Model gene regulatory networks under mutation-selection balance

Marcin Zagórski

Institute of Physics, Jagiellonian University, Kraków, Poland

in collaboration with
Zdzislaw Burda, Andre Krzywicki and Olivier C. Martin

26 November 2010

Outline

- What are gene regulatory networks (GRNs)?
- Biological aspects of GRNs
- GRN model
- Main results
- Conclusions and outlook

Genomes

Number of protein-coding genes

E. coli	$\sim$	4000
M. tuberculosis	$\sim$	4000
yeast	$\sim$	6000
fruit fly	$\sim$	14000
human	$\sim$	22000

Genomes

Number of protein-coding genes

E. coli	$\sim$	4000
M. tuberculosis	$\sim$	4000
yeast	$\sim$	6000
fruit fly	$\sim$	14000
human	$\sim$	22000

It seems, we are not even an order of magnitude better than bacteria...

Gene regulation

Number of regulatory genes grows faster than linearly with the total number of genes.

Gene regulation

Number of regulatory genes grows faster than linearly with the total number of genes.

Gene products

- structural proteins
- enzymes
- transcription factors (TFs) - serve only to activate (deactivate) other genes.

Gene regulation

Number of regulatory genes grows faster than linearly with the total number of genes.

Gene products

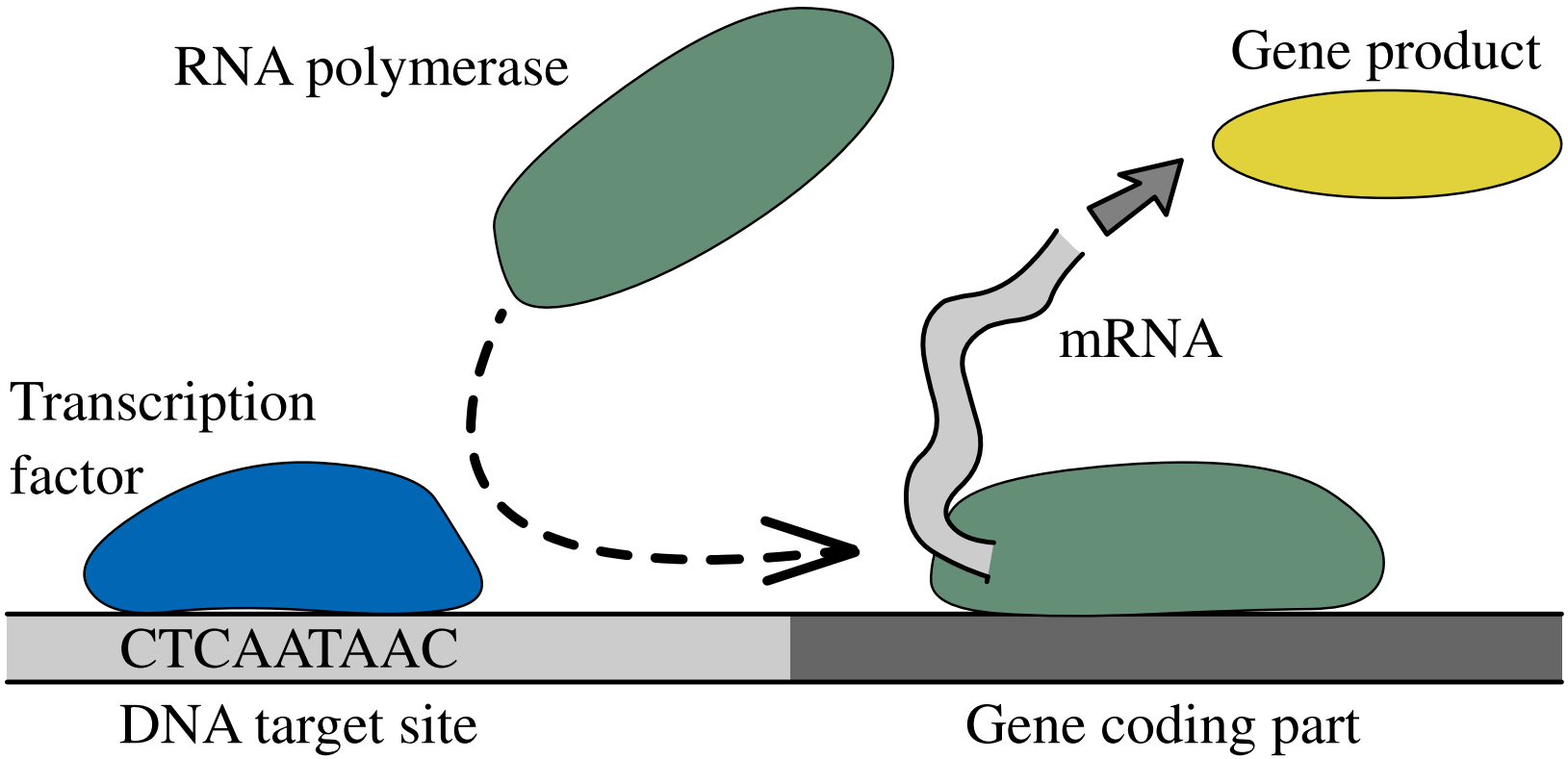
- structural proteins
- enzymes
- transcription factors (TFs) - serve only to activate (deactivate) other genes.

Gene regulatory network

GRN consists of interactions mediated by TFs (edges) which are products of certain genes (nodes).

Activators

TF bound to DNA target site in promoter region can enhance protein production rate.



Transcriptional Networks

Biological features

- a given gene is generally influenced by a "small" number of other genes
- some genes have effect on multiple target genes (pleiotropic effect)
- GRNs seem to be robust to random change (e.g., to environmental fluctuations or to mutations)

In terms of network terminology, GRNs are sparse with narrow in-degree distribution and broad out-degree distribution.

Phenotype $\mathbf{S} = (S_1, S_2, \ldots, S_N)$
$S_i \in [0, 1]$

Genotype $\mathbf{W}$ is matrix $N \times N$

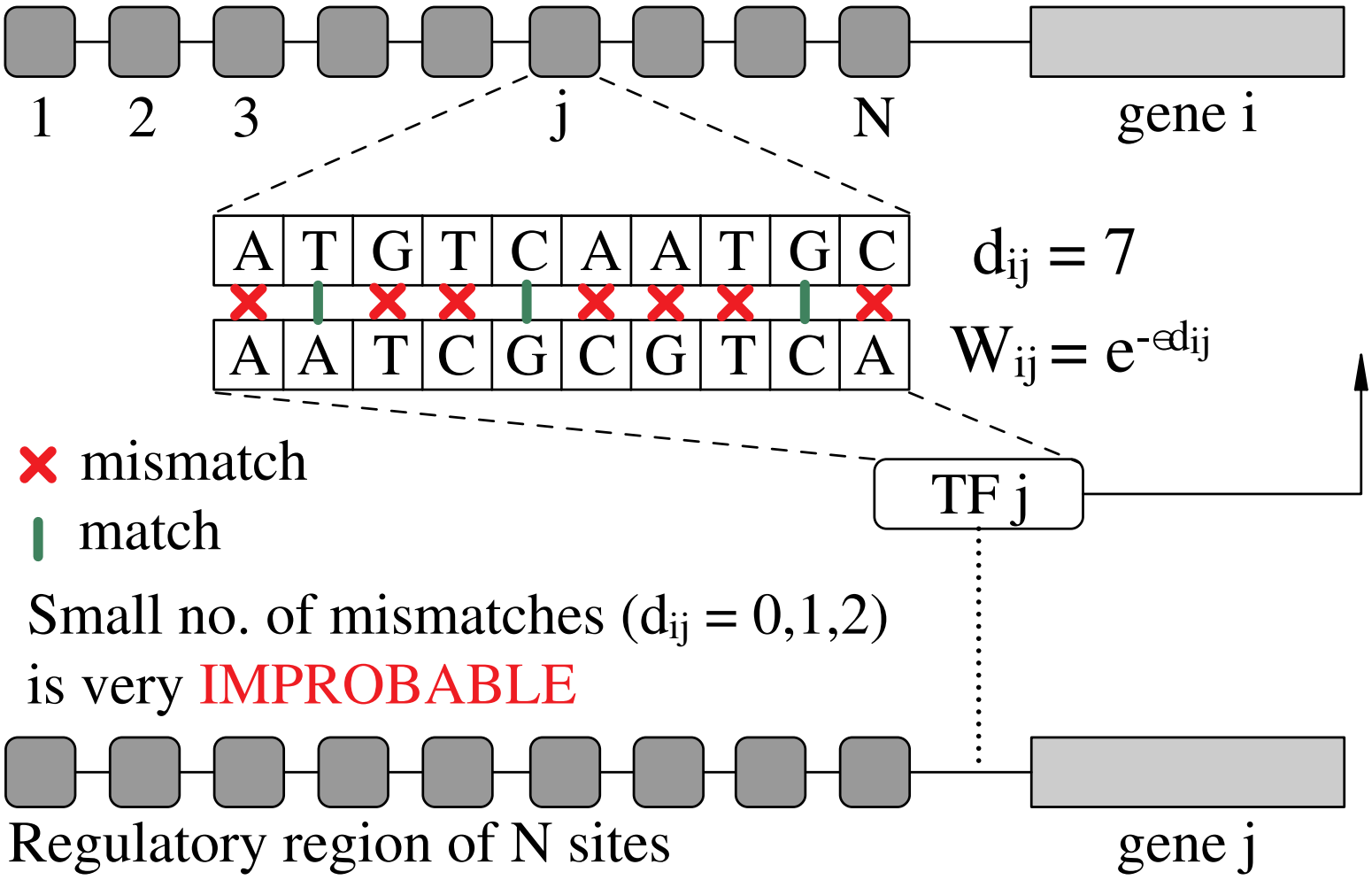
Dynamics $\mathbf{S}(t+1) = G(\mathbf{S}(t), \mathbf{W})$

$$\underbrace{\mathbf{S}(0)}_{\mathbf{S}^{(initial)}} \xrightarrow{\mathbf{W}} \mathbf{S}(1) \xrightarrow{\mathbf{W}} \ldots \xrightarrow{\mathbf{W}} \underbrace{\mathbf{S}(t) \xrightarrow{\mathbf{W}} \mathbf{S}(t+1)}_{\text{fixed point phenotype}}$$

Fitness $F(\mathbf{S}) = \exp(-f \sum_i |S_i - S_i^{(target)}|)$

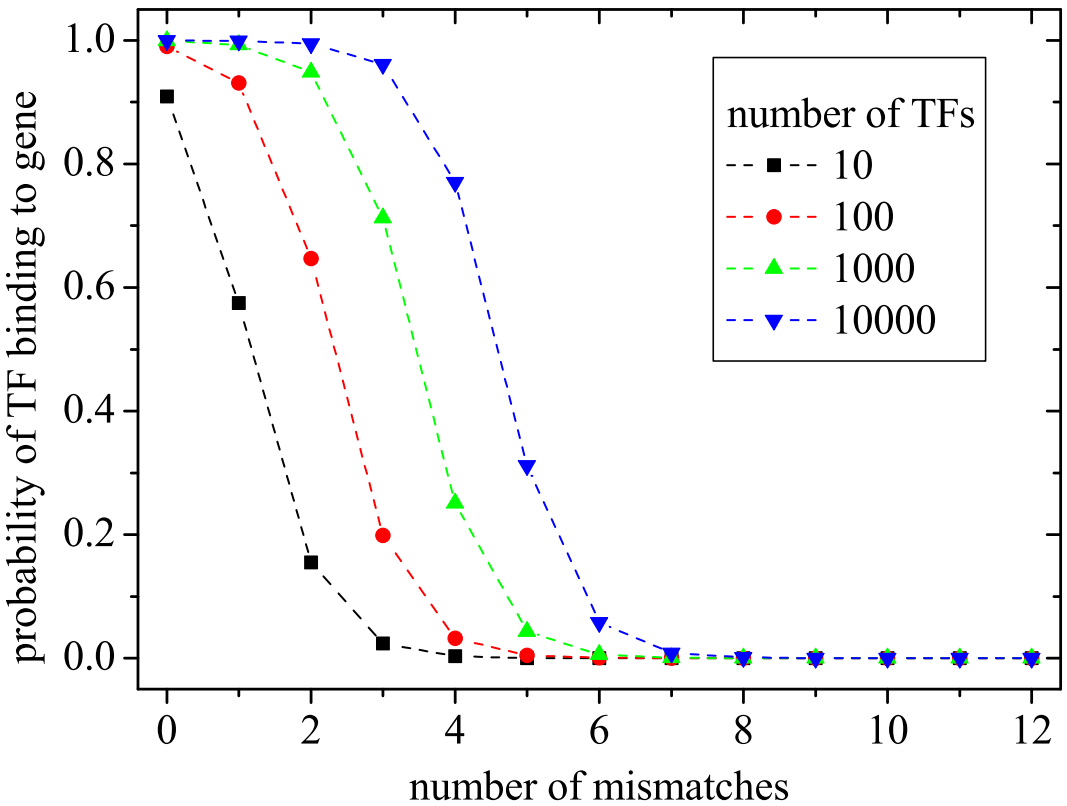
- S_i a level of gene i expression
- W_{ij} interaction between gene i and TF j
- $\mathbf{S}^{(target)}$ corresponds to cell's function
- we sample space of genotypes which lead from $\mathbf{S}^{(initial)}$ to $\mathbf{S}^{(target)}$

GRN realization



Specific binding

$$p_{ij} = \frac{1}{1 + \exp\left(\varepsilon d_{ij} - \log nS_j\right)} = \frac{1}{1 + 1/W_{ij} n_j}$$

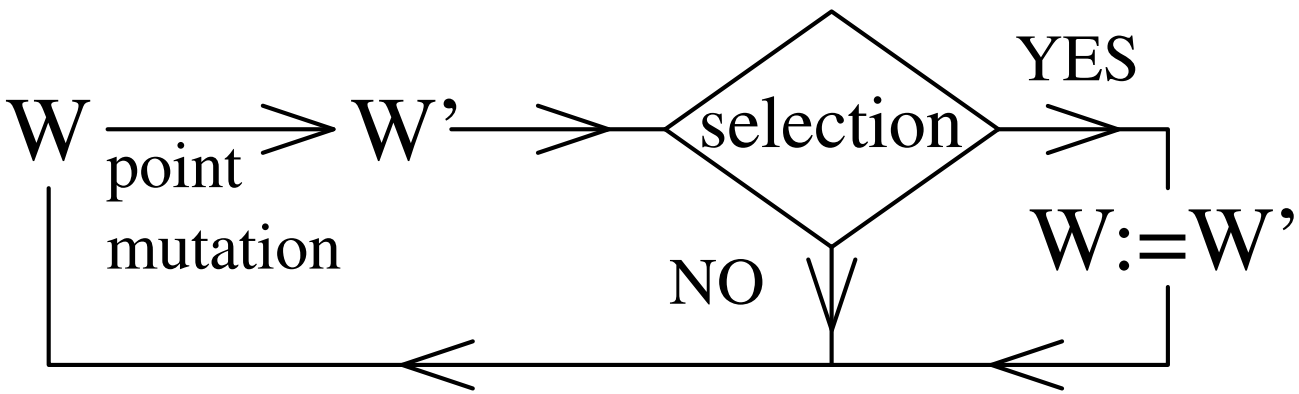


Expression pattern dynamics

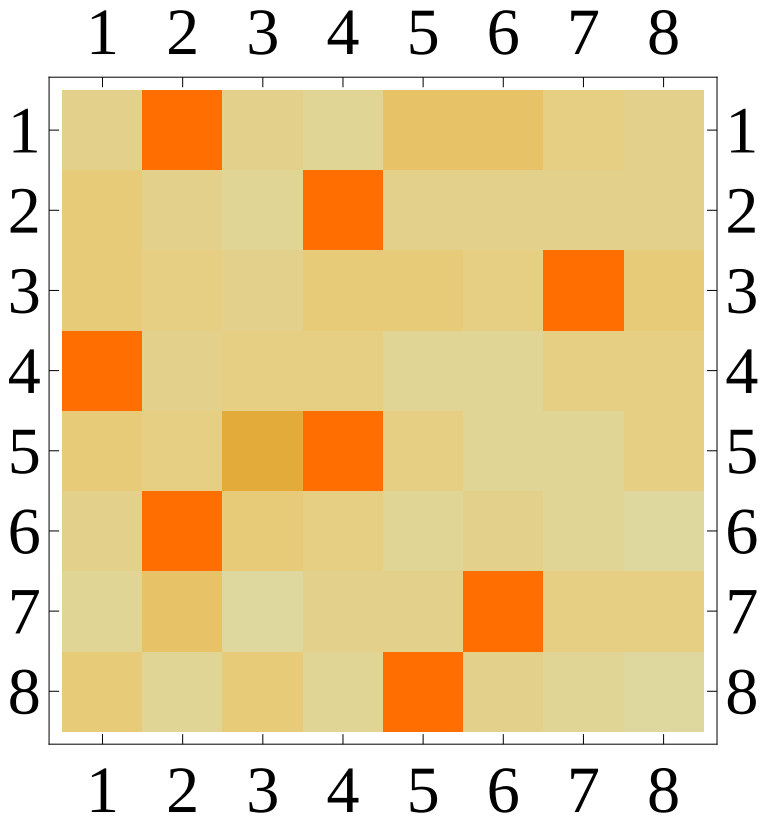
$$S_i(t+1) = 1 - \prod_j (1 - p_{ij}(t))$$

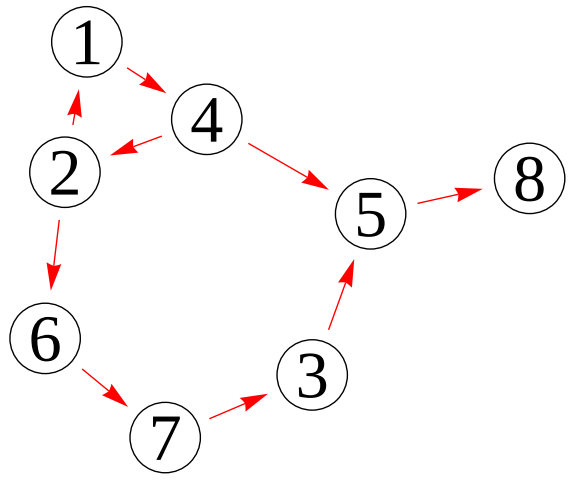
$$p_{ij}(t) = \frac{1}{1 + \exp\left(\varepsilon d_{ij} - \log n S_j(t)\right)}$$

Metropolis sampling of space of viable genotypes.



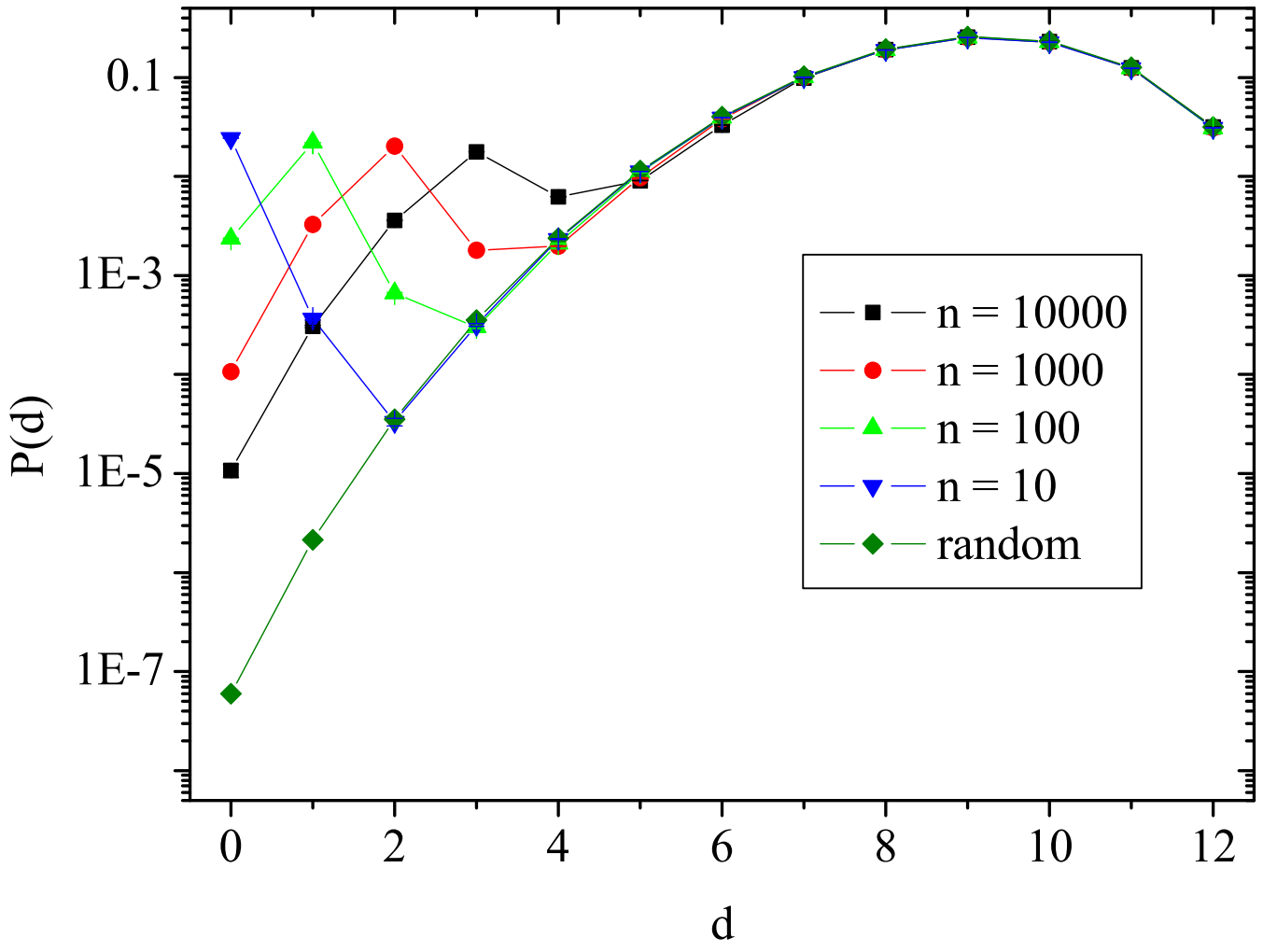
GRN as a directed weighted graph



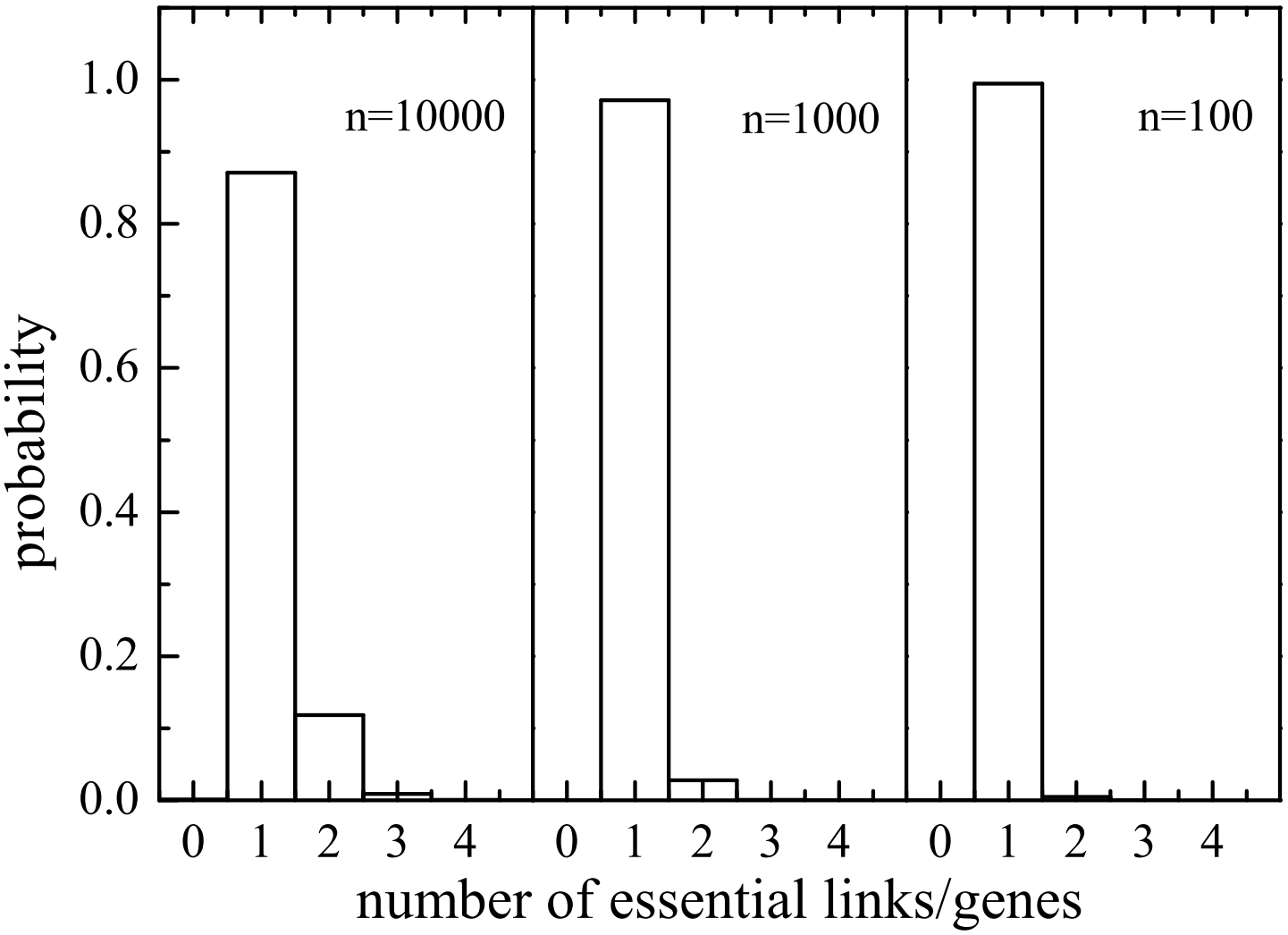


Matrix **W** with corresponding GRN (only strong interactions).

Mismatch distribution



Essential interactions



Conclusions and outlook

We explore ensemble of GRNs having one biological function

Conclusions and outlook

We explore ensemble of GRNs having one biological function

- resulting networks are as sparse as possible being compatible with their function

Conclusions and outlook

We explore ensemble of GRNs having one biological function

- resulting networks are as sparse as possible being compatible with their function
- in- and out-degree distribution qualitatively agree with those observed in real GRNs

Conclusions and outlook

We explore ensemble of GRNs having one biological function

- resulting networks are as sparse as possible being compatible with their function
- in- and out-degree distribution qualitatively agree with those observed in real GRNs
- GRNs are robust to point mutations

Conclusions and outlook

We explore ensemble of GRNs having one biological function

- resulting networks are as sparse as possible being compatible with their function
- in- and out-degree distribution qualitatively agree with those observed in real GRNs
- GRNs are robust to point mutations
- the results are valid for parameters in the biologically relevant window

Conclusions and outlook

We explore ensemble of GRNs having one biological function

- resulting networks are as sparse as possible being compatible with their function
- in- and out-degree distribution qualitatively agree with those observed in real GRNs
- GRNs are robust to point mutations
- the results are valid for parameters in the biologically relevant window

In order to explore GRNs having two or more functions (more than one target phenotype) one has to introduce repressors. Preliminary results suggest the above still holds.

Thank you for your attention