Genomics Education Partnership
thegep.org

P A T H W A Y S

Pathways Project

Analyzing the evolution of metabolic and signaling pathway genes

Katie M. Sandlin
Last Update: 01/19/2025

1

Overview

- The Pathways Project uses network analysis approaches to better understand the **evolution and function of biological pathways**.
- The current focus is on annotating genes within the insulin signaling pathway across the *Drosophila* genus.

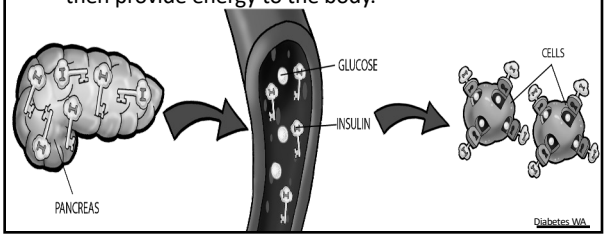
2

Insulin

3

Insulin's Function

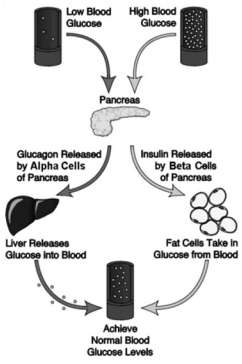
- Insulin works like a key, unlocking the doors of the cells in your body to let glucose in.
- Once insulin opens the cells' doors, glucose can move from the blood into the cells, where it can then provide energy to the body.



Diabetes WA

4

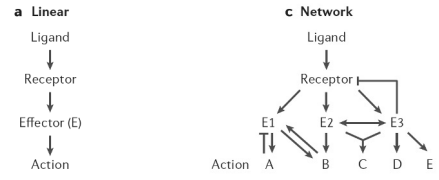
Low vs. High Blood Glucose



endocrineweb.com

5

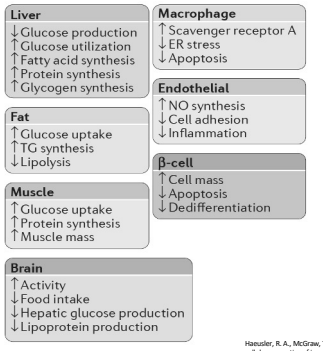
What are Signaling Pathways?



Taniguchi, C. M., Emanuelli, B., & Kahn, C. R. (2006). Critical nodes in signalling pathways: insights into insulin action. *Nature reviews. Molecular cell biology*, *7*(2), 85–96.

6

What are the Effects of Insulin Signaling?



Haeusler, R. A., McGraw, T. E., & Accili, D. (2018). Biochemical and cellular properties of insulin receptor signalling. *Nature reviews. Molecular cell biology*, *19*(1), 31–44.

7

Why analyze evolution of insulin signaling pathway?

"We cannot understand properly a function if not trying to **understand history of its origin.**"

~ L.A. Orbeli

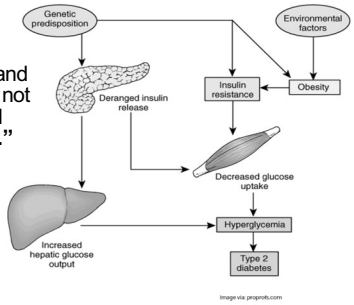


Image via proprofs.com

8

Network Architecture

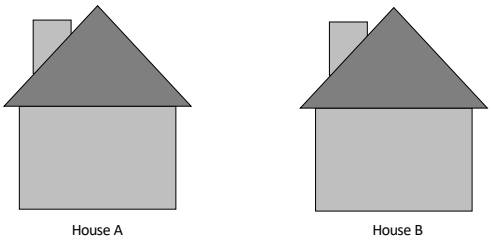
9

Jog Your Memory

- What is a phenotype? Give some examples.

10

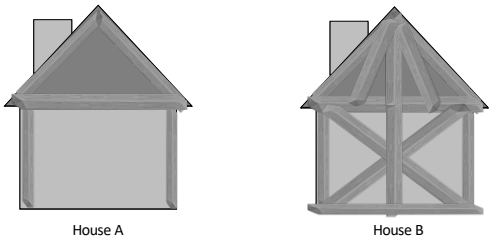
What's different about the phenotype of House A and House B?



House A House B

11

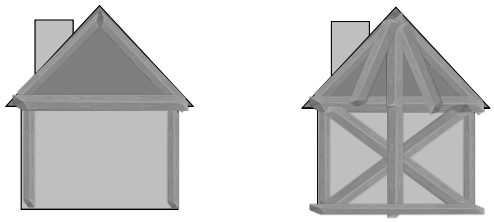
What's different about the architecture of House A and House B?



House A House B

12

How might their architectural differences impact each house's ability to evolve in order to adapt to changing conditions (i.e., which house would be easier to change)?

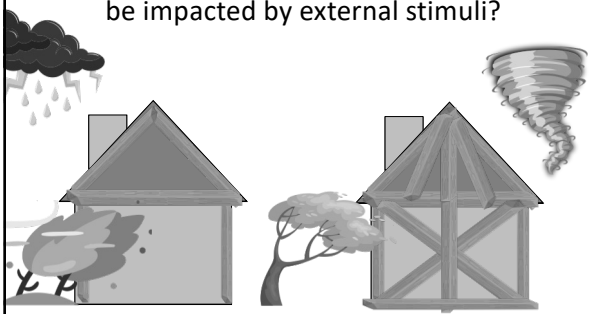


House A
highly evolvable

House B
resistant to evolution

13

How might their architectural differences be impacted by external stimuli?



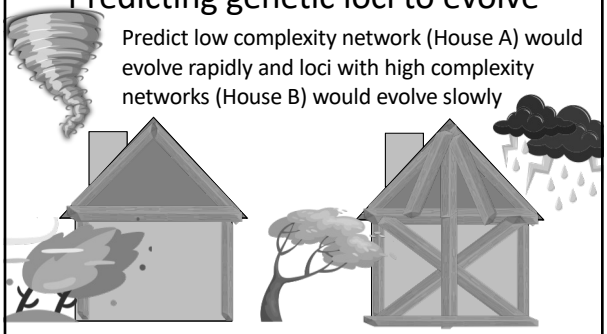
House A
easily perturbed

House B
robust

14

Predicting genetic loci to evolve

Predict low complexity network (House A) would evolve rapidly and loci with high complexity networks (House B) would evolve slowly



House A
highly evolvable
easily perturbed

House B
resistant to evolution
robust

15

Southwest Airlines canceled 62% of its flights during the 2022 Holiday Travel Season

The Denver Gazette

Millions of Southwest passengers experience historic holiday travel debacle

The New York Times

The Southwest Airlines Meltdown

In the last 10 days of 2022, the airline canceled as many flights as it had in the previous 10 months. What went wrong?

REUTERS®

U.S. senators press Southwest to answer questions on holiday meltdown

16

Think, Pair, then Share

- Why do you think Southwest Airlines had to cancel so many of its flights when other large airlines such as United, Delta, and American didn't?
- Take 30 seconds to think about your answer to the above, and then share it with your neighbor.

17

Southwest Airlines

- Southwest Airlines uses a point-to-point route model which allows passengers in smaller cities (e.g., Birmingham, AL) to fly directly to their destination.
- Delta Airlines uses a hub-and-spoke route model in which passengers from smaller cities must travel to a central hub airport (e.g., Atlanta, GA) to change planes before flying to their destination.

18

Think, Pair, then Share

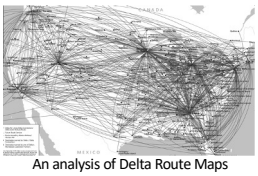
- Which of these models do you think is more robust (i.e., could withstand perturbations* like extreme weather conditions in part of the country)?

*a deviation of a system, moving object, or process from its regular or normal state or path, caused by an outside influence

19

What defines the architecture of a network?

- Nodes
  - Airports
- Edges
  - Flight paths



An analysis of Delta Route Maps

- Architecture can be changed by
  - Adding/subtracting nodes
    - New airports
  - Adding/subtracting edges
    - New flight path between two airports

Image: https://blogs.cornell.edu/info2040/2015/09/11/an-analysis-of-delta-route-maps/

20

In the context of biological networks…

- Nodes
  - Genes/proteins
- Edges
  - Interaction between two genes/proteins

An analysis of Delta Route Maps

- Architecture can be changed by
  - Adding/subtracting nodes
    - Gene duplicates
  - Adding/subtracting edges
    - New transcription factor binding site in a gene's promoter

Image: https://blogs.cornell.edu/info2040/2015/09/11/an-analysis-of-delta-route-maps/

21

Review

- In our airline example, what were Nodes vs. Edges?
- What biological examples of Nodes vs. Edges did we use?

22

Hub-and-Spoke Model

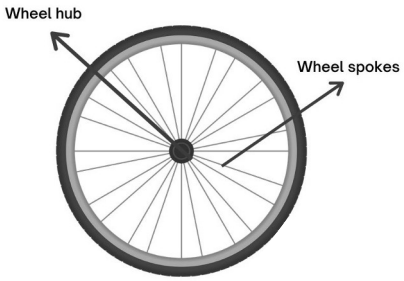
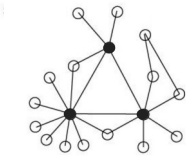


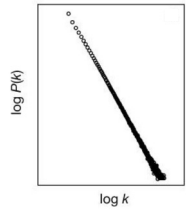
Image: https://nowboarding.changiairport.com/discover-changi/changi-airport-cafe-hub---spoke.html/

23

Hub-and-Spoke Network Architecture



In the hub-and-spoke network, most nodes have only a few links, but a few nodes, called **hubs** (red), have a very large number of links.

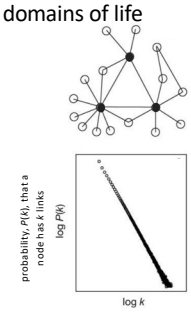


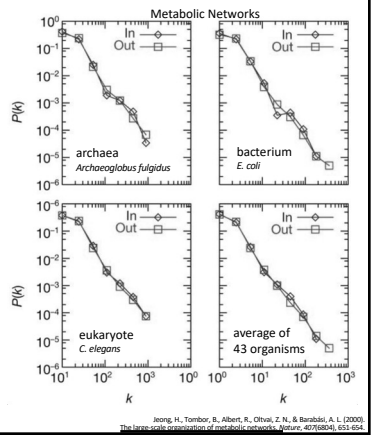
The network connectivity can be characterized by the probability, $P(k)$, that a node has k links.

Jeong, H., Tombor, B., Albert, R., Oltvai, Z. N., & Barabási, A. L. (2000). The large-scale organization of metabolic networks. *Nature*, 407(6804), 651-654.

24

Most biological networks are hub-and-spoke across the domains of life





Jeong, H., Tombor, B., Albert, R., Oltvai, Z. N., & Barabási, A. L. (2000). The large-scale organization of metabolic networks. *Nature*, 407(6804), 651-654.

25

Guiding Questions of the Pathways Project

- How does network structure determine robustness of function in response to perturbation (within-species timescales)?
- How does network structure influence and reflect the process of evolution (between-species timescales)?

26

Rates of Molecular Evolution

27

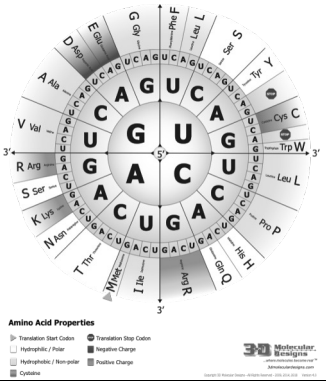
Measuring Rate of Molecular Evolution

- Number of differences in nucleotide sequence between two species
  - More differences → more/faster evolution
- How do we distinguish between neutral and adaptive evolution?
  - **Substitution Rates:**
    - If difference in nucleotide sequence **leads to a change** in the amino acid, what type of substitution is this?
      - ❑non-synonymous ("non-synonym")
    - If difference in nucleotide sequence **doesn't lead to a change** in the amino acid, what type of substitution is this?
      - ❑synonymous ("synonym")

28

Check For Understanding

- What **type of substitution** would this be? Would it change the chemical properties of the amino acid? If so, how?
  1. GCG → GCA
  2. UUG → UCG
  3. GGA → GUA



29

Measuring Rate of Molecular Evolution

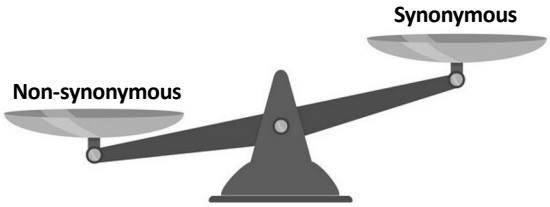
- Substitution Rates:
  - Non-synonymous substitution rate (d_N)
    - difference in amino acid → assumed to be non-neutral
  - Synonymous substitution rate (d_S)
    - no change in amino acid → assumed to be neutral
- d_N/d_S = ratio of non-synonymous to synonymous nucleotide substitutions (also called Ka/Ks; ω)

30

$d_N/d_S = 1$
gene evolving neutrally

31

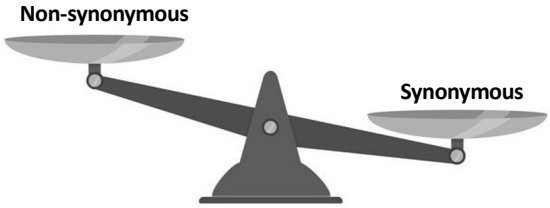
$d_N/d_S < 1$
gene experiencing selective constraint



the proportion of new mutations that are strongly deleterious are removed by purifying selection

32

$d_N/d_S > 1$
gene experiencing adaptive evolution



evolutionary changes that are adaptive to a particular environment

33

- d_N/d_S = ratio of non-synonymous to synonymous nucleotide substitutions (also called Ka/Ks; ω)
  - $d_N/d_S = 1$ - gene evolving neutrally
  - $d_N/d_S < 1$ - gene experiencing selective constraint
  - $d_N/d_S > 1$ - gene experiencing adaptive evolution

34

More complex measures consider other factors

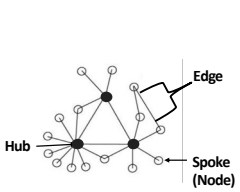
- Sliding windows
- Codon Bias: preferential use of certain codons for particular amino acids
- McDonald-Kreitman (MK) test
  - excess rare polymorphism → positive selection
- compare ratio of within species polymorphic (nonsynon:synon) to between species differences (nonsynon:synon)
- Tajima's D
- Using linkage disequilibrium to identify selective sweeps

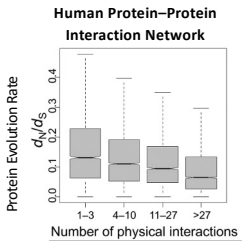
35

Impact of network architecture on evolution of protein-coding genes

36

Rate of protein evolution correlates with the number of interactions with other proteins (proteins that are spokes evolve faster)



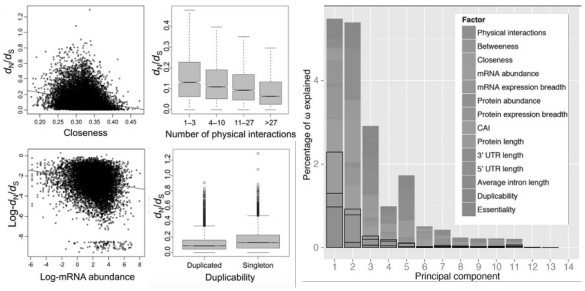


Jeong, H., Tombor, B., Albert, R., Oltvai, Z. N., & Barabási, A. L. (2000). The large-scale organization of metabolic networks. *Nature*, 407(6804), 651-654.

Alvarez-Ponce, D., Feyertag, F., & Chakraborty, S. (2017). Position Matters: Network Centrality Considerably Impacts Rates of Protein Evolution in the Human Protein-Protein Interaction Network. *Genome biology and evolution*, 9(6), 1742–1756.

37

Rates of Protein Evolution Partially Correlated with Many Factors



- Relationship between rates of protein evolution and several factors in the human protein–protein interaction network
- Same patterns hold for the human signal transduction network, flies, worms, and yeast!

Alvarez-Ponce, D., Feyertag, F., & Chakraborty, S. (2017). Position Matters: Network Centrality Considerably Impacts Rates of Protein Evolution in the Human Protein-Protein Interaction Network. *Genome biology and evolution*, 9(6), 1742–1756.

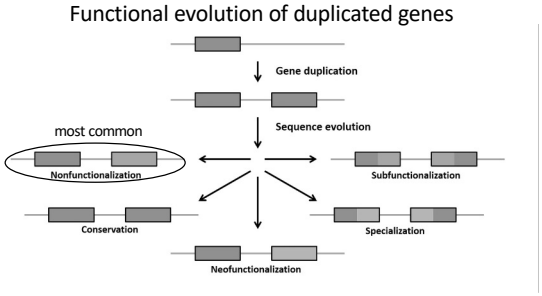
38

Gene Duplication

39

Linking Genetic Evolution to Physiological Networks

Functional evolution of duplicated genes



Lynch, M., & Conery, J. S. (2000). The evolutionary fate and consequences of duplicate genes. *Science (New York, N.Y.)*, 290(5494), 1151–1155.

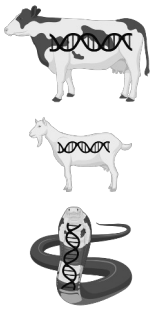
40

Regulatory Region Evolution

41

How to build a beast?

- For most of the 20th Century…
  - "Cows had cow molecules and goats had goat molecules and snakes had snake molecules, and it was because they were made of cow molecules that a cow was a cow." — Francois Jacob
- But, most genes are **highly conserved across species**…
- Most variation in phenotype is caused by changes in the **timing and location of gene expression**
- So, how does regulation of gene expression evolve?

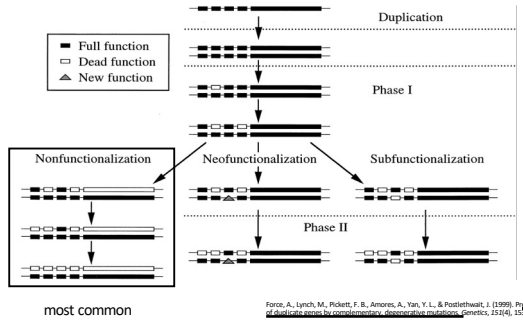


Katie Created with BioRender.com

42

Evolution of Cis-Regulatory Elements

duplication-degeneration-complementation (DDC) model



most common

Force, A., Lynch, M., Pickett, F. B., Amores, A., Yan, Y. L., & Postlethwait, J. (1999). Preservation of duplicate genes by complementary, degenerative mutations. *Genetics*, *151*(4), 1531–1545.

Lynch, M., & Conery, J. S. (2000). The evolutionary fate and consequences of duplicate genes. *Science (New York, N.Y.)*, *290*(5494), 1151–1155.

43

Importance of synteny in assigning orthology

44

Homologs, Orthologs, and Paralogs

- Homology: sequences or structures which share a common ancestor
- Two segments of DNA can have shared ancestry (be homologous) because of either a speciation event (orthologs) or a duplication event (paralogs).

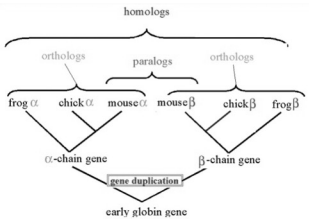


Image: BLAST Glossary: orthologs https://www.ncbi.nlm.nih.gov/books/NBK62057/

45

Defining Synteny

- Classic Definition ("Global" synteny): two "genetic entities" (e.g., genes) on the same chromosome
- "Local" Synteny: conservation of the order and orientation of genes in the vicinity of a target gene - This is what the GEP uses!

Myers PZ. (2008). Synteny: Inferring ancestral genomes. Nature Education 1(1):47.

46

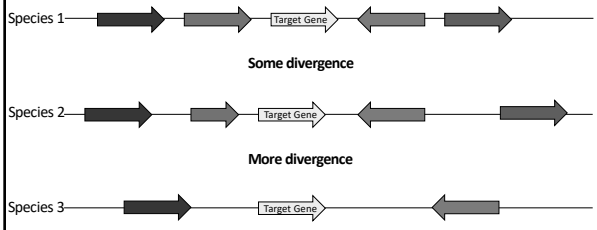
Defining Synteny cont.

- As organisms evolve, changes occur in their genomes (deletions, insertions, etc.)
- Over time, these changes can disrupt synteny
  - Two genetic entities may no longer be on the same chromosome
  - The order and orientation of genes relative to a target gene is no longer conserved
- Generally, we hypothesize that no changes have taken place, because this is the simplest explanation
  - Evolutionarily, this concept is known as "parsimony"
  - More commonly, this is known as Occam's Razor: "the simplest explanation should be preferred over more complex theories"

47

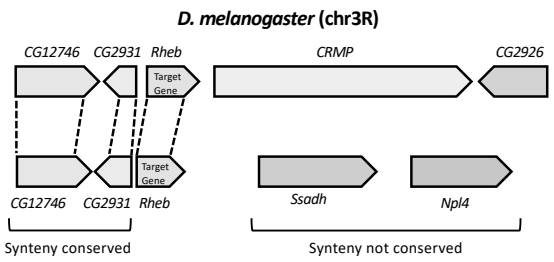
Identifying the Ortholog via Synteny

- Synteny: the order and orientation of genes tends to be conserved



48

What does synteny look like in the GEP?



49