

# Системная биология – сети

М.Гельфанд  
«Сравнительная геномика»

БиБи 4 курс

## разные сети

- белок-белковые взаимодействия
- регуляторные сети (фактор-ген)
- метаболические

# свойства сетей

- $N$  = количество вершин
- распределение степеней вершин  
 $P(k)$  = вероятность того, что у случайно взятой вершины будет  $k$  ребер
- средняя длина пути между вершинами  $L$

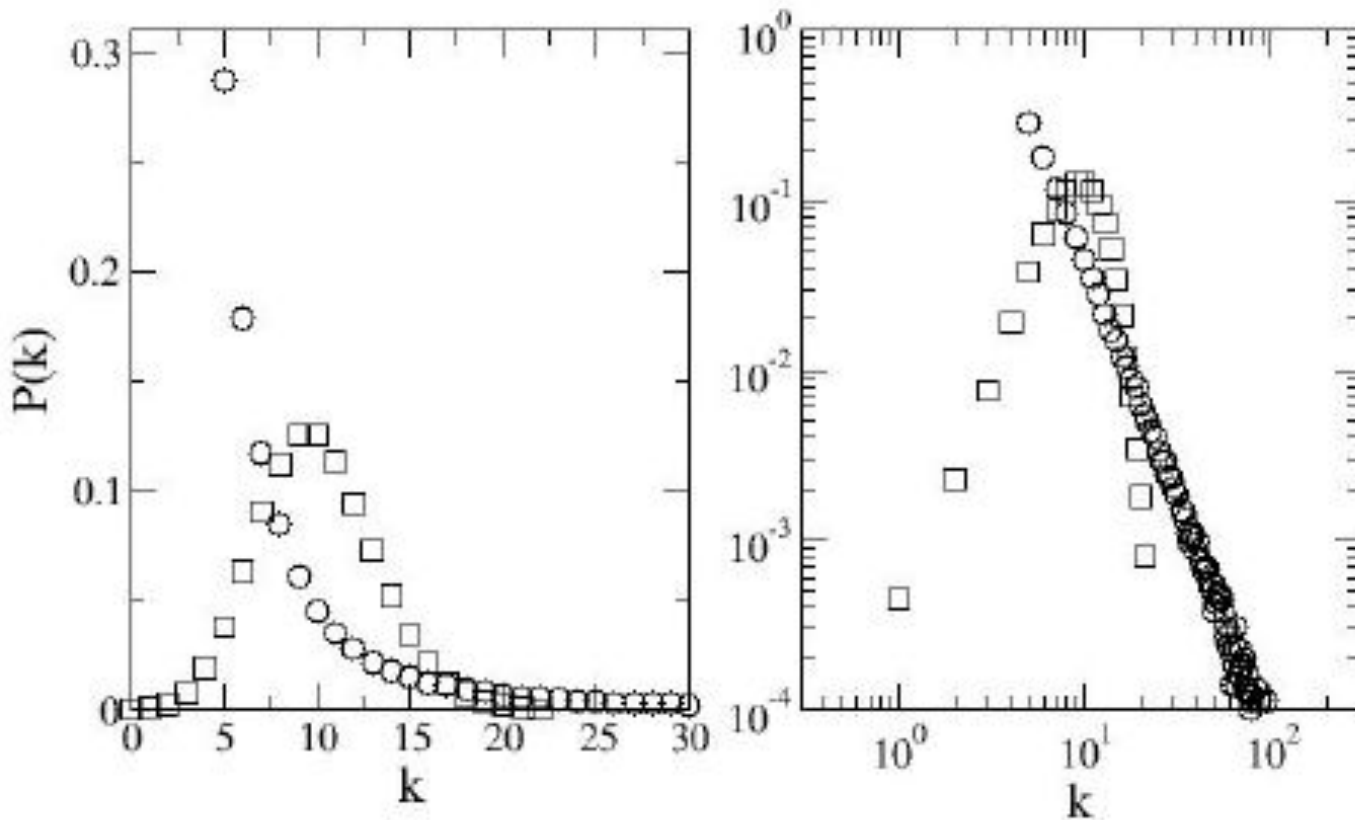
# случайная сеть

- пуассоновское распределение  
 $P(k) = \exp(-\lambda) \lambda^k / k!$
- Теорема Эрдеша-Реньи: фазовый переход – возникновение гигантской компоненты
- средняя длина пути  $\sim \log N$

# scale-free network

- $P(k) \sim k^{-\gamma}$
- $\gamma > 3$  – ничего особенного
- $2 < \gamma < 3$  – hubs, иерархия
- $\gamma = 2$  большой hub, соединенный с большой долей вершин
- При  $\gamma < 3$  удаление случайной вершины не разрушает сеть, удаление hub'a – разрушает
- средняя длина пути (при  $2 < \gamma < 3$ )  $\sim \log \log N$

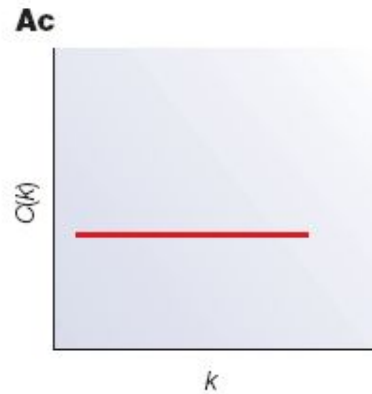
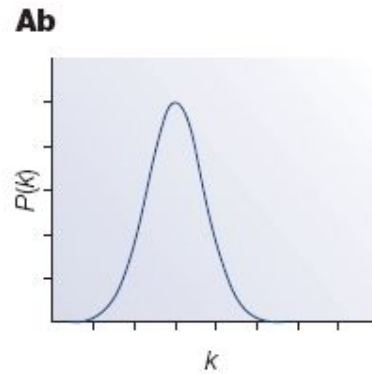
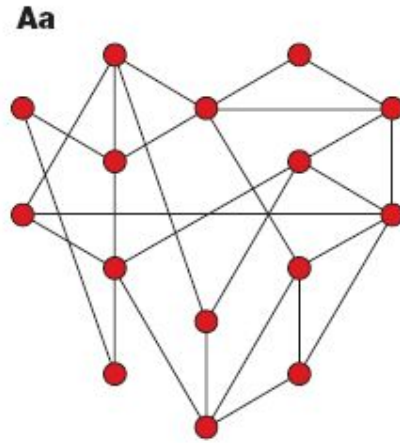
# Random and scale-free $P(k)$ (linear and log scales)



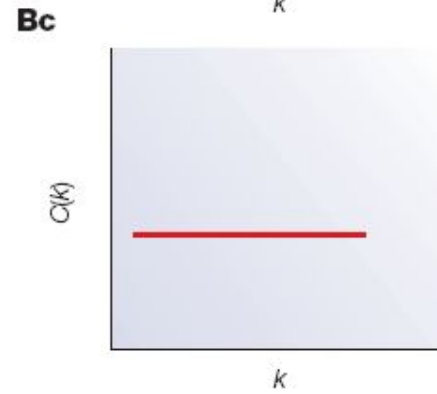
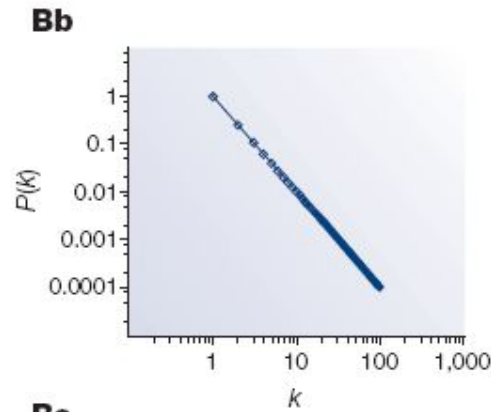
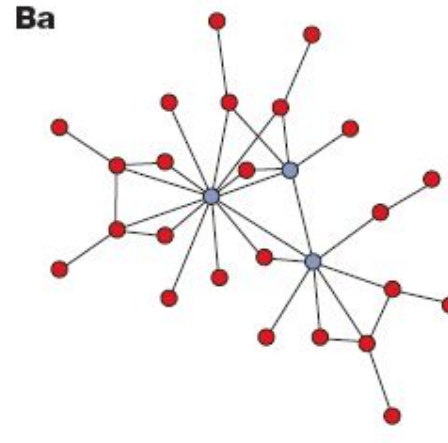
# Коэффициент кластеризации

- Мера связи между соседями данной вершины

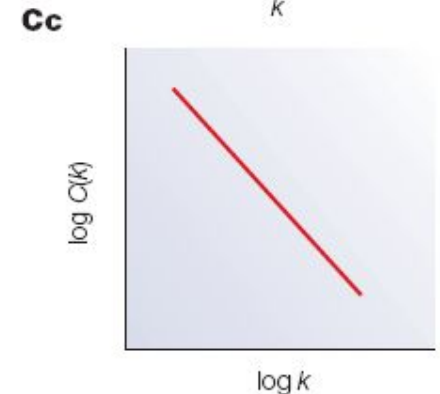
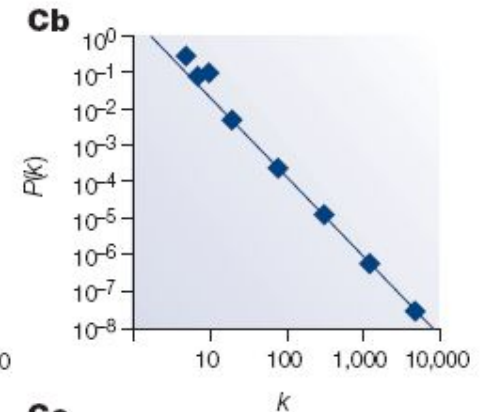
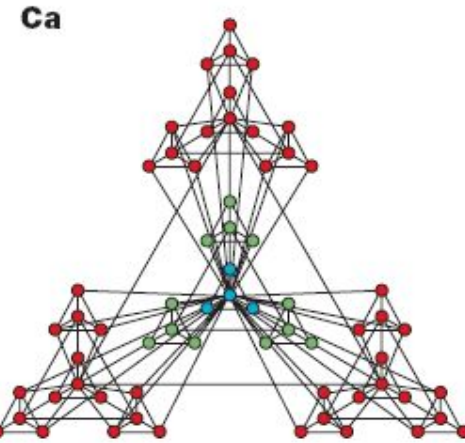
**A** Random network



**B** Scale-free network



**C** Hierarchical network

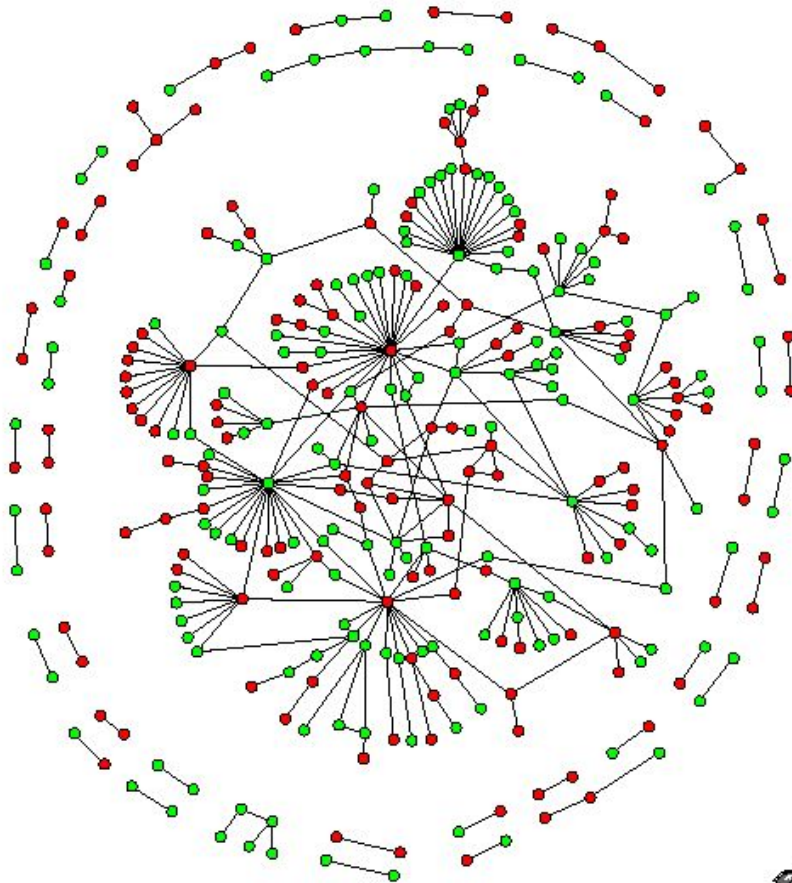


## примеры

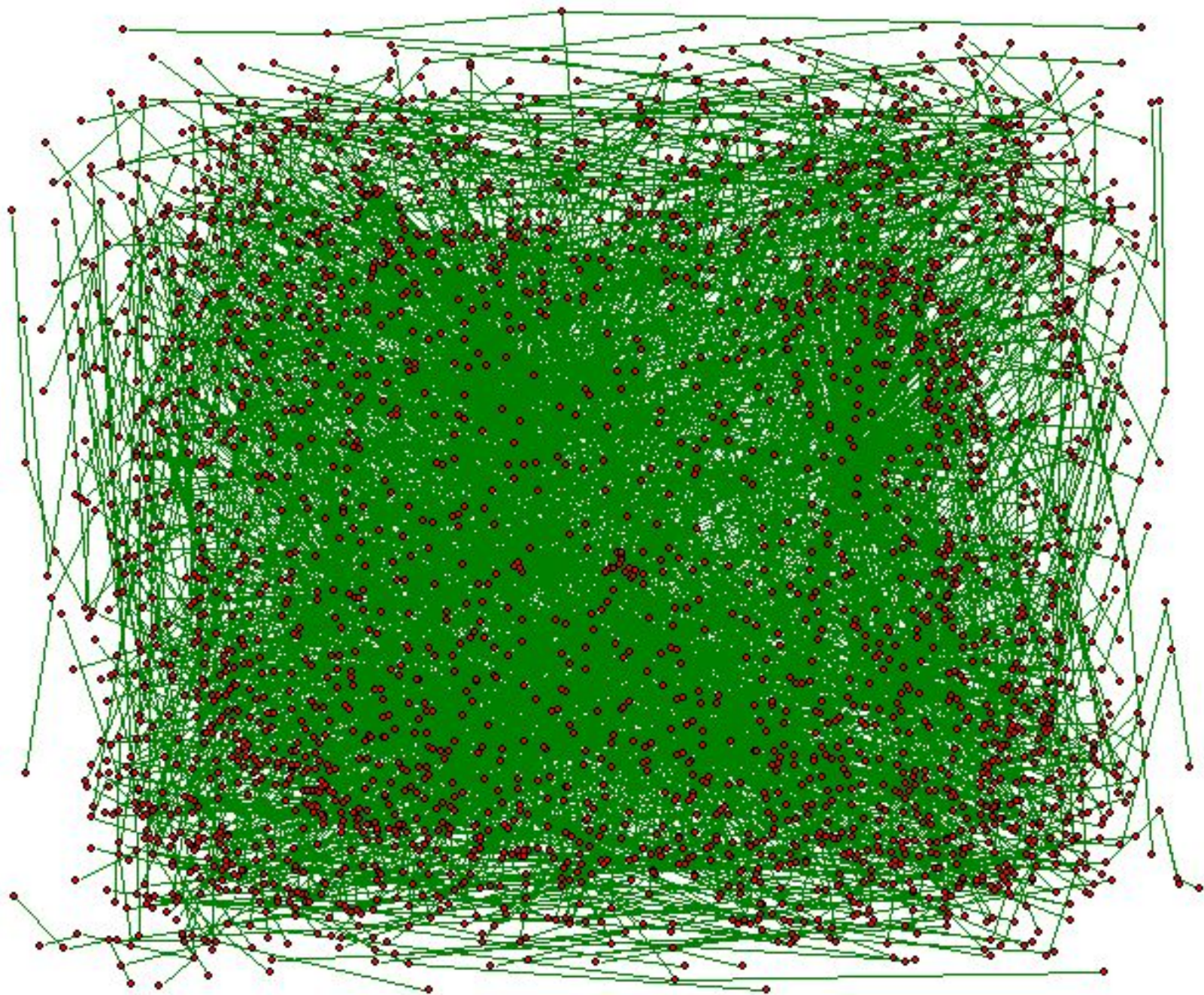
- белок-белковые взаимодействия
- синтетические летали
- регуляция транскрипции
- метаболические сети



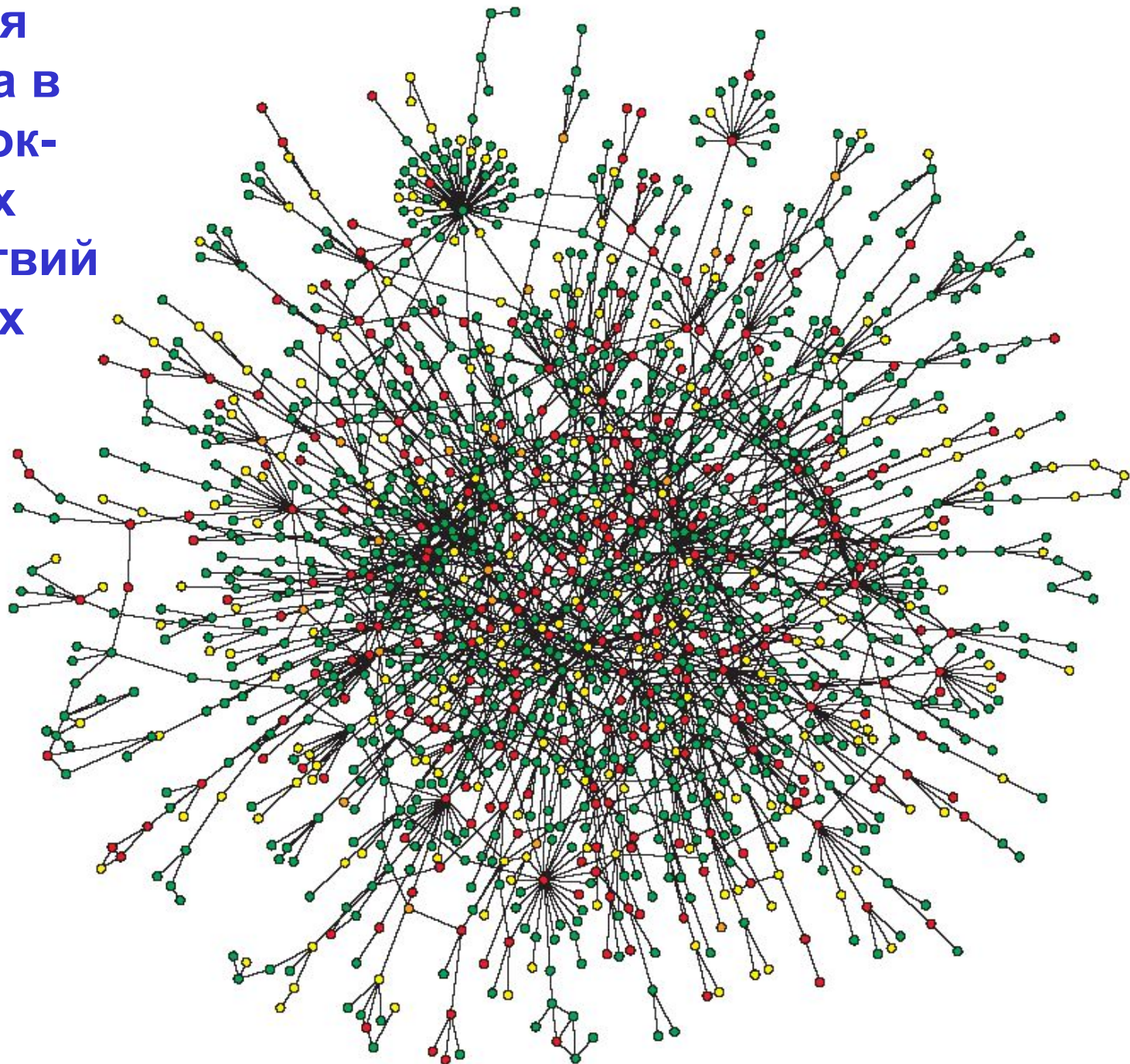
# Yeast protein interaction network



- Data from the high-throughput two-hybrid experiment (T. Ito, *et al.* PNAS (2001) )
- The full set containing 4549 interactions among 3278 yeast proteins
- 87% nodes in the largest component
- The highest connected protein interacts with 285 others!
- Figure shows only nuclear proteins

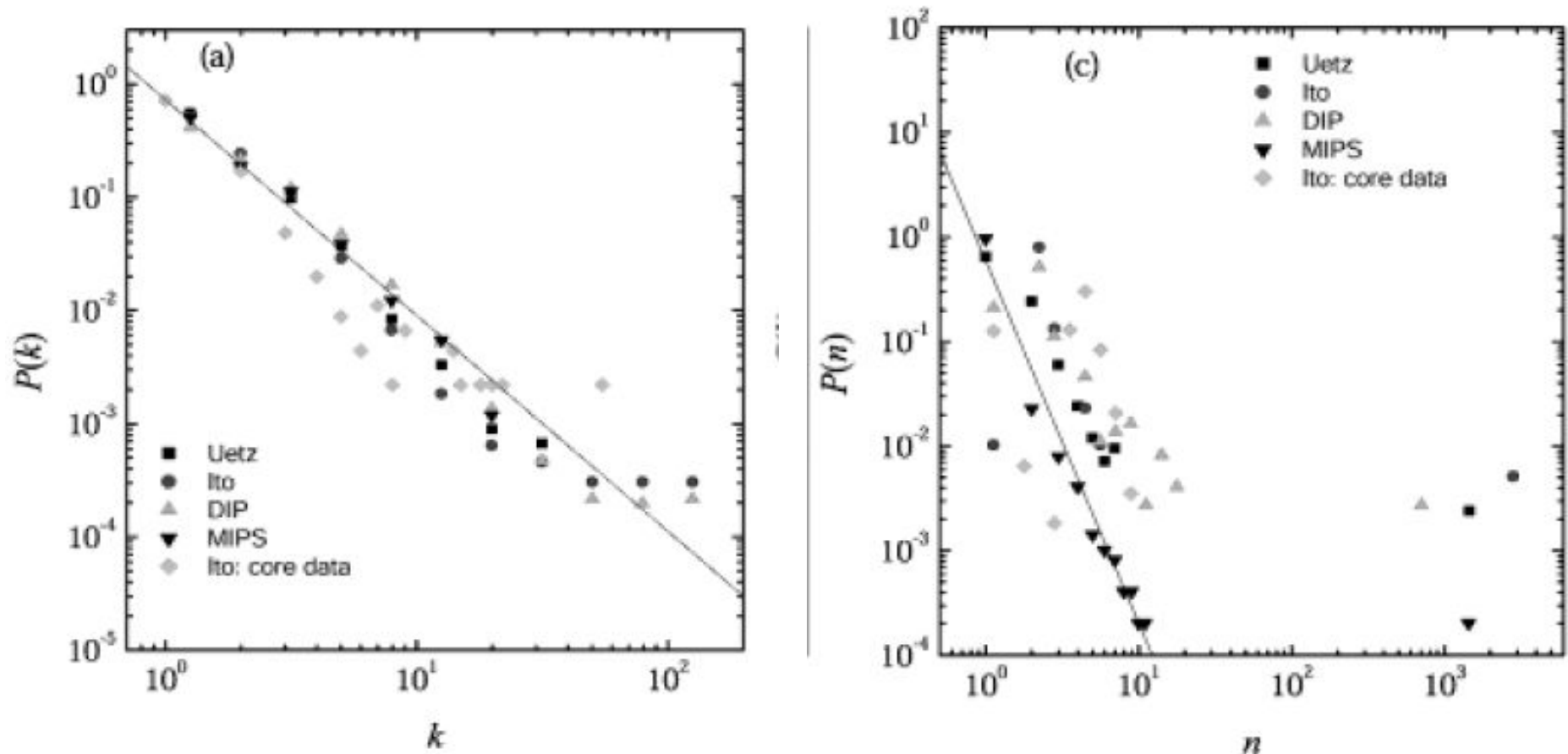


# Гигантская компонента в графе белок-белковых взаимодействий в дрожжах

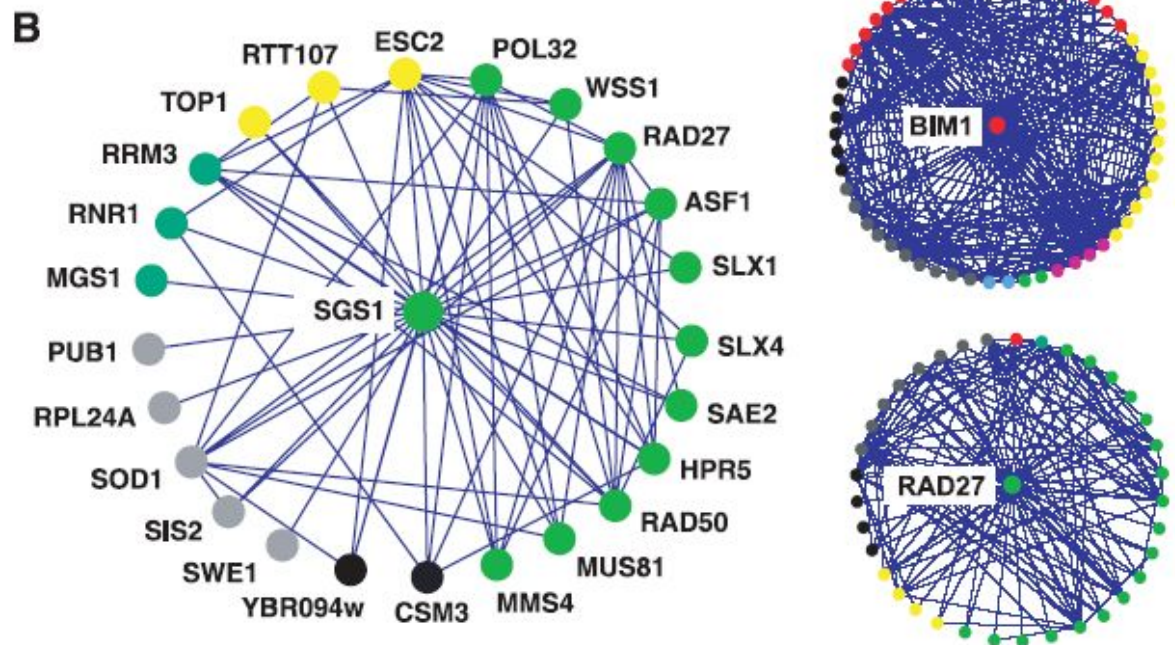
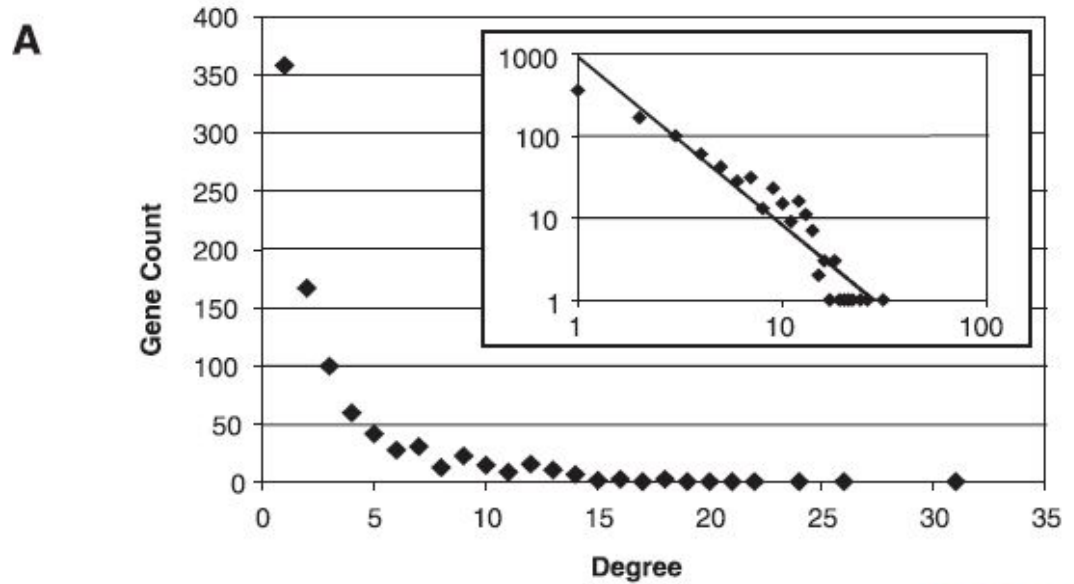


- Красный – летальная мутация
- Оранжевый – медленный рост
- Желтый – неизвестно
- Зеленый – нелетальная мутация

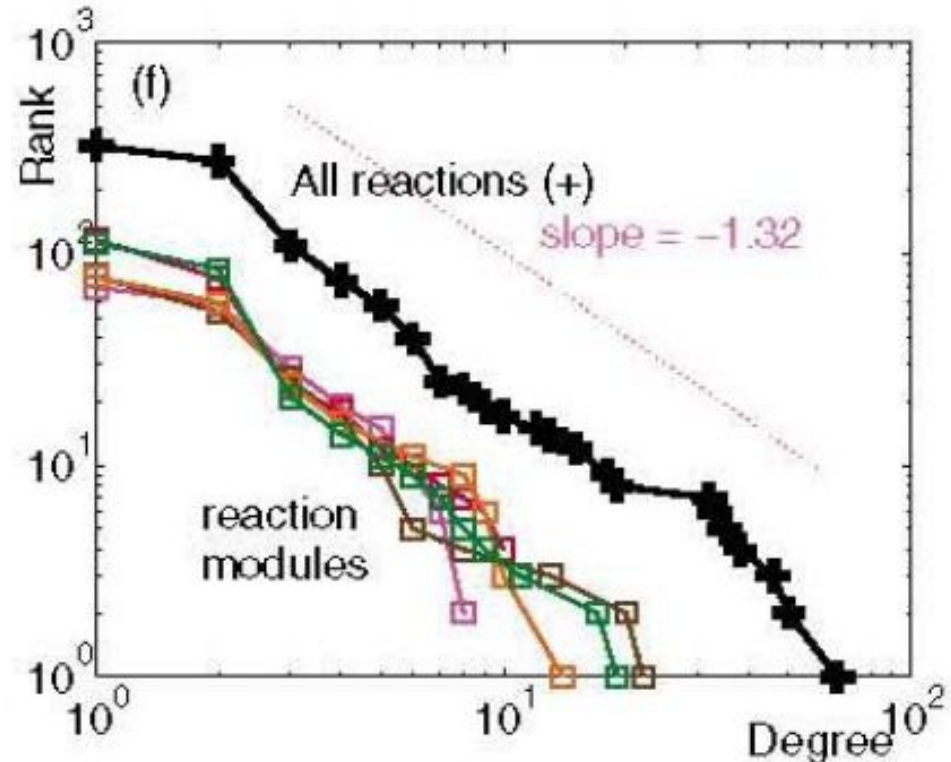
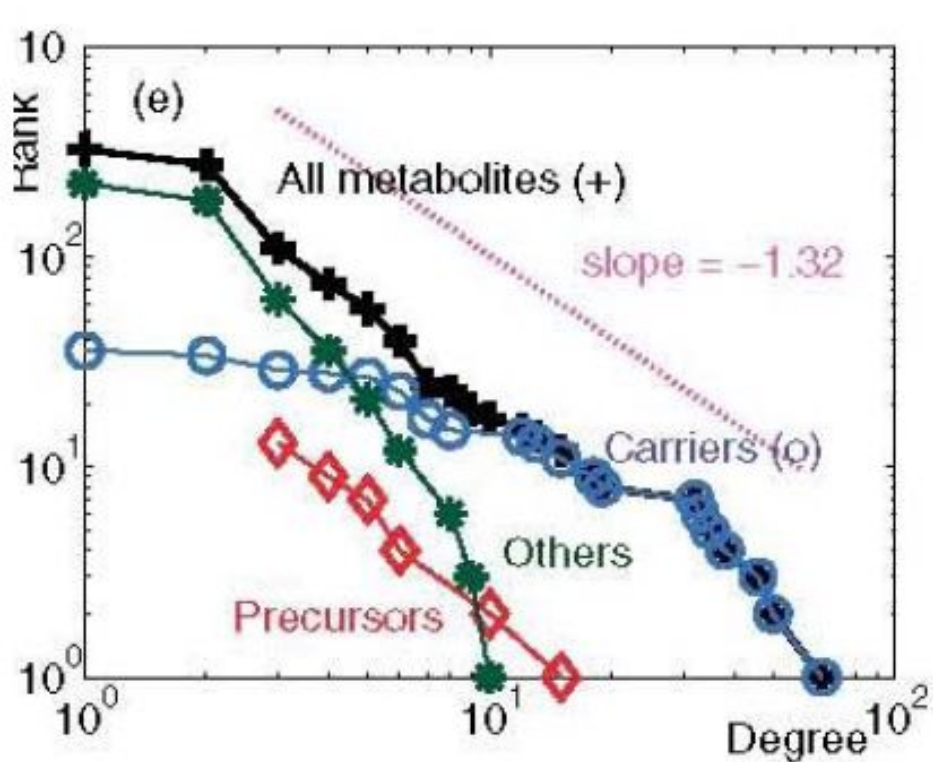
# Белок-белковые взаимодействия в дрожжах: $P(k)$ и размеры связных компонент



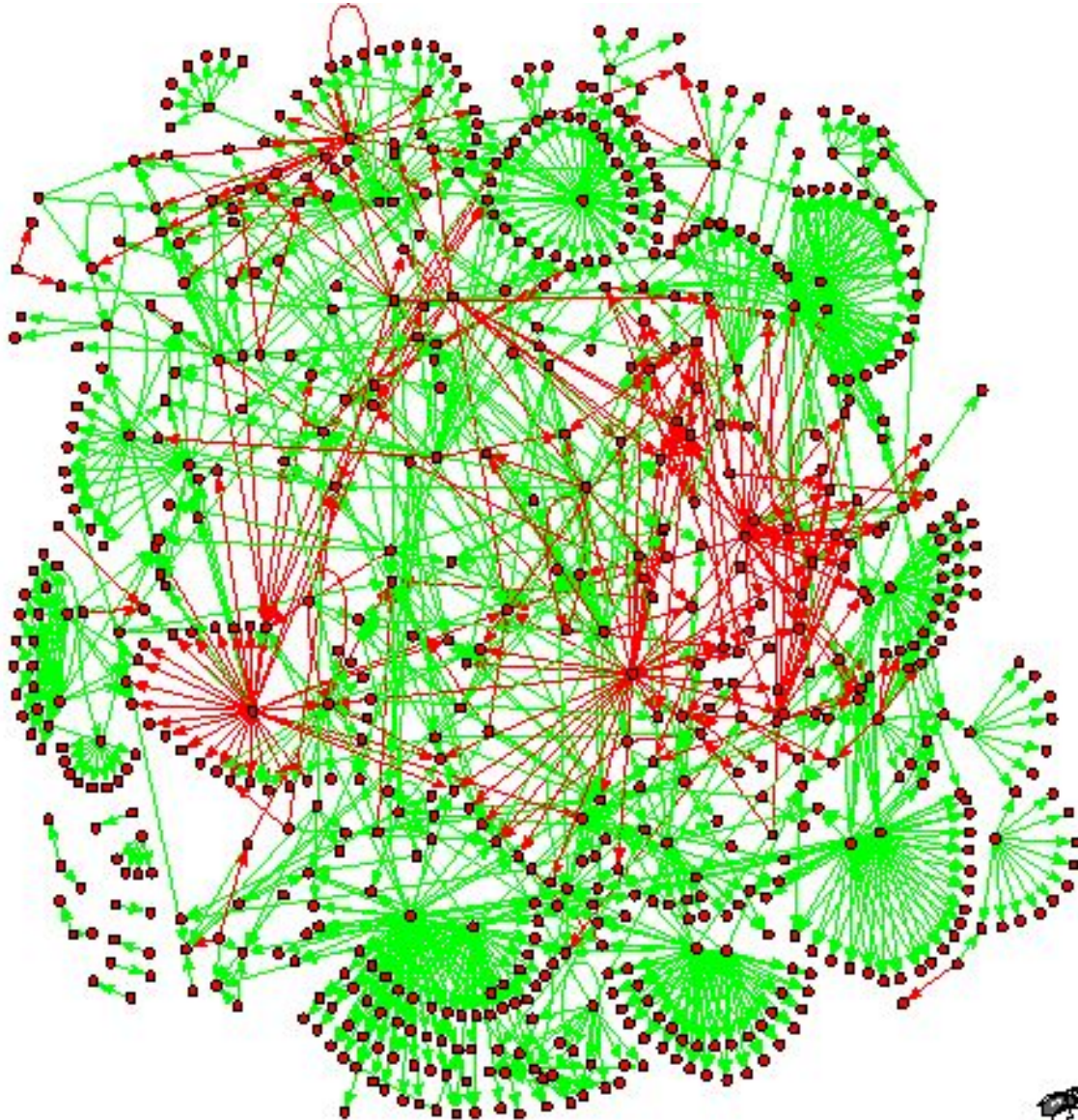
# Synthetic lethals in yeast



# Rank vs. degree for metabolites and reactions in *Helicobacter pylori*

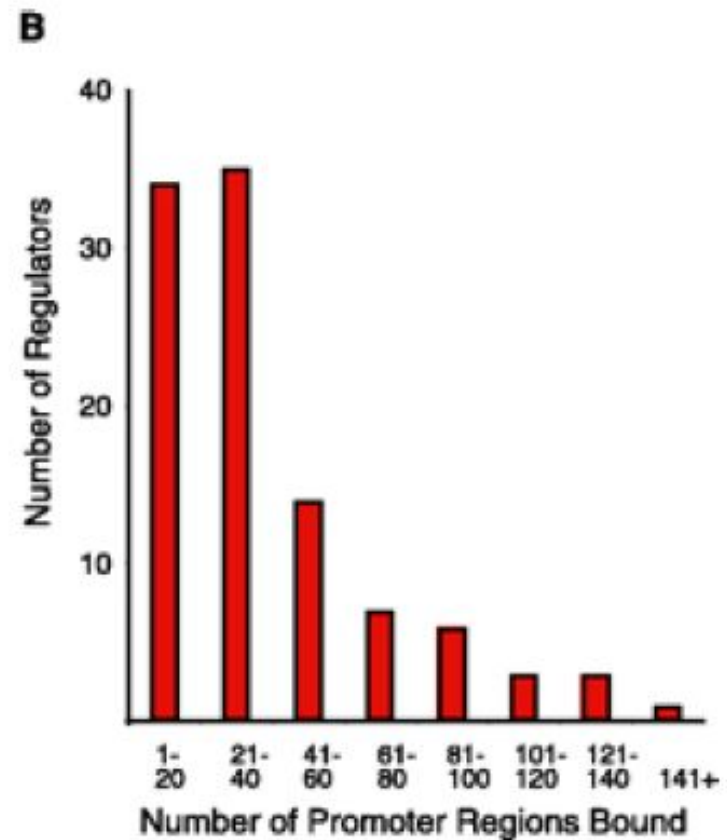
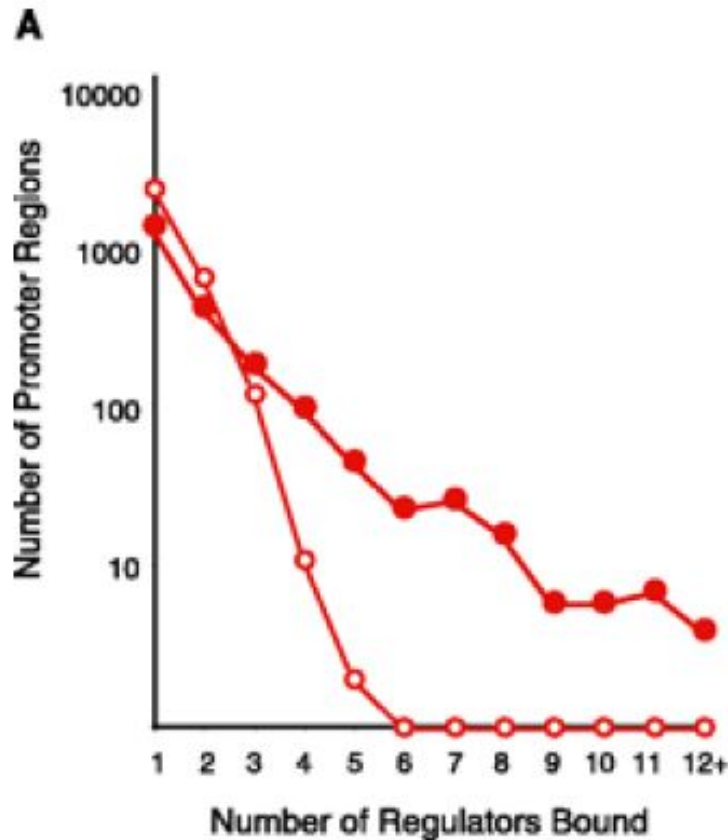


# Transcription regulatory network in baker's yeast



- Downloaded from the YPD database: 1276 regulations among 682 proteins by 125 transcription factors (10 regulated genes per TF)
- Part of a bigger genetic regulatory network of 1772 regulations among 908 proteins
- Positive to negative ratio 3:1
- Broader distribution of out-degrees (up to 72) and more narrow of in-degrees (up to 21)

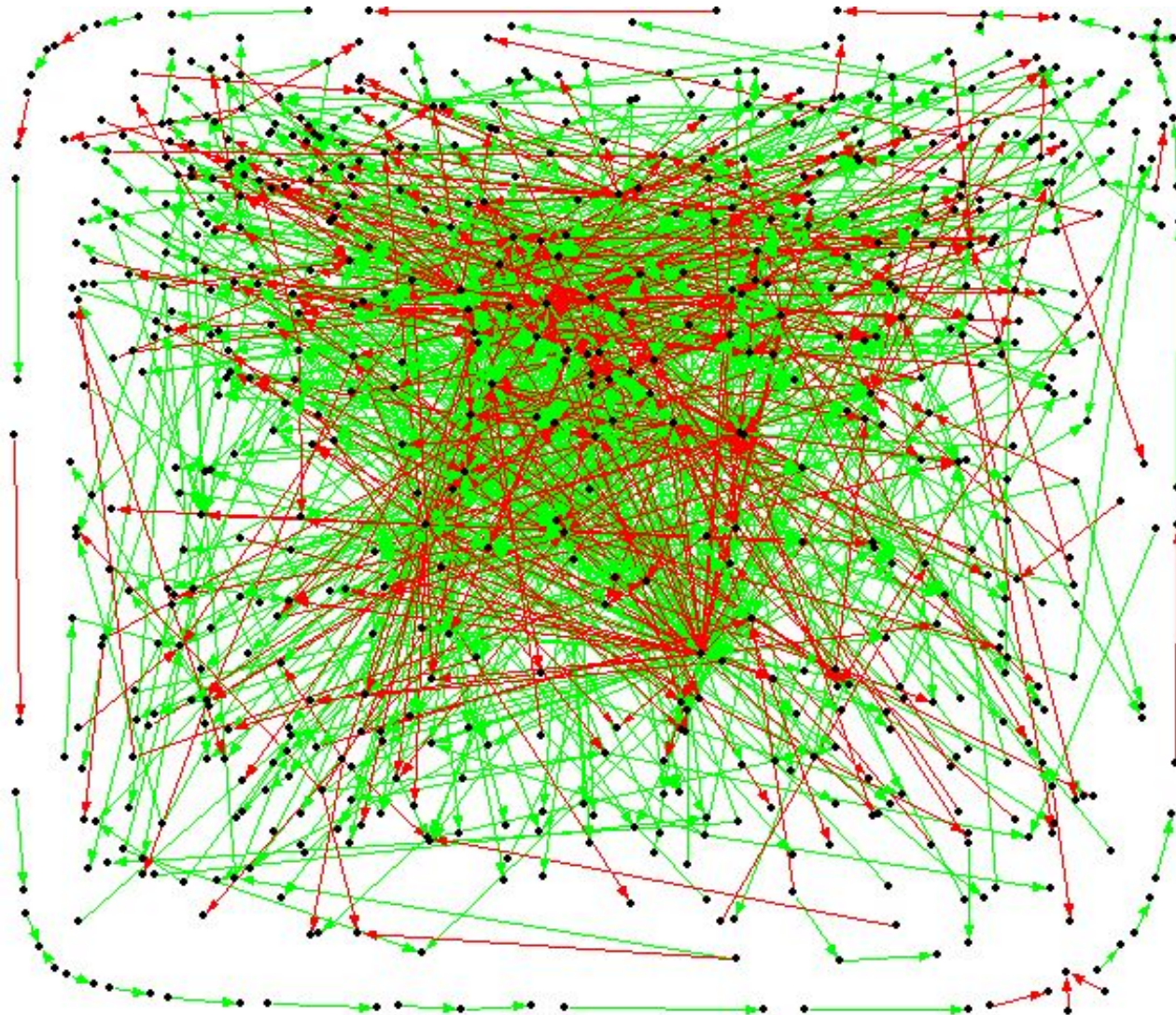
# регуляция транскрипции (дрожжи, ChIP-chip)



- A: in-degree (относительно регулируемых генов): гистограмма (в полулогарифмических координатах) количества промоторов с заданным числом регуляторов – экспоненциальное распределение (у большинства генов мало регуляторов). Пустые кружки – случайный граф
- B: out-degree (относительно факторов): гистограмма количества факторов, связывающих заданное количество промоторов – scale-free

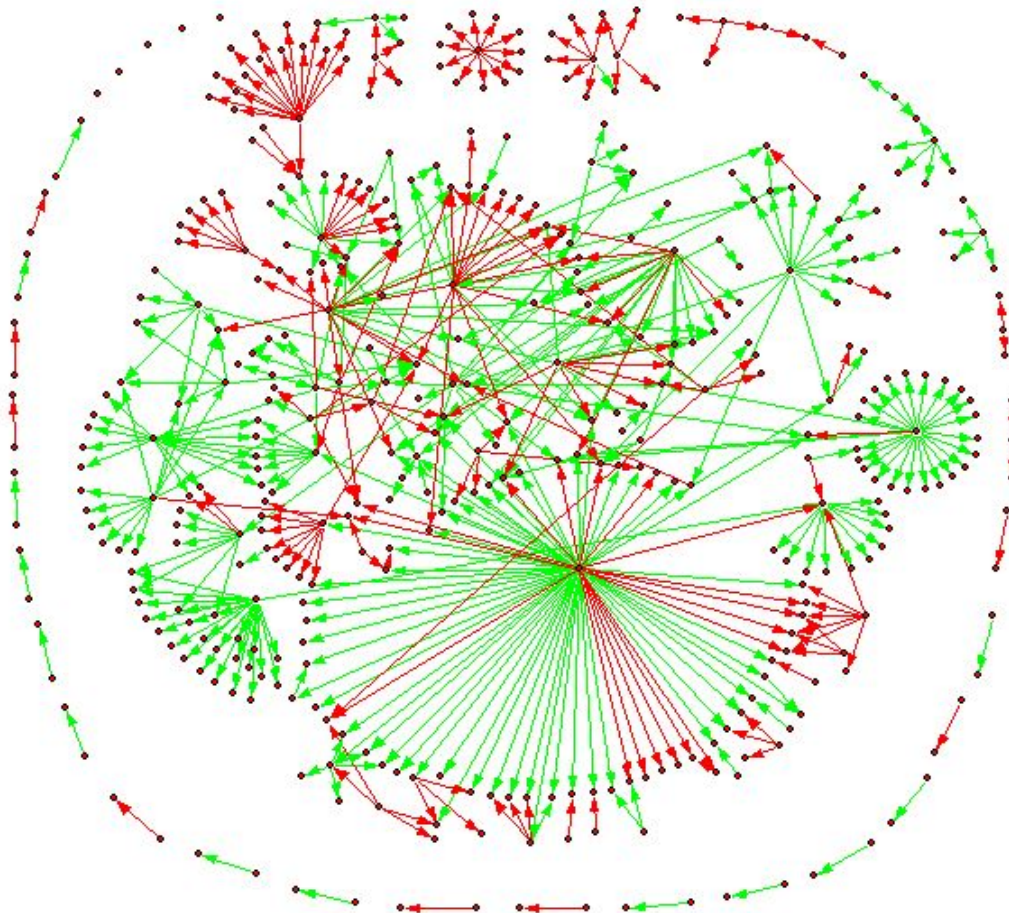


# Transcription regulatory network in *Homo Sapiens*



- Data courtesy of **Ariadne Genomics** obtained from the literature search: **1449** regulations among **689** proteins
- Positive to negative ratio is **3:1 (again!)**
- Broader distribution of out-degrees (up to **95**) and more narrow of in-degrees (up to **40**)

# Transcription regulatory network in *E. coli*



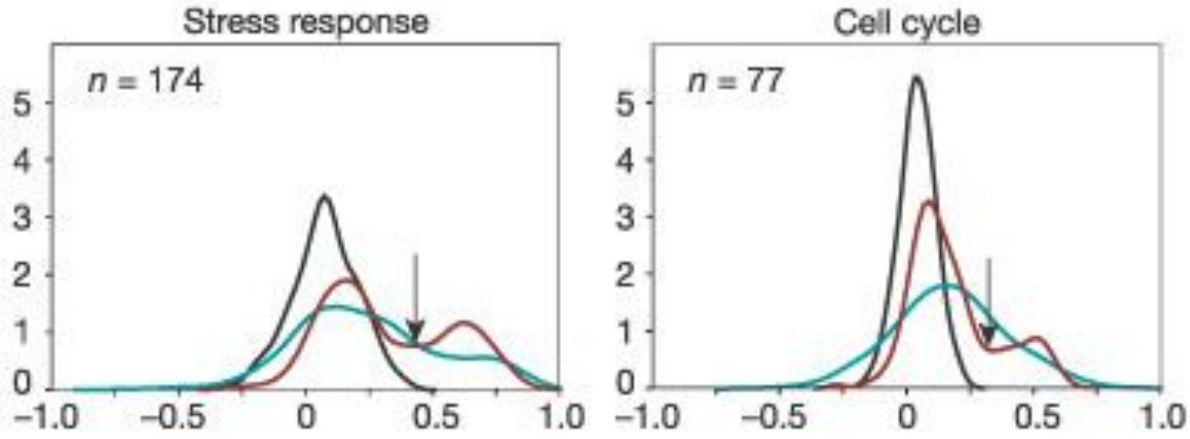
- Data (courtesy of **Uri Alon**) was curated from the Regulon database: **606** interactions between **424** operons (by 116 TFs)
- Positive to negative ratio is **3:2** (**different from eukaryots!**)
- Broader distribution of out-degrees (up to **85**) and more narrow of in-degrees (only up to **6** !)

# зависимость физиологических и геномных свойств от топологии

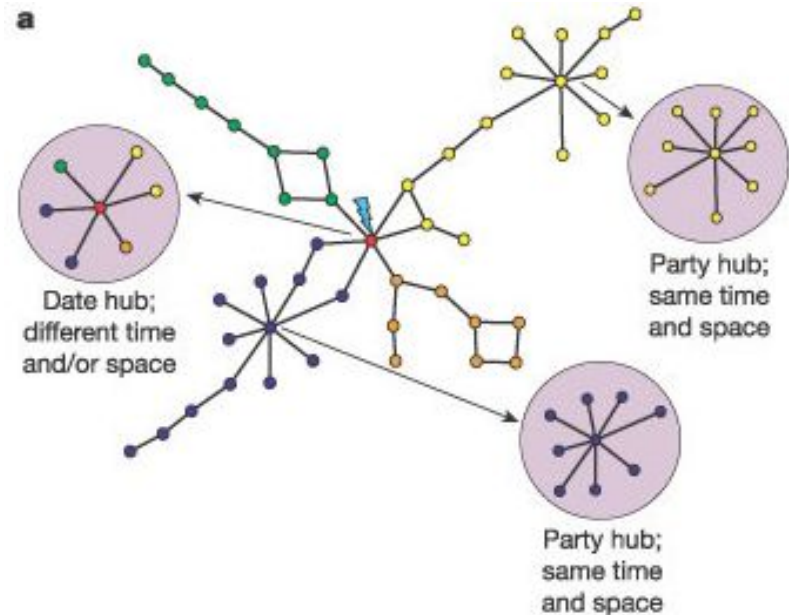
- дрожжи:
  - ~10% genes with <5 links are essential
  - >60% genes with >15 links are essential
- гены с большим числом связей
  - с большей вероятностью имеют ортологов в многоклеточных эукариотах
  - ближе к ортологам из *C. elegans*

# party hubs и date hubs

- Бимодальное распределение корреляций уровня экспрессии
  - Красный: hubs
  - Голубой: non-hubs
  - Черный: случайный граф

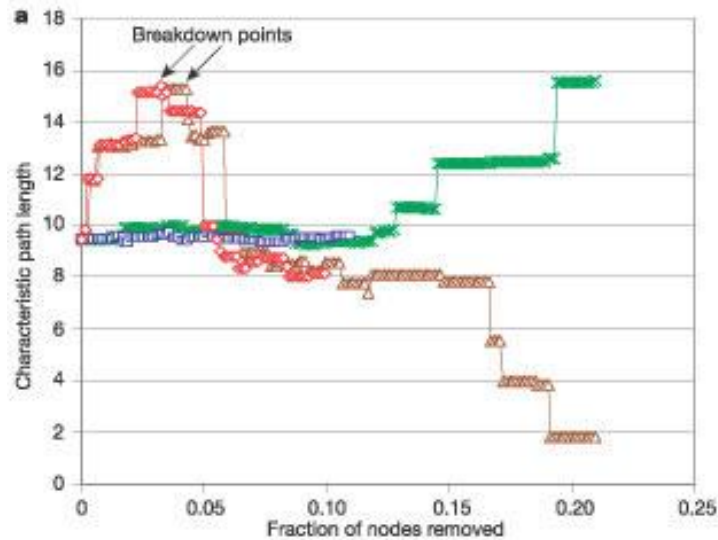


- Party hubs: сам и соседи ко-экспрессируются (комплексы)
- Date hub: нет корреляции в уровнях экспрессии (сигнальные пути)



# Устойчивость к атаке (распадение гигантской компоненты)

ОСНОВА СЕТИ  
– party hubs



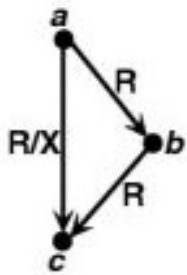
- Красный: атака на party hubs
- Коричневый: атака на все хабы
- Голубой: атака на date hubs
- Зеленый: атака на случайные белки

# МОТИВЫ

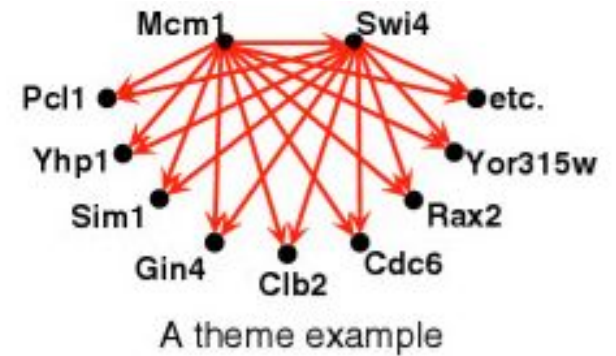
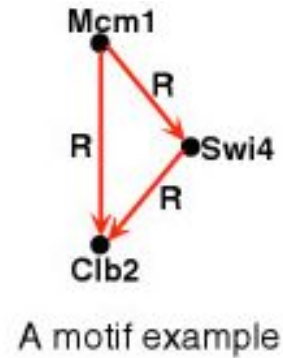
- КЛИКИ
  - много в графах белок-белковых взаимодействий (масс-спек. анализ комплексов – по определению)
- подграфы фиксированной структуры, встречающиеся существенно чаще, чем в случайном графе (с теми же свойствами)

# Регуляторный каскад

Motif set A

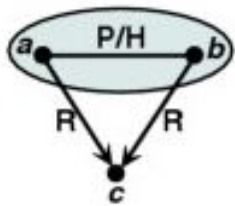


$N_{real}$	$4.7 \times 10^2$	$3.0 \times 10^1$
$N_{rand}$	$(2.6 \pm 0.5) \times 10^2$	$5.4 \pm 3.2$

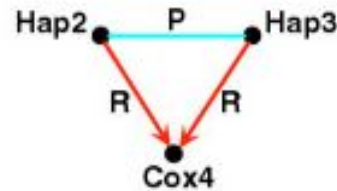


- R – транскрипционная регуляция
- X – ко-экспрессия

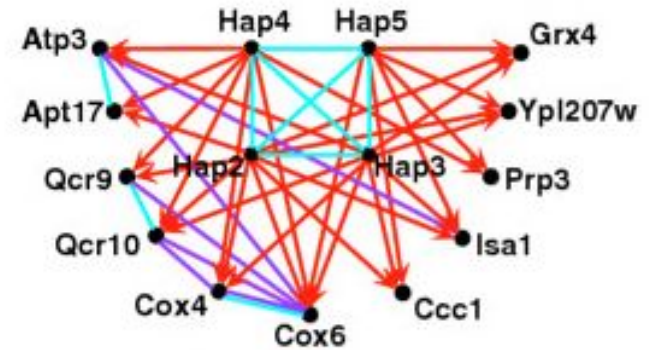
Motif set B



$N_{real}$	$1.3 \times 10^2$	$6.1 \times 10^2$
$N_{rand}$	$3.3 \pm 3.7$	$(8.0 \pm 2.3) \times 10^1$



A motif example



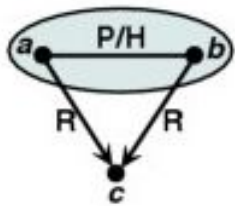
A theme example

- R – транскрипционная регуляция
- P – белок-белковое взаимодействие
- H – гомология

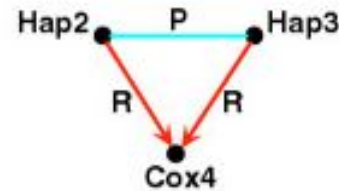


# Субъединицы факторов транскрипции

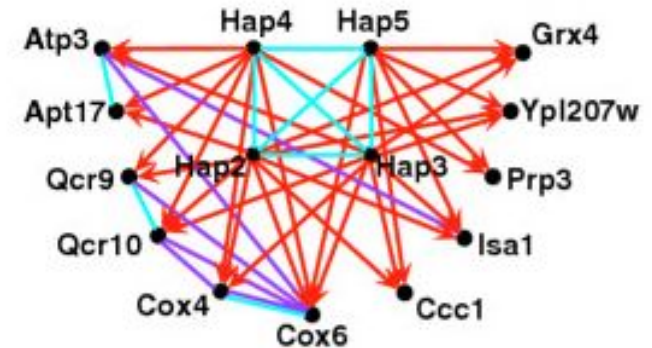
Motif set B



$N_{real}$	$1.3 \times 10^2$	$6.1 \times 10^2$
$N_{rand}$	$3.3 \pm 3.7$	$(8.0 \pm 2.3) \times 10^3$



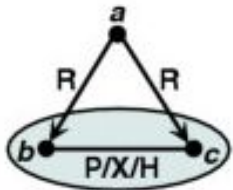
A motif example



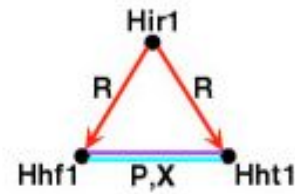
A theme example

- R – транскрипционная регуляция
- P – белок-белковое взаимодействие
- H – гомология

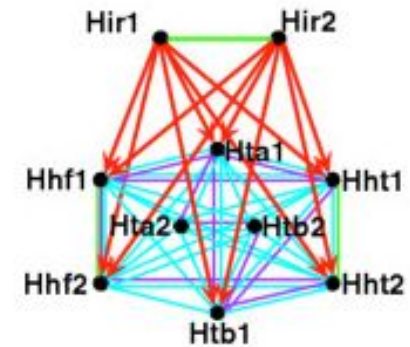
Motif set C



	C1	C2	C3
$N_{\text{real}}$	$5.9 \times 10^3$	$3.5 \times 10^3$	$1.9 \times 10^3$
$N_{\text{rand}}$	$(5.4 \pm 0.5) \times 10^2$	$(2.7 \pm 0.3) \times 10^2$	$(5.3 \pm 0.5) \times 10^2$



A motif example

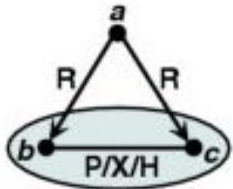


A theme example

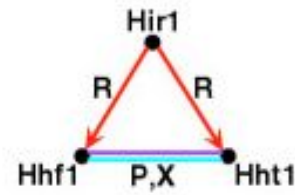
- R – транскрипционная регуляция
- P – белок-белковое взаимодействие
- X – ко-экспрессия
- H – гомология

# Регулоны

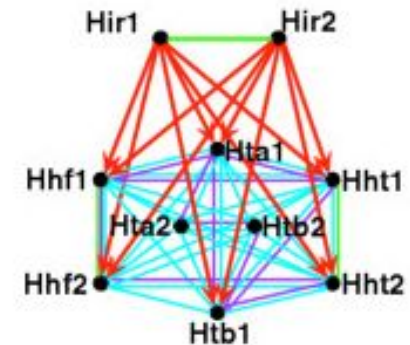
Motif set C



	C1	C2	C3
$N_{\text{real}}$	$5.9 \times 10^3$	$3.5 \times 10^3$	$1.9 \times 10^3$
$N_{\text{rand}}$	$(5.4 \pm 0.5) \times 10^2$	$(2.7 \pm 0.3) \times 10^2$	$(5.3 \pm 0.5) \times 10^2$



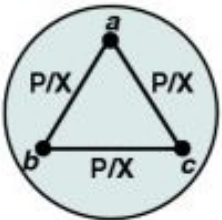
A motif example



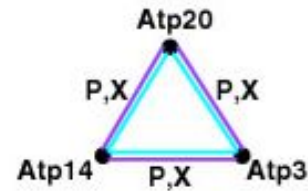
A theme example

- R – транскрипционная регуляция
- P – белок-белковое взаимодействие
- X – ко-экспрессия
- H – гомология

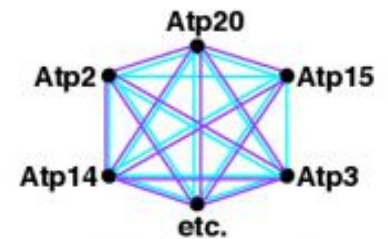
Motif set D



	D1	D2	D3	D4
$N_{real}$	$5.7 \times 10^5$	$9.9 \times 10^4$	$6.7 \times 10^4$	$1.2 \times 10^6$
$N_{rand}$	$(1.1 \pm 0.0) \times 10^5$	$(8.2 \pm 0.3) \times 10^3$	$(5.2 \pm 0.2) \times 10^3$	$(2.7 \pm 0.1) \times 10^4$



A motif example

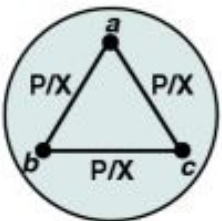


A theme example

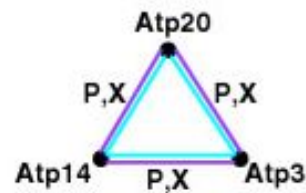
- P – белок-белковое взаимодействие
- X – ко-экспрессия

# Ко-экспрессия в комплексах

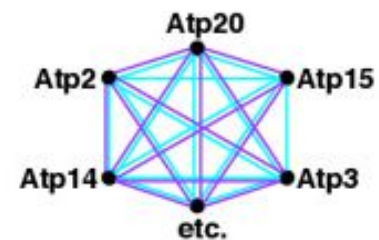
Motif set D



	D1	D2	D3	D4
$N_{real}$	$5.7 \times 10^5$	$9.9 \times 10^4$	$6.7 \times 10^4$	$1.2 \times 10^6$
$N_{rand}$	$(1.1 \pm 0.0) \times 10^5$	$(8.2 \pm 0.3) \times 10^3$	$(5.2 \pm 0.2) \times 10^3$	$(2.7 \pm 0.1) \times 10^4$



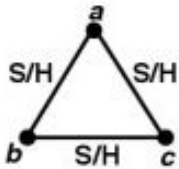
A motif example



A theme example

- P – белок-белковое взаимодействие
- X – ко-экспрессия

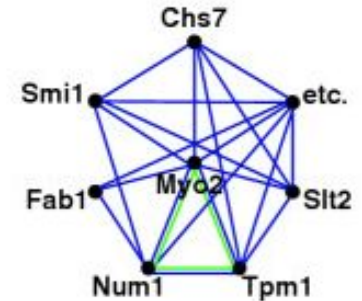
Motif set E



	E1	E2	E3	E4
$N_{real}$	$2.7 \times 10^3$	$9.8 \times 10^2$	$3.2 \times 10^3$	$5.6 \times 10^5$
$N_{rand}$	$(1.7 \pm 0.1) \times 10^3$	$(3.8 \pm 0.4) \times 10^2$	$(1.3 \pm 0.1) \times 10^3$	$(1.0 \pm 0.2) \times 10^5$



A motif example

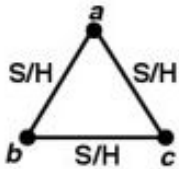


A theme example

- S – синтетические летали (слабость)
- H – гомология

# Взаимозаменяемость паралогов (?)

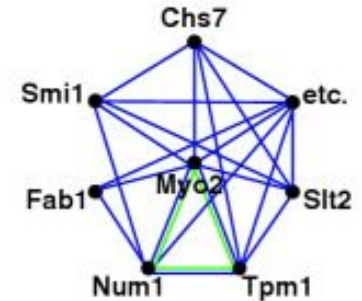
Motif set E



	E1	E2	E3	E4
$N_{real}$	$2.7 \times 10^3$	$9.8 \times 10^2$	$3.2 \times 10^3$	$5.6 \times 10^5$
$N_{rand}$	$(1.7 \pm 0.1) \times 10^3$	$(3.8 \pm 0.4) \times 10^2$	$(1.3 \pm 0.1) \times 10^3$	$(1.0 \pm 0.2) \times 10^5$



A motif example



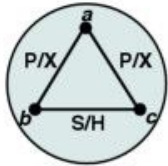
A theme example

- S – синтетические летали (слабость)
- H – гомология

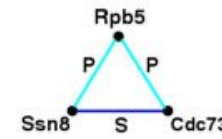
# Компенсаторные комплексы (?)

(f)

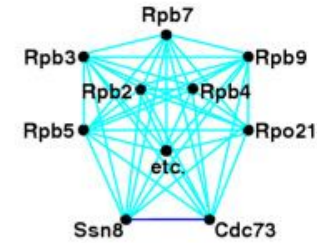
Motif set F



	F1	F2	F3	F4	F5	F6
$N_{real}$	$1.2 \times 10^3$	$2.7 \times 10^2$	$2.8 \times 10^2$	$4.1 \times 10^4$	$1.1 \times 10^4$	$4.4 \times 10^4$
$N_{rand}$	$(7.6 \pm 0.7) \times 10^2$	$(1.5 \pm 0.3) \times 10^2$	$(1.3 \pm 0.2) \times 10^2$	$(1.1 \pm 0.0) \times 10^4$	$(2.0 \pm 0.1) \times 10^3$	$(2.4 \pm 0.1) \times 10^3$



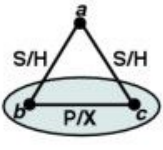
A motif example



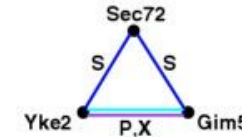
A theme example

(g)

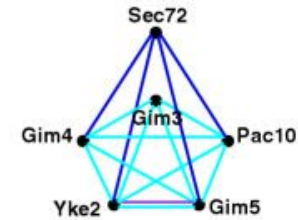
Motif set G



	G1	G2	G3	G4	G5	G6
$N_{real}$	$2.0 \times 10^3$	$7.2 \times 10^2$	$3.0 \times 10^4$	$2.8 \times 10^2$	$2.5 \times 10^2$	$4.3 \times 10^4$
$N_{rand}$	$(2.4 \pm 0.3) \times 10^2$	$(3.5 \pm 0.3) \times 10^2$	$(1.2 \pm 0.1) \times 10^4$	$(7.0 \pm 1.5) \times 10^1$	$(1.2 \pm 0.2) \times 10^2$	$(4.0 \pm 0.2) \times 10^3$



A motif example



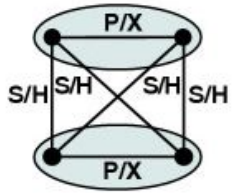
A theme example

- S – синтетические летали (слабость)
- H – гомология
- P – белок-белковое взаимодействие
- X – ко-экспрессия



# Четверные мотивы: взаимозаменяемость

(a)



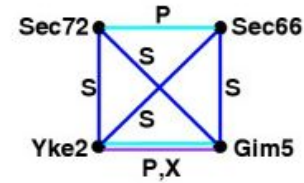
(b)

					etc.
$N_{real}$	$3.5 \times 10^2$	$6.7 \times 10^1$	$1.6 \times 10^1$	$3.8 \times 10^1$	
$N_{rand}$	$0.16 \pm 0.50$	$0.13 \pm 0.39$	$1.1 \pm 1.4$	$5.9 \pm 4.1$	

**Key**

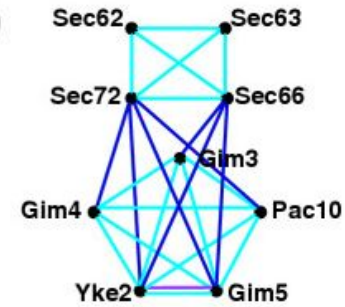
- S: synthetic sickness or lethality
- H: sequence homology
- X: correlated expression
- P: stable physical interaction

(c)



A motif example

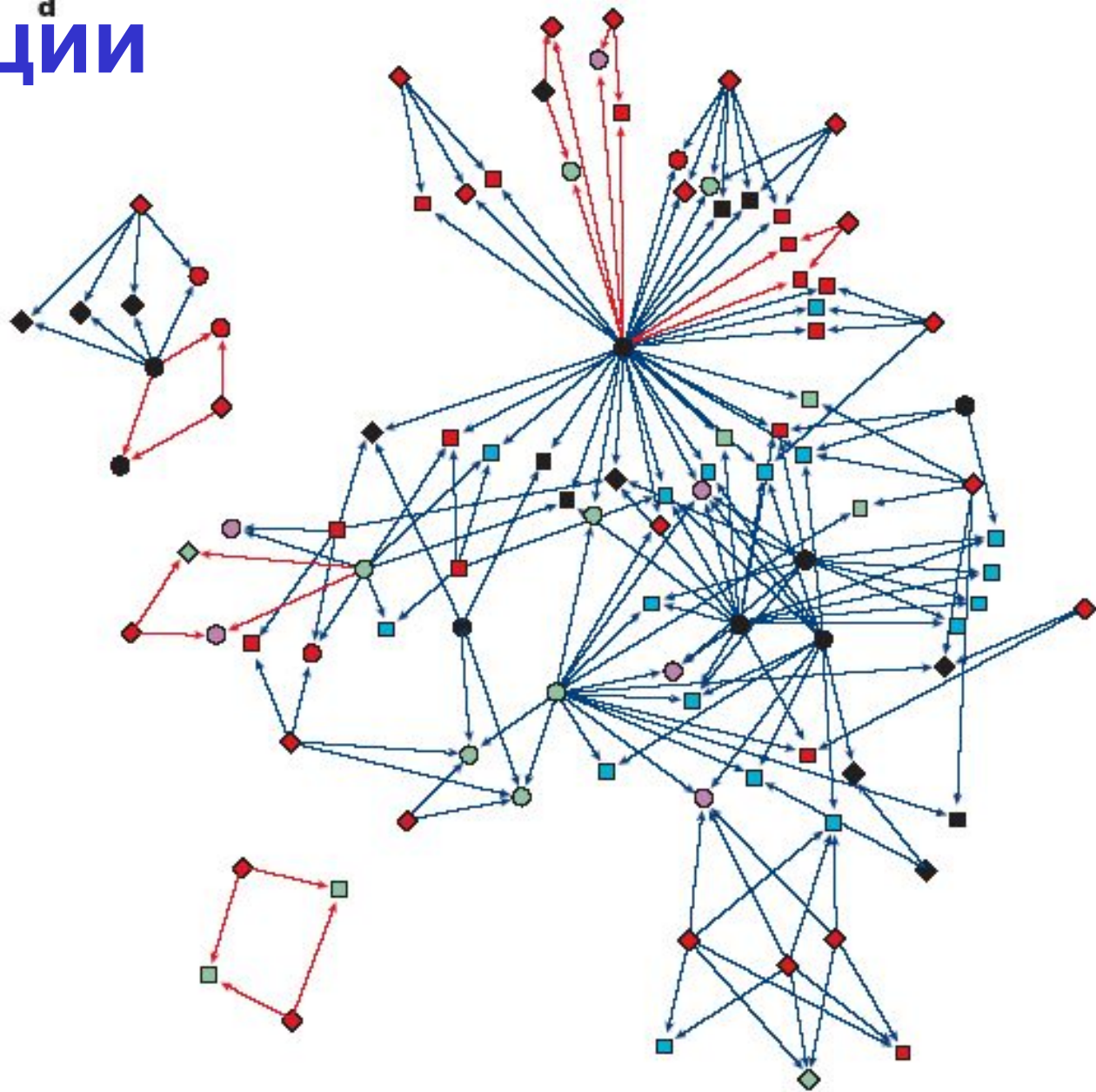
(d)



A theme example

# Регуляция транскрипции<sup>d</sup> в *E.coli*

- Почти все  
“bi-fan”  
МОТИВЫ  
СВЯЗАНЫ  
друг с  
другом



# ЭВОЛЮЦИЯ

- rich get richer
- дубликации
- случайные рождения/исчезновение ребер

