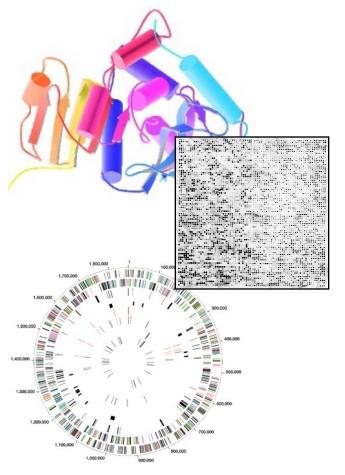
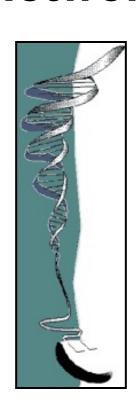
#### **Biomedical Data Science:**

### Analysis of Network Topology D – Central Network Points







Mark Gerstein, Yale University gersteinlab.org/courses/452 (last edit in spring '21)

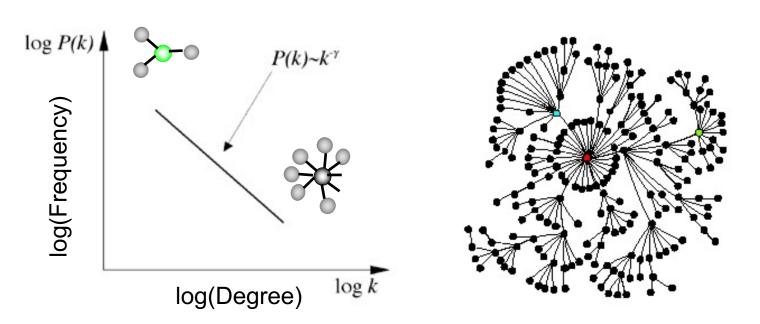
#### **Network Topology**

### Measures of Centrality: Hubs

# (c) M Gerstein, gerstein.info/talks

#### Scale-free networks in Biology

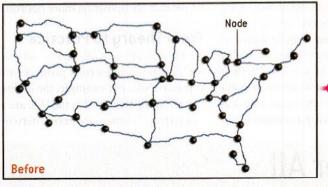
#### Power-law distribution

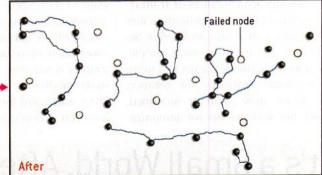


**Hubs** dictate the structure of the network

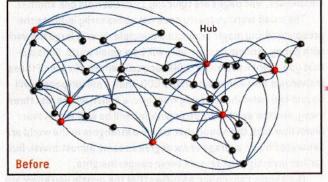
[Barabasi]

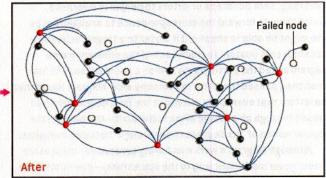
#### Random Network, Accidental Node Failure



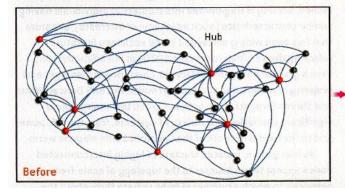


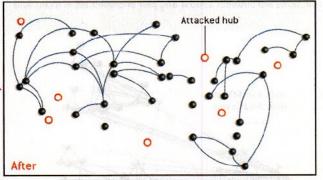
Scale-Free Network, Accidental Node Failure





Scale-Free Network, Attack on Hubs

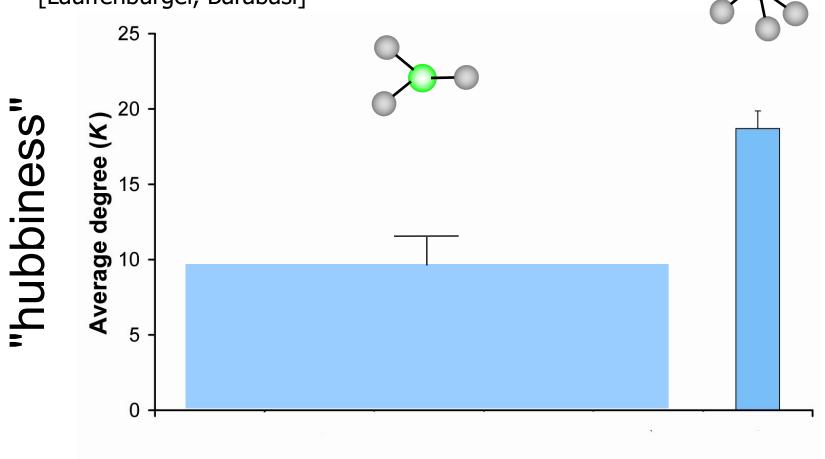




#### Knocking Out Nodes in Scale-free and Random Networks

Integrate gene essentiality data with protein interaction network. Perhaps hubs represent vulnerable points?

[Lauffenburger, Barabasi]



[Yu et al., 2003, TIG]

Non- Essential

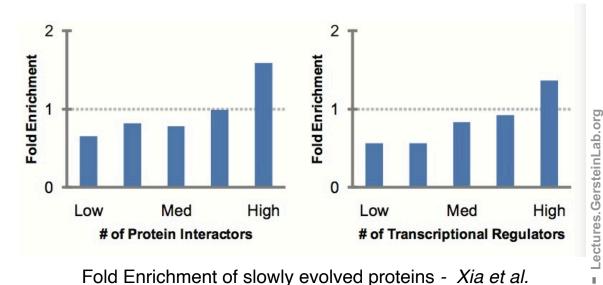
Essential

- Notably, the most highly conserved proteins were highly connected in *E.coli* protein interaction network. (E.coli)
  - Butland et al. 2004, Nature Connectivity of well-conserved proteins in network is negatively correlated with their rate of evolution.
- (Saccharomyces cerevisiae)
  - Fraser et al. 2002, Science - Fraser et al. BMC Evol. Biol.

(2003)

- Proteins that have a more central position evolve more slowly and are more likely to be essential for survival. (Saccharomyces cerevisiae; C. elegans; Drosophila)
  - Hahn et al. 2005, Mol Biol Evol
- More miRNA regulation of a target gene associated with lower dN/dS (r = -.21) (M. musculus) - Cheng et al. 2009, BMC Genomics
- Slowly evolving proteins tend to have more interaction partners. (Saccharomyces cerevisiae) - Xia et al. 2009, Plos Comput Biol

**More Connectivity, More Constraint:** A theme borne out in many studies



Fold Enrichment of slowly evolved proteins - Xia et al.

# Pos. sel. v. deg. centrality: ρ = -.06, P < 1.2e-6</li> Effect is independent of any bias due to gene expression differences Update w. 1000G Phase I SNP dens. v. centrality: ρ = -.1, p< 2.2e-16</li>

Rapid Evolution in the

interaction network takes

place at the periphery

High likelihood of positive selection

Lower likelihood of positive selection

Not under positive

selection

No data about positive selection

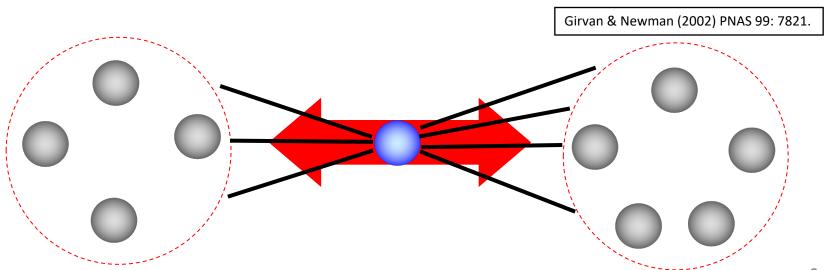
#### **Network Topology**

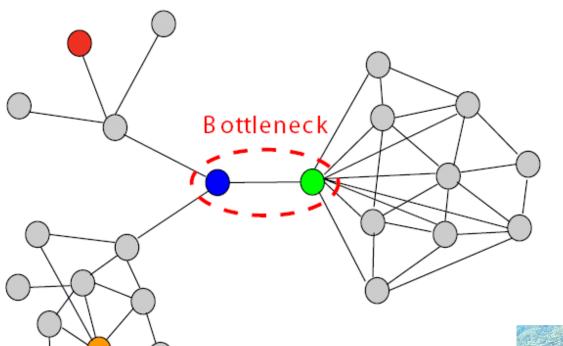
# Other Measures of Centrality besides Hubs: Bottlenecks

### Another measure of Centrality: Betweenness centrality

Betweenness of a node is the number of shortest paths of pairs of vertices that run through it -- a measure of information flow.

Freeman LC (1977) Set of measures of centrality based on betweenness. Sociometry 40: 35–41.



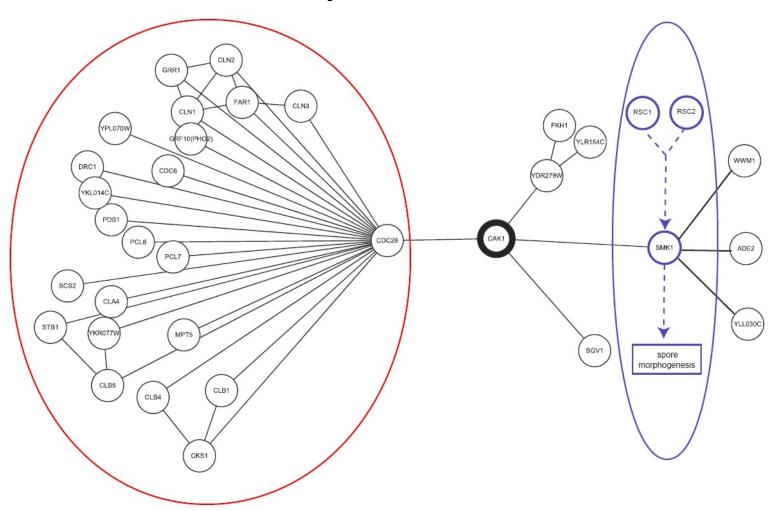


#### **Bottlenecks & Hubs**

- Hub-bottleneck **node**
- Non-hub-bottleneck **node**
- Hub-non-bottleneck **node**
- Non-hub-non-bottleneck node



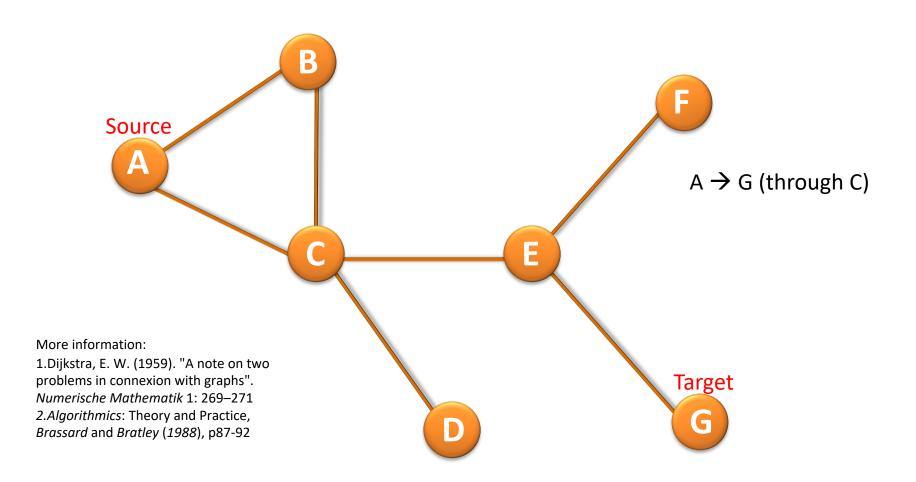
### Bottleneck bridging between processes



Regulation of mitotic cell cycle

MAP Kinase pathway regulating spore morphogenesis

#### Shortest Paths: The Dijkstra Algorithm



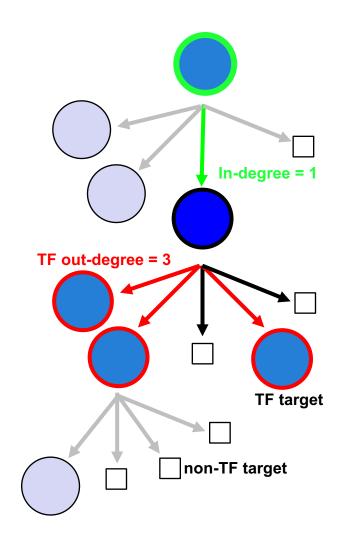
#### **Network Topology**

# Other Measures of Centrality besides Hubs: Hierarchy

#### VICE PRESIDENT MICHAEL BECK BECK JOHN BRANDON FENGER VP, IPHONE SALES JENNIFER VP, CHANNEL **VPAPPLE** JOHN COUCH BAILEY JAFAN SALES VP, EDU CATION SALES **VP, ONLINE** WILLIAM MICHAEL CULBERT FREDERICK RULFILLM ENT ARCHITECTURE RITA LANE STEVE ZADESKY KATIE VP, OPERATIONS VP, IPHONE/IPOD DESIGN JOEL Podolny VP, HR TIMOTHY DAVID TUPMAN SABIH KHAN VP, HAR DWARE COOK CHIEF OFFRATING OFFICER VP, OPERATIONS ENGINEERING. JEFFREY WILLIAMS SYLDFERATIONS IPHONE/IPOO MANSFIELD SVP. MAC MARINGARE BIGINEERINS DEIRDRE OBRIEN DAN RICCIO VP, OPERATIONS VP, IPAD JONATHAN SEWELL SVP (ENERAL COUNSEL JERRY JOHN THERMULT MCODUGAL VP, GLOBAL VP, RETAIL SECURITY. PETER RONALD SHEG CILLEY BETSY RAFAEL OPPENHEIMER SYR CRIEF FINANCIAL OFFICER JOHN SON SVE REIAIL VEVIOED APPS CEO VP. CONTROLLER ROGER ROSNER VP. **GARYWIPFLER** ANDY MILLER VP. Mobile Advertising/Vao PRODUCTIVITY **VP.TREASURER** VP. CREATIVE DIRECTOR APPS SCOTT SCHILLER SVP. MARKETING FORSTALL SVP. NS SUFTMARE HENRI LAMIRALIX **GREG JOSWIAK** VP, ENGINEERING, ICES APPS WP.IPHONE MARKETING ISABELGE MAKE MICHAELTCHAO VEIOS WIRELESS VP, IPAO SOFTWARE MARKETING P MAC SIDETWA ENGINET RING KIM VOR RATH DAVIO MODDY VP. PROGRAM VP MAC MARKETING MANAGEMENT JEFF ROBBIN RON DIXAMOTO VP, CONSUMER VP. DEVELOPER APPS RELATIONS **BRIAN CROLL** MAX PALEY VP.MAC VPAUDIO/VIDEO BUOTRIBBLE SOF WARE PATIENCE VP, SOFTWARE VP, CORE OS TECHNOLOGY

#### Network Hierarchy

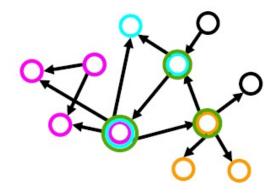
Network
Stats to
Identify
Hierarchy



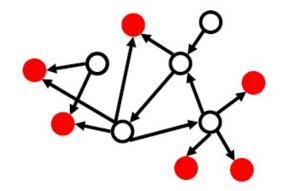
Hierarchy Height Statistic = (normalized TF Out deg. – In deg.)

## Determination of "Level" in Regulatory Network Hierarchy with Breadth-first Search

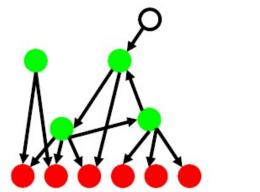
I. Example network with all 4 motifs



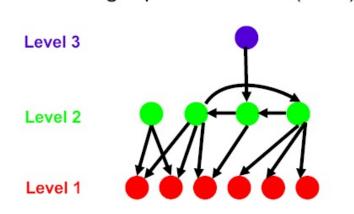
II. Finding terminal nodes (Red)



III. Finding mid-level nodes (Green)



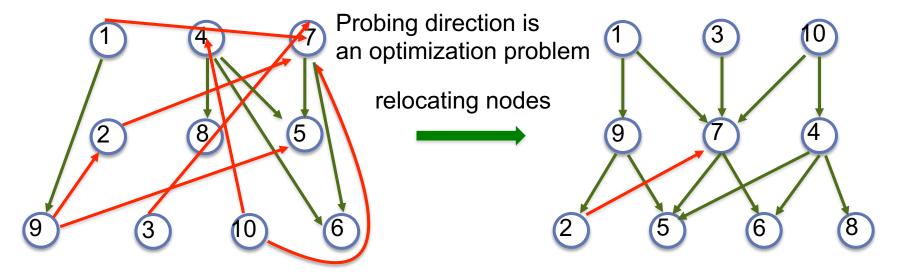
IV. Finding top-most nodes (Blue)



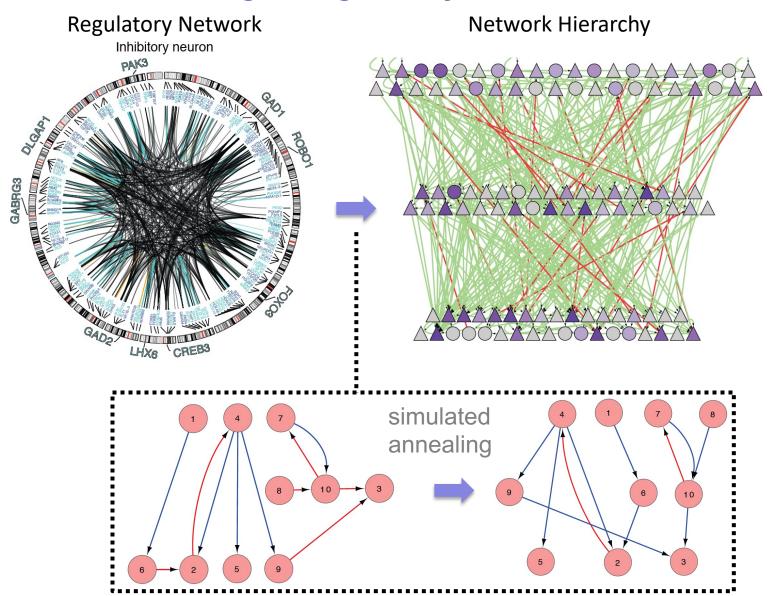
Level 1

16 - Lectures.GersteinLab.org

# Using Simulated Annealing to Globally Minimize the Number of Upward Pointing Edges



#### **Transforming a Regulatory Network into a Hierachy**



### Mid-level of the hierarchy has many high-connectivity bottlenecks

