

Transcription regulation:  
a genomic network

**Nicholas Luscombe**

Laboratory of **Mark Gerstein**  
Department of Molecular Biophysics  
and Biochemistry  
Yale University

**Gerstein lab:** Haiyuan Yu  
**Snyder lab:** Christine Horak  
**Teichmann lab:** Madan Babu

**proteomics**

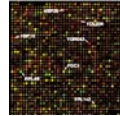
Understand **Proteins**,  
through analyzing **populations**

**Structures** (Motions, Composition)

**Functions** (Locations, Interactions)

**Evolution** (Pseudogenes)

Work on a diverse range of data



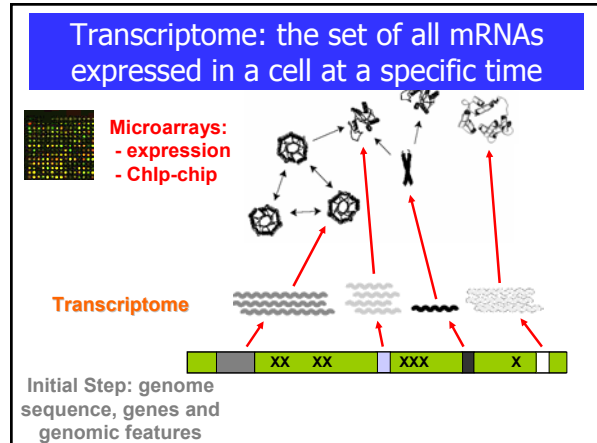
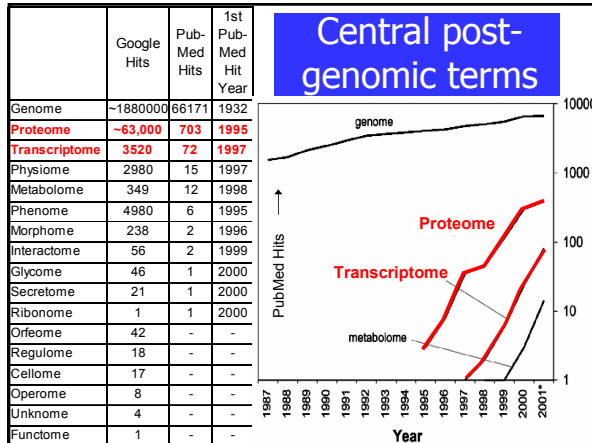
Microarrays



Sequences



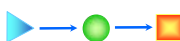
Structures



Overview



1. ChIP-chip experiments for 11 transcription factors - SBF & MBF



2. Computational studies of yeast regulatory network

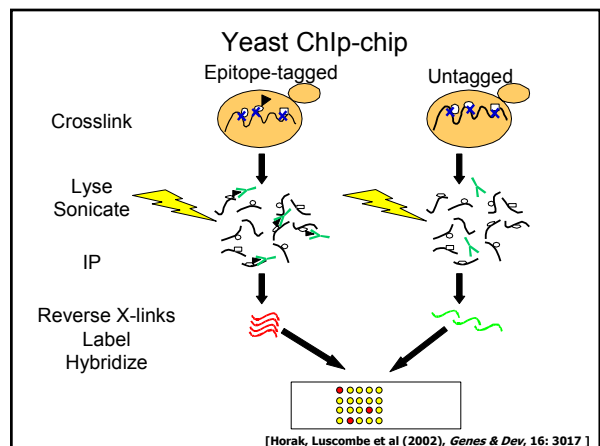
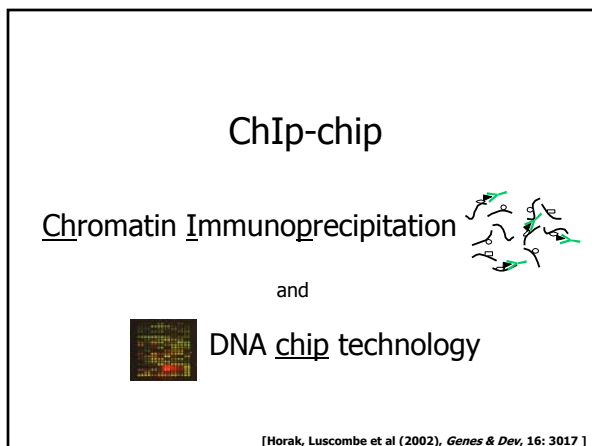
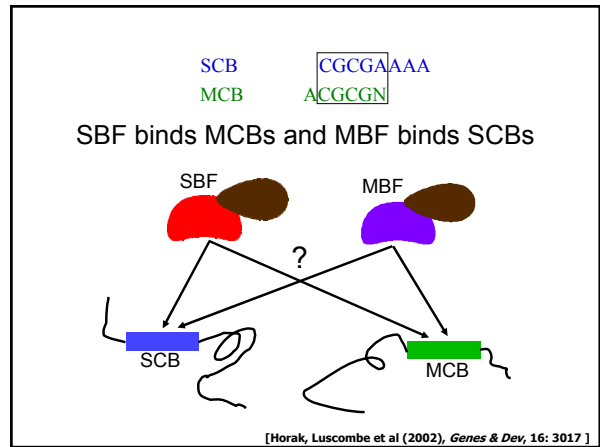
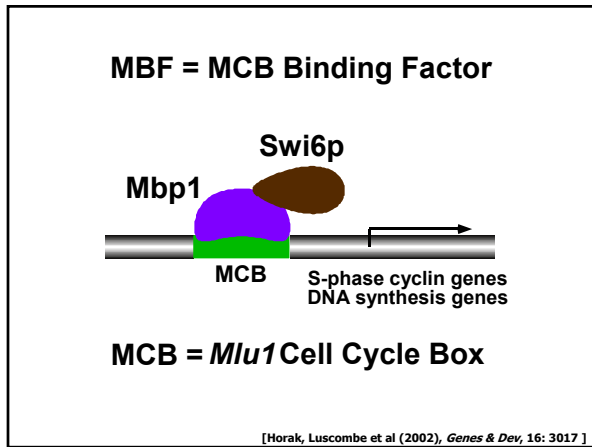
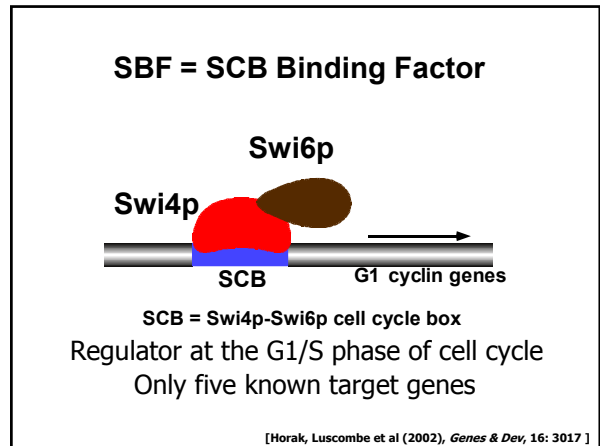
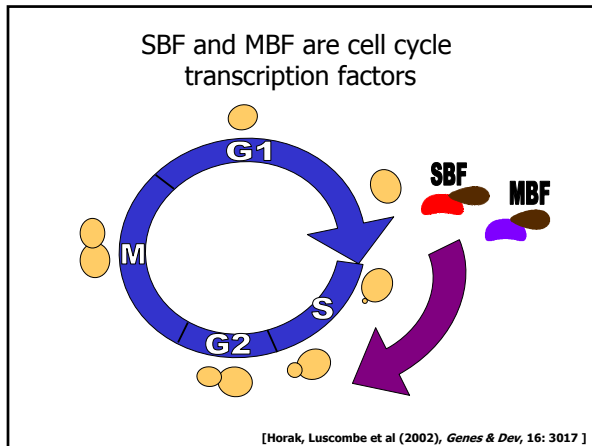
Overview

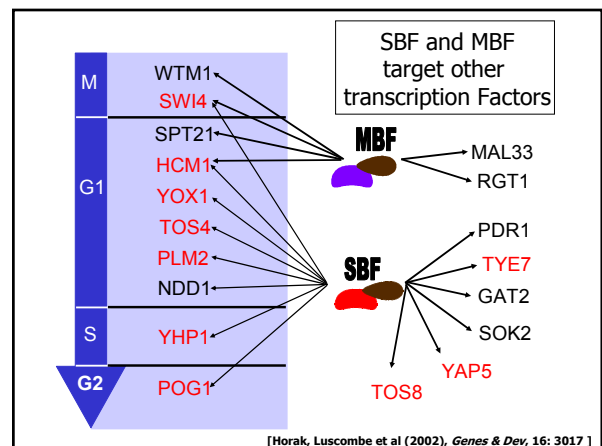
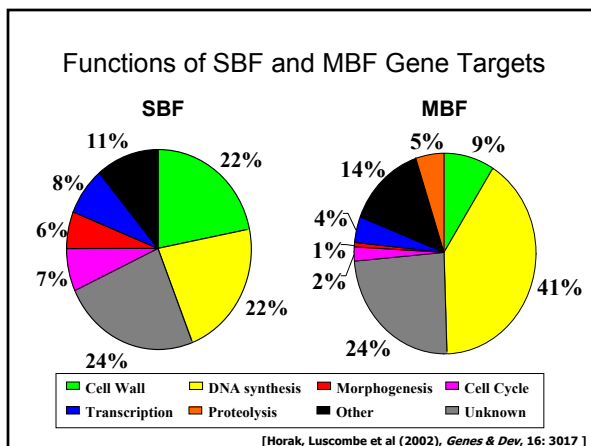
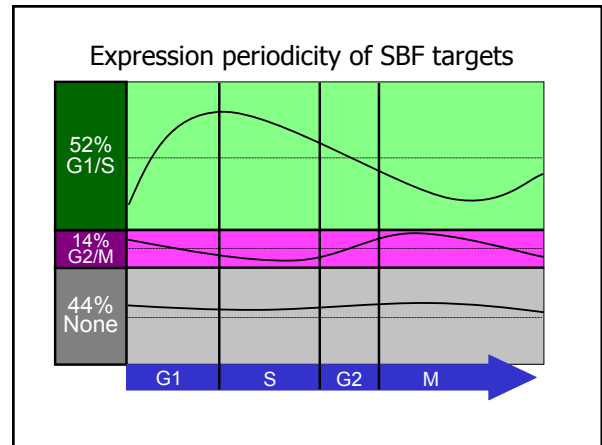
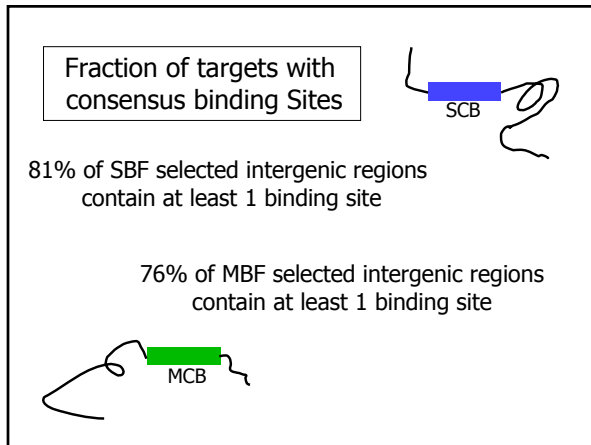
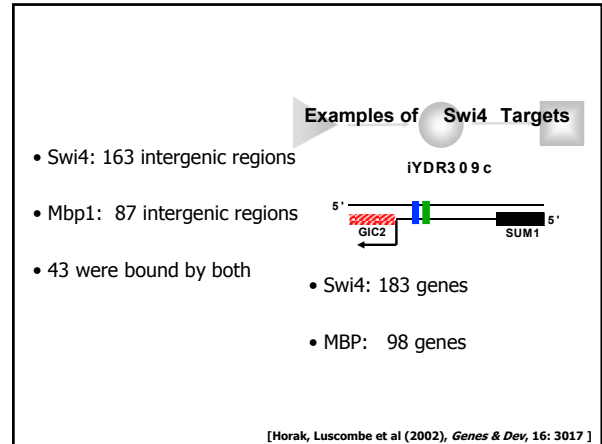
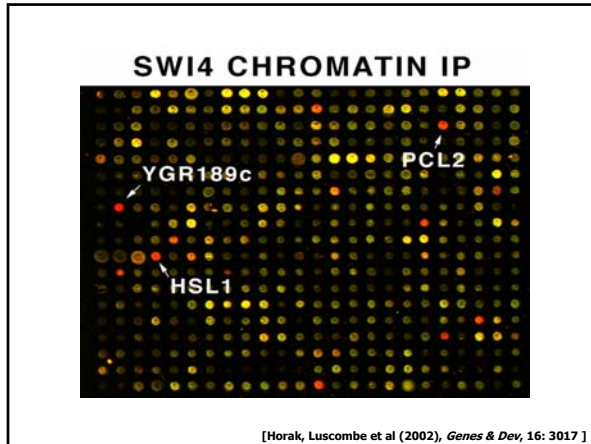


1. ChIP-chip experiments for 11 transcription factors - SBF & MBF

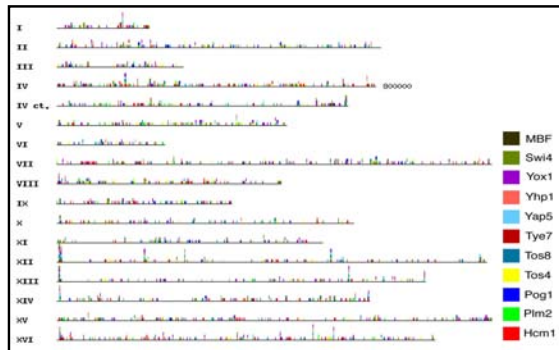


2. Computational studies of yeast regulatory network

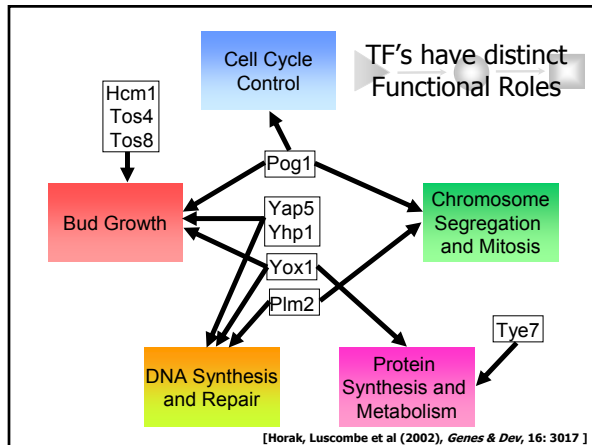
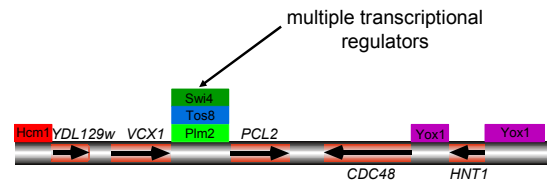




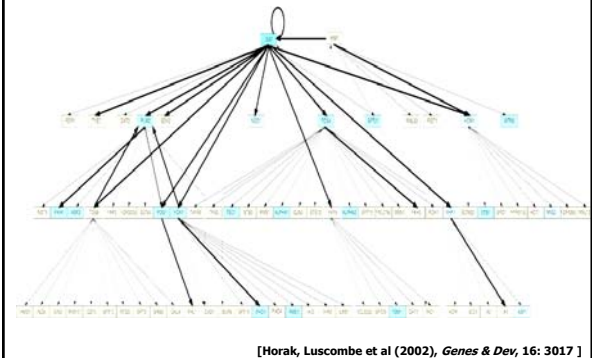
## Genomic distribution of TF binding sites



## Genomic distribution of TF binding sites



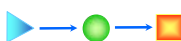
## Yeast transcriptional regulatory network



## Overview



1. ChIP-chip experiments for 11 transcription factors - SBF & MBF



2. Computational studies of yeast regulatory network

## Comprehensive TF dataset

	Dataset	Authors	URL	# of genes	# of regulations
Manual collection	TRANSFAC	Wingender, E., et al. 2001	<a href="http://transfac.gbf.de/TRANSFAC/">http://transfac.gbf.de/TRANSFAC/</a>	288	356
	Keiper's dataset	Gueltrin, N., et al. 2002	<a href="http://www.yeastgenome.org/journal/v31/iss1/supplinfo/ng873_S1.html">http://www.yeastgenome.org/journal/v31/iss1/supplinfo/ng873_S1.html</a>	477	906
ChIP-chip experiments	ChIP-chip data by Snyder's lab	Horak, C. E., et al. 2002	<a href="http://array.mbb.yale.edu/yeast/transcription/download.html">http://array.mbb.yale.edu/yeast/transcription/download.html</a>	1560	2124
	ChIP-chip data by Young's lab	Lee, T. I., et al. 2002	<a href="http://web.mit.edu/young/regulator_network/">http://web.mit.edu/young/regulator_network/</a>	2416	4358

142 transcription factors  
 3,420 target genes  
 7,074 regulatory interactions

[Yu, Luscombe et al (2003), *Trends Genet*, 19: 422]

### Comprehensive Yeast TF network

- Very complex network
- But we can simplify:
  - Topological measures
  - Network motifs

[Barabasi, Alon]

### Comprehensive Yeast TF network

- Very complex network
- But we can simplify:
  - Topological measures
  - Network motifs

[Barabasi, Alon]

### Topological measures

Indicate the gross topological structure of the network

Degree      Path length      Clustering coefficient

[Barabasi]

### Topological measures

Number of incoming and outgoing connections

Degree

Incoming degree = 2.1  
→ each gene is regulated by ~2 TFs

Outgoing degree = 49.8  
→ each TF targets ~50 genes

[Barabasi]

### Scale-free distribution of outgoing degree

Regulatory hubs  
>100 target genes  
Dictate structure of network

- Most TFs have few target genes
- Few TFs have many target genes

[Barabasi]

### Topological measures

Number of intermediate TFs until final target

Indicate how immediate a regulatory response is

Average path length = 4.7

Path length = 1

[Barabasi]

### Topological measures

Ratio of existing links to maximum number of links for neighbouring nodes

Measure how inter-connected the network is

**Average coefficient = 0.11**

4 neighbours  
1 existing link  
6 possible links

Clustering coefficient  
=  $1/6 = 0.17$

[Barabasi]

### Comprehensive Yeast TF network

- Very complex network
- But we can simplify:
  - Topological measures
  - Network motifs

[Alon]

### Network motifs

Regulatory modules within the network

SIM      MIM      FBL      FFL

[Alon]

### SIM = Single input motifs

..... ECM22    STB1    SPO1    YPR013C .....

[Alon; Horak, Luscombe et al (2002), *Genes & Dev*, 16: 3017]

### MIM = Multiple input motifs

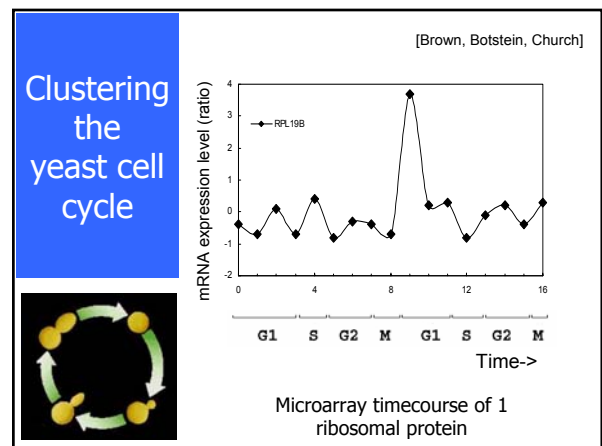
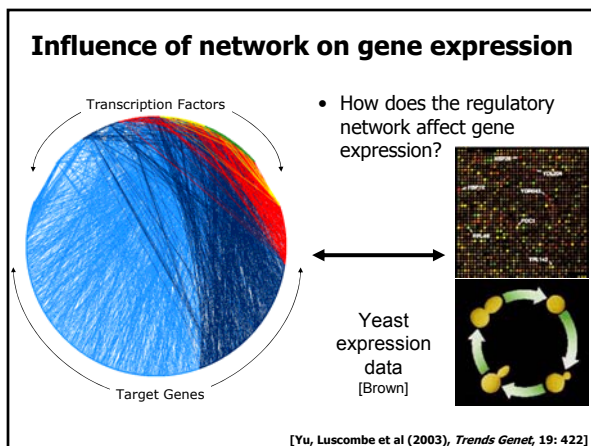
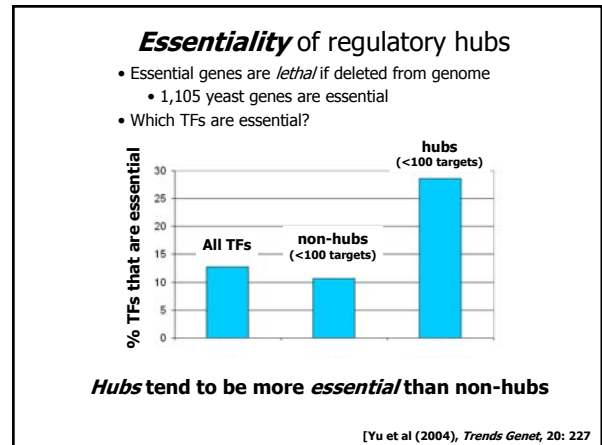
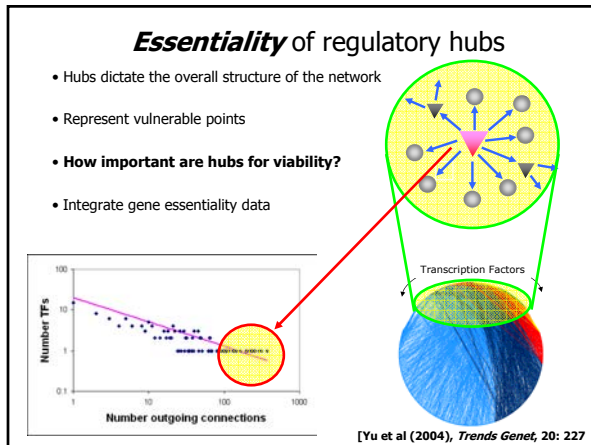
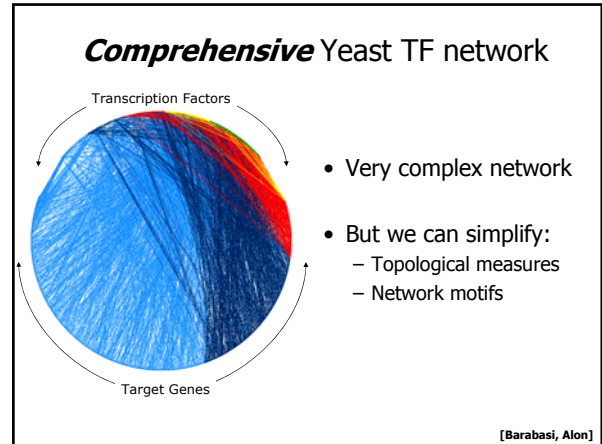
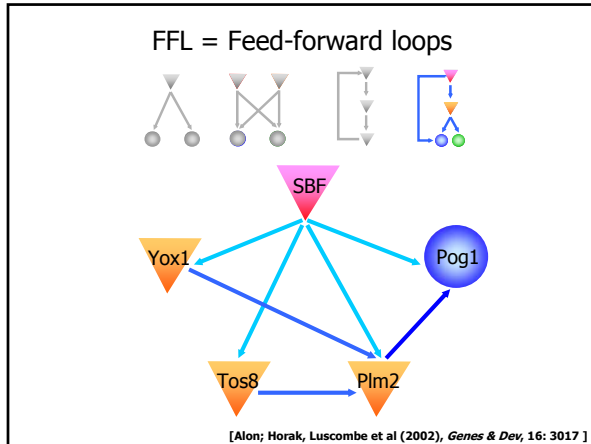
SPT21      HCM1

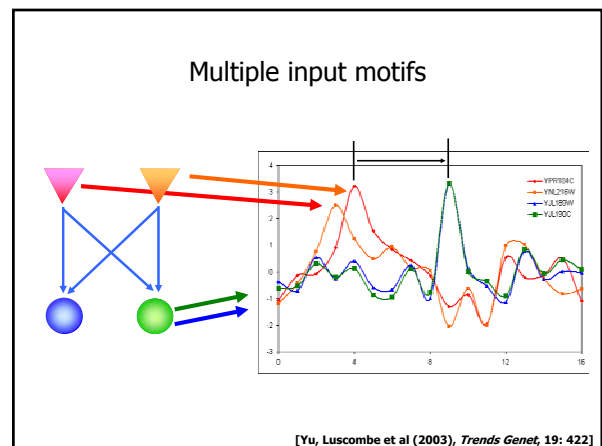
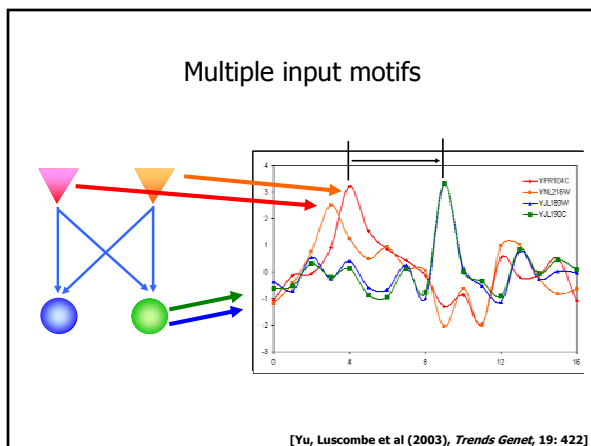
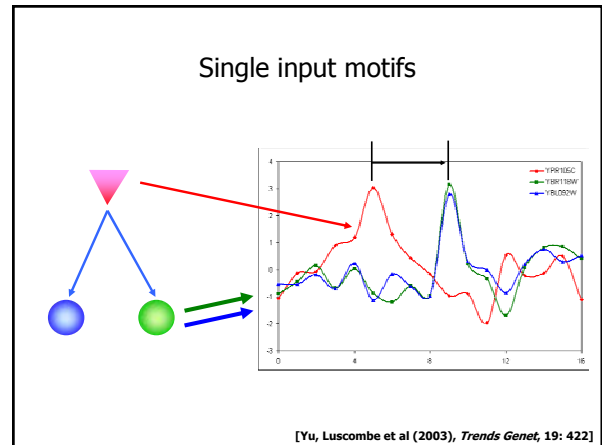
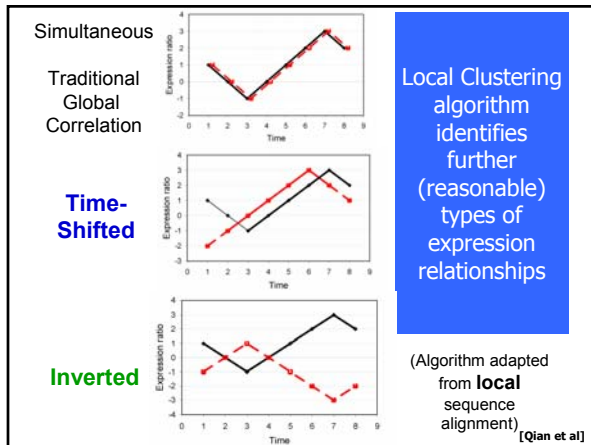
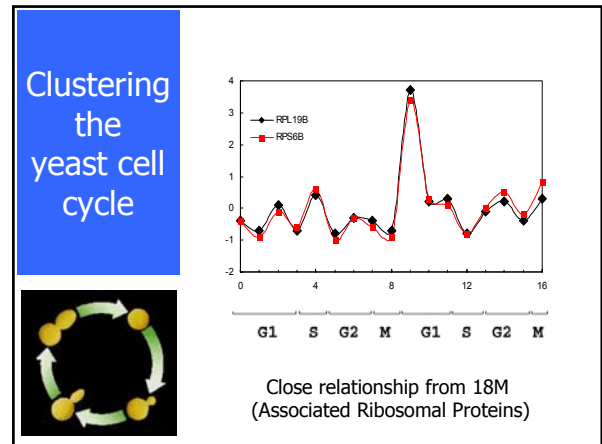
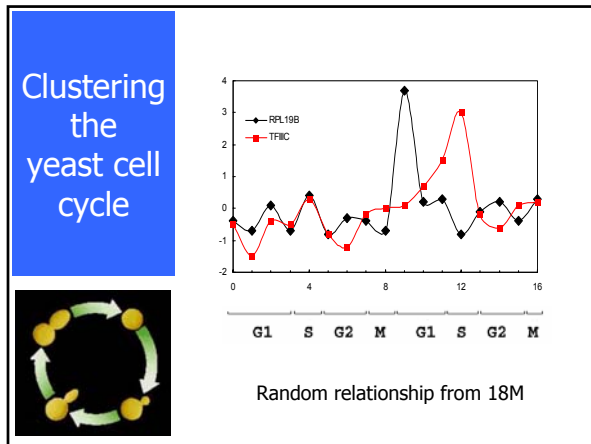
[Alon; Horak, Luscombe et al (2002), *Genes & Dev*, 16: 3017]

### FBL = Feed-back loops

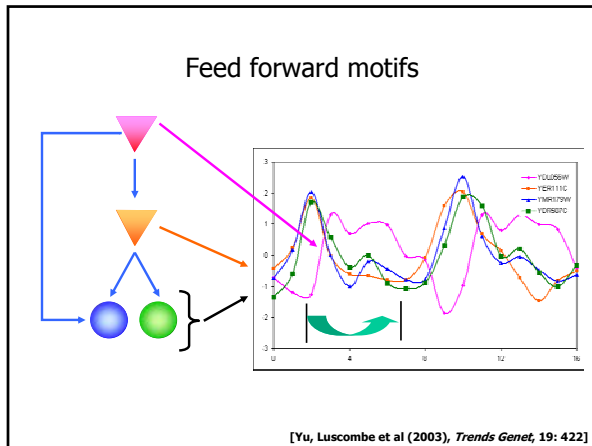
Tos4

[Alon; Horak, Luscombe et al (2002), *Genes & Dev*, 16: 3017]

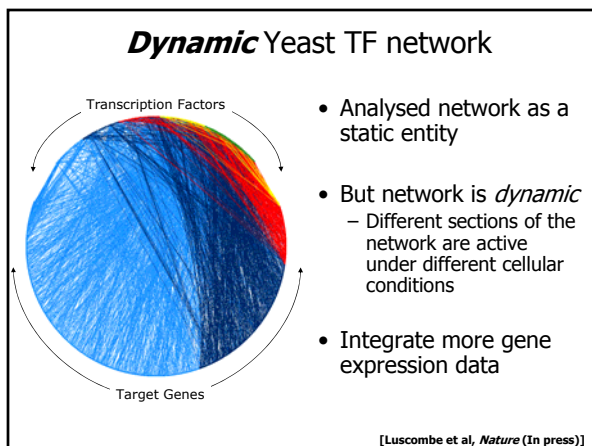
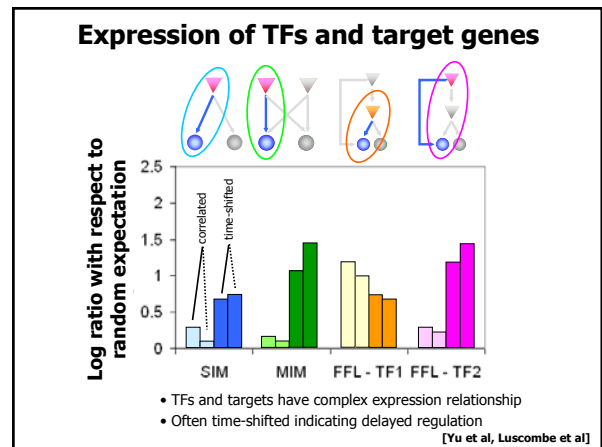
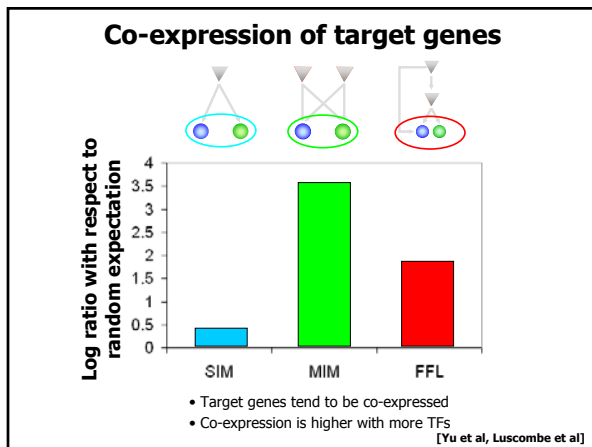








- ### Summary of expression relationships
- Target genes tend to be co-expressed
  - Co-expression is higher when more TFs are involved
  - TFs and target genes have complex relationships
  - Often expression profiles are time-shifted
- [Yu, Luscombe et al (2003), *Trends Genet*, 19: 422]



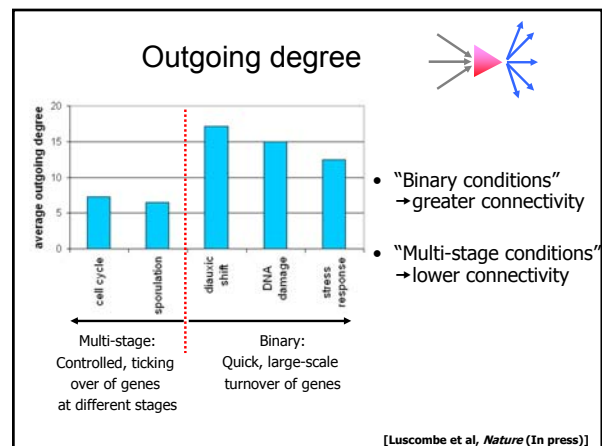
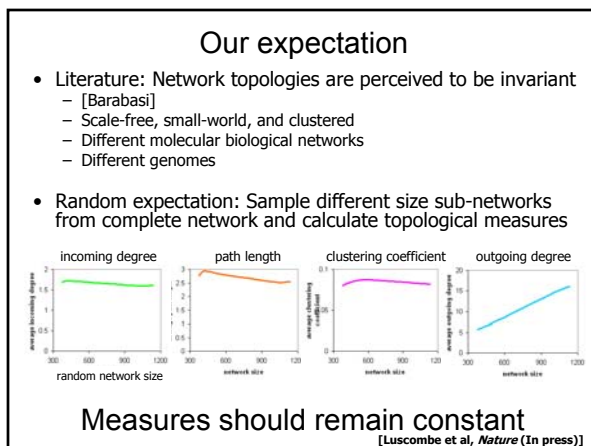
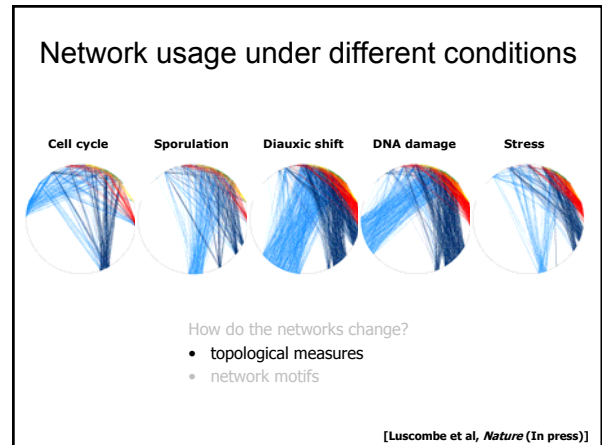
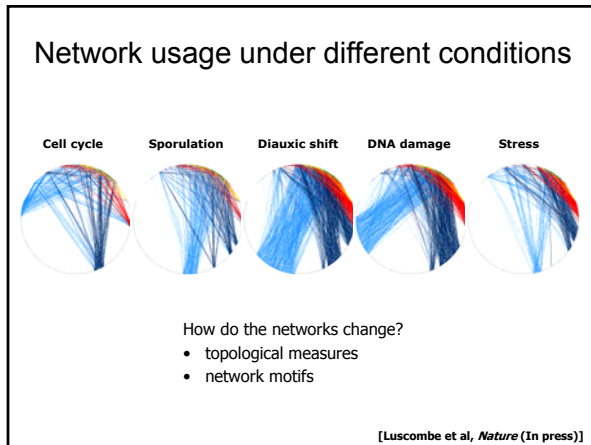
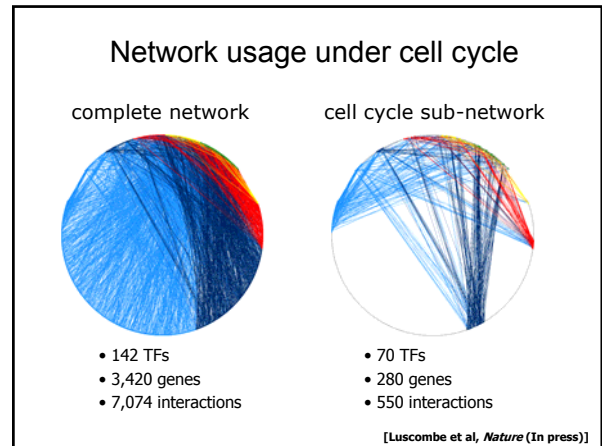
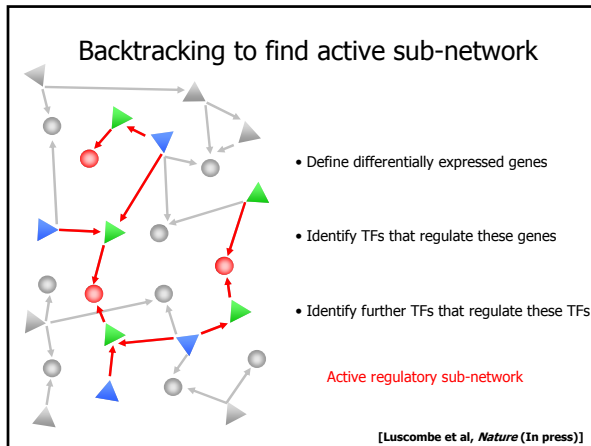
### Gene expression data

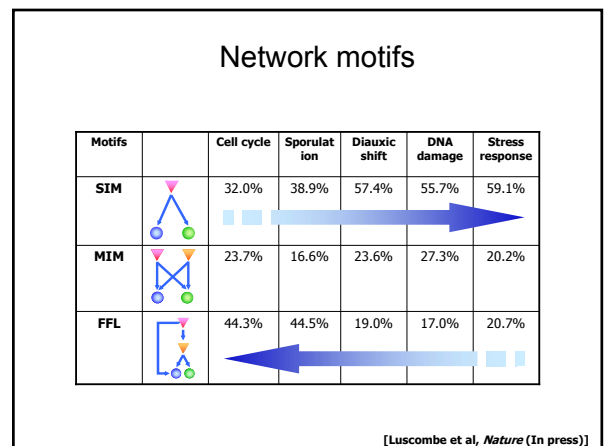
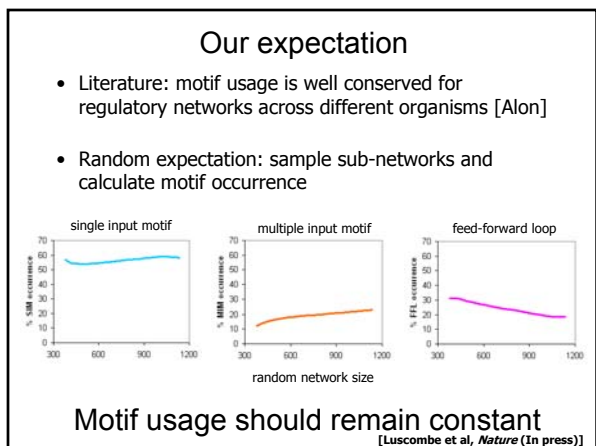
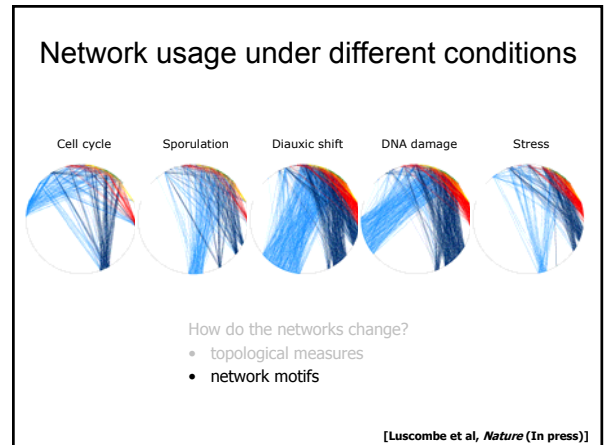
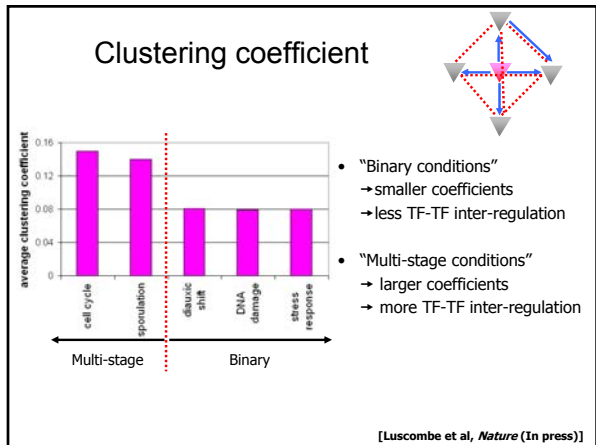
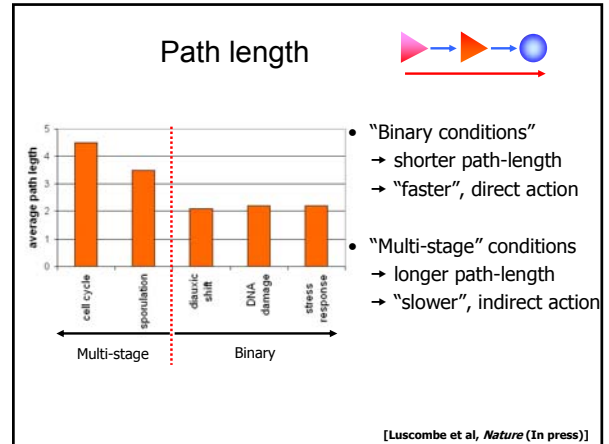
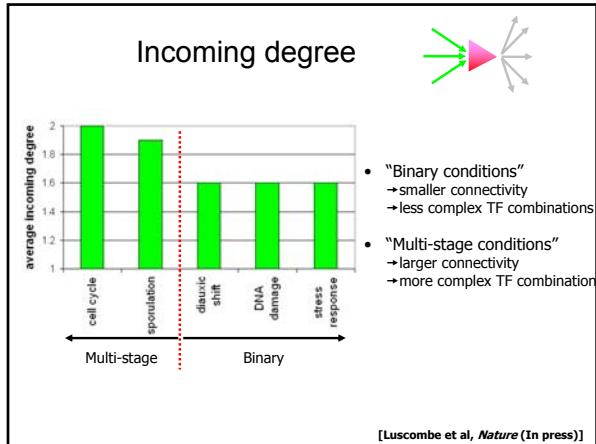
- Genes that are differentially expressed under five cellular conditions

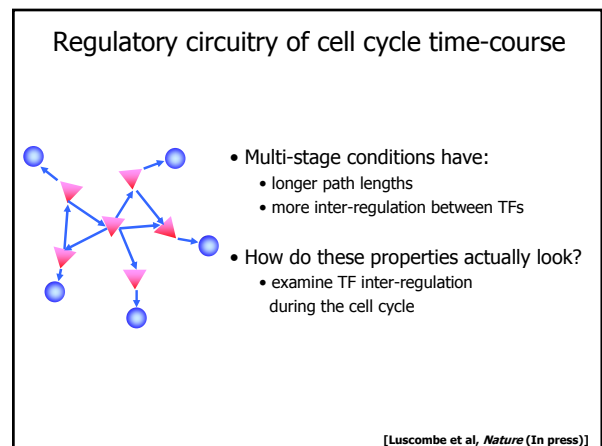
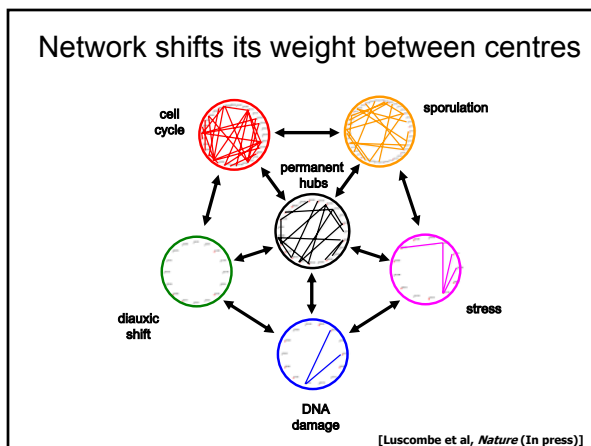
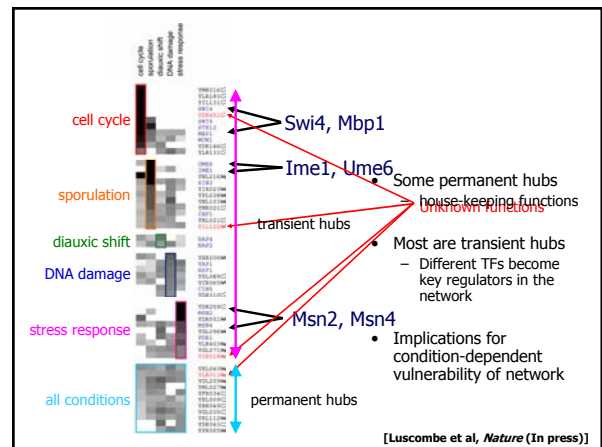
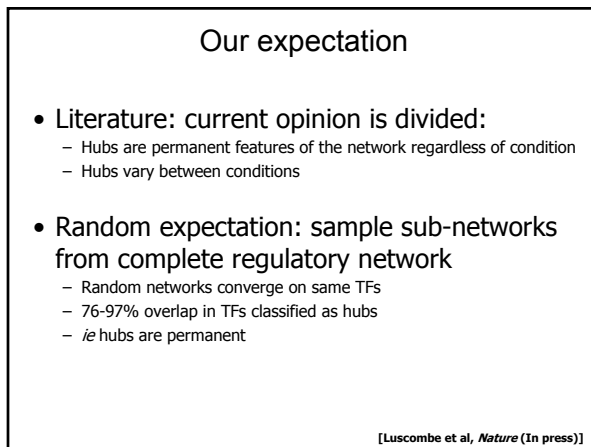
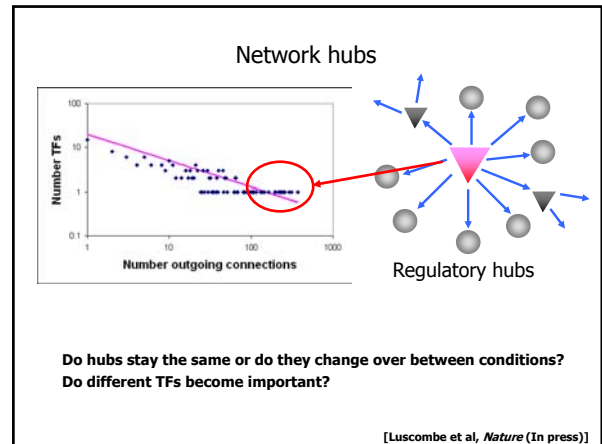
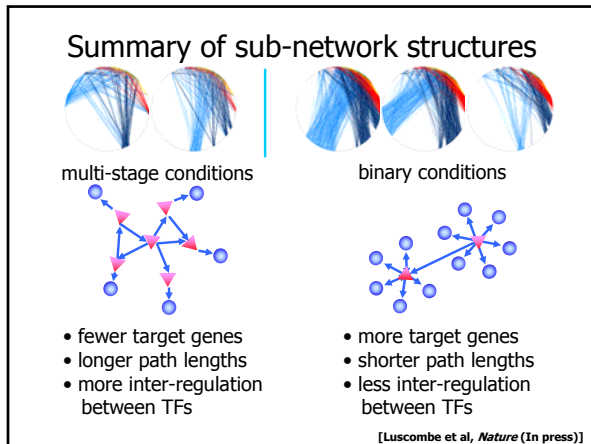
Cellular condition	No. genes
Cell cycle	437
Sporulation	876
Diauxic shift	1,876
DNA damage	1,715
Stress response	1,385

- Assume these genes undergo transcription regulation

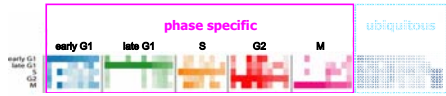
[Luscombe et al, *Nature* (In press)]







## Regulatory circuitry of cell cycle time-course



Phase-specific TFs show **serial** inter-regulation

→ drive cell cycle forward through time-course

[Luscombe et al, *Nature* (In press)]

## Regulatory circuitry of cell cycle time-course



### 2. parallel inter-regulation

Ubiquitous and phase-specific TFs show **parallel** inter-regulation

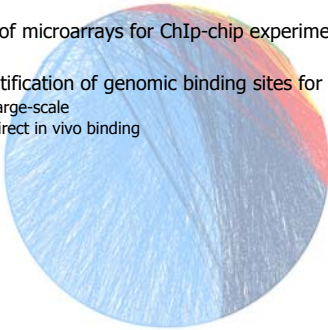
Many ubiquitous TFs are permanent **hubs**

→ Channel of communication between house-keeping functions and cell cycle progression

[Luscombe et al, *Nature* (In press)]

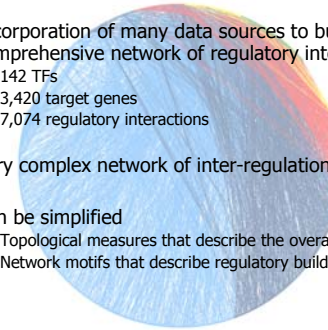
## Summary 1

- Use of microarrays for ChIP-chip experiments
- Identification of genomic binding sites for 11 TFs
  - Large-scale
  - Direct in vivo binding



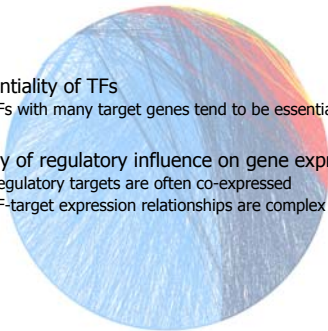
## Summary 2

- Incorporation of many data sources to build a comprehensive network of regulatory interactions
  - 142 TFs
  - 3,420 target genes
  - 7,074 regulatory interactions
- Very complex network of inter-regulation
- Can be simplified
  - Topological measures that describe the overall structure
  - Network motifs that describe regulatory building blocks



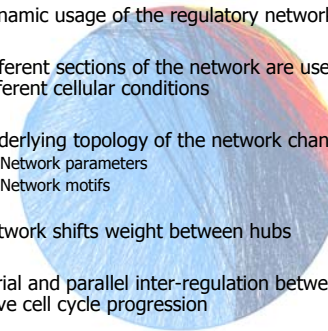
## Summary 3

- Essentiality of TFs
  - TFs with many target genes tend to be essential
- Study of regulatory influence on gene expression
  - Regulatory targets are often co-expressed
  - TF-target expression relationships are complex



## Summary 4

- Dynamic usage of the regulatory network
- Different sections of the network are used in different cellular conditions
- Underlying topology of the network changes
  - Network parameters
  - Network motifs
- Network shifts weight between hubs
- Serial and parallel inter-regulation between TFs to drive cell cycle progression



## Snyder group and collaborators



**Michael Snyder**

Christine Horak

Lara Umansky

Kevin madden

Susana Vidan

Ghia Euskirchen

Rebecca Goetsch

John Rinn

Madhuparna Roychowdry

Gillian Hooker

Mike McEvoy

Stacy Piccirillo

**Sherman Weissman**

Milind Mahajan

**Patrick Brown**

Vishwanath Iyer

**David Botstein**

Charles Scafe

**Kenneth Nelson**

Yale Microarray Facility

## Gerstein group



**Mark Gerstein**

Haiyuan Yu

**Sarah Teichmann**

Madan Babu

