

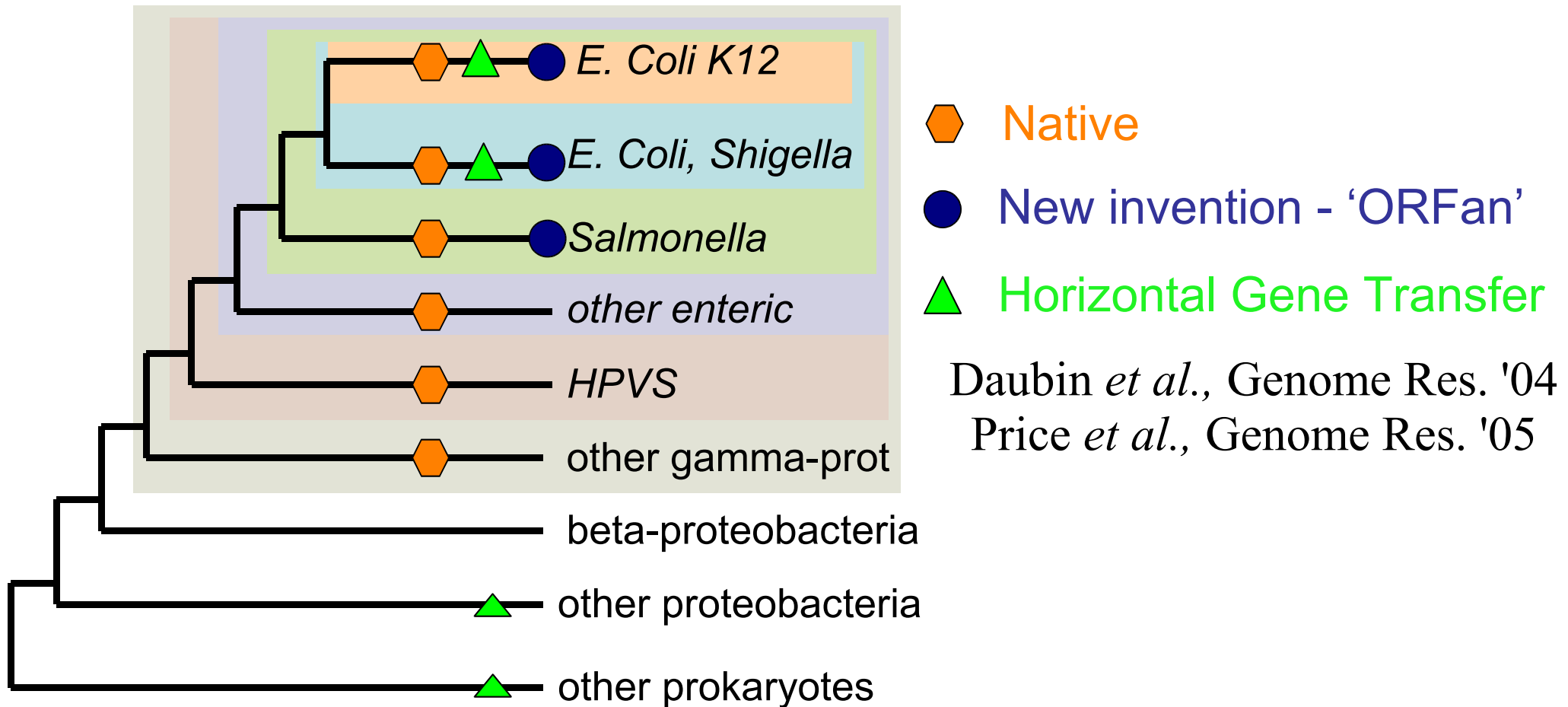
# The Regulation of Horizontally Transferred Genes

Morgan Price  
Arkin group  
February 2006

# Questions

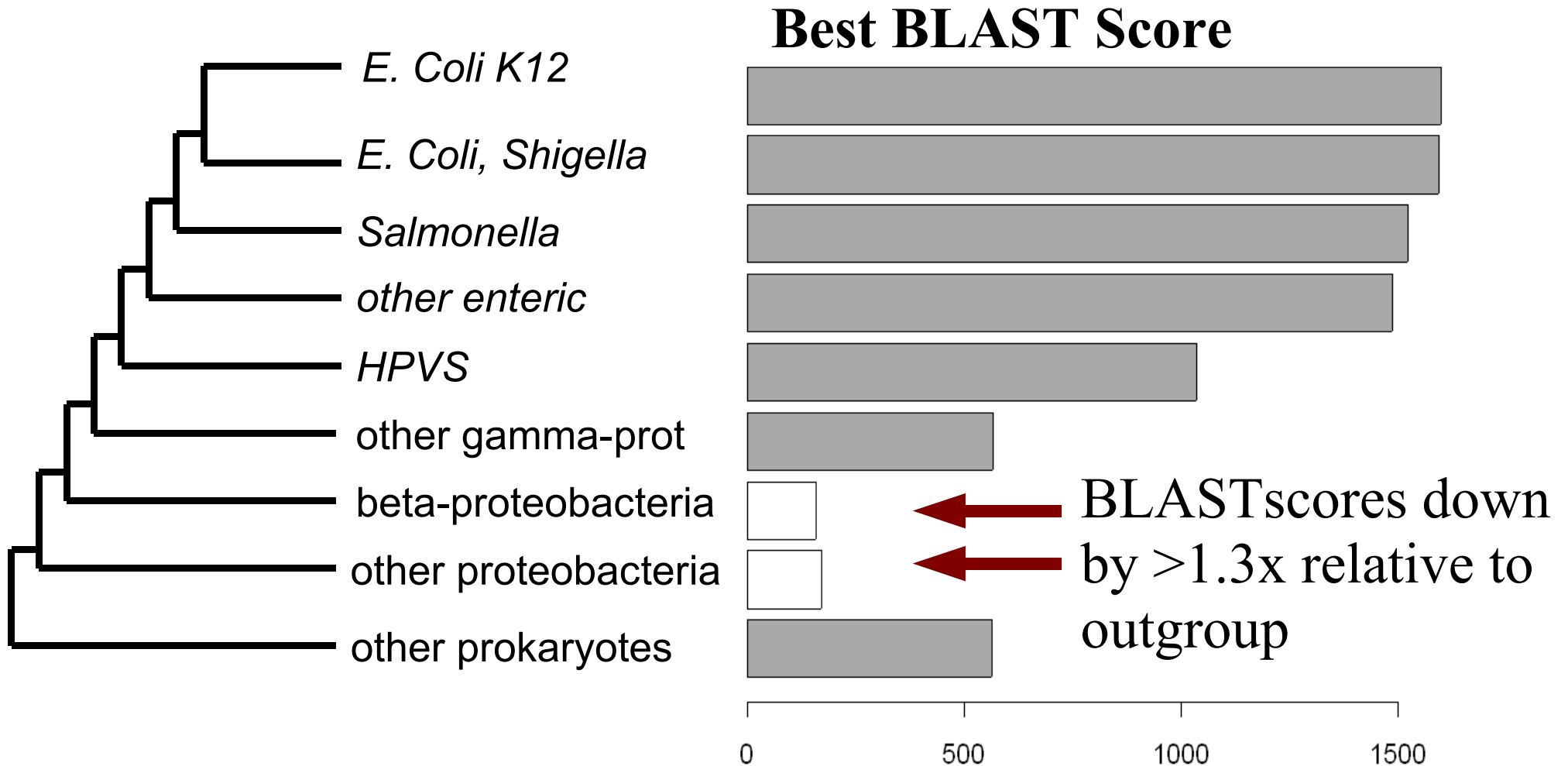
- Bacteria evolve by acquiring (& losing) genes
  - e.g. new metabolic capabilities
- How do these genes get regulated?
  - Function immediately in host?
  - Evolve new regulation?
    - How do new regulators arise?

# Identifying HGT Genes in *E. coli* K12



- Examined presence & absence of gene families (COGs)
  - HGT if in outgroup after 2 consecutive absences
  - found 368 HGT genes from 4,300 ORFs

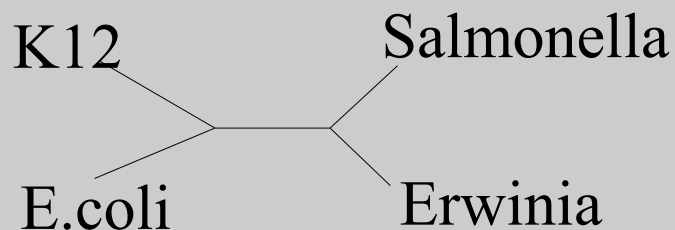
# Identifying *More* HGT Genes



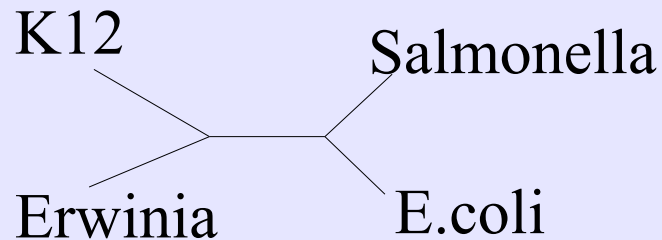
- Find 1,229 HGT genes, but are they genuine?

# Rejecting Alternate Explanations

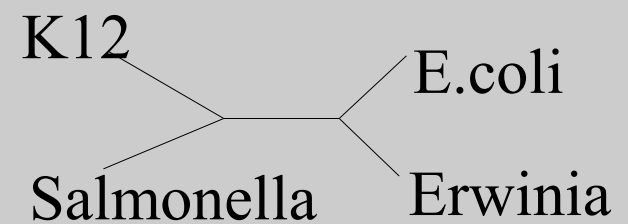
- Loss in multiple groups
  - but usually  $>2$  losses if examine species tree
- Rapid evolution in intervening taxa
  - need to test phylogenetic relationships for these genes
  - e.g. if transferred from Erwinia to E.coli K12:



**6% Native**  
0.1% confident



**92% HGT**  
**71% confident\***



**3% Weird**  
0.1% confident

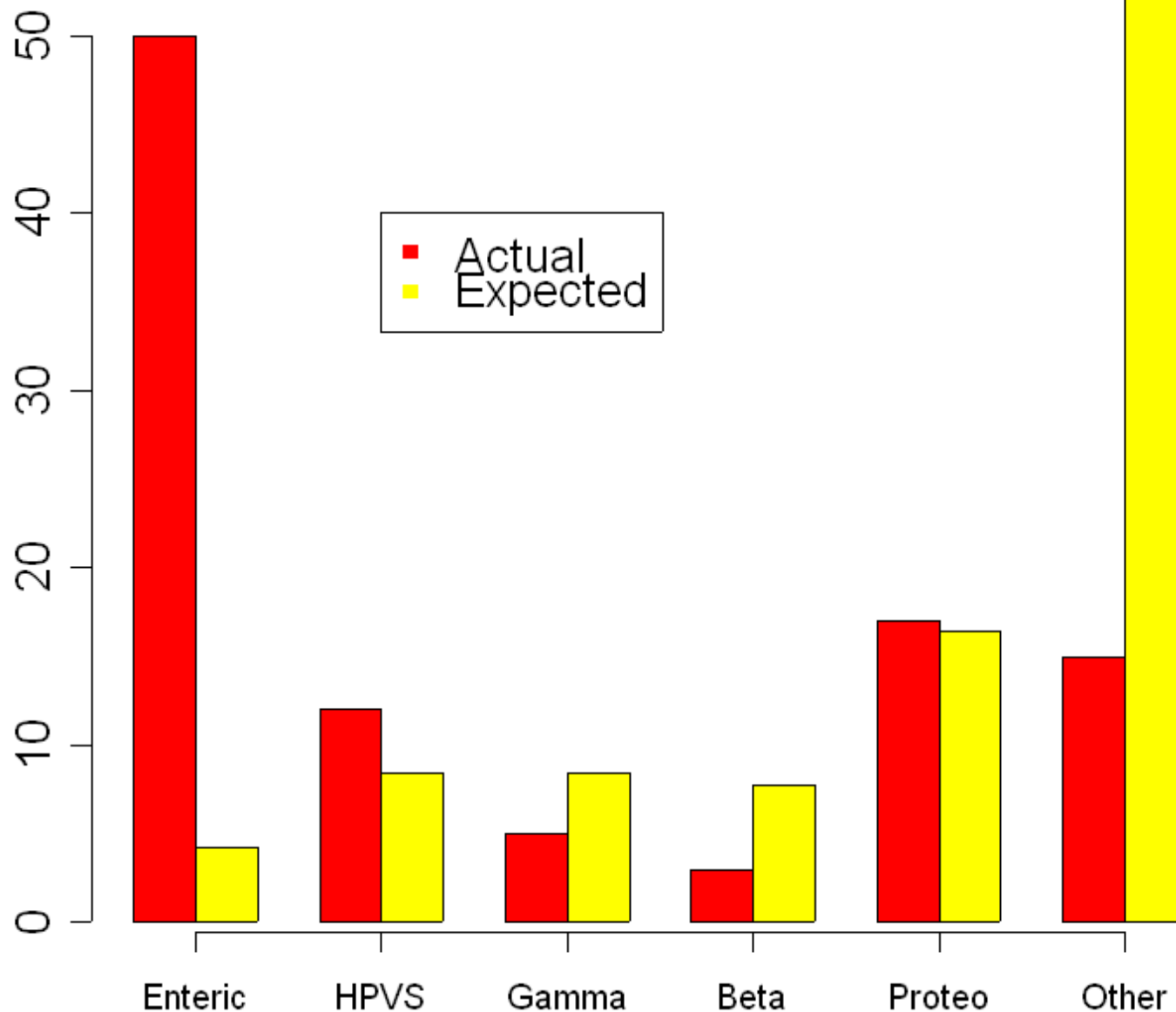
\*  $P < 0.05$ , Shimodaira-Hasegawa test of multiple trees

# The Source of HGT Genes

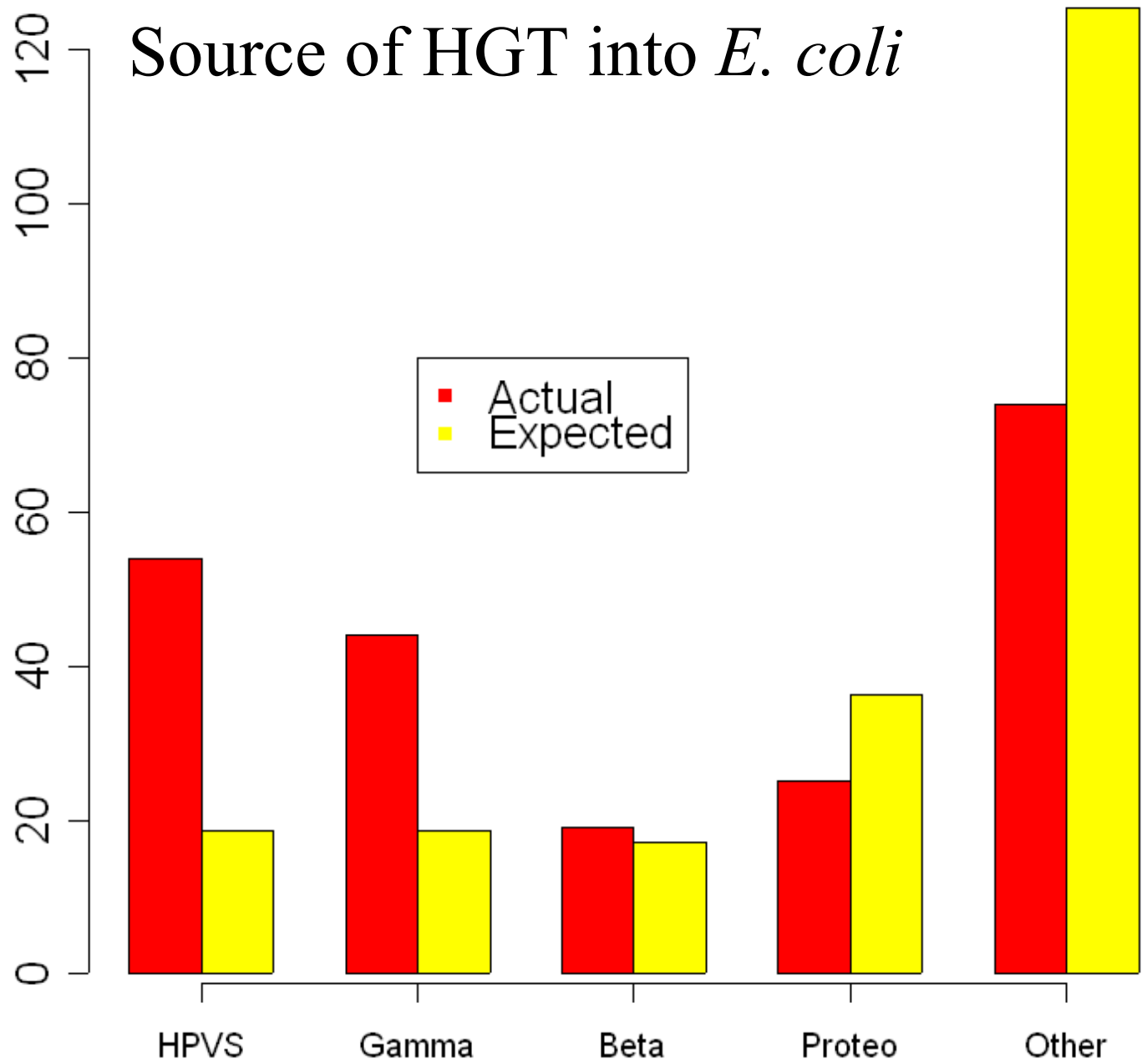
- Directionality problem
  - also, often cannot root trees
- Usually see homologs in a wide array of unrelated bacteria, so assume (mostly) importation
- Consider the best hit in the outgroups as the most likely source
- See *Firmicutes* as a major source, and also:

# HGT Between Related Genomes

Source of HGT into *E. coli* K12



# HGT Between Related Genomes





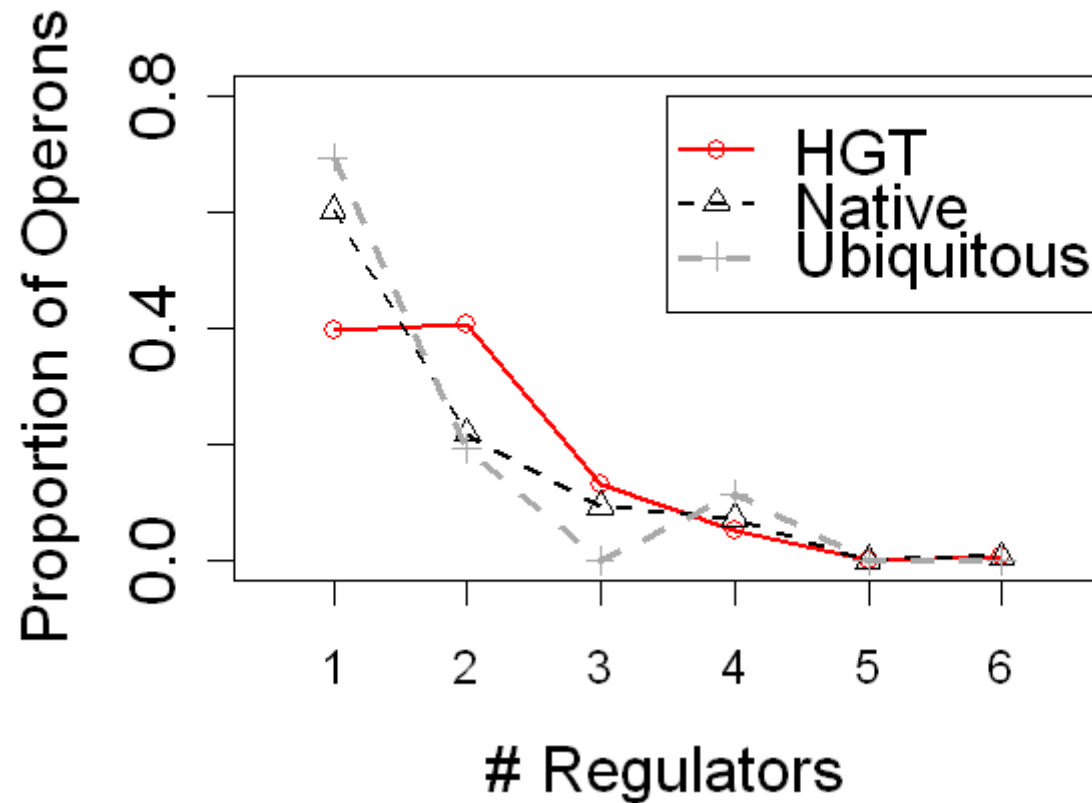
# HGT Between Related Genomes

- Could be loss in two lineages?
  - But most examples require multiple losses
  - Homologs in distant species show that these genes are prone to HGT
- Lerat et al (PLoS Bio 2005) reported this is rare
  - used fewer genomes?
- Genes in related organisms may be more compatible with the new host
  - regulatory environment
  - adjoining metabolic pathways? (Pal *et al.* 2005)

# How Are HGT Genes Regulated?

- Examine ColiNet (Shen-Orr *et al.* 2002)
  - transcription factors and operons, from the literature
- Complex regulation
- Neighbor regulation

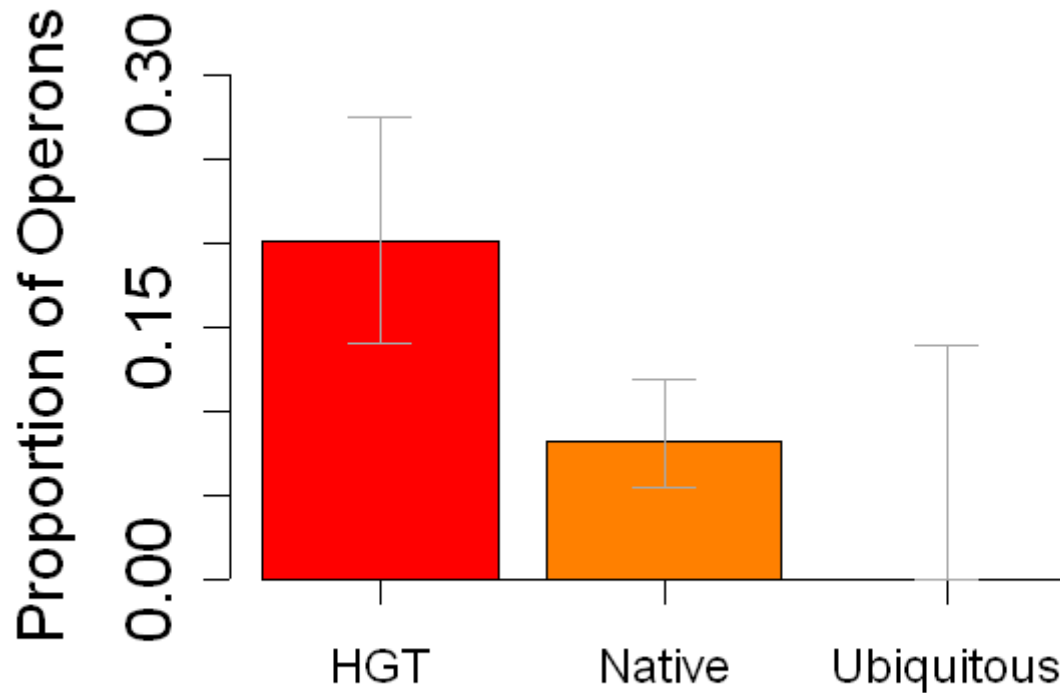
# Complex Regulation of HGT Genes



HGT > Native  
 $p < 0.0002$   
Wilcoxon test

- HGT operons often regulated by CRP & another TF
  - Fits high TF content of large bacterial genomes?
- Doesn't consider other regulation
  - stringent response, attenuators, translational control

# Neighbor Regulation of HGT Genes



HGT > Native  
P < 0.005  
Fisher exact test

# Where Does Neighbor Regulation Come From?

- Transfer together

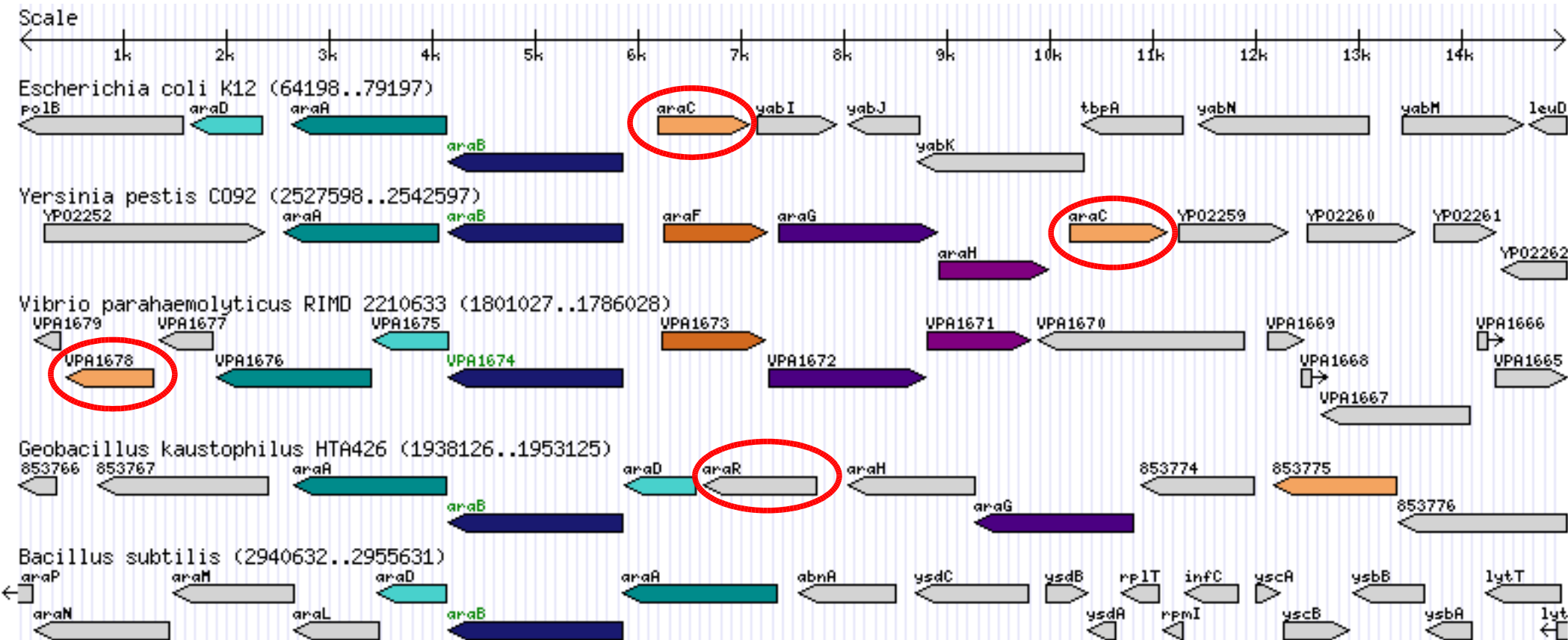
cynR:cynTSX fecIR:fecABCDE\* feaR:feaB\*  
ebgR:ebgAC\* araC:araBAD\*

- Assemble from pieces that are HGT from disparate organisms

atoSC:atoDAEB hacR:hcaA<sub>1</sub>A<sub>2</sub>CBD caiF:caiTABCDE  
celD:celABCDF leuO:leuLABC

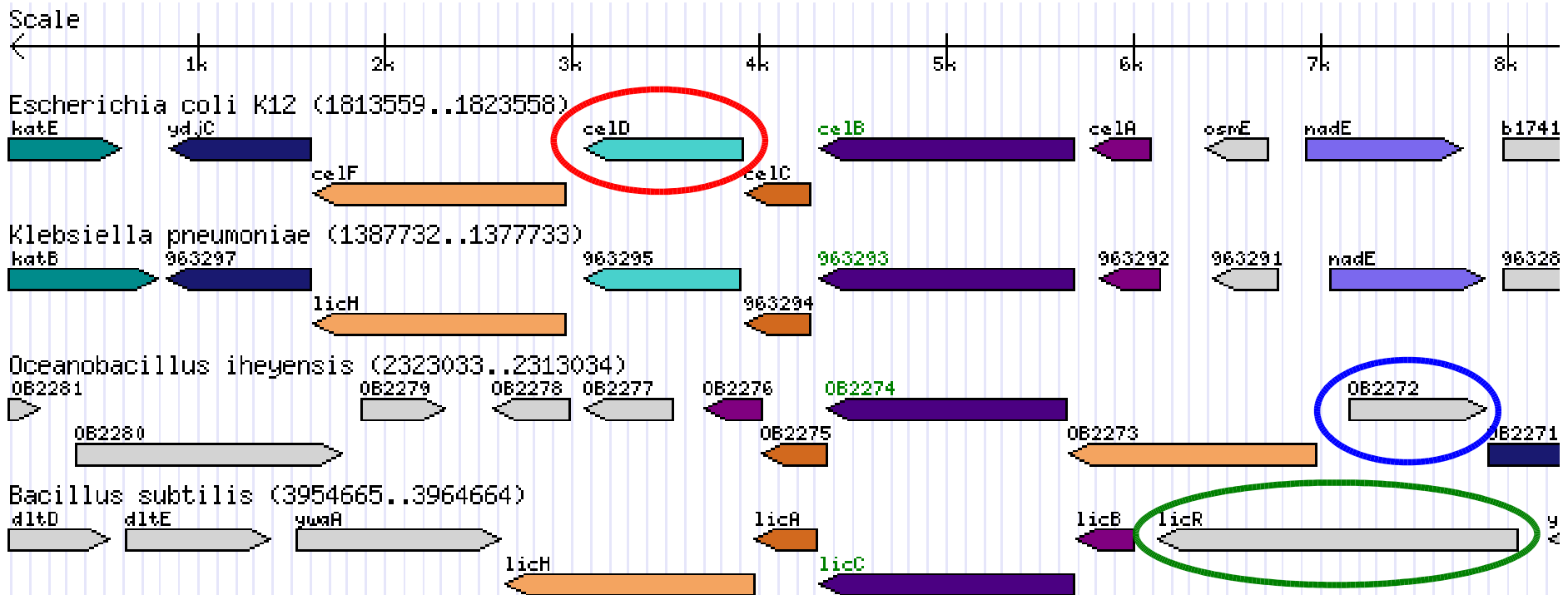
\* Transferred between related bacteria

# Arabinose Utilization



Transfer across phyla involves change in regulator?

# Cellobiose/ $\beta$ -glucoside Utilization



Transfer across phyla involves change in regulator?

# Why Neighbor Regulation?

- Seen in bacteria; not in yeast (Hershberg *et al.* 2005)
- TF localization?
  - only 30 seconds to search
  - rare for “important” genes, global regulators
  - TFs often regulate only the neighboring operon

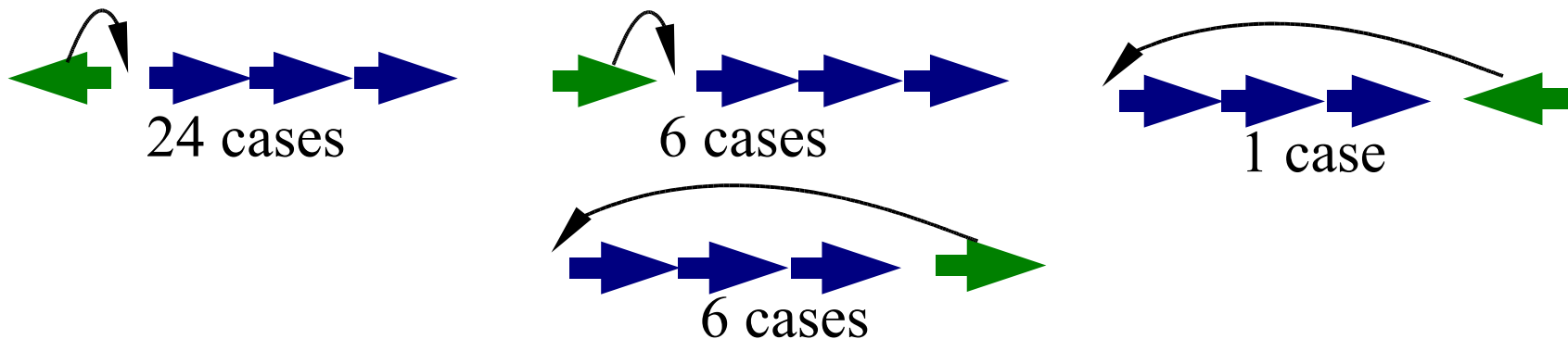


# “Selfish Operons” Revisited

- Repeated HGT selects for proximity of genes?
  - originally proposed to explain operons
    - but native genes often form new operons w/o HGT
    - non-HGT genes tend to be in operons
  - could explain neighbor regulation
    - rare for non-HGT genes
- But “assembly” involves genes from different taxa
  - 618 adjacent HGT pairs; 326 acquired together
  - 12% (73/618) have good distant hits in the same genome
    - random pairs of HGT genes also give 12%

# Why Neighbor Regulation?

- Plasmids & conjugative transposons
  - new catabolic capabilities, antibiotic resistance
- Divergent shared site facilitates selection for a new regulator



# Horizontal Transfer Allows the Prediction of Neighbor Regulation

- Motif-finding won't work for these TFs
  - many regulate just one operon
- Apply the gene neighbor method to TFs

*E. coli* K12  
*Erwinia* ...  
**fecI-fecR**

*Pseudomonas*  
**fecI-fecR**

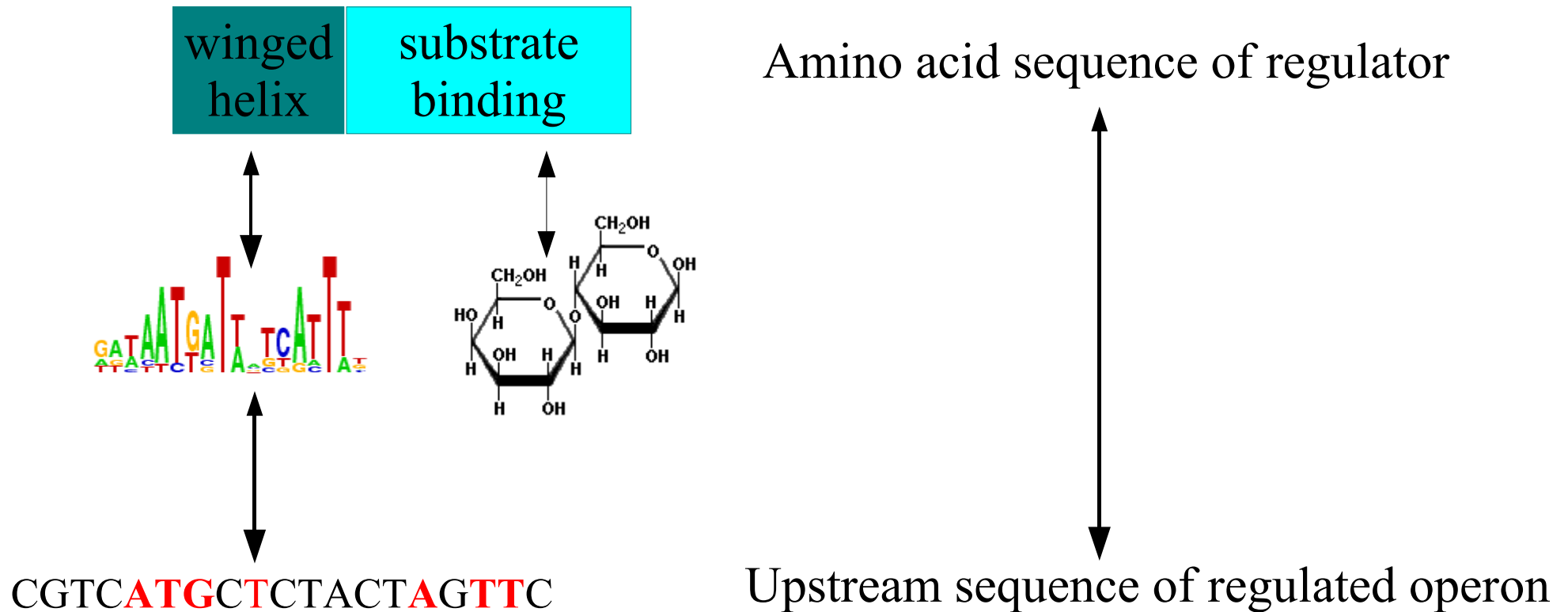
*Burkholderia*  
**fecI-fecR**

# Horizontal Transfer Allows the Prediction of Neighbor Regulation

- Results
  - Sensitivity 42% Specificity 41%
  - Focus on HGT events to improve specificity?
  - Sensitivity limited b/c of recent “assembly”
- Still not predicting the inducer

# How Do New Regulators Evolve?

- Looks like several things need to happen at once

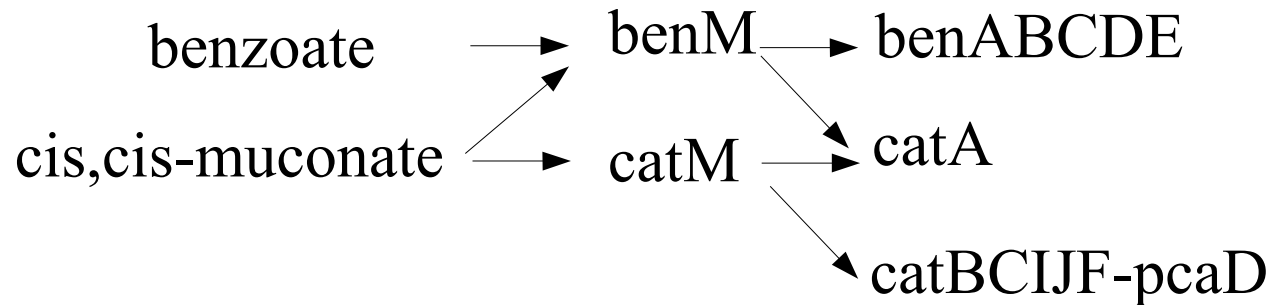


# Model For Evolving a New TF

1. New metabolic pathway (constitutive; on a plasmid)
2. Existing TF responds to a similar inducer (fortuitous)
3. Selected for a binding site (& operon formation?)
4. TF duplicates or moves to plasmid
5. One duplicate evolves to specialize in
  - the operon's binding site
  - the new inducer
6. Plasmid integrates into chromosome
7. Most plasmid genes are lost; operon & TF remain

# Duplication & Divergence of a Regulator

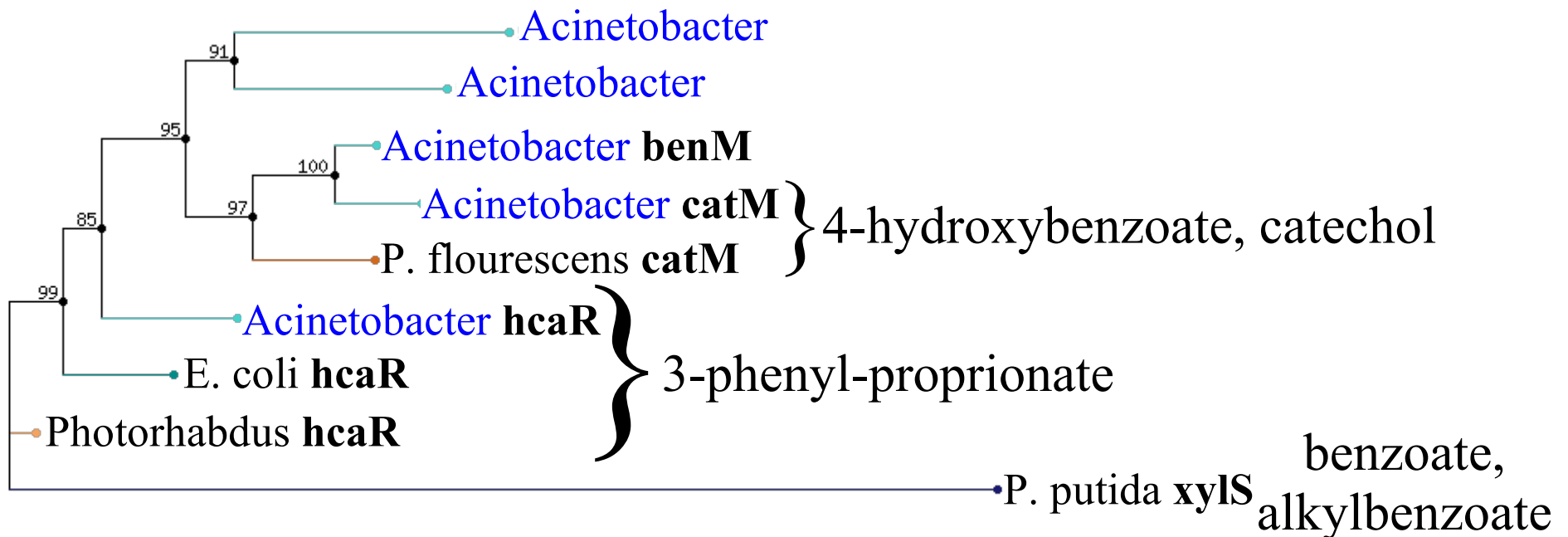
- Acinetobacter ADP1 *benM*/*catM*
- Near each other, near the *ben/cat* operons



Collier, Gaines,  
Neidle JB98

# Duplication & Divergence of a Regulator

- *benM* arose from *hcaR*
  - 3-phenylpropionate utilization
- *benM* duplicated to give *catM*



0.1



# Conclusions

- Acquiring genes with regulators is common
  - These still need to work with global regulators (e.g. CRP)
    - could explain why HGT from relatives is common
- Co-transfer allows prediction of neighbor regulation
  - no motifs required
- HGT genes often have new regulators, also acquired
  - from disparate sources
  - evolution in plasmids leads to neighbor regulation?
  - new regulators form from a pre-existing TF with a similar inducer?