add



Reverse



Add



Add

Convergence of MCMC simulation

Burn-in: T MCMC steps

Sampling: T MCMC steps

How large do we have to choose T ?

Convergence of MCMC simulation

Burn-in: T MCMC steps

Sampling: T MCMC steps

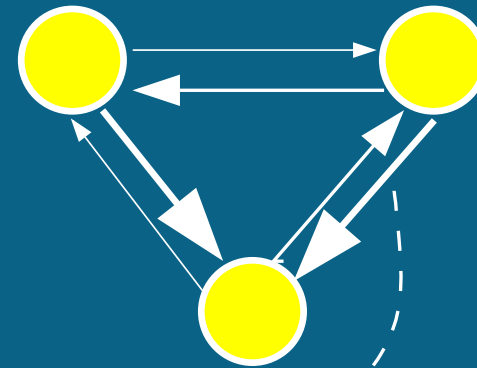
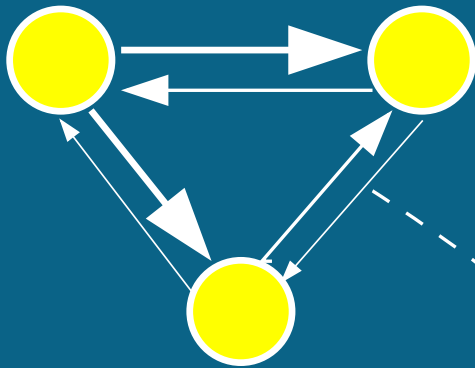
How large do we have to choose T ?

Repeat MCMC simulations from different initializations

Scatter plots of posterior probabilities of edges

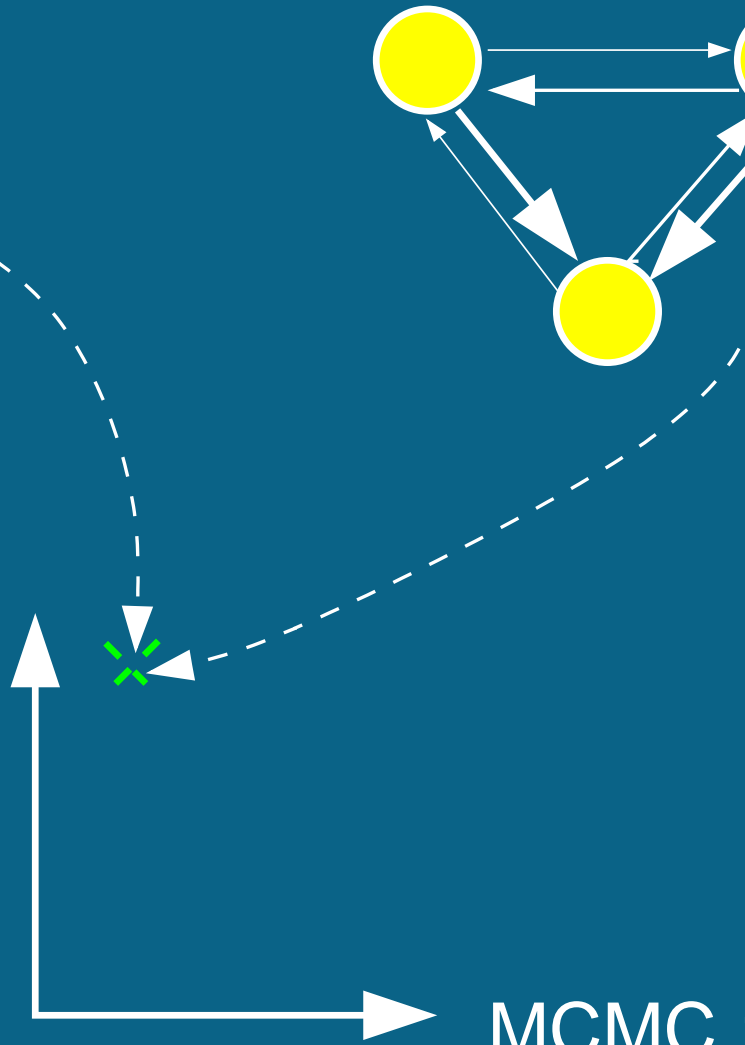
MCMC simulation 1

MCMC simulation 2

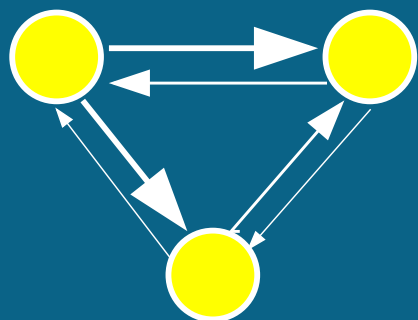


MCMC 2

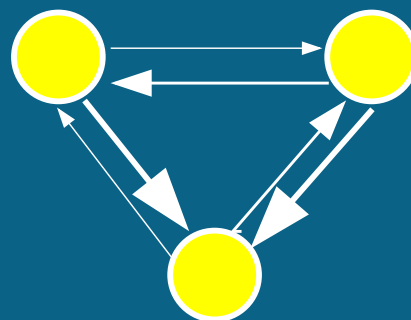
MCMC 1



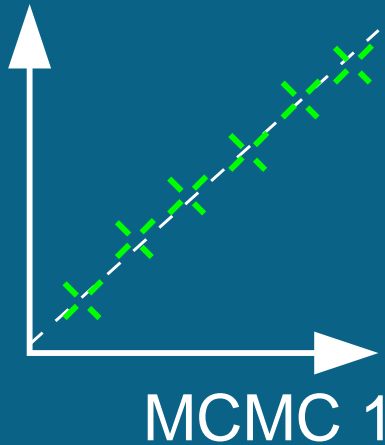
MCMC simulation 1



MCMC simulation 2

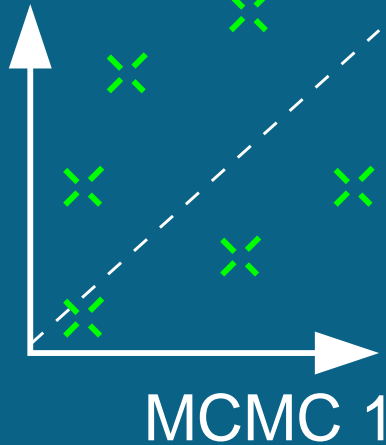


MCMC 2



T infinite

MCMC 2



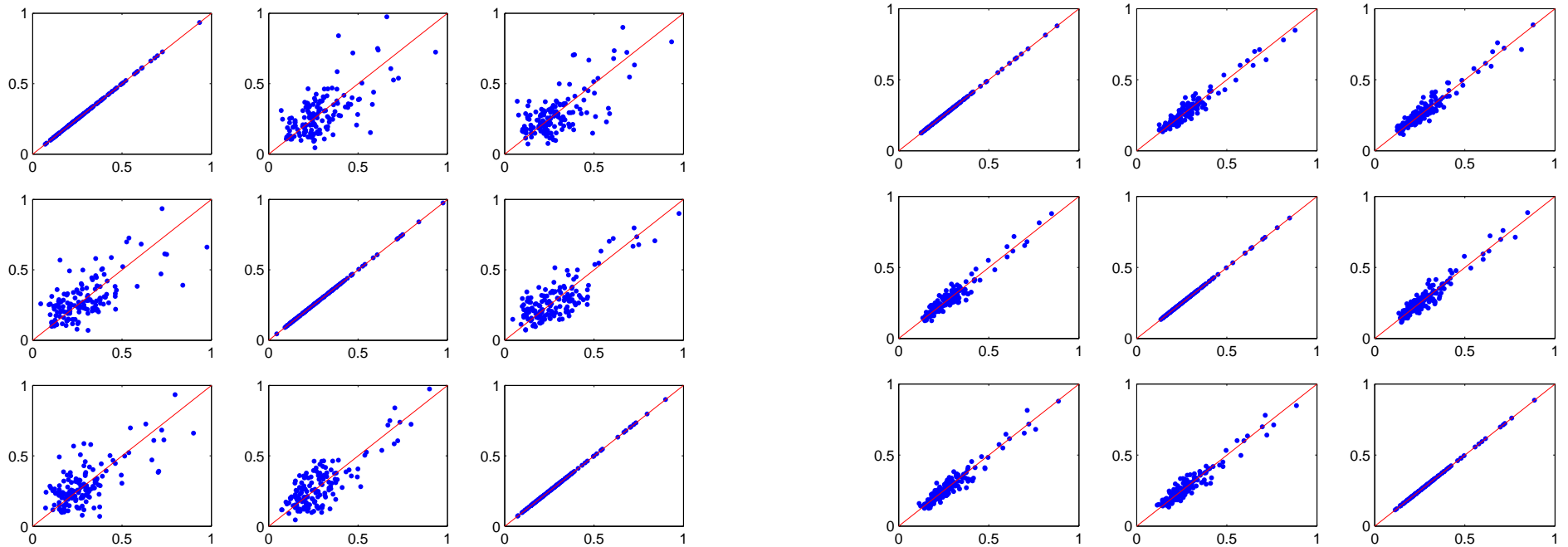
T too short

MCMC 2

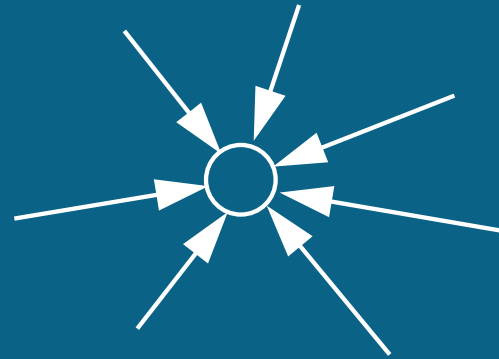
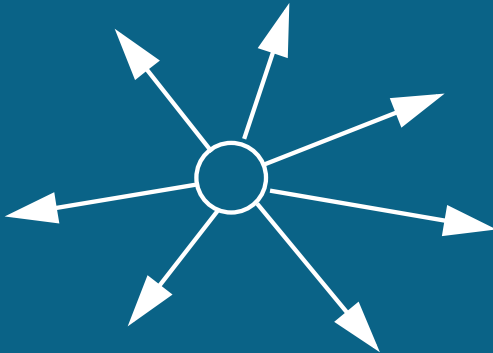
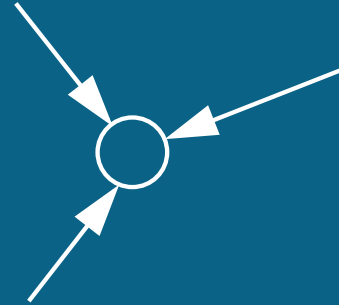
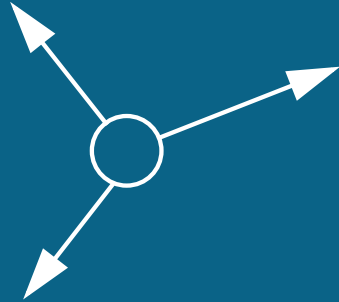


T long enough

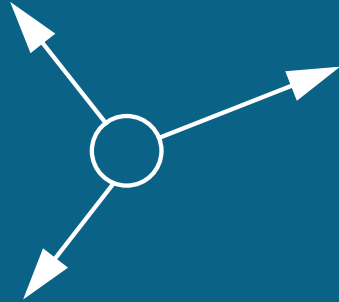
MCMC steps: **10,000** (left) versus **100,000** (right)



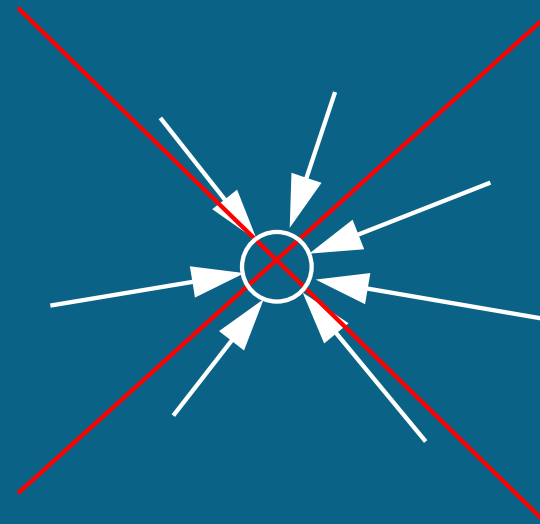
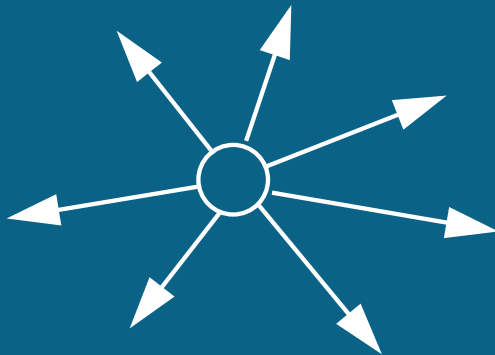
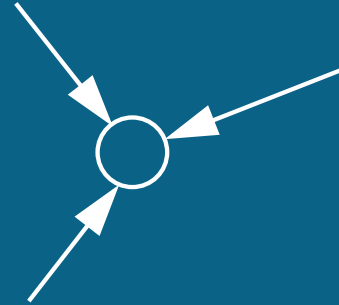
The structure prior $P(\mathcal{M})$



Fan-out unrestricted



Fan-in restricted



not permissible

$P(\mathcal{M})$ uniform over structures

or

$P(\mathcal{M})$ uniform over cardinalities of parent sets

$P(\mathcal{M})$ uniform over structures

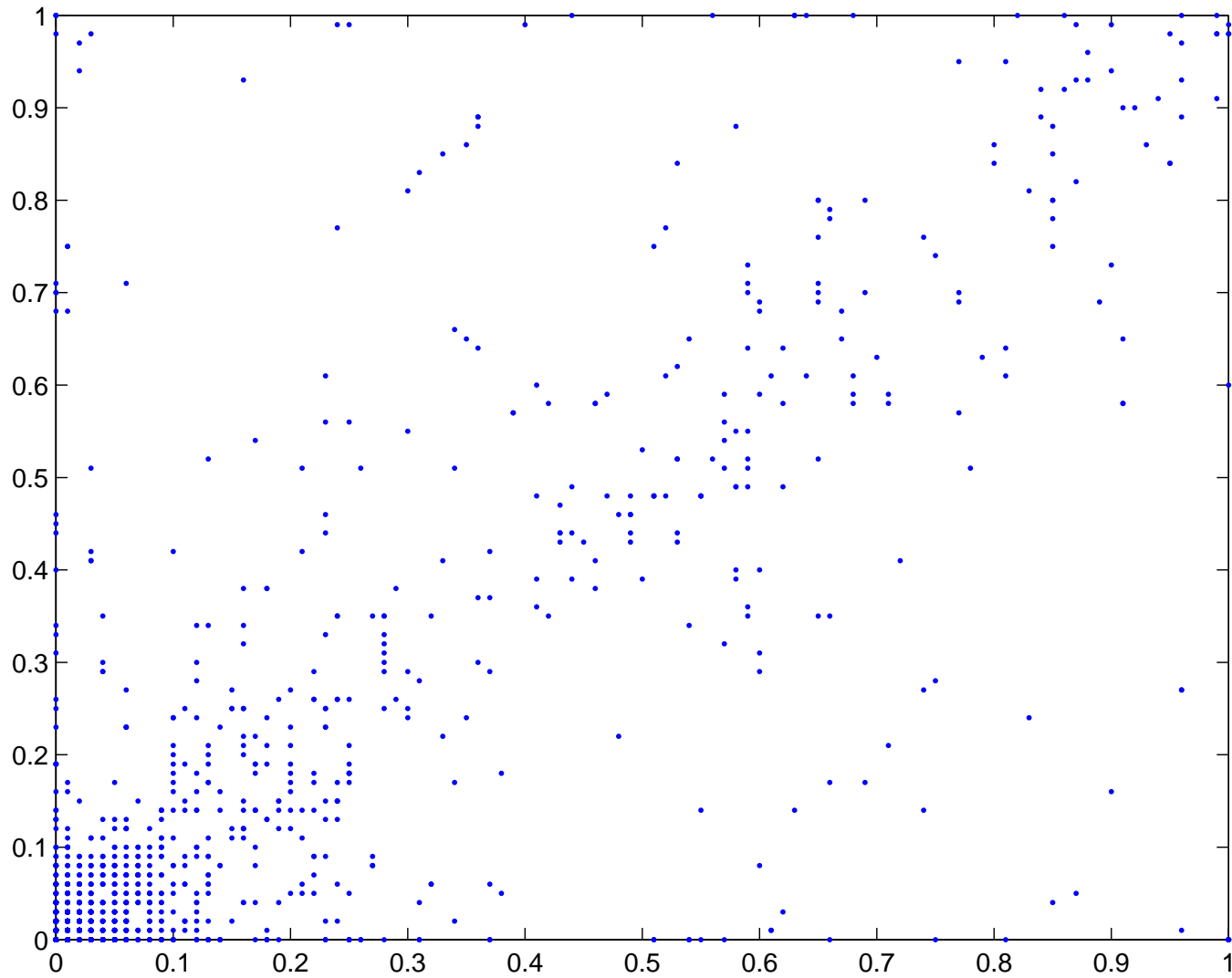
or

$P(\mathcal{M})$ uniform over cardinalities of parent sets

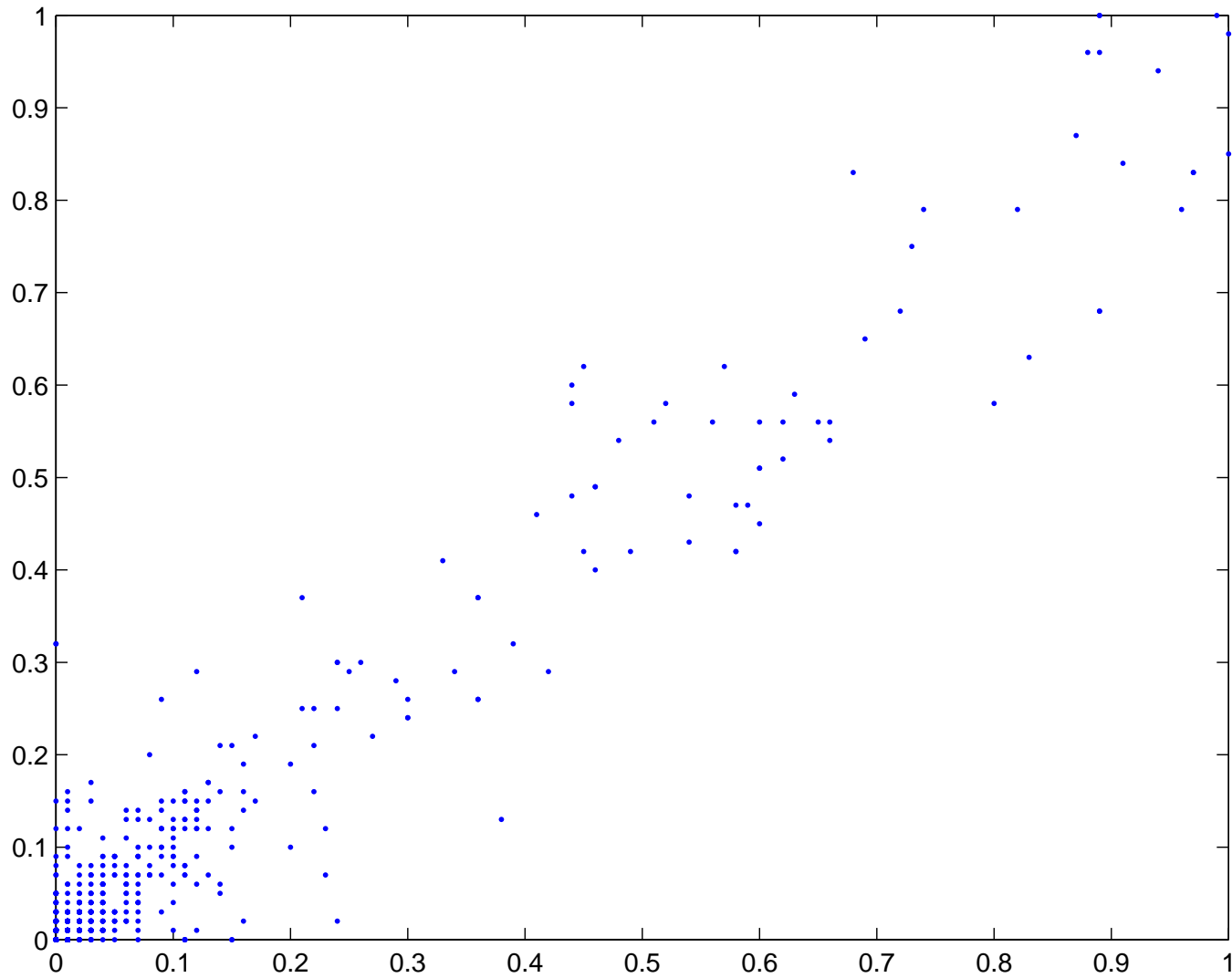
Study with Marco Grzegorzczuk and Wolfgang Urfer

- 100 genes
- 60 kidney cancer patients
- MCMC: 100,000 Metropolis-Hastings steps

$P(\mathcal{M})$ uniform over structures

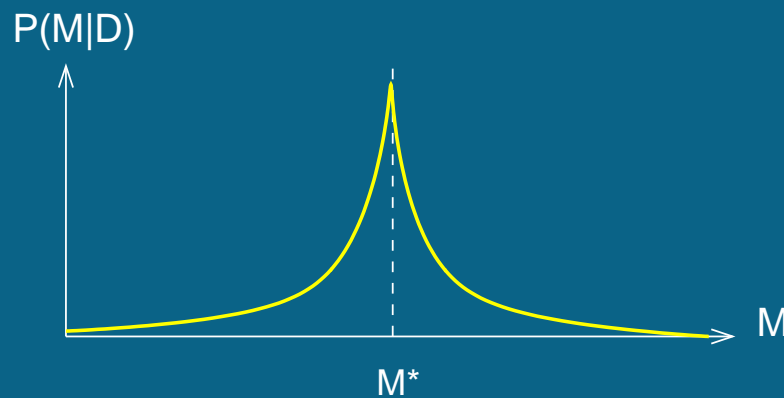


$P(\mathcal{M})$ uniform over cardinalities

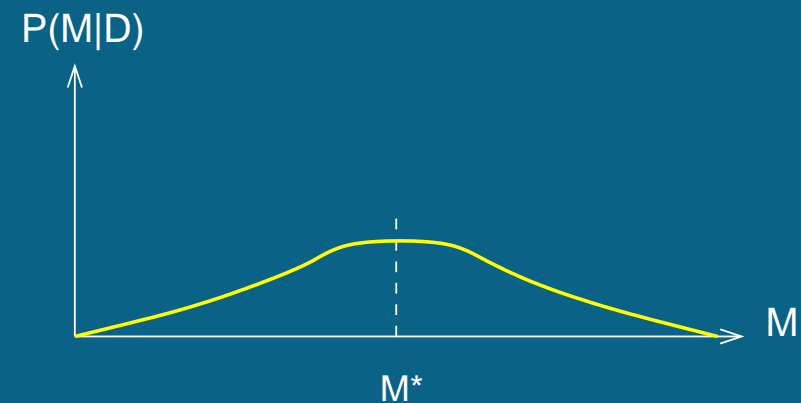


Problem: Statistical significance of the networks

- **Complex models:** Transcript levels of hundreds of genes.
- **Sparse data:** Typically a few dozen samples.



Large data set D:
Best network structure M^* well defined



Small data set D:
Intrinsic uncertainty about M^*

- Posterior probability $P(M|D)$ diffuse: **Global network** inference is **meaningless**.

Solution: Focus on **features** and **subnetworks**

Feature: Indicator variable for a property of interest,
e.g.: Are X and Y close neighbours in the network?

$$f(M) = \begin{cases} 1 & \text{if } M \text{ satisfies the feature} \\ 0 & \text{otherwise} \end{cases}$$

Solution: Focus on **features** and **subnetworks**

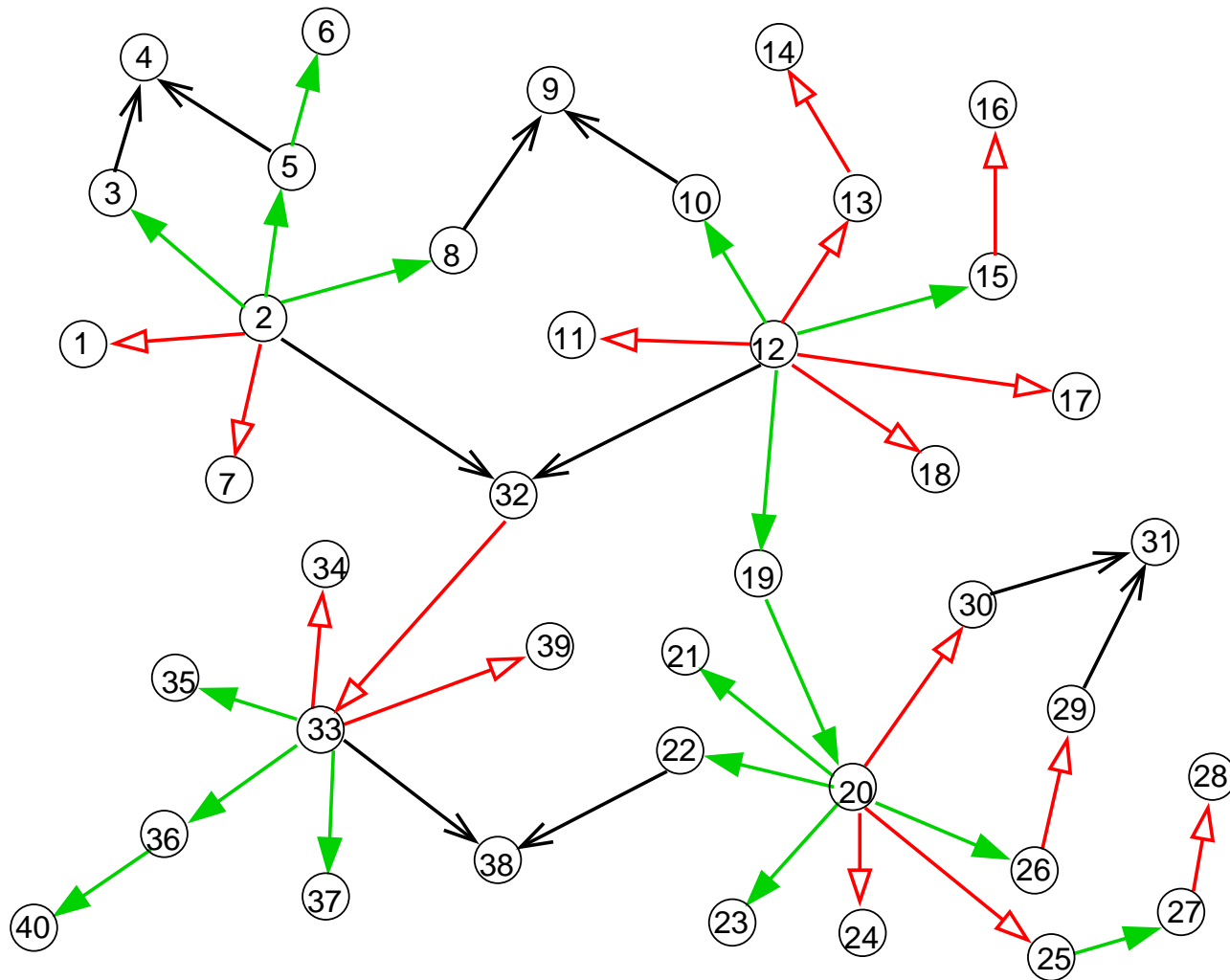
Feature: Indicator variable for a property of interest,
e.g.: Are X and Y close neighbours in the network?

$$f(M) = \begin{cases} 1 & \text{if } M \text{ satisfies the feature} \\ 0 & \text{otherwise} \end{cases}$$

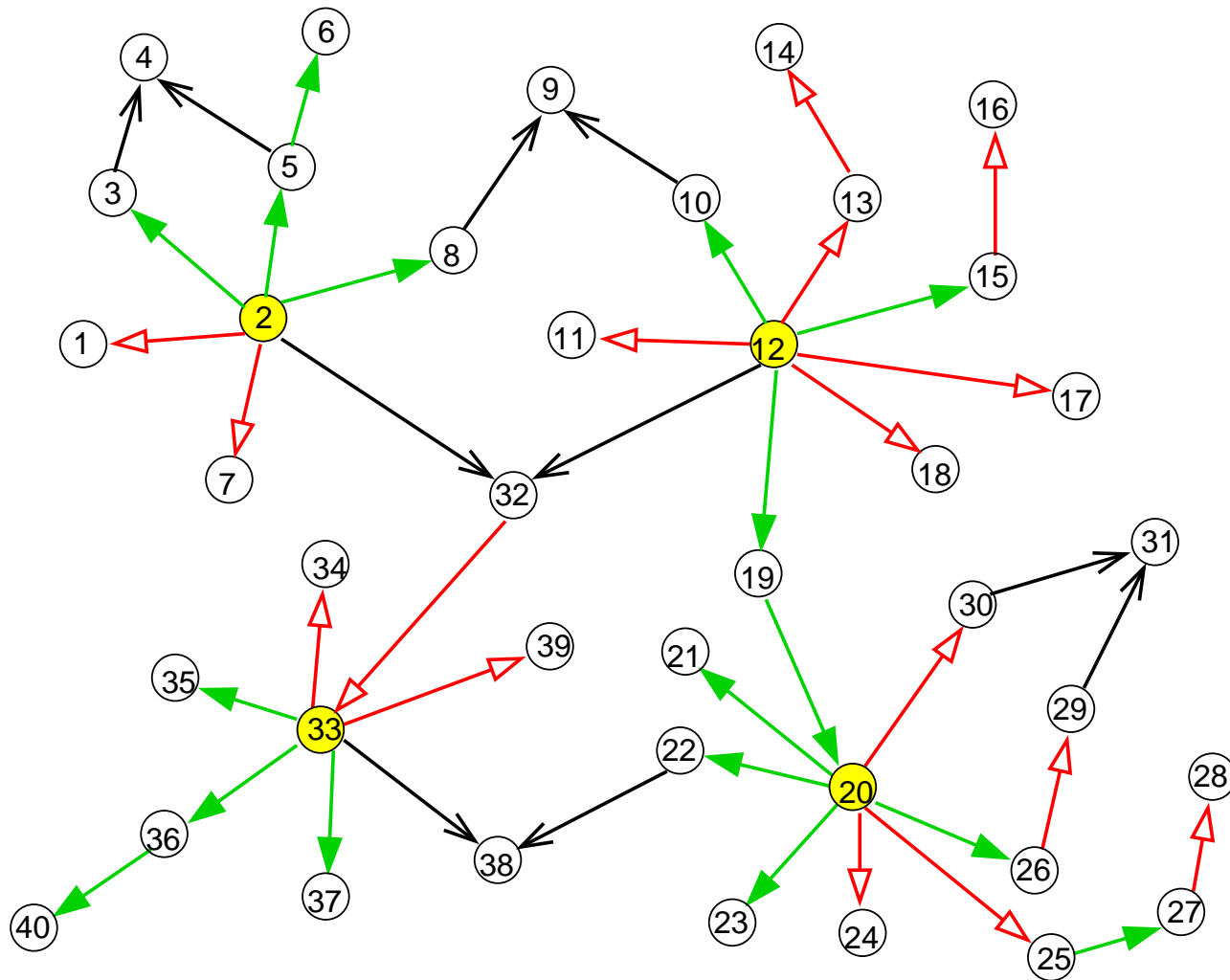
Posterior probability of features: $P(f|D) = \sum_M f(M)P(M|D)$

assumed to be sufficiently informative.

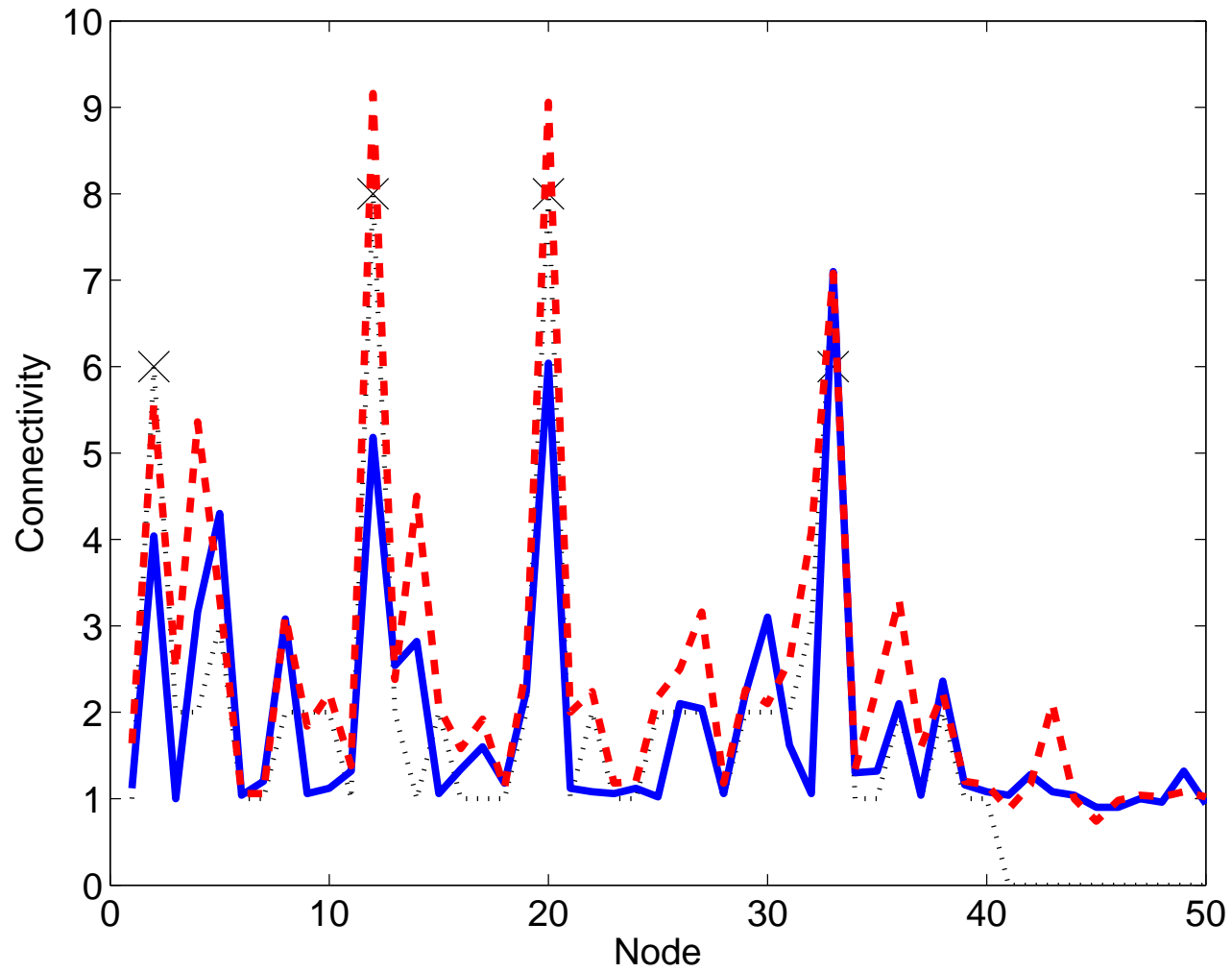
Model network, data set size: $N = 50$



Model network, data set size: $N = 50$



Predicted connectivity spectrum



Outline of the talk

- Recapitulation: Bayesian networks
- Reverse engineering:
Learning networks from data
- **Application to the yeast cell cycle**
- Estimating the accuracy of inference

Experimental Results

- Friedman, Linial, Nachman, Pe'er (2000)
Journal of Computational Biology 7 (3/4): 601-620
<http://www.cs.huji.ac.il/labs/compbio/expression/#papers>
- Pe'er, Regev, Elidan, Friedman (2001)
Bioinformatics S1: 215-224
<http://www.cs.huji.ac.il/labs/compbio/ismb01/>
- Spellman, Sherlock, Zhang, Iyer, Anders, Eisen, Brown, Botstein, Futcher (1998)
Molecular Biology of the Cell 9 (12) :3273-97
<http://cellcycle-www.stanford.edu/>

- **Yeast** cell cycle (*S. cerevisiae*).
- Six time series under different experimental conditions, altogether **76 gene expression measurements**.
- **800 genes**.
- No biological **prior knowledge**.
- Do not take into account the **temporal aspect** of the measurements. Introduce an additional root node representing the cell cycle phase.
- **Discretization**: Underexpressed (-1), normal (0), overexpressed (1).

Order relations

- Is A an **ancestor** of B in all the networks of a given equivalence class?
- Does the **network** contain a **directed path** from A to B ?
Indication that A might be a **causal ancestor** of B .

Order relations

Confidence in X being an ancestor of Y :

$$P(X \rightarrow Y | D)$$

Dominance score of X : $\sum_Y P(X \rightarrow Y | D)$

Genes with high dominance scores are **indicative** of potential **causal** sources of the cell cycle process.

Order relations

Confidence in X being an ancestor of Y :

$$P(X \rightarrow Y | D)$$

Dominance score of X : $\sum_Y P(X \rightarrow Y | D)$

Genes with high dominance scores are **indicative** of potential **causal** sources of the cell cycle process.

Finding: Only a few genes dominate the order.

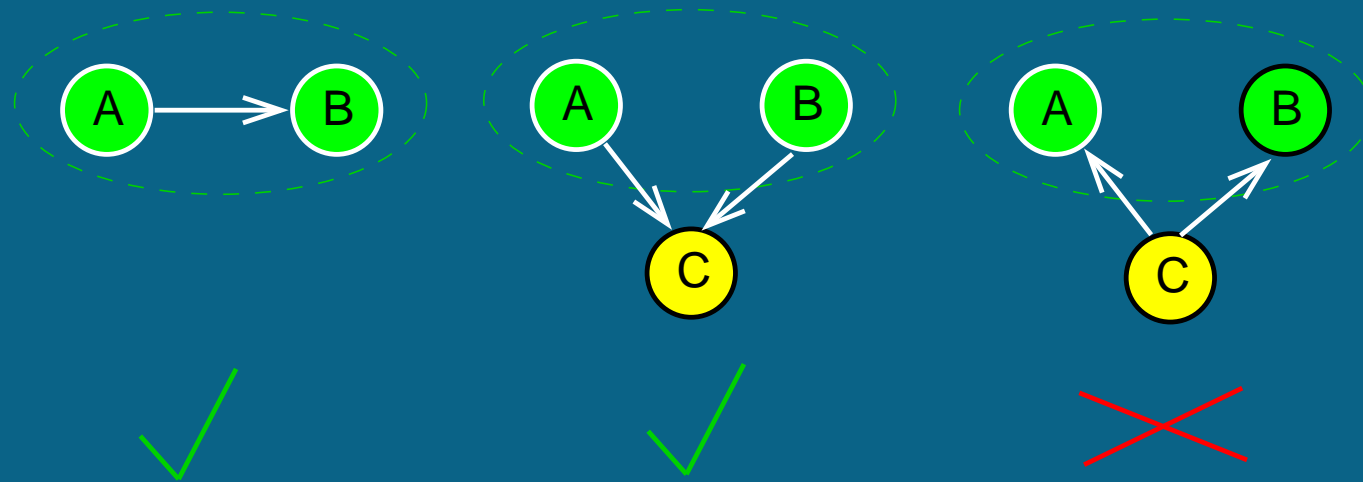
Dominant genes in the ordering relations

CLN1	Role in cell cycle start
CLN2	Role in cell cycle start
CDC5	Cell cycle control , required for exit from mitosis
RAD53	Cell cycle control : checkpoint function
RFA2	Involved in nucleotide excision repair
PLO30	Required for DNA replication and repair
MSH6	Required for mismatch repair in mitosis and meiosis

DNA repair is associated with **transcription initiation**: DNA areas which are more active in transcription are also repaired more frequently.

Markov neighbours

- Variables that are not separated by any other measured variable in the domain.



- Indication that two genes are related in some **joint biological interaction or process**.
- **Parent-child:** One gene regulating another.
- **Spouse relations:** Two genes co-regulating another.

Markov relations

$P(X \leftrightarrow Y|D)$: Indication that genes are **functionally related**.

- Most Markov pairs: **Intracluster pairings** with high correlation in their expression.
- **But:** Genes where $P(X \leftrightarrow Y|D)$ is high and correlation is low.

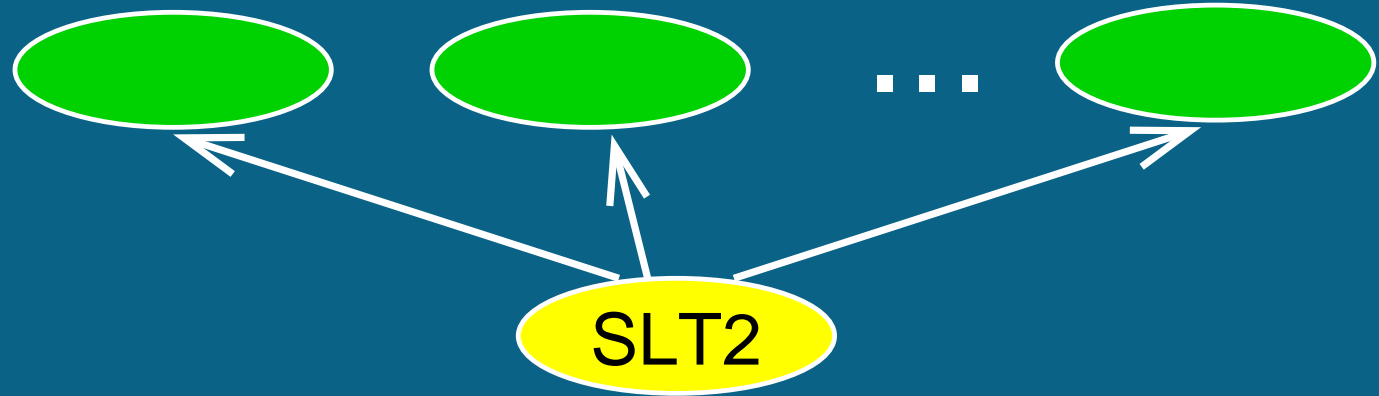
FAR1	Role in a mating type switch
ASH1	Role in a mating type switch
LAC1	GPI transport protein
YNL300W	Modified by GPI
SAG1	Induces the mating process
MF-ALPHA-1	Participates in the mating process

Separator relations

and

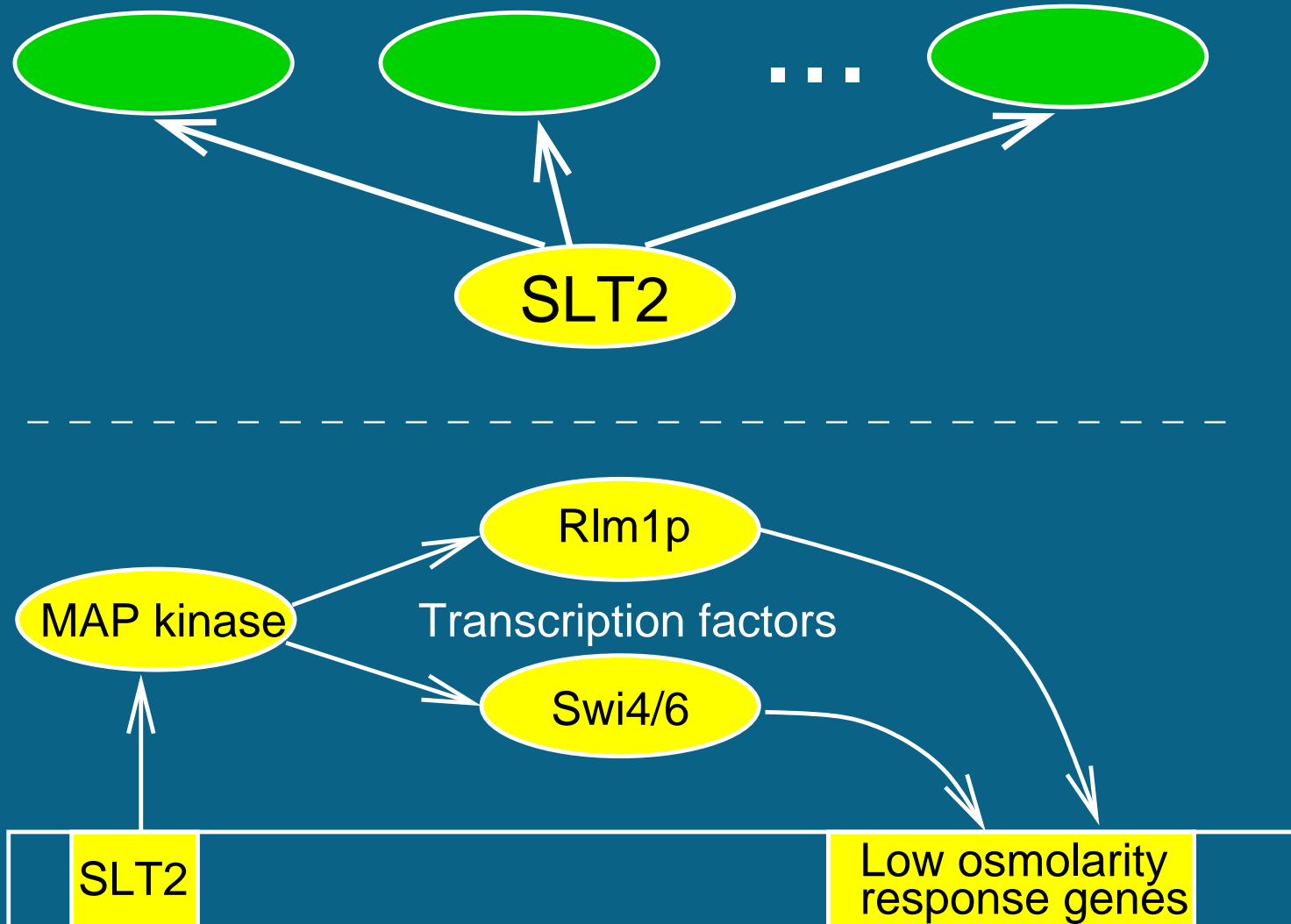
subnetworks

Low osmolarity response genes



	SLT2		Low osmolarity response genes	
--	------	--	-------------------------------	--

Low osmolarity response genes

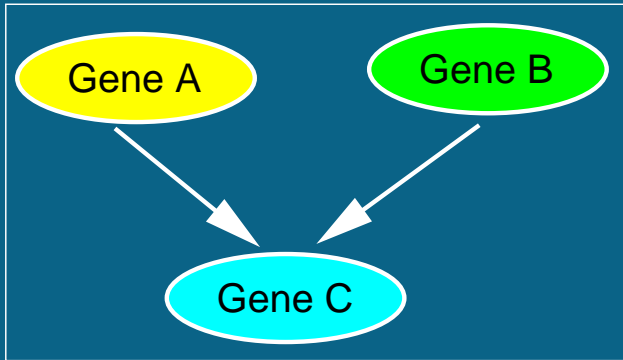


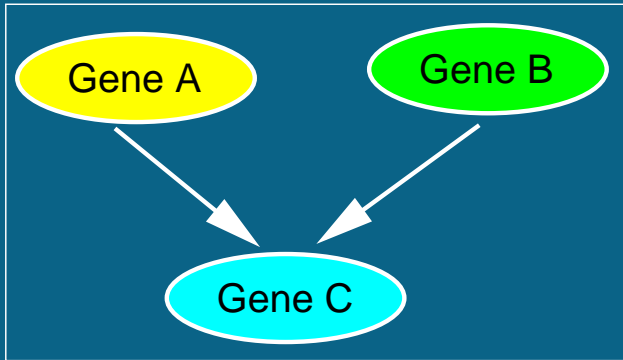
Outline of the talk

- Recapitulation: Bayesian networks
- Reverse engineering:
Learning networks from data
- Application to the yeast cell cycle
- **Estimating the accuracy of inference**

Dirk Husmeier (2003)

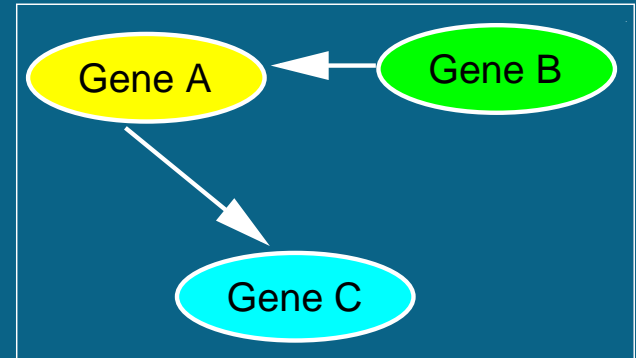
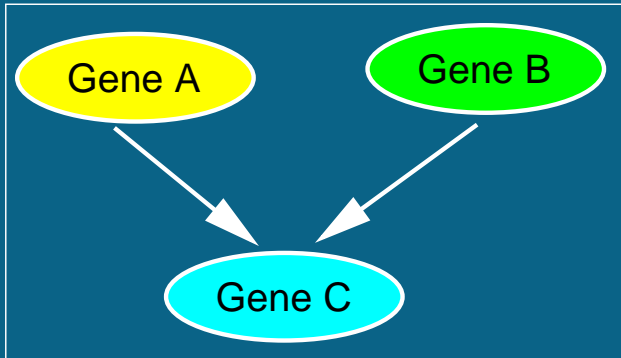
Bioinformatics 19, 2271-2282





generate

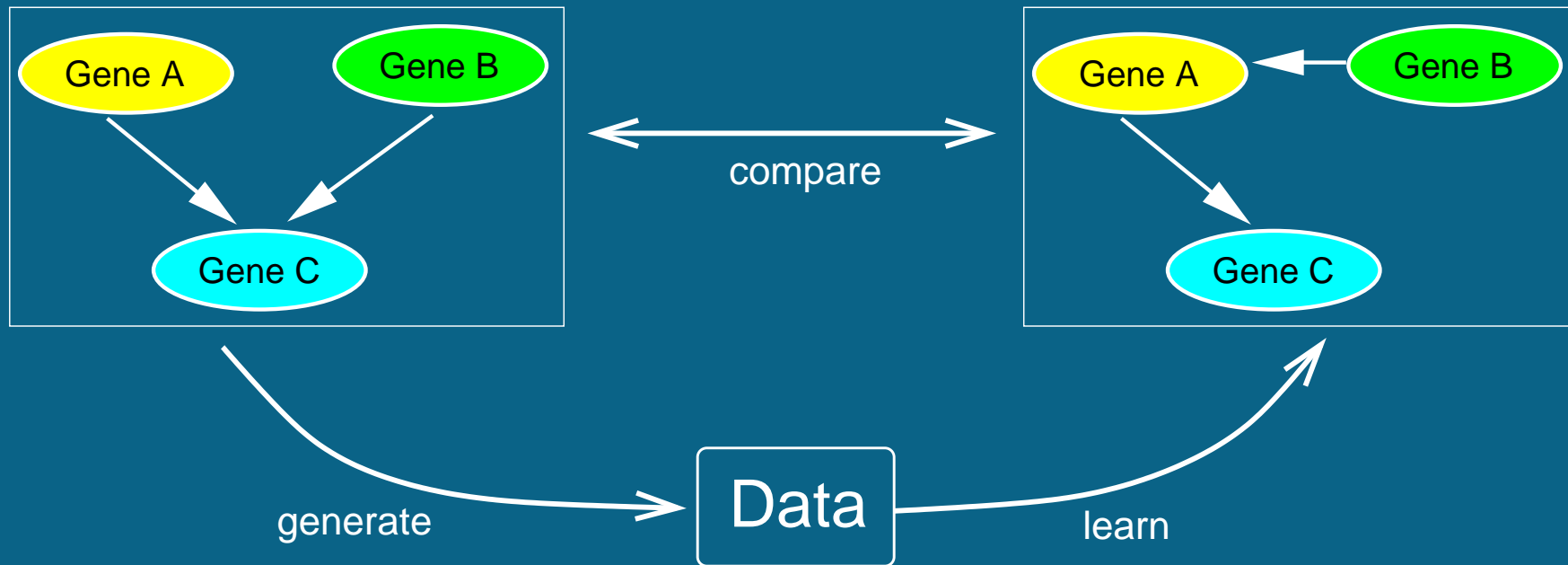
Data

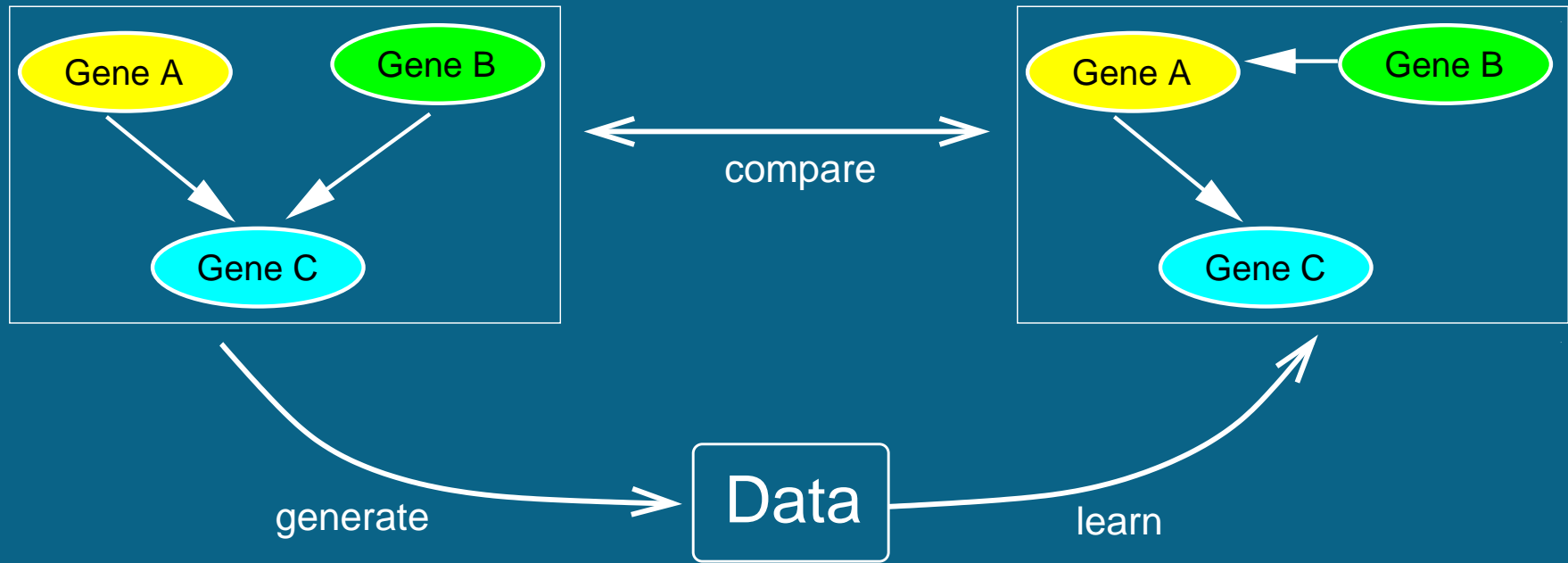


generate

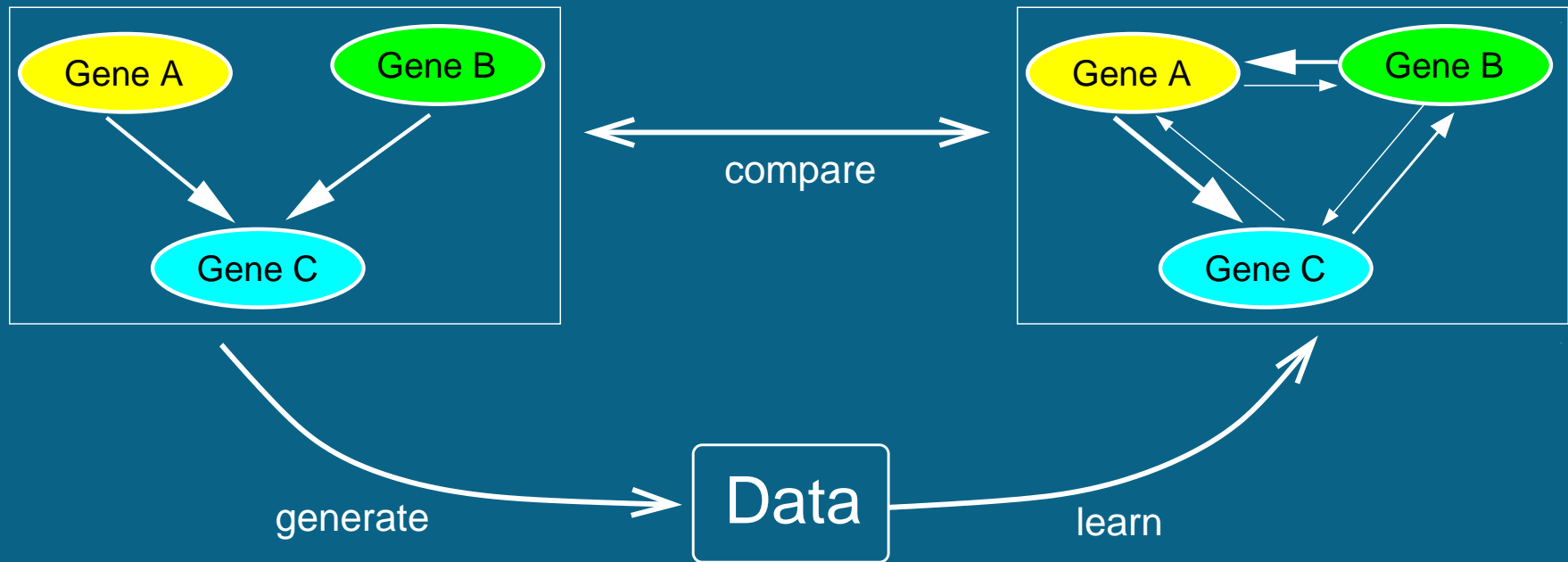
Data

learn

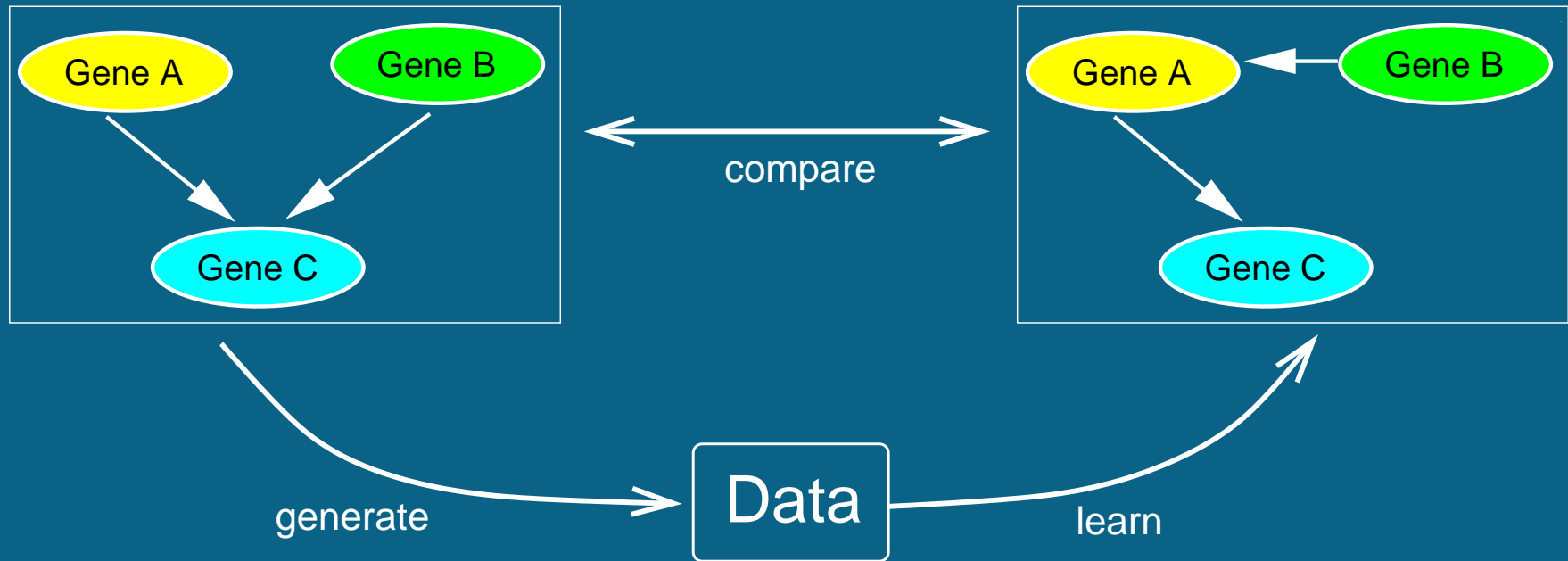




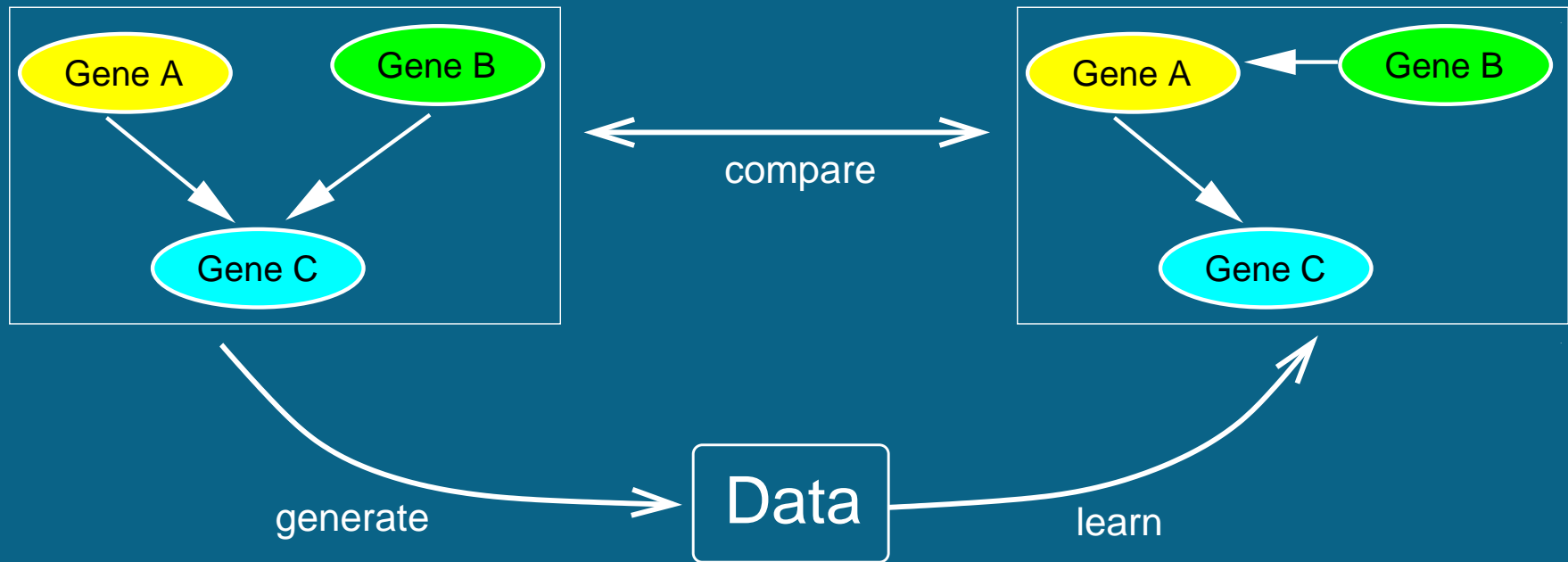
Deterministic inference



Probabilistic inference



Thresholding

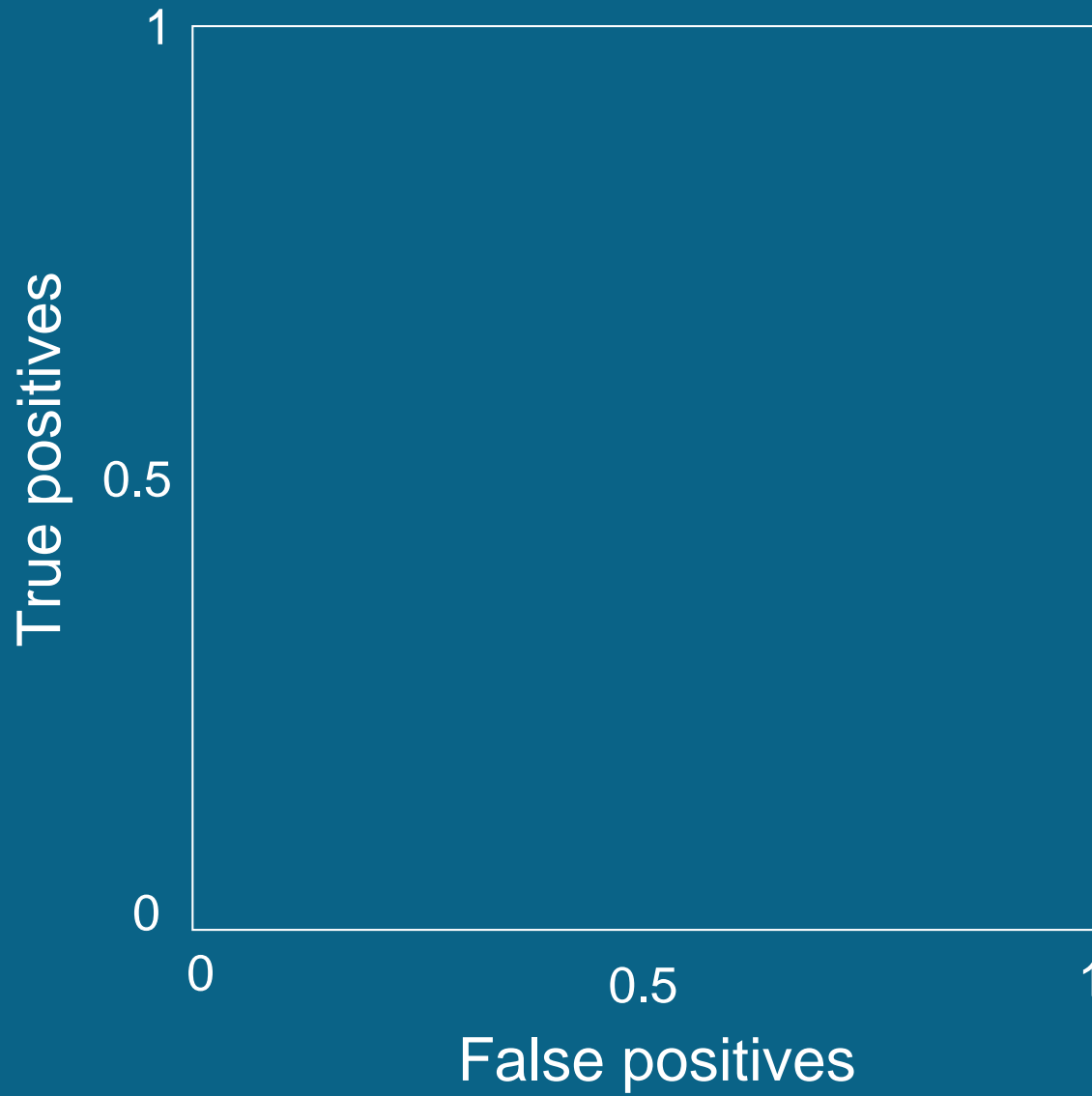


Thresholding

True positives

False positives

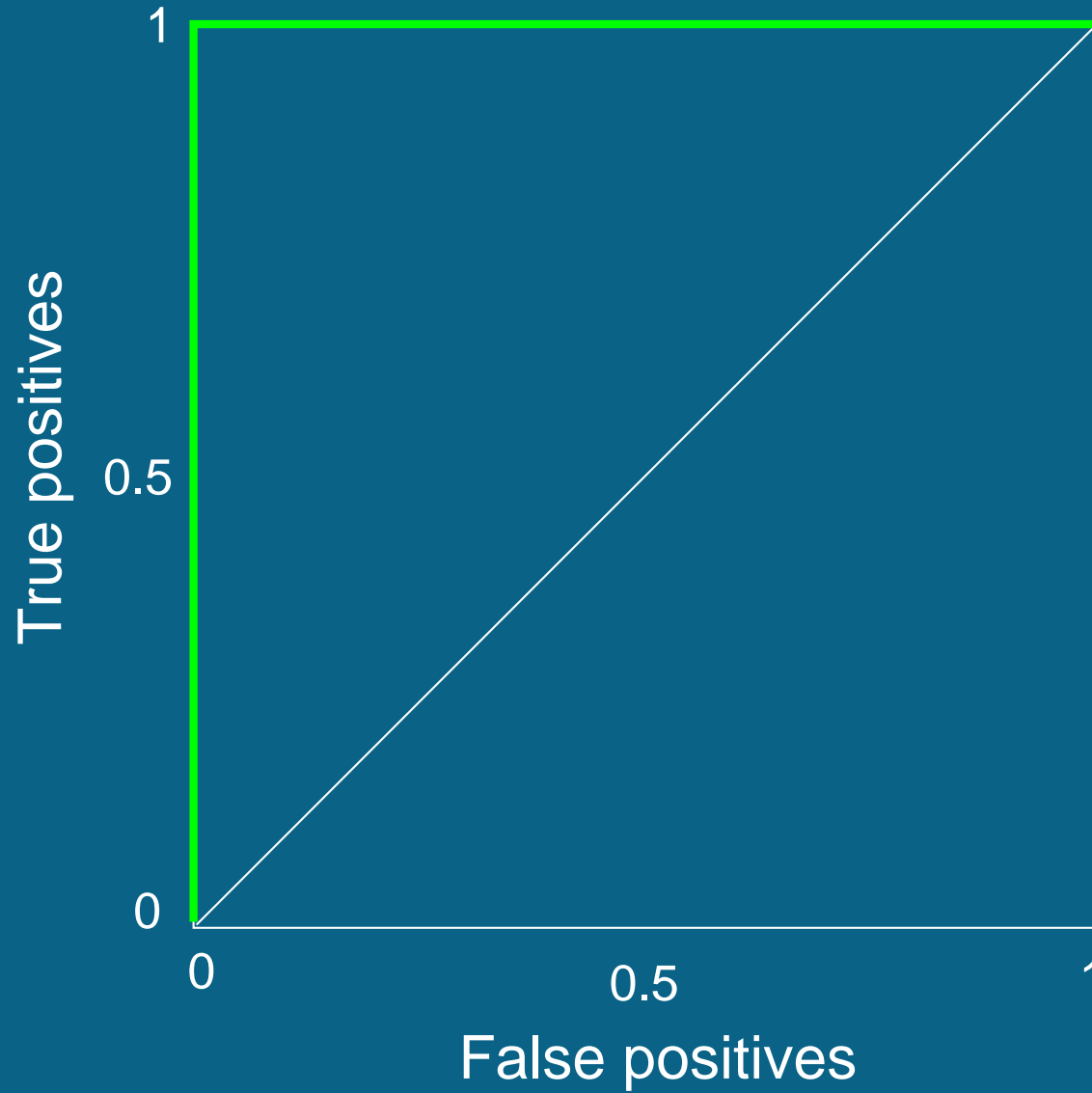
ROC curve



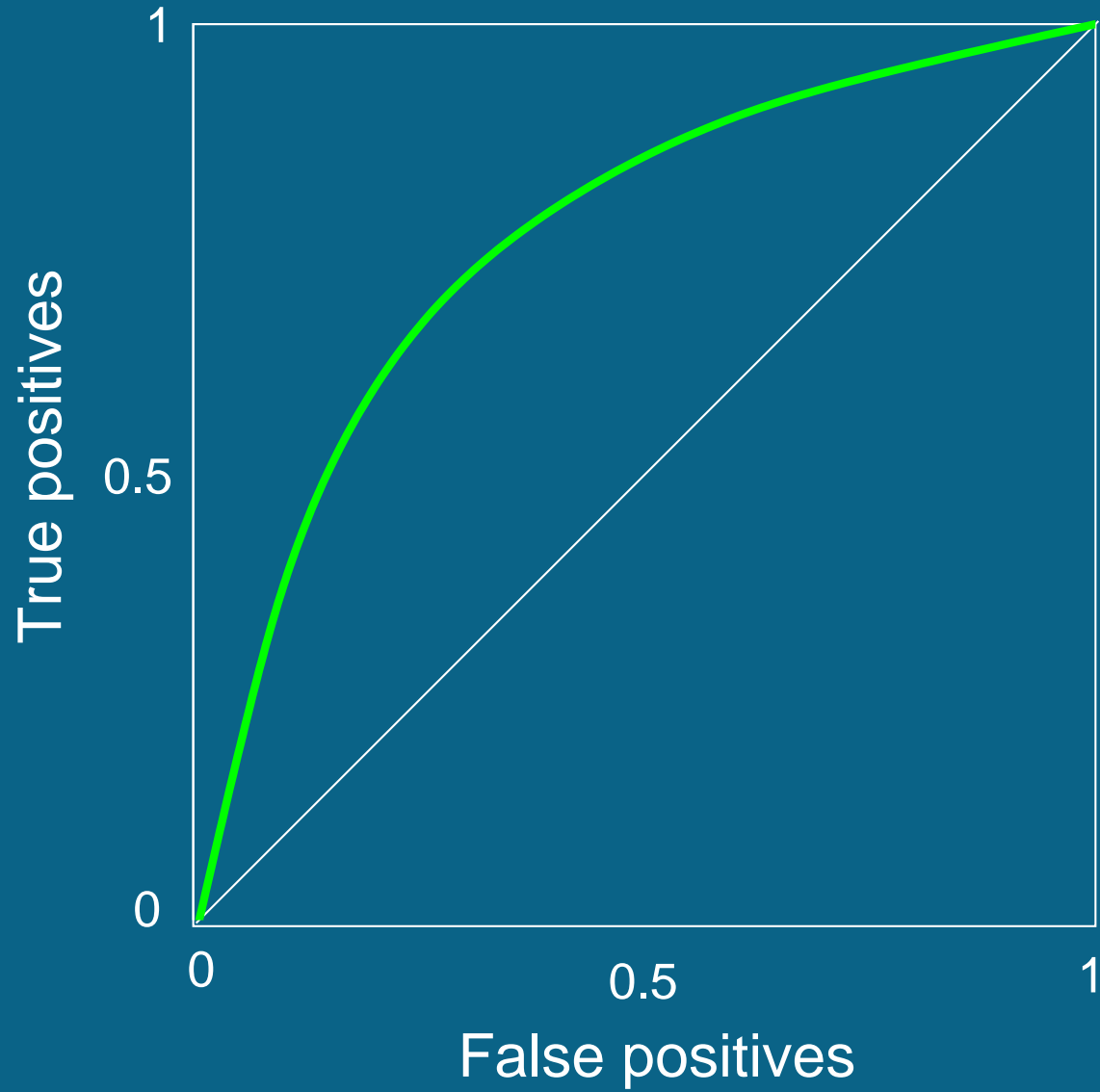
Random predictor



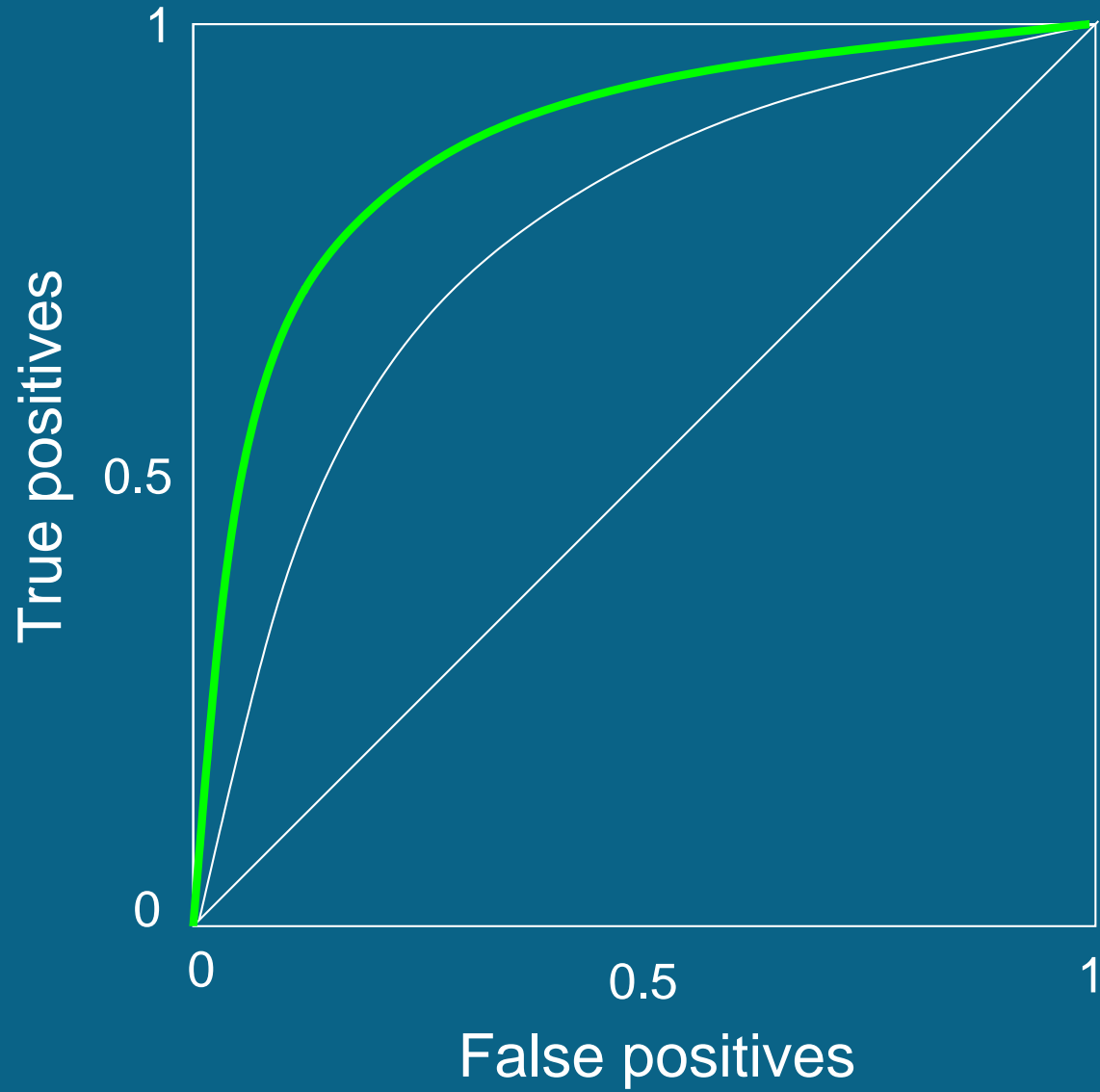
Perfect predictor



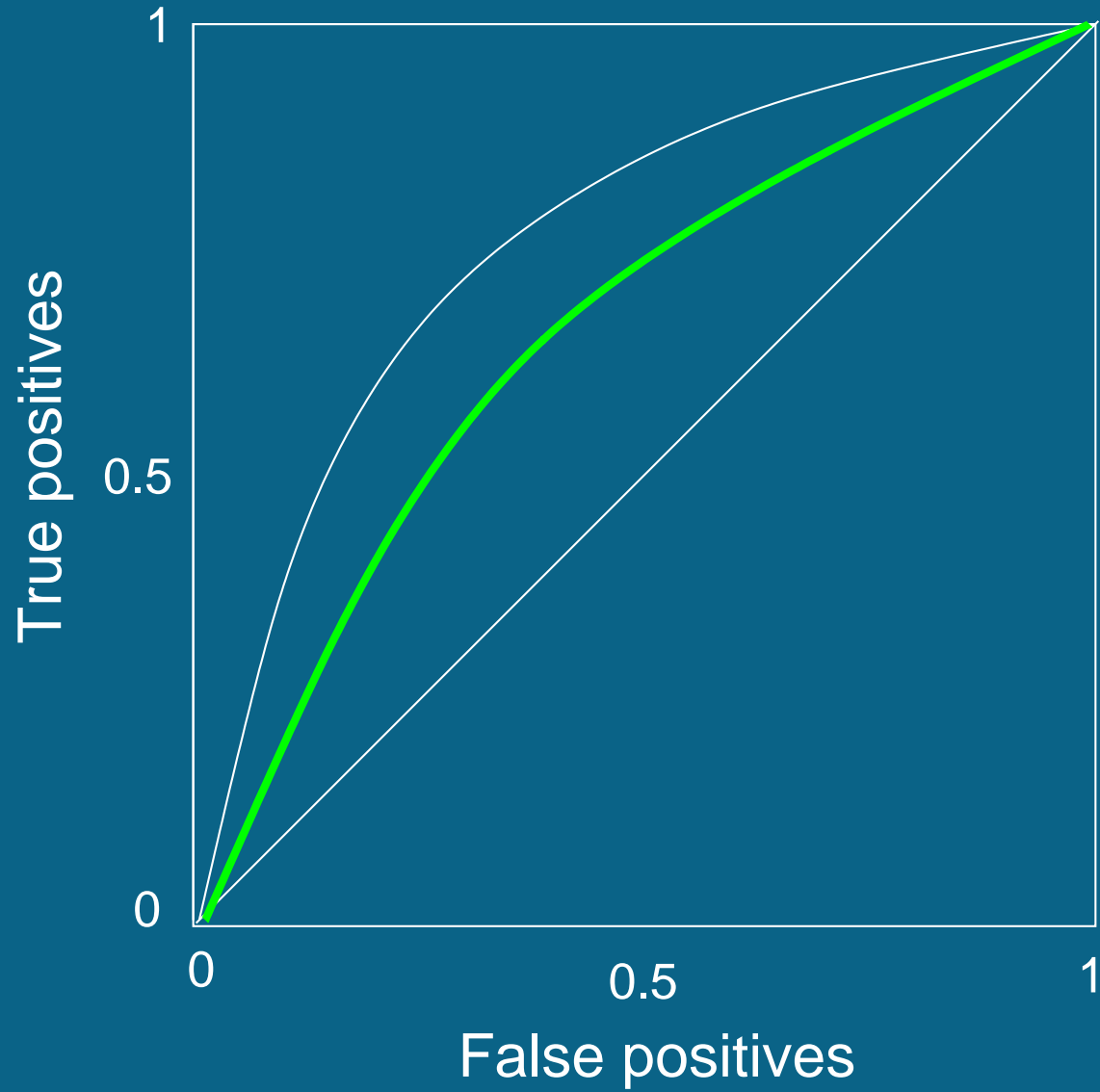
Realistic predictor



Better predictor

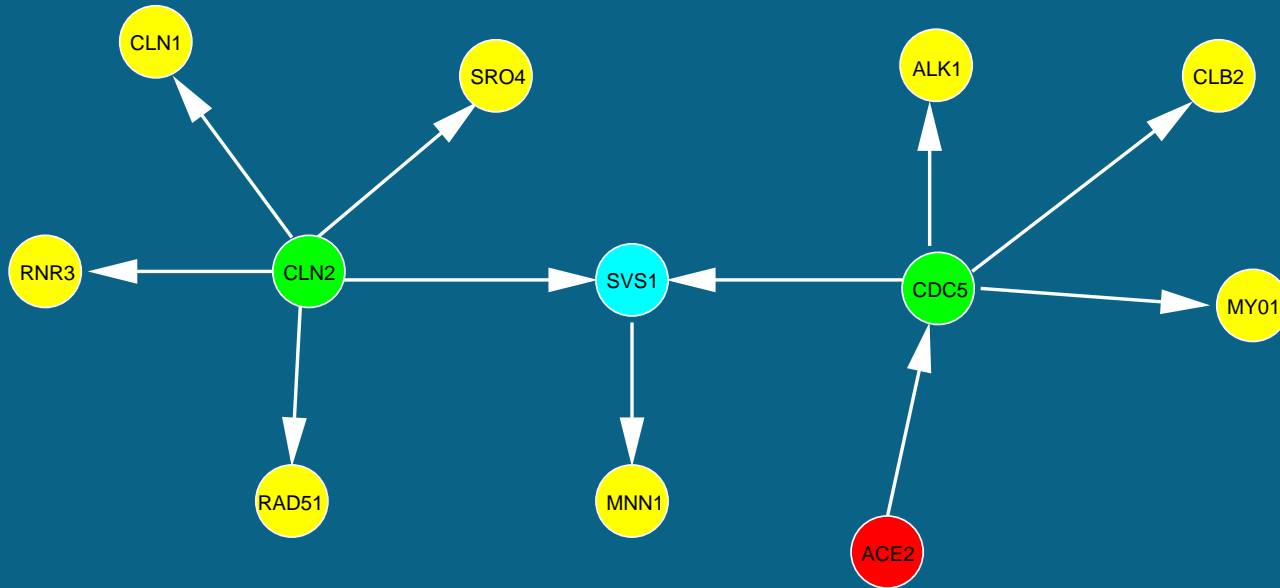


Poorer predictor



Data: binary

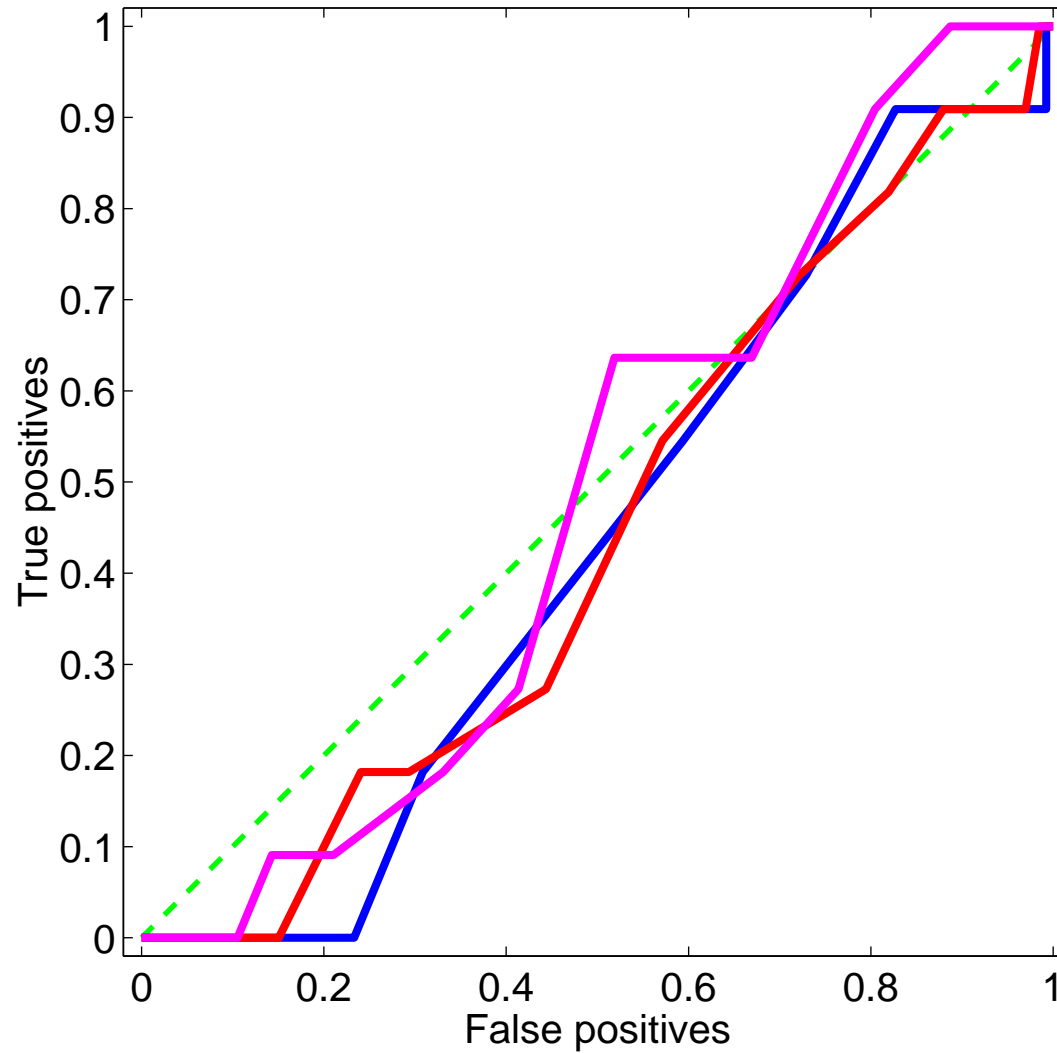
Model:



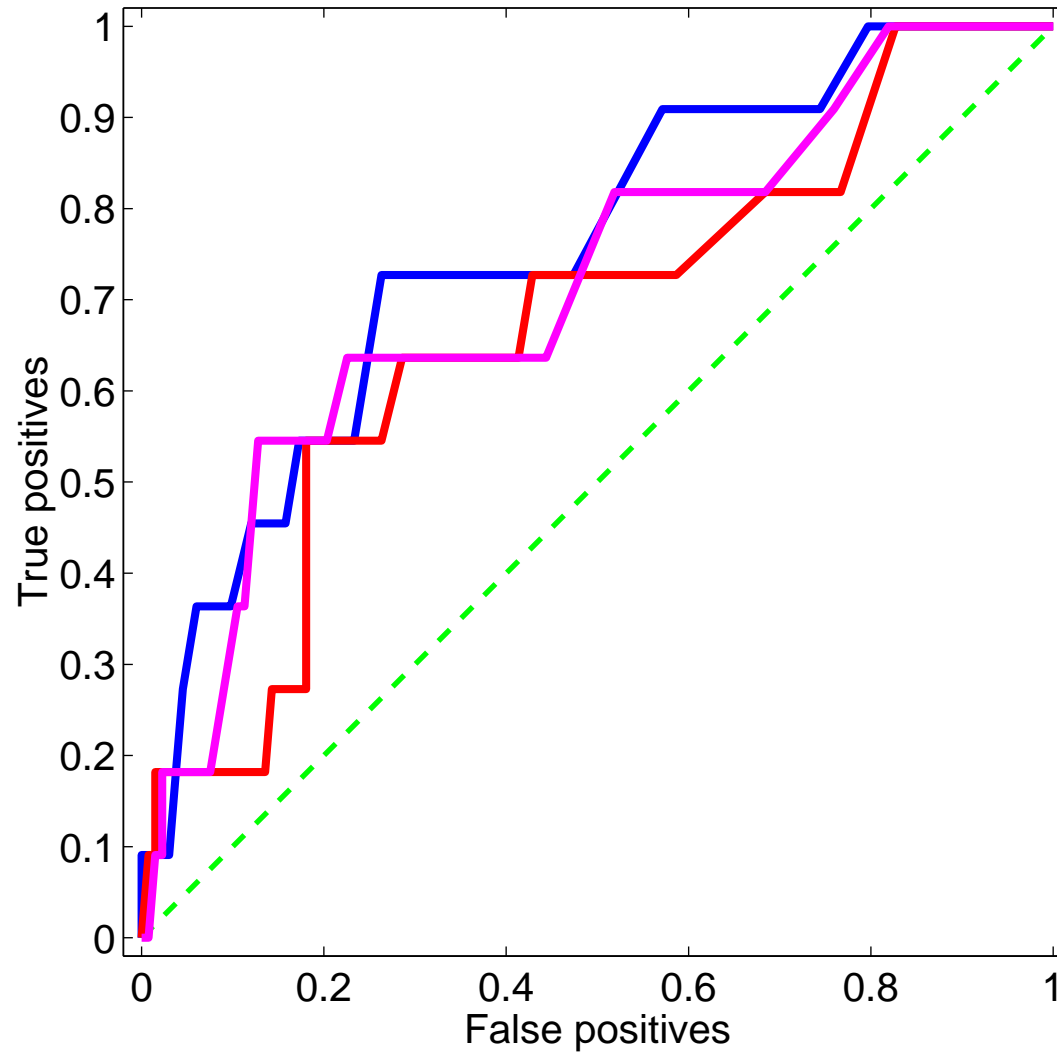
Parameters:

Noisy boolean: $P \in \{0.1, 0.9\}$

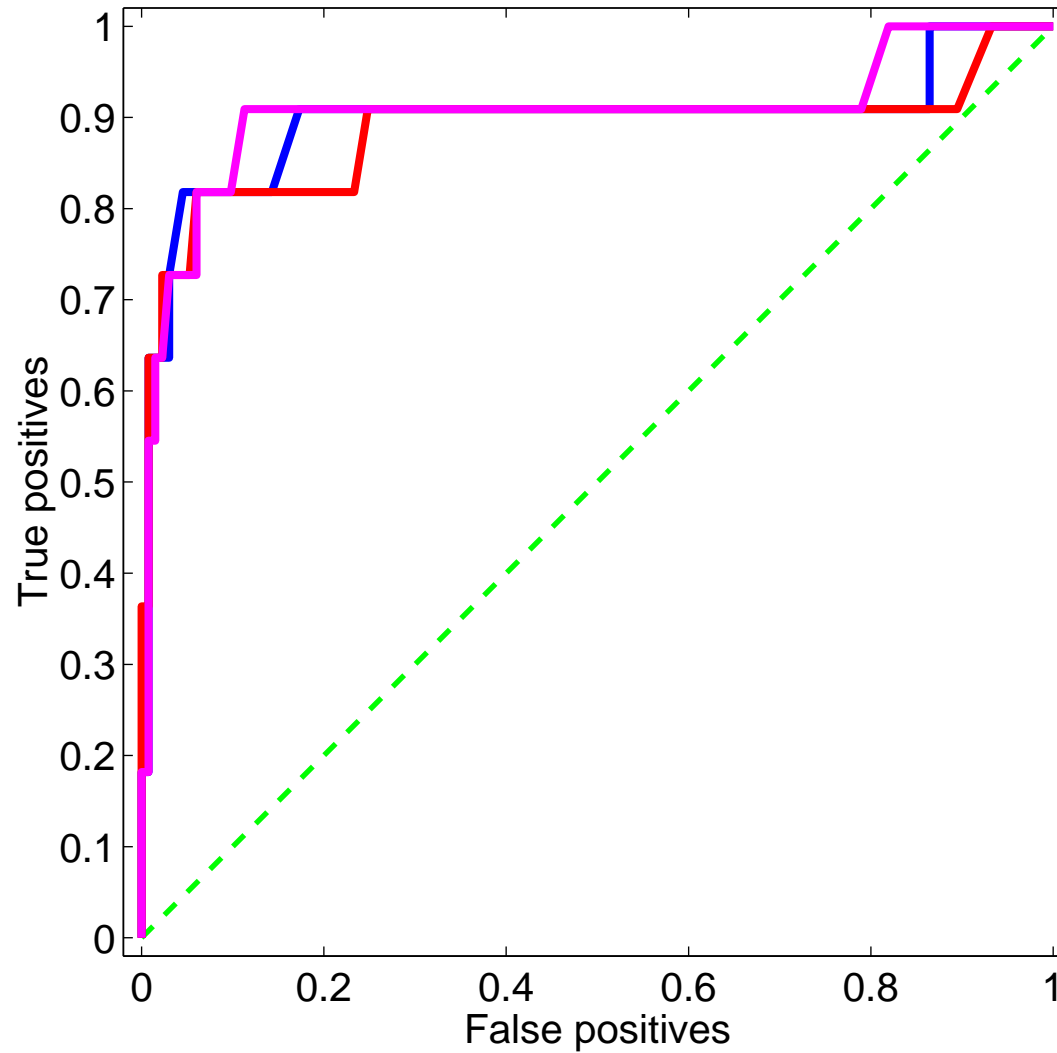
ROC curve: Sample size= 3



ROC curve: Sample size= 6

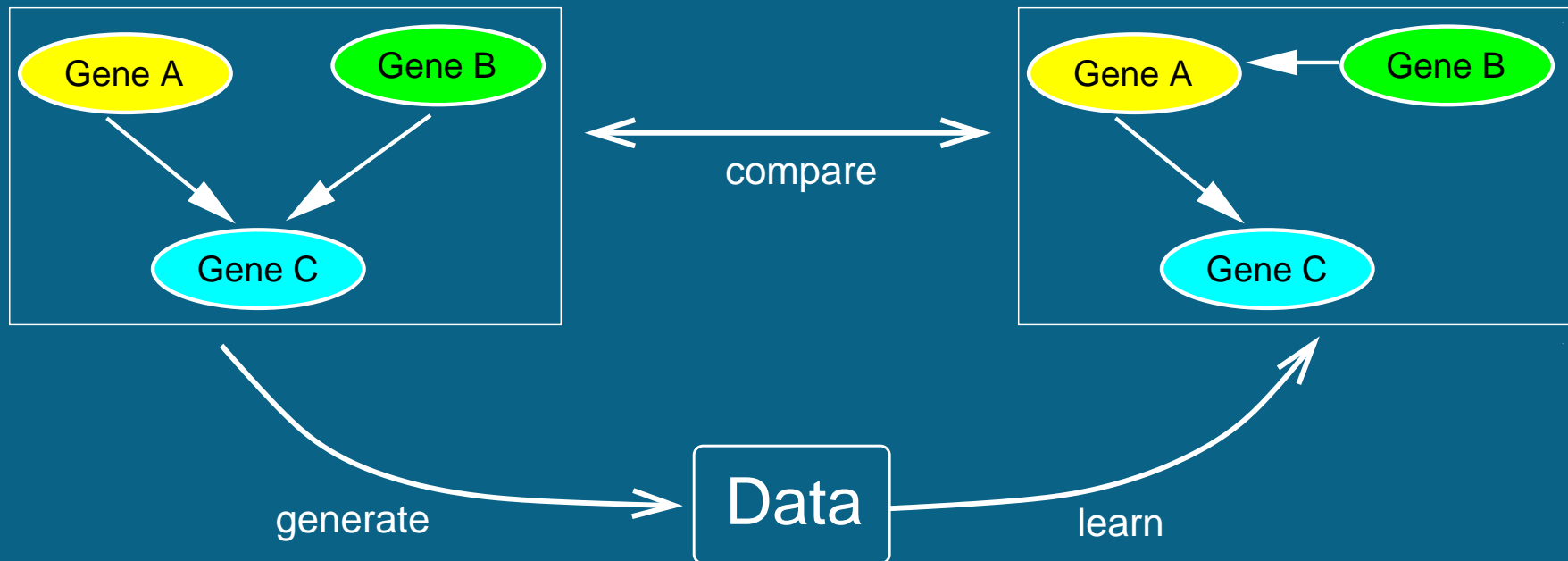


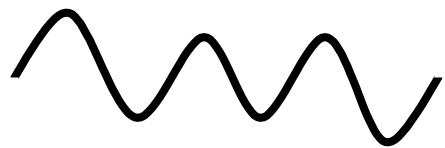
ROC curve: Sample size= 12



Disadvantage:

Unrealistic, **no mismatch** between the model used for **data generation** and the model used for **inference**.

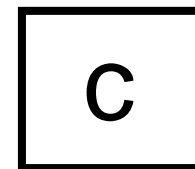
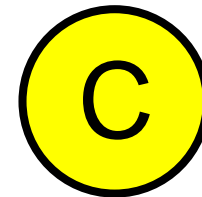




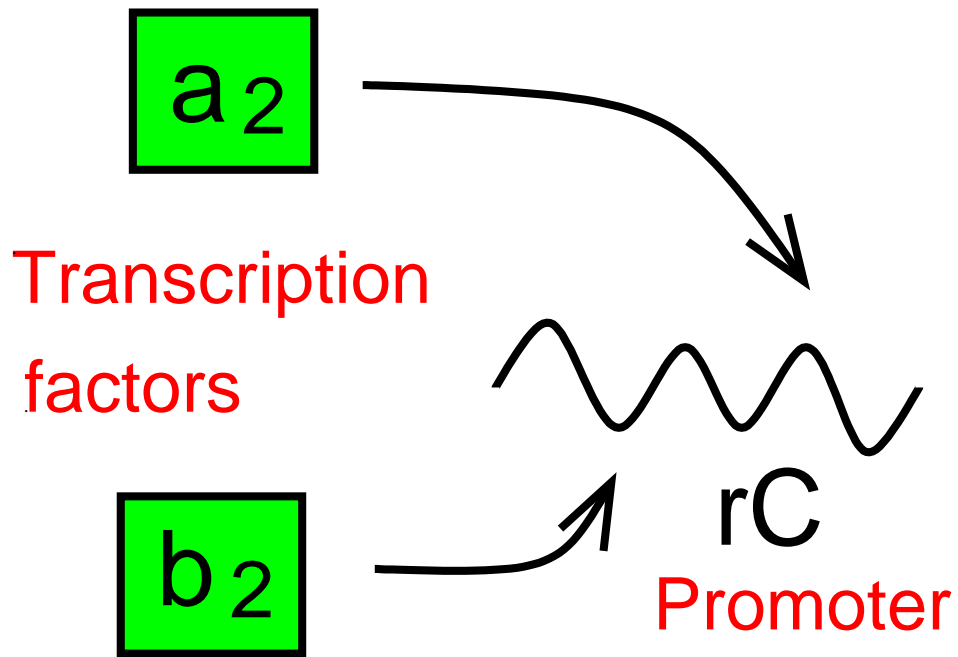
rC

Promoter

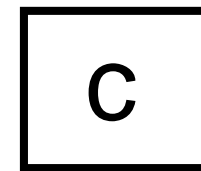
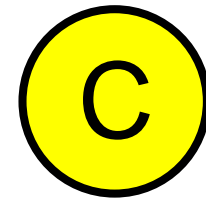
mRNA



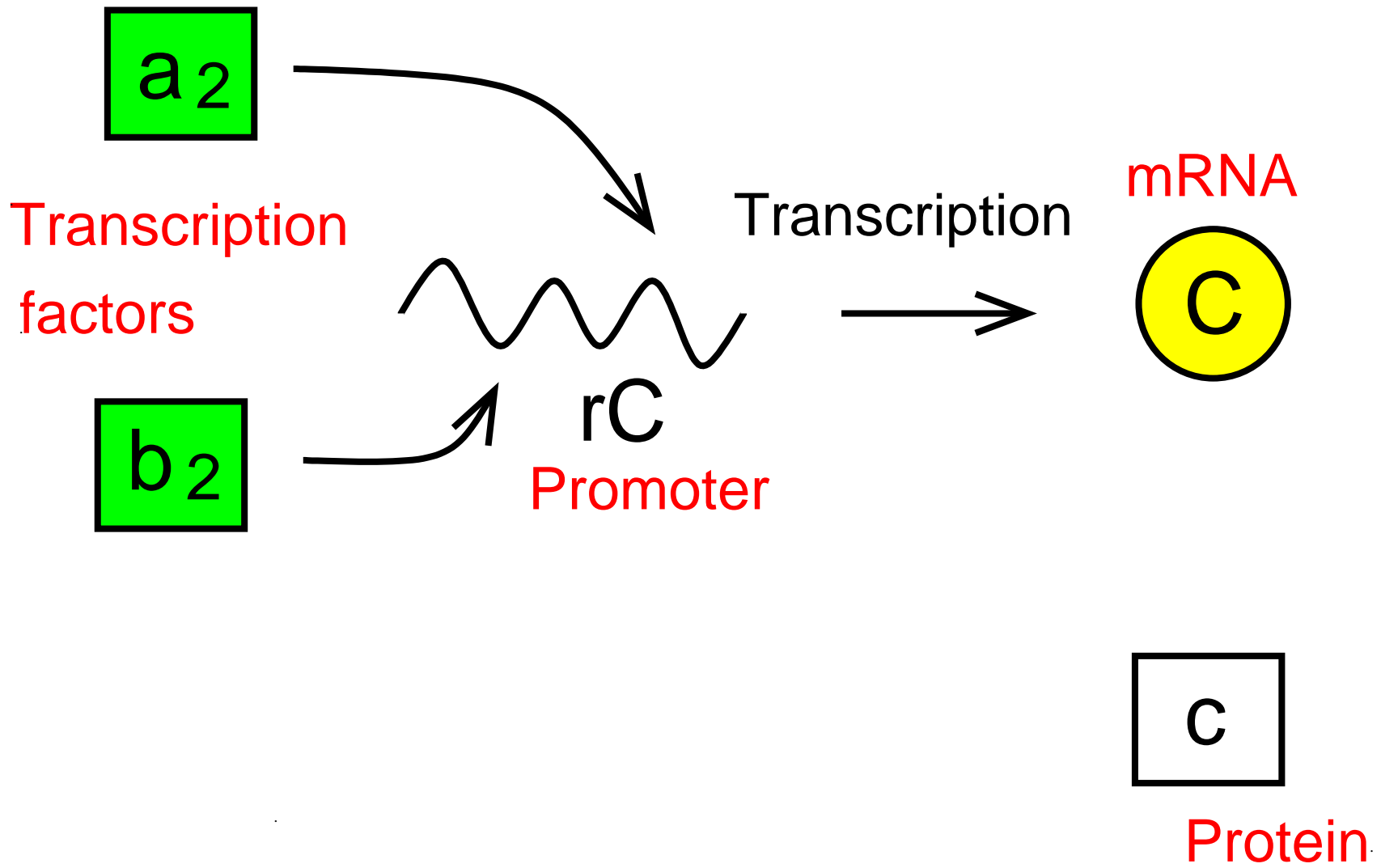
Protein

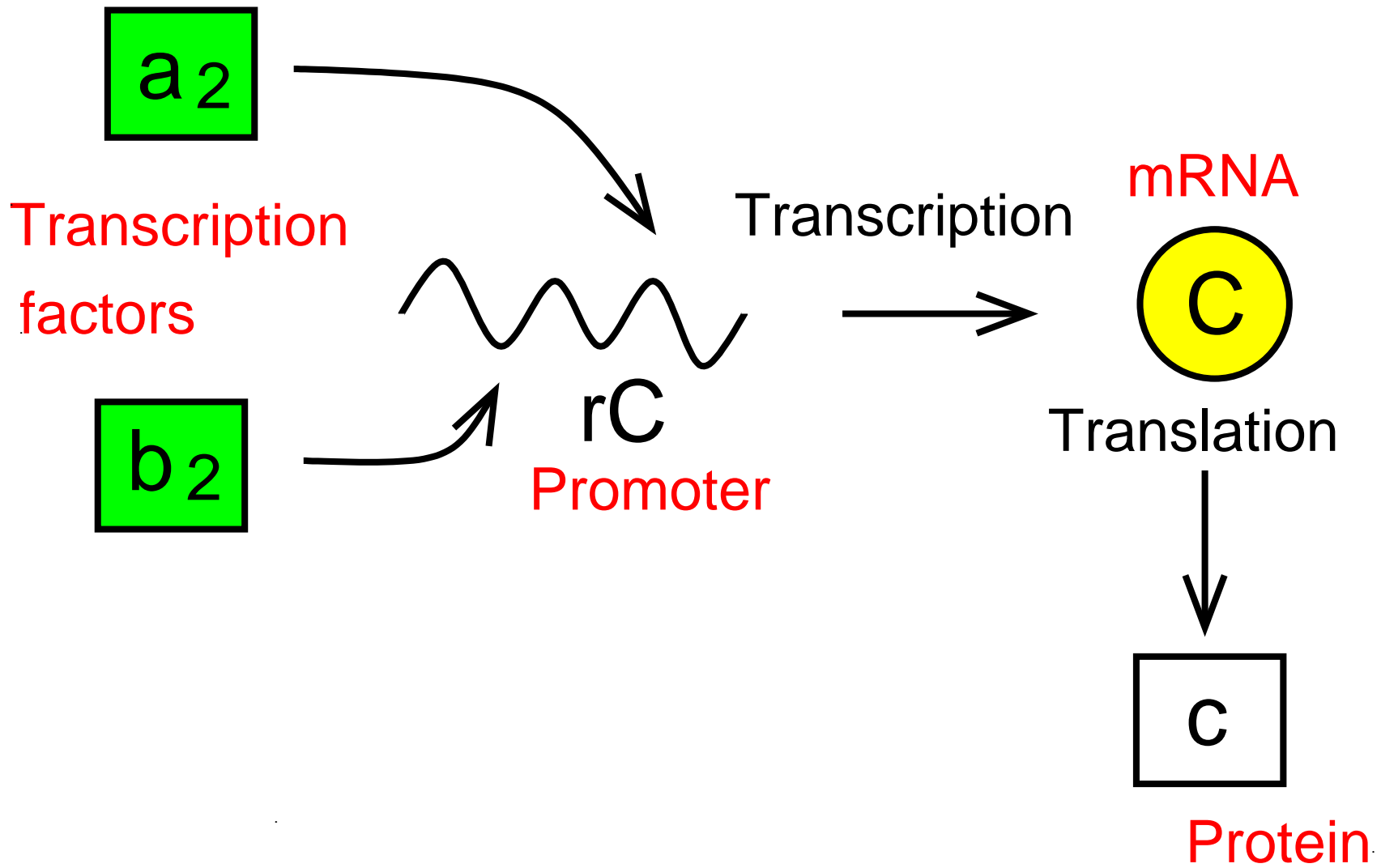


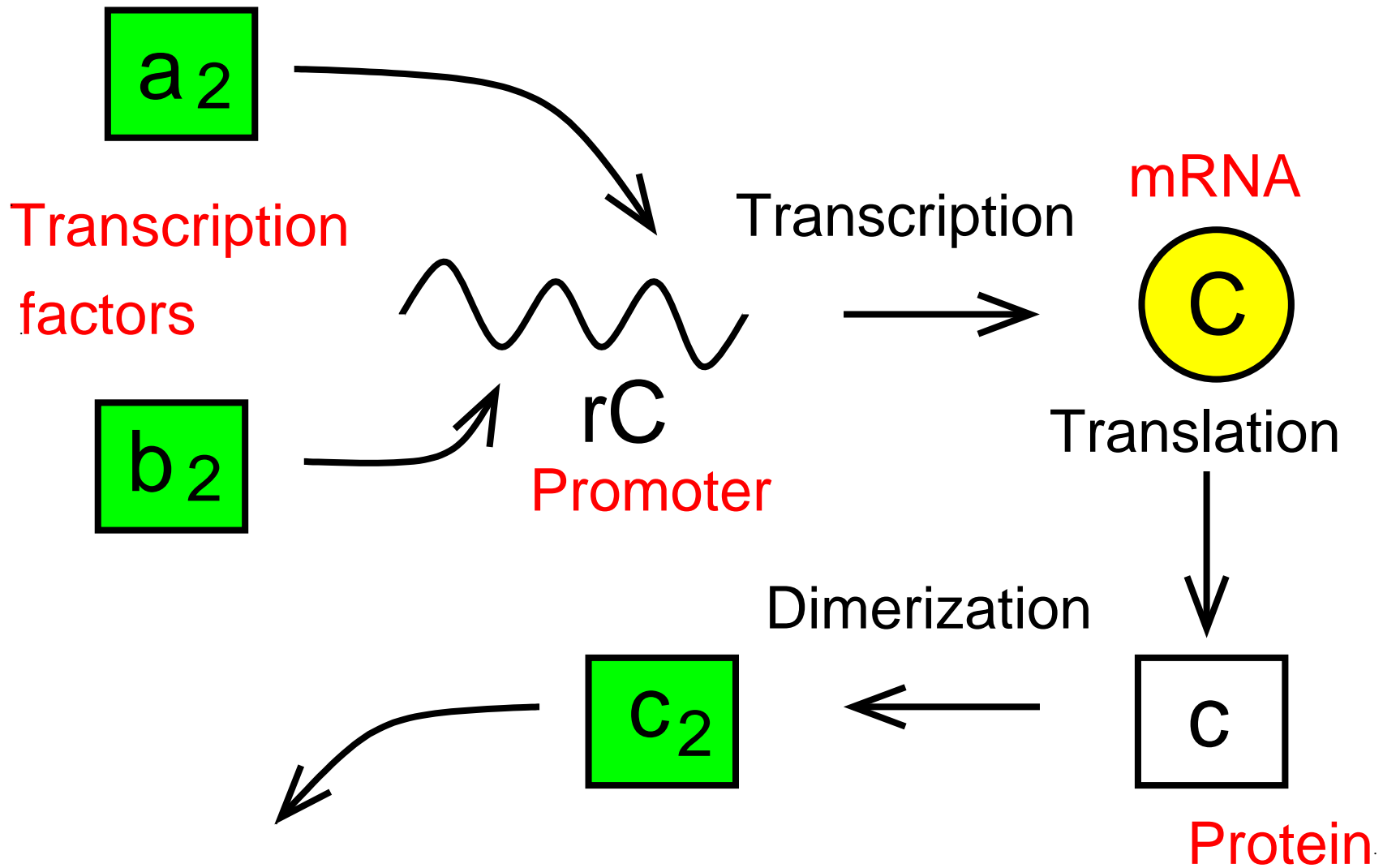
mRNA

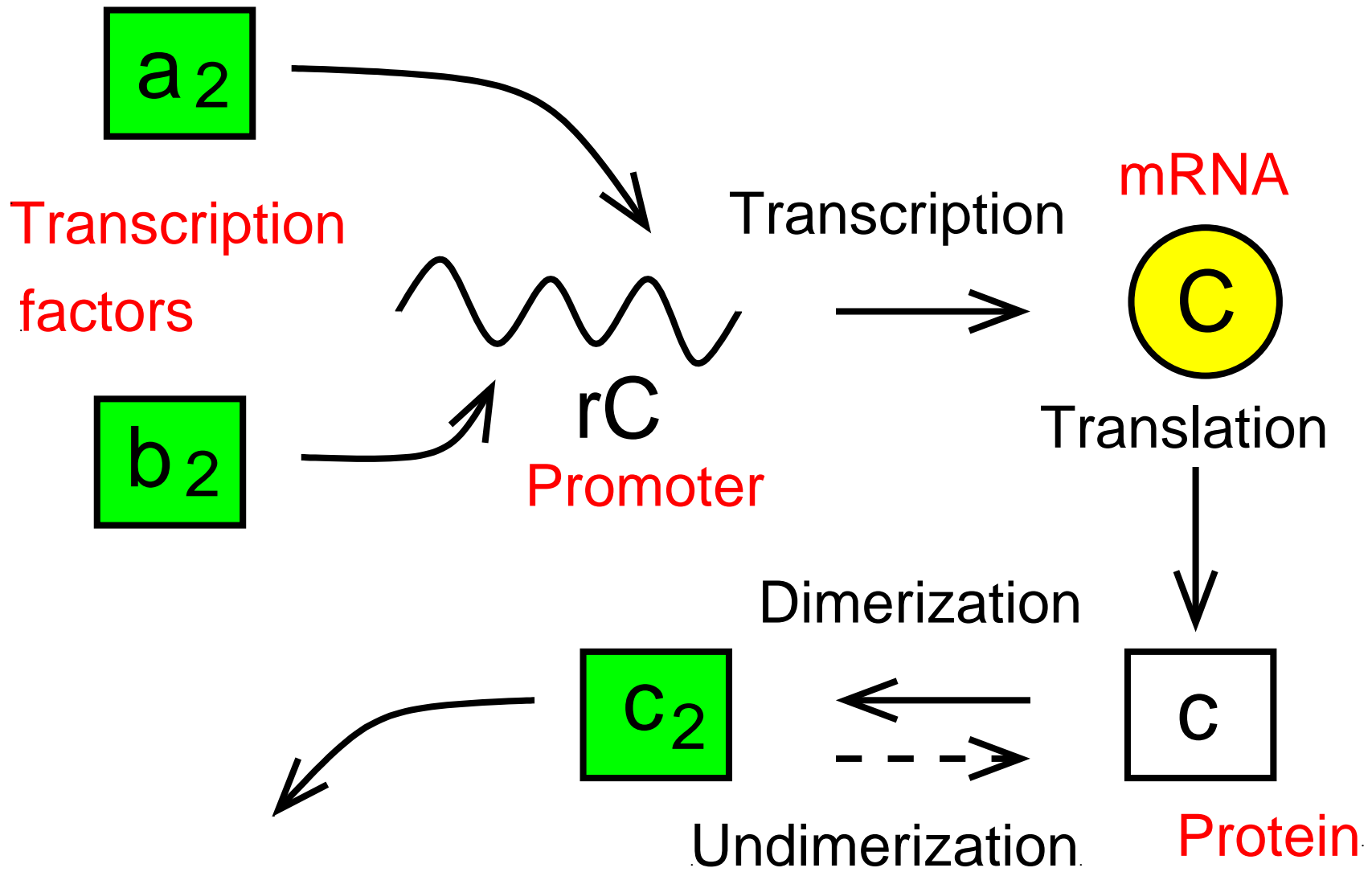


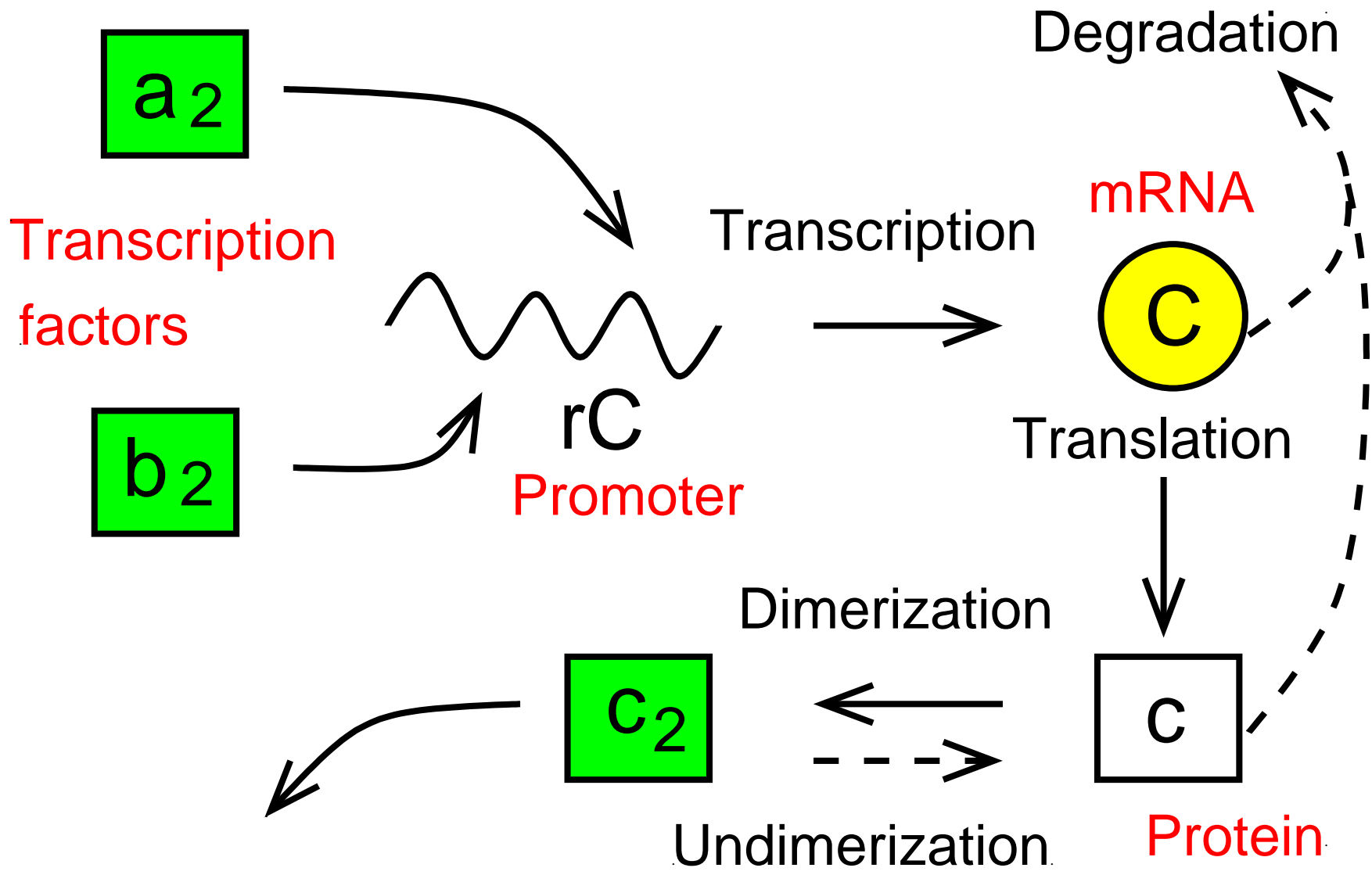
Protein

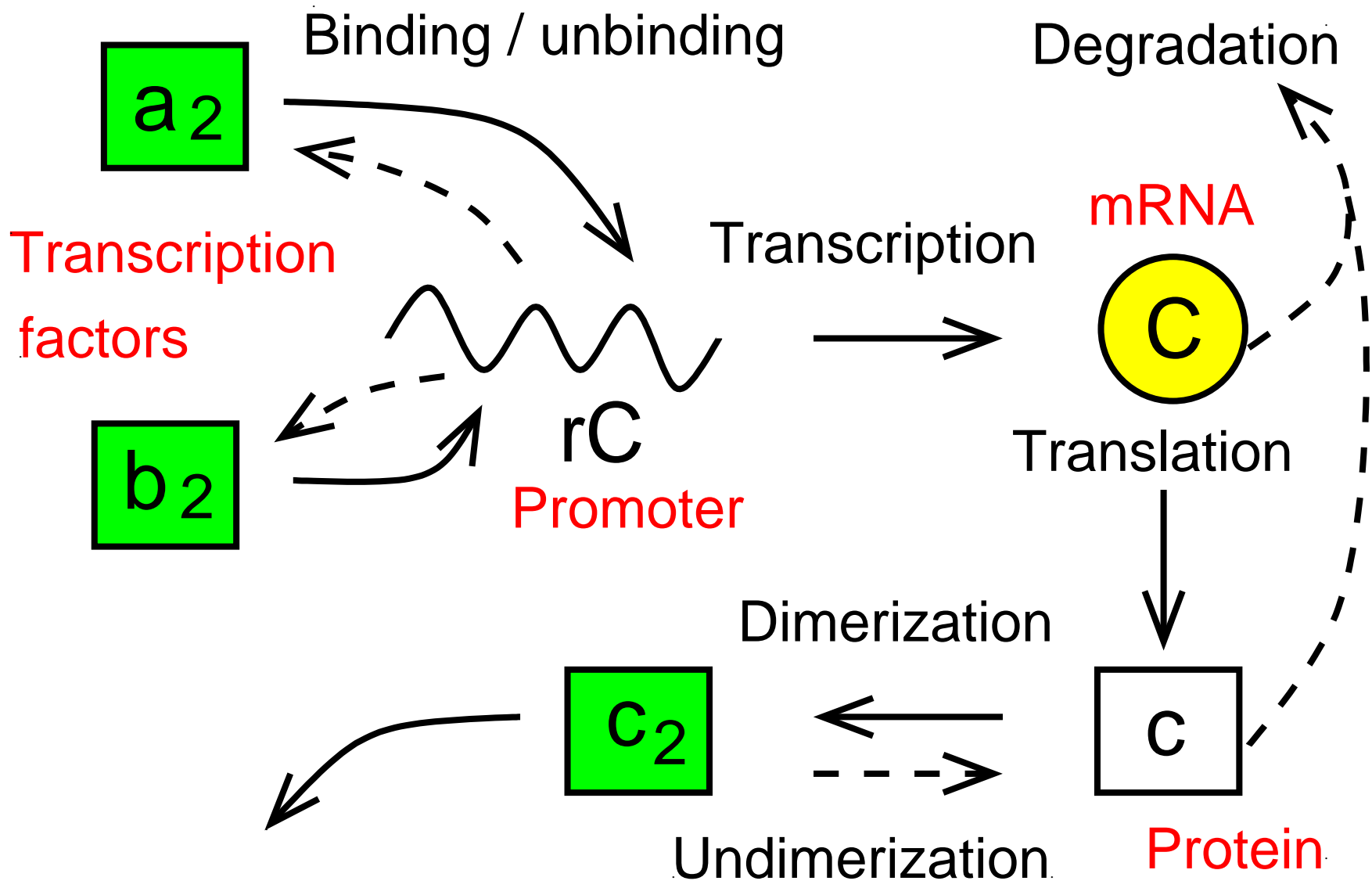










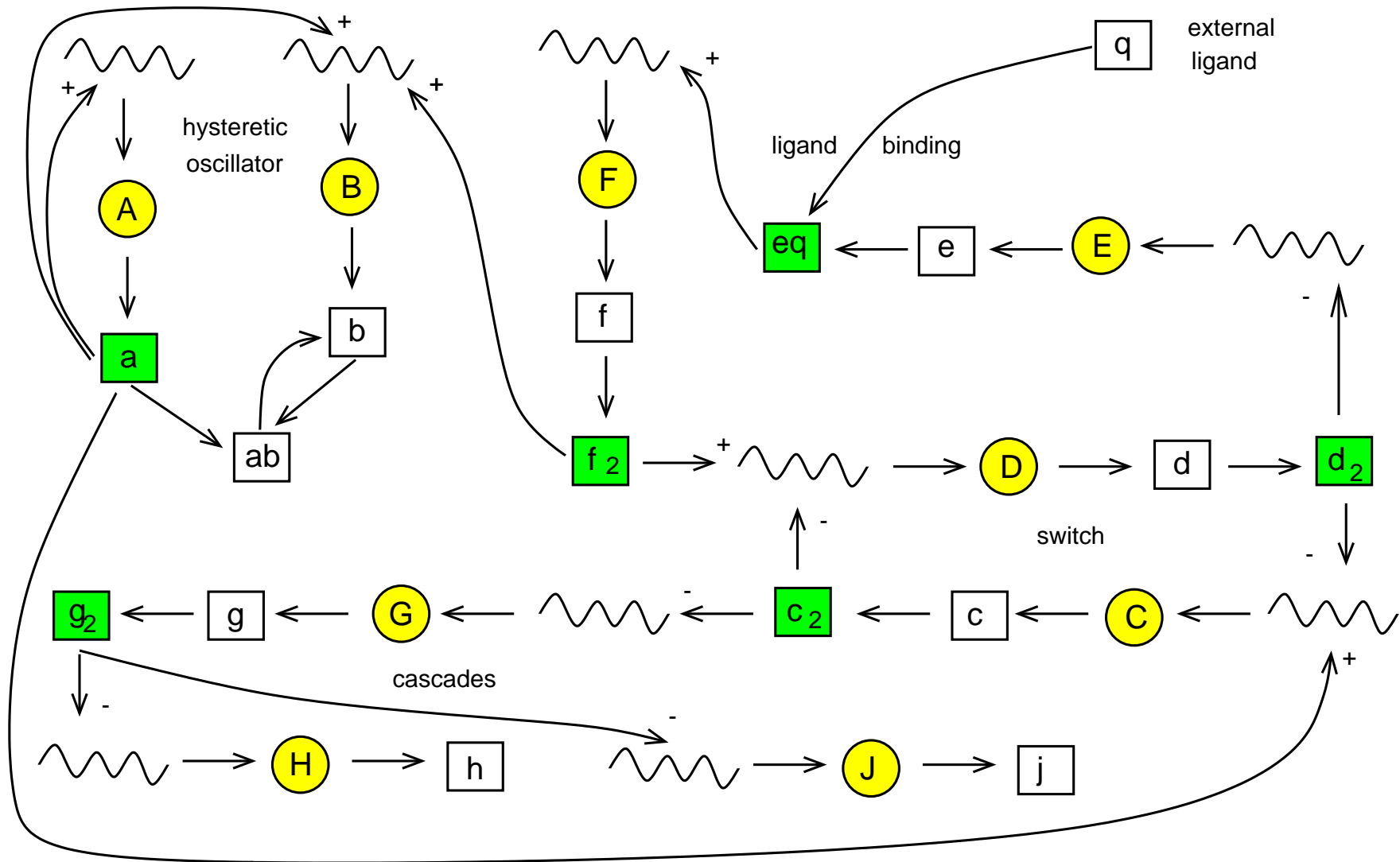


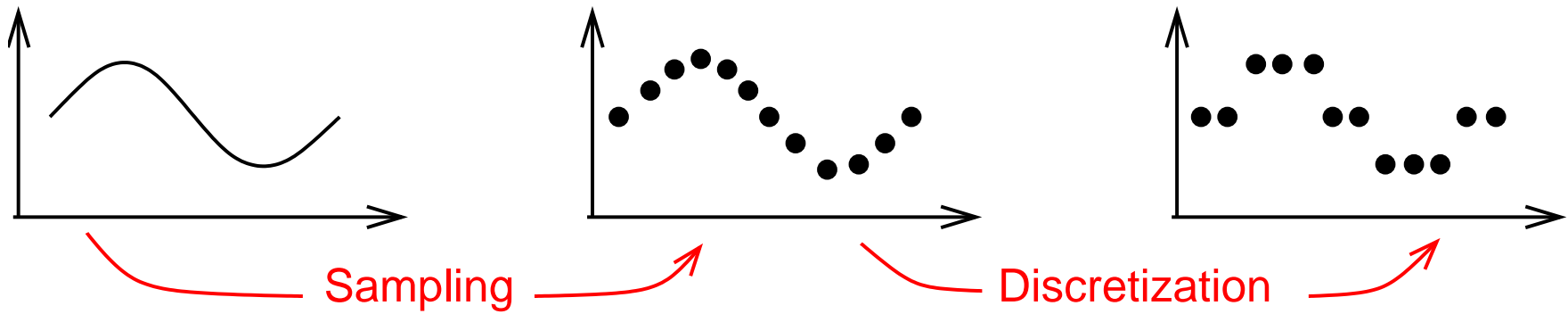
$$\frac{d}{dt}[a_2.rC] = \lambda_{a_2.rC}^+[a_2][rC] - \lambda_{a_2.rC}^-[a_2.rC]$$

$$\frac{d}{dt}[C] = \lambda_{rC}[rC] + \lambda_{a_2.rC}[a_2.rC] + \lambda_{b_2.rC}[b_2.rC] - \lambda_C[C]$$

$$\frac{d}{dt}[c] = \lambda_{Cc}[C] - \lambda_c[c]$$

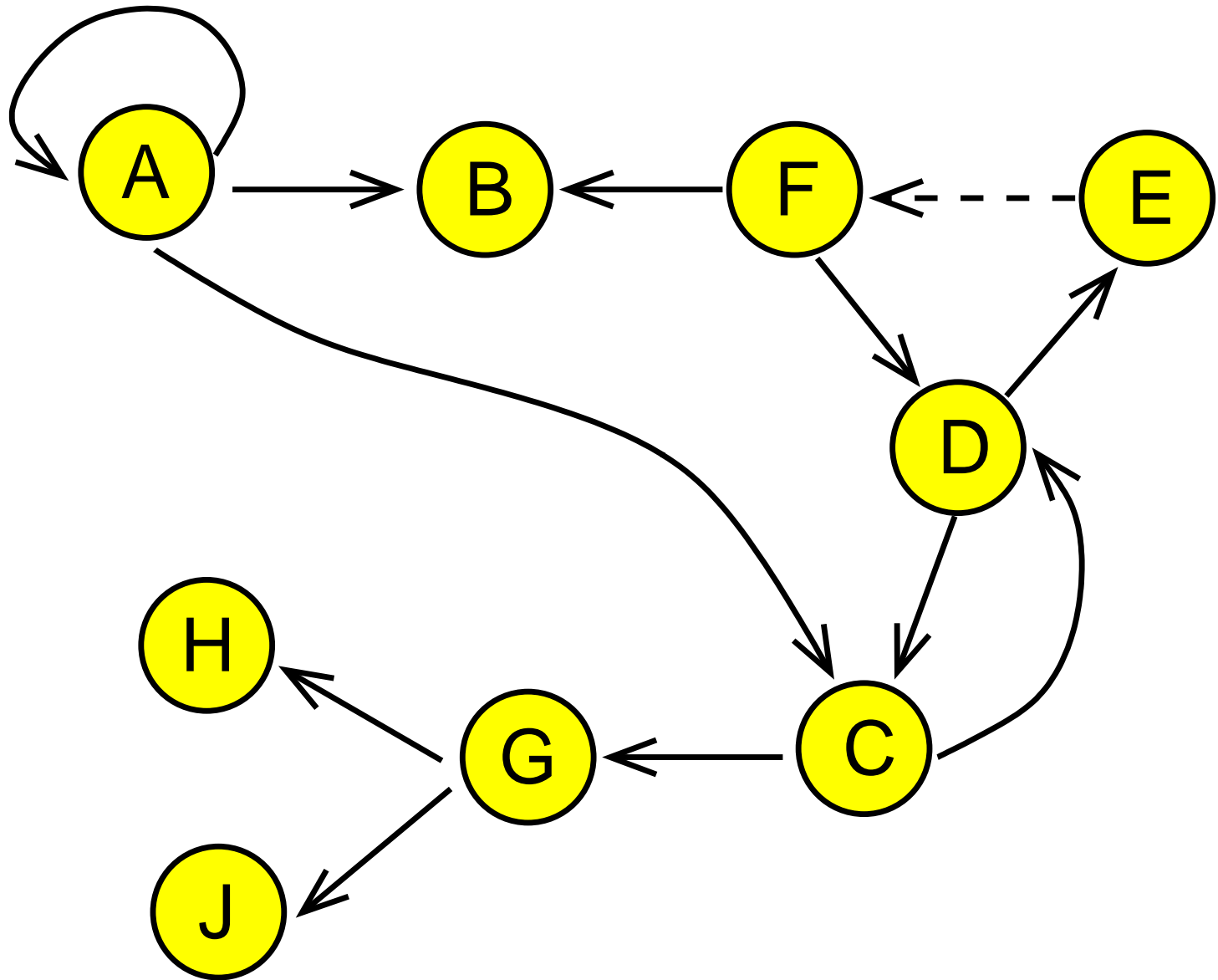
$$\frac{d}{dt}[c_2] = \lambda_{cc}^+[c]^2 - \lambda_{cc}^-[c_2]$$

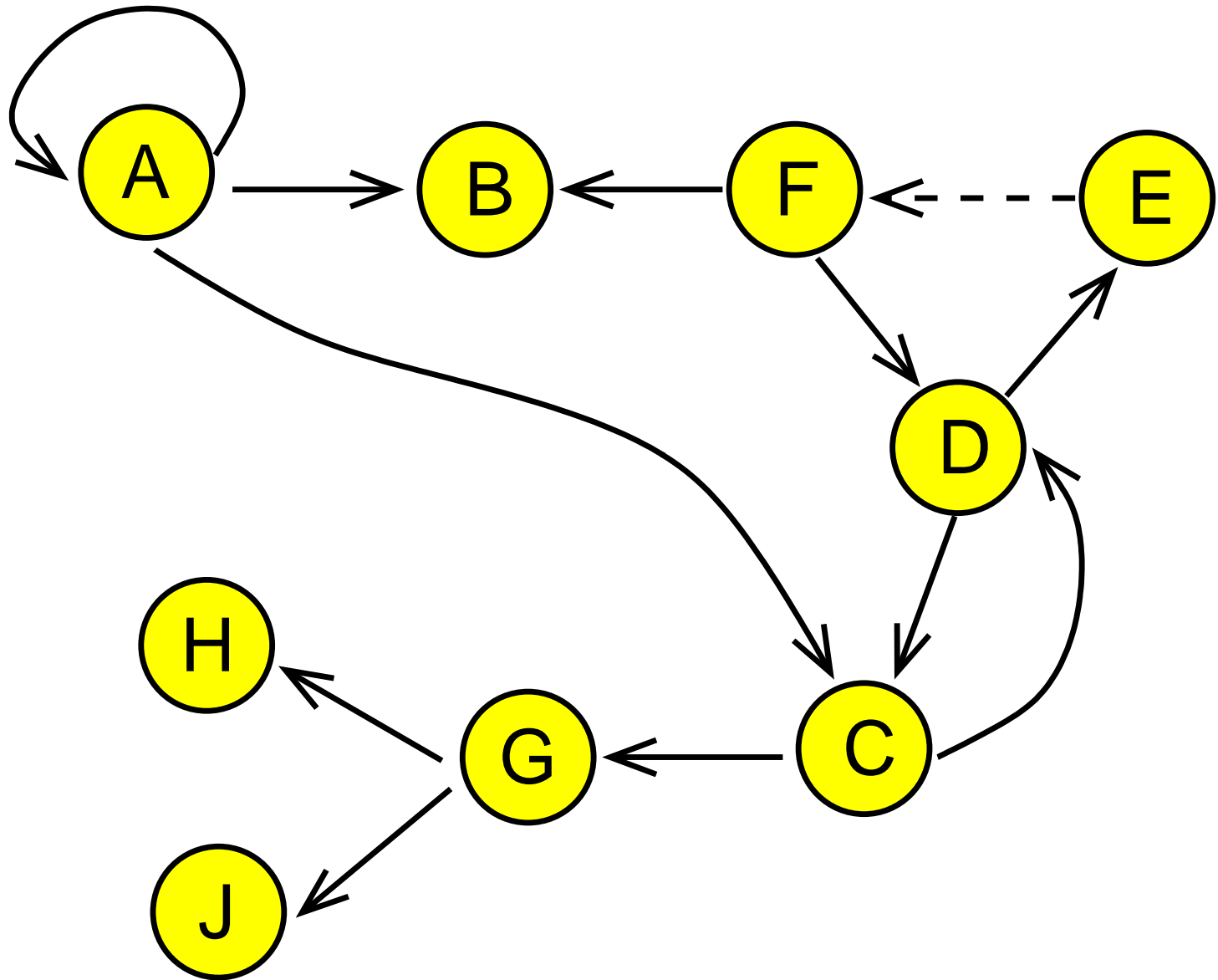


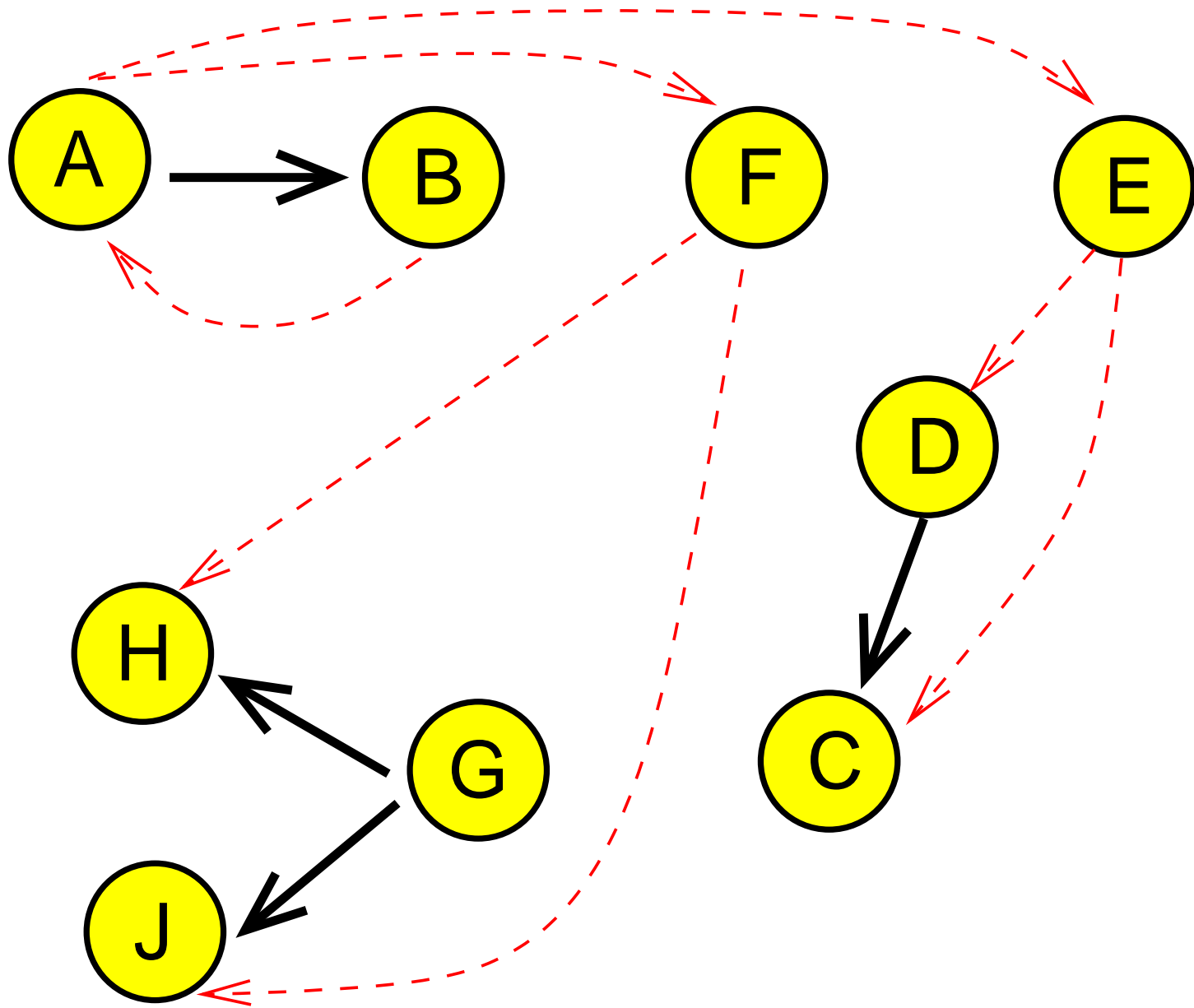


12 time points

Recover the **true genetic network**
with
reverse engineering.







Simulation Experiments

Ligand injection for 10 minutes.

Equilibrium

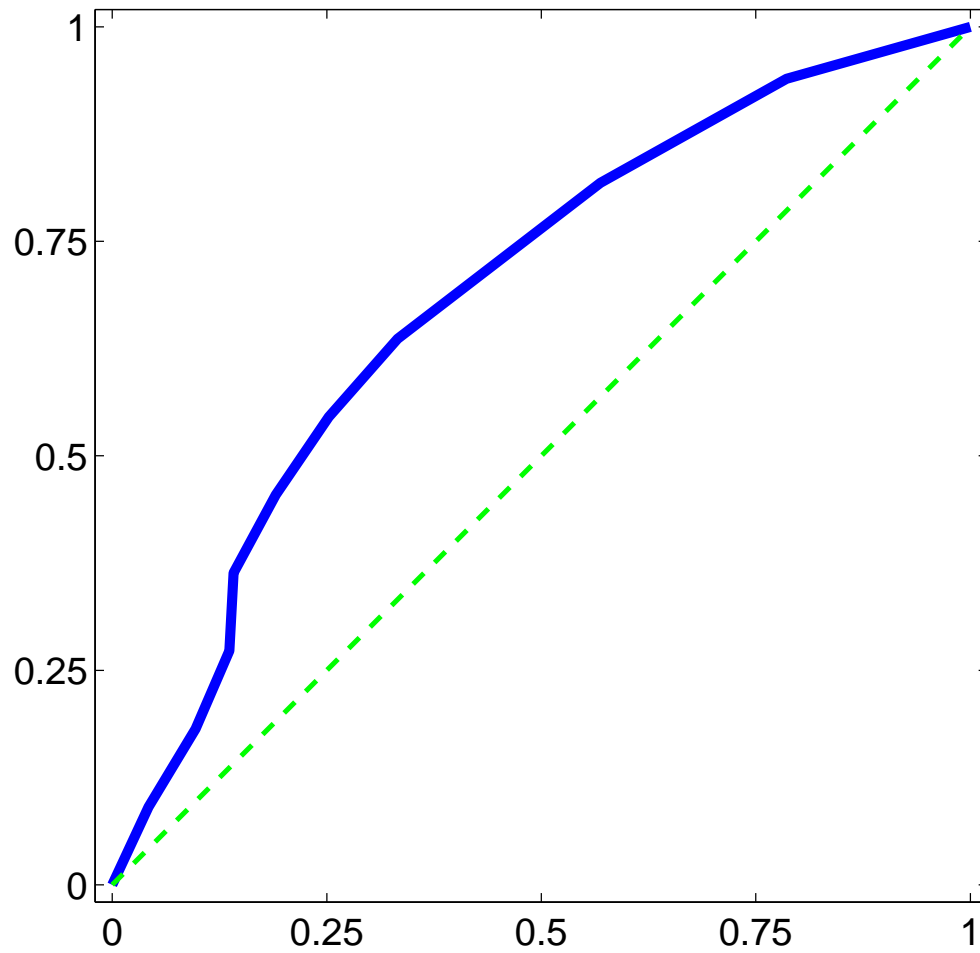
12 data points collected over 4000 min
in equi-distant intervals.

Disequilibrium

12 data points collected over 500 min
in equi-distant intervals.

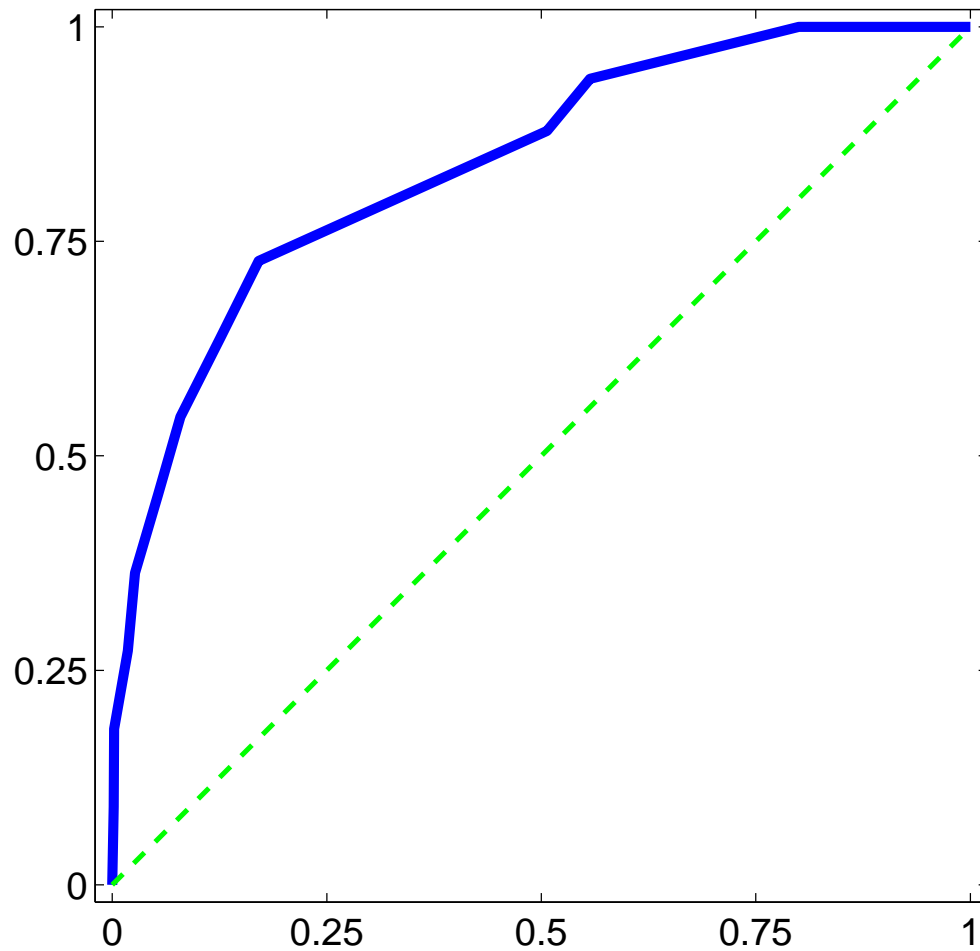
ROC curve: Equilibrium

True positives (vertical axis) \longleftrightarrow False positives (horizontal axis)

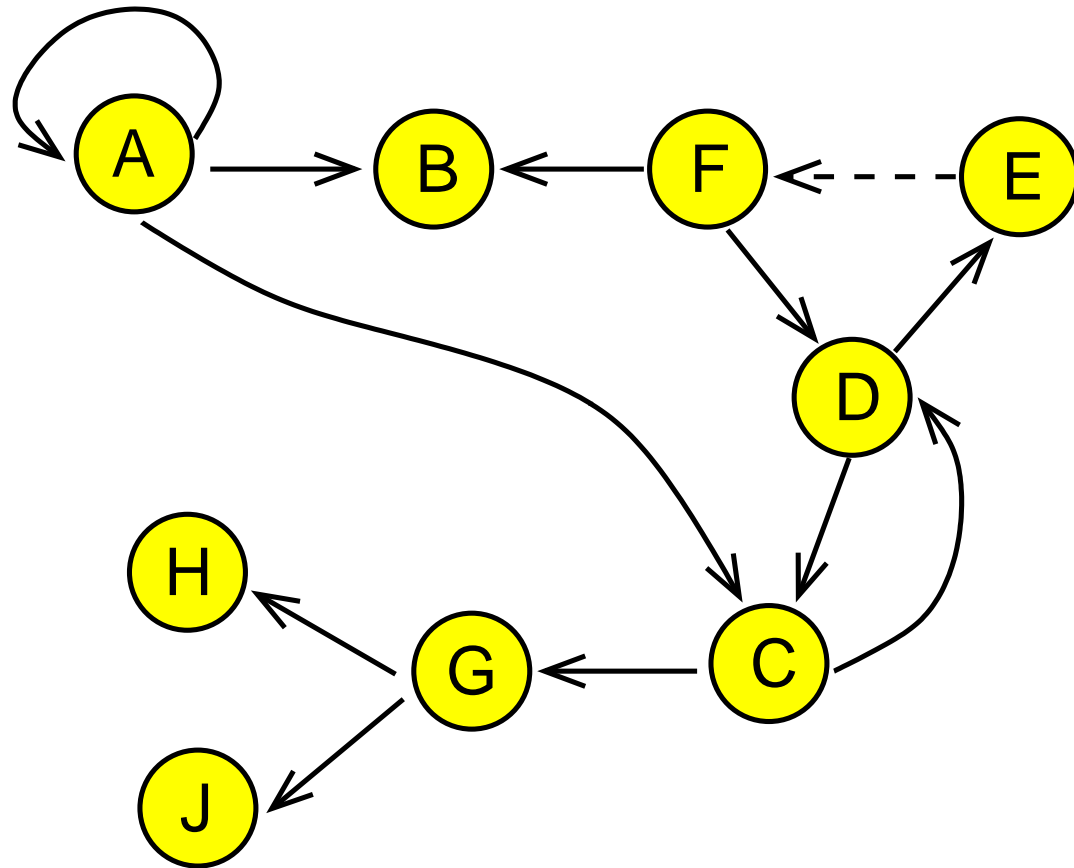


ROC curve: Disequilibrium

True positives (vertical axis) \longleftrightarrow False positives (horizontal axis)



Structure Prior



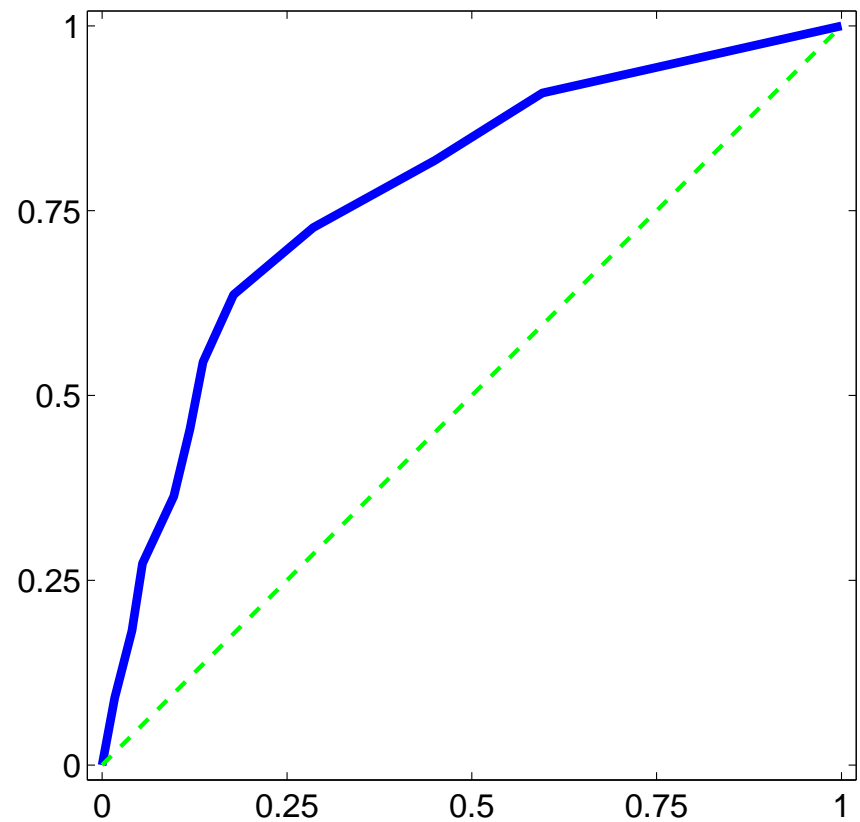
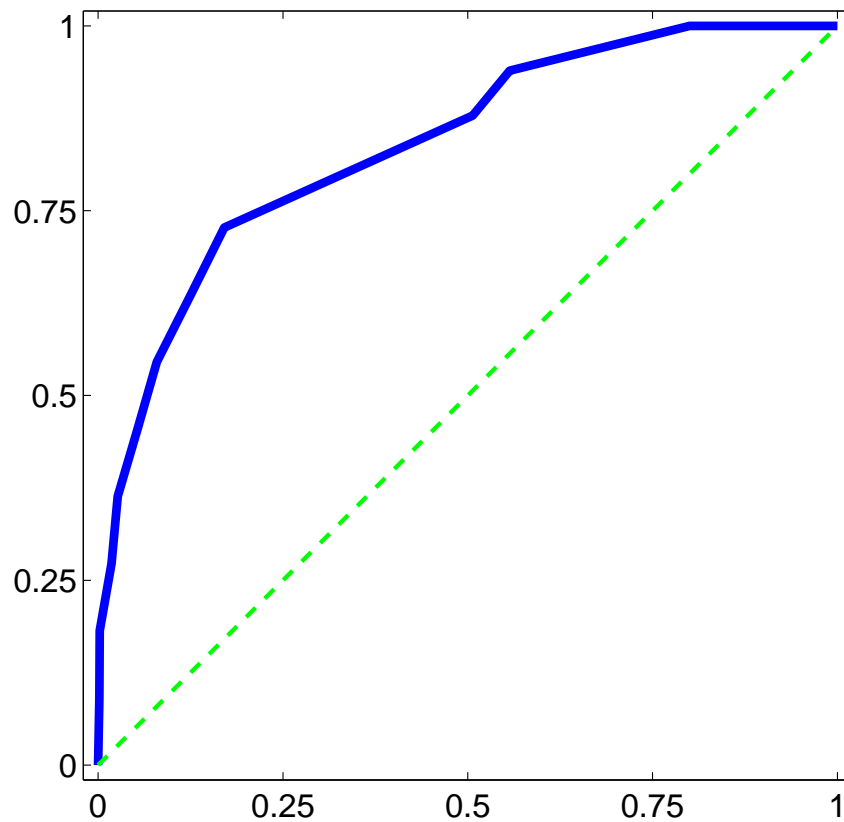
Max fan-in = 2, 3, 4

ROC curves

True positives (vertical axis) \longleftrightarrow False positives (horizontal axis)

Left: max fan-in = 2

Right: max fan-in = 3

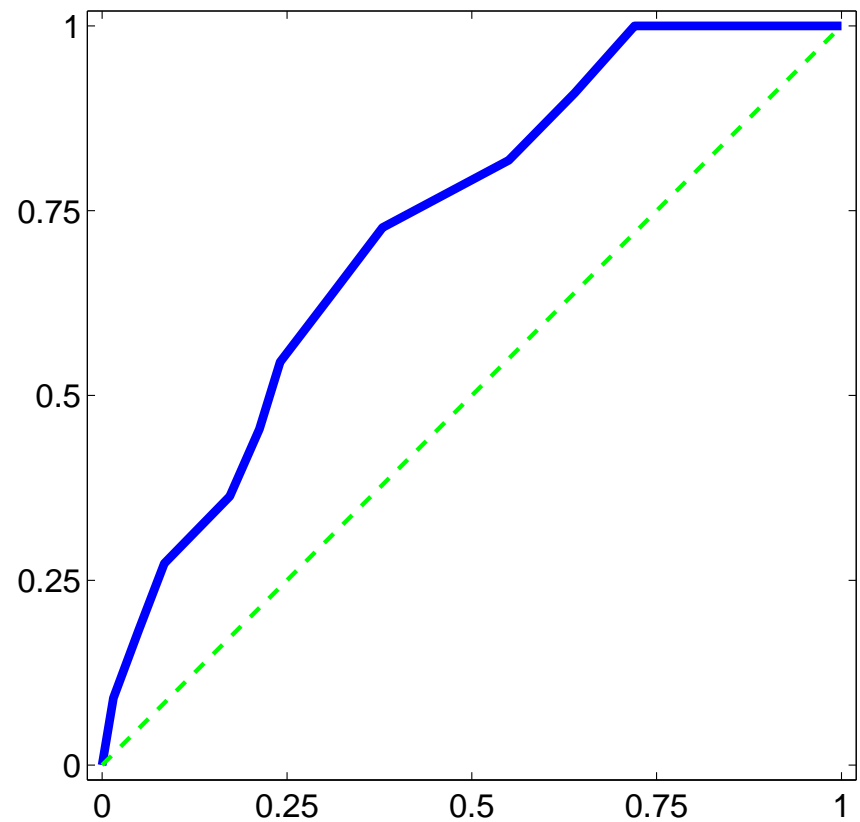
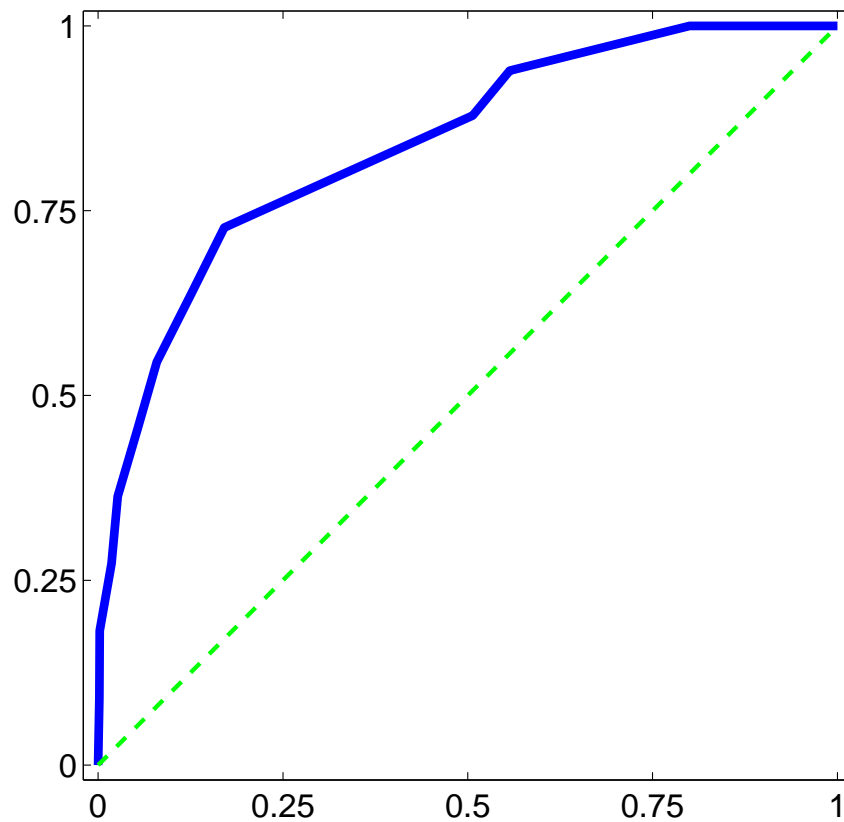


ROC curves

True positives (vertical axis) \longleftrightarrow False positives (horizontal axis)

Left: max fan-in = 2

Right: max fan-in = 4



Sequence information

$$\frac{P(y \rightarrow rX | r \in B[y])}{P(y \rightarrow rX | r \notin B[y])} = 2$$

$y \rightarrow rX$ denotes the event that transcription factor y binds to the promoter r upstream of gene X , and $B[y]$ is the set of (known) binding motifs for y .

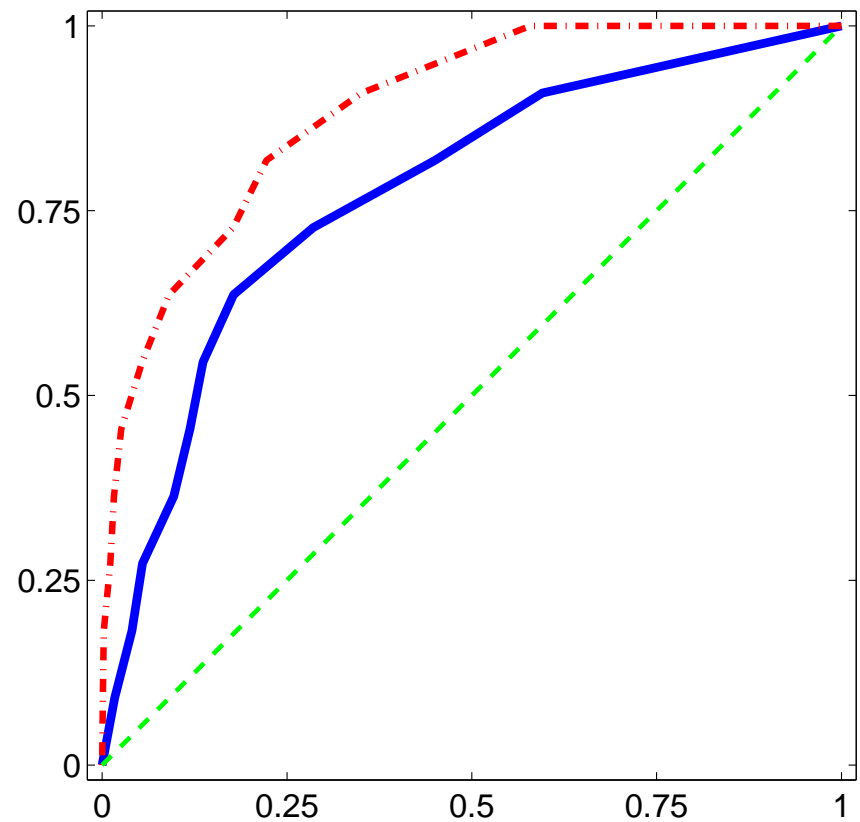
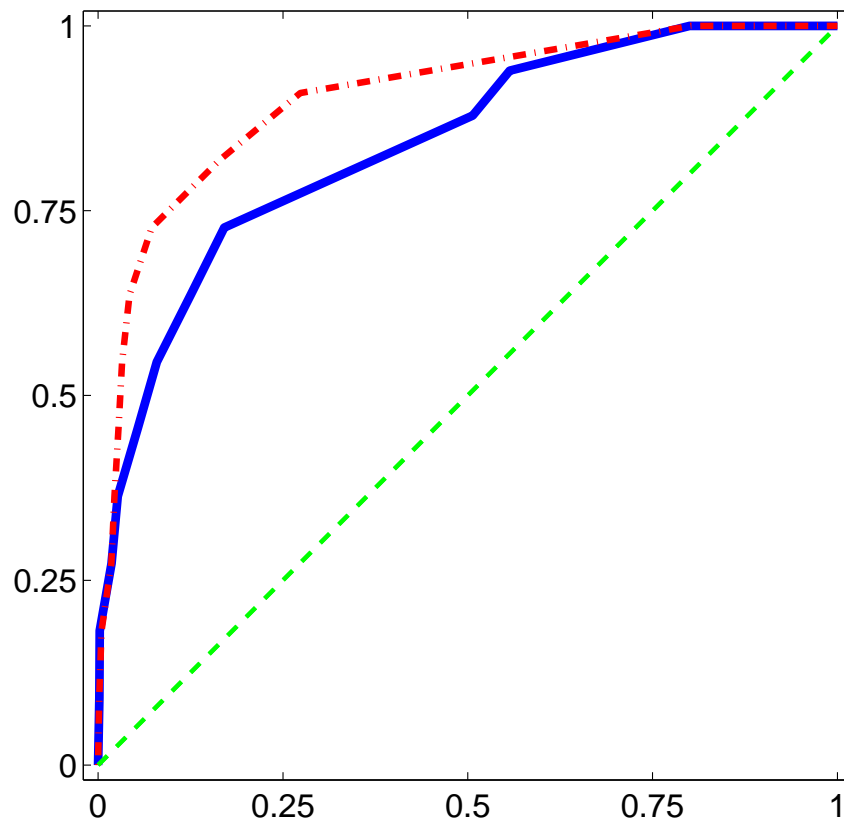
In words: The equation expresses that on identifying a binding motif for transcription factor y in the upstream region of gene X , this transcription factor is twice as likely to bind to X than in the absence of such a motif.

ROC curves

True positives (vertical axis) \longleftrightarrow False positives (horizontal axis)

Left: max fan-in = 2

Right: max fan-in = 3



Conclusions

Conclusions

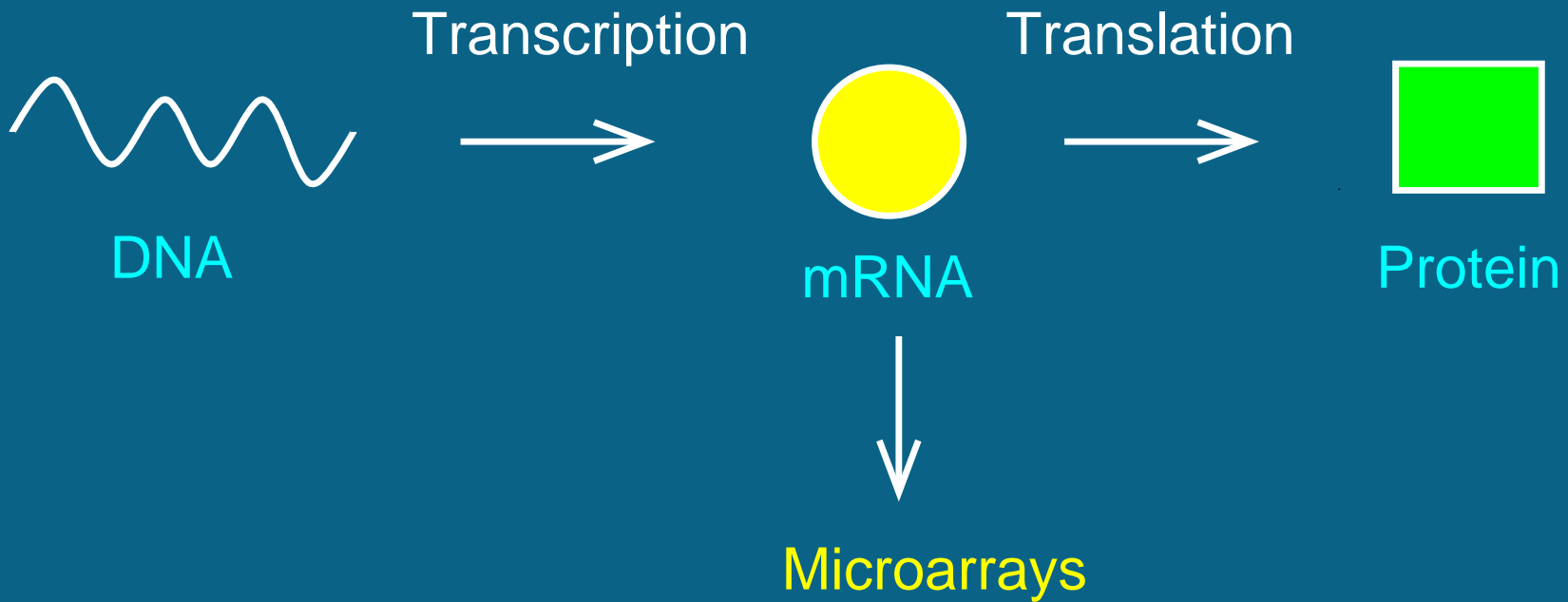
- Learning the **global network** → impossible
- Intrinsic **uncertainty** due to lack of data

Conclusions

- Learning the **global network** → impossible
- Intrinsic **uncertainty** due to lack of data
- Inference **local substructures** possible, but obscured by noise.

Conclusions

- Learning the **global network** → impossible
- Intrinsic **uncertainty** due to lack of data
- Inference **local substructures** possible, but obscured by noise.
- Biologically realistic **priors** important
- **Integrating** post-genomic **data**.



Acknowledgments

Marco Grzegorzczuk
Isabelle Grimmenstein
Wolfgang Urfer

Dortmund University