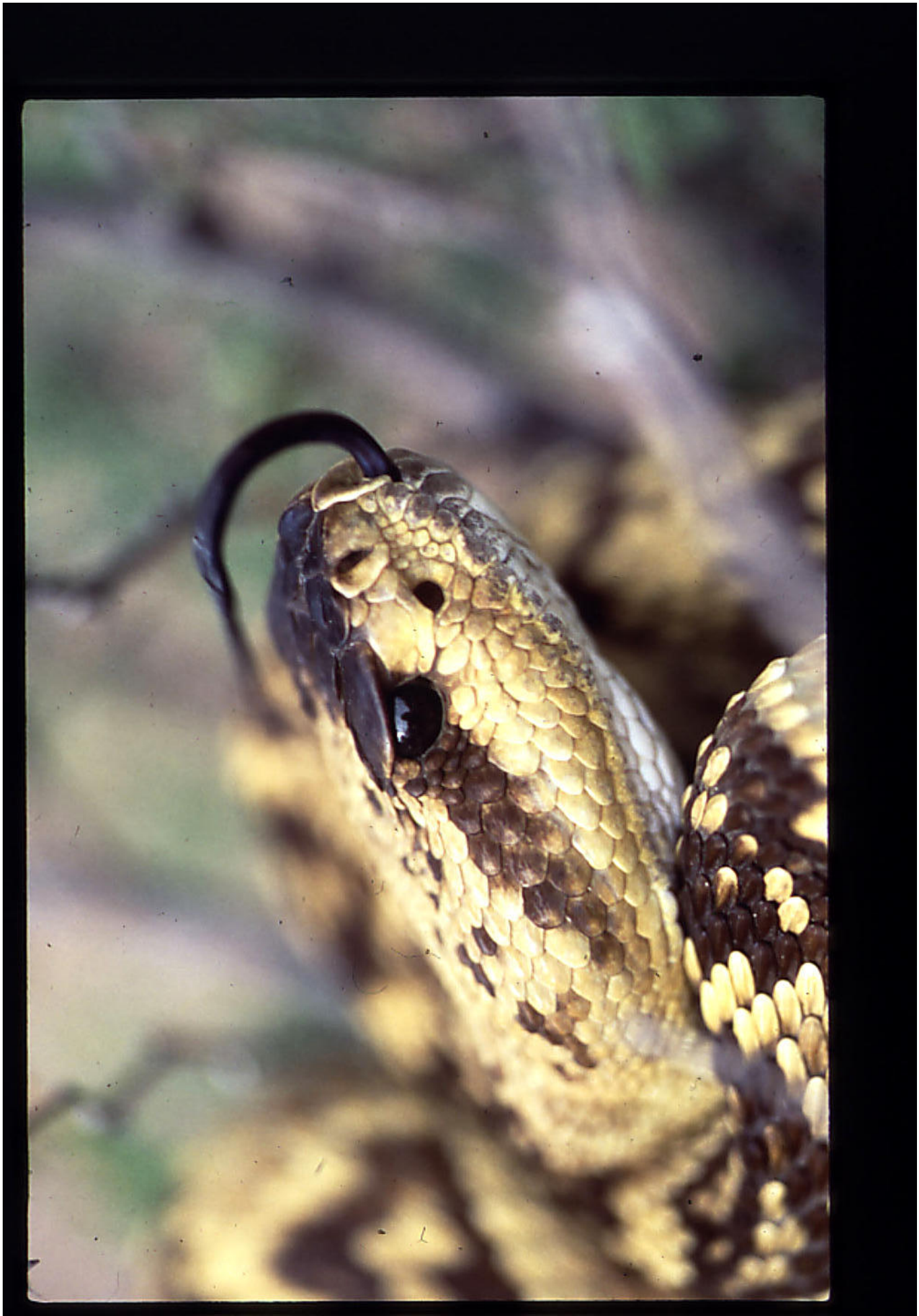
[01] Dr. D. Pollock, LSU (ITP 5/11/01) Genomic Biodiversity: Molecular Evolution Meets Functional Genom



A Case for Evolutionary Genomics and Genomic Biodiversity

- David D. Pollock, Biological Sciences, Louisiana State University (LSU), and Theoretical Division, Los Alamos National Laboratory (LANL)
- Jonathan A. Eisen, The Institute for Genomic Research (TIGR)
- Norman A. Doggett, Genomics Group, LANL
- Michael P. Cummings, Marine Biological Laboratory (MBL), Woods Hole

Evolutionary Genomics and Genomic Biodiversity

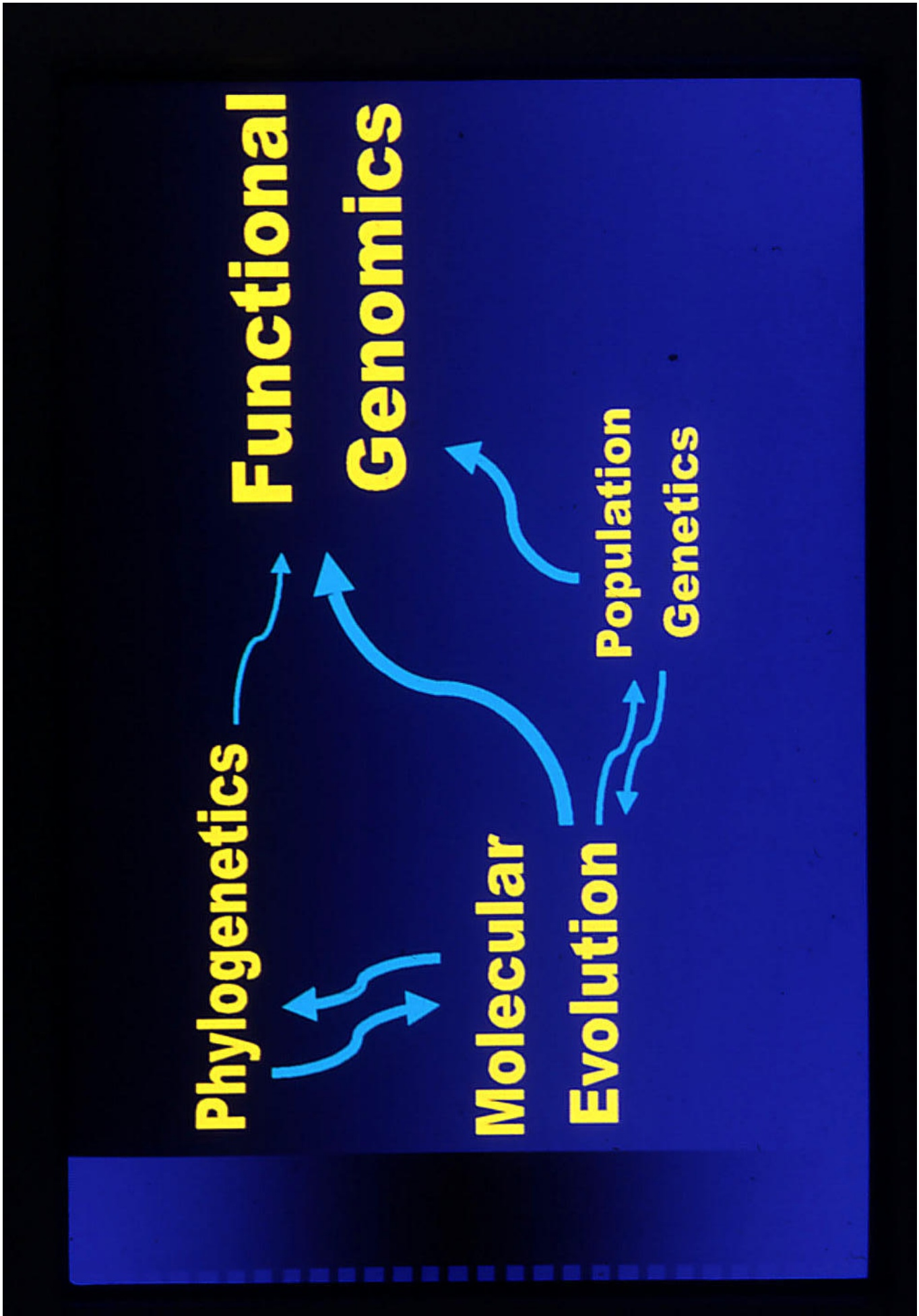
- **Evolutionary Genomics:** Application of Large-Scale Genomic Technology and Strategies in Molecular Evolution
- **Genomic Biodiversity:** Intense Sampling of Organisms at Different Taxonomic Levels for Large Genomes or Genomic Regions
- Relation to Comparative Genomics, Functional Genomics

Overview

- Justification of Evolutionary Genomics and Genomic Biodiversity on a Large Scale
- Opportunities for Increased Efficiency and Interaction with Genome Centers
- A Demonstration: The Vertebrate Mitochondria Project

Why Evolutionary Genomics and Genomic Biodiversity?
Phylogenetic Analysis
Molecular Evolution
Functional Genomics
Population Genetics

[06] Dr. D. Pollock, LSU (ITP 5/11/01) Genomic Biodiversity: Molecular Evolution Meets Functional Genom



Features of genome-based strategies
Multiple genes and taxa simultaneously
Large scale, coordinated, multi-group project
Shotgun, high throughput, high redundancy
Shift costs from labor to automated equipment
High efficiency
Human researchers focus on scientific design and analysis

Primary Products

- Complete sequencing of large regions from many taxa
- More accurate phylogenetic reconstruction
- Better estimation of evolutionary processes
- Estimation of changes in evolutionary processes
- Link sequence evolution to structure and function
- Improve accuracy of functional genomics predictions

Better Phylogenetic Analysis with Greater Taxon Density

- Better Estimation of Evolutionary Models
- Better Reconstruction of Topological Relationships
- Better Estimation of Divergence Times (Branch Lengths)

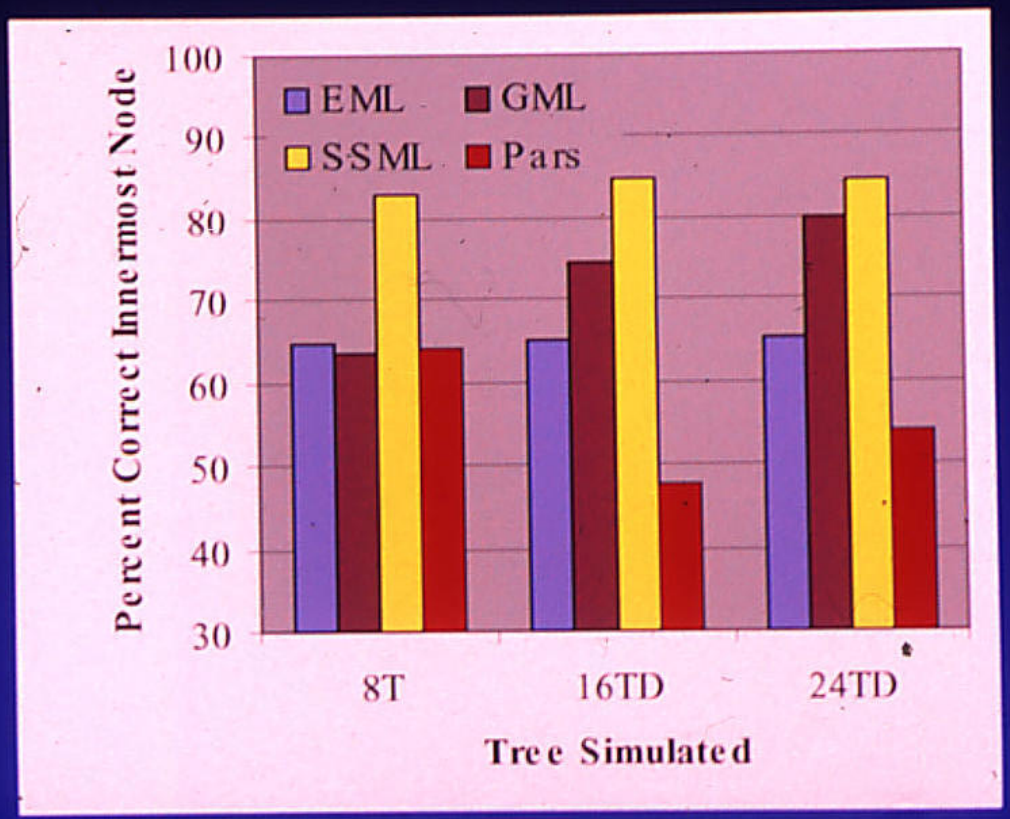
Proportion of Trees Identical to the Genome Tree: Contiguous Sites

[11] Dr. D. Pollock, LSU (ITP 5/11/01) Genomic Biodiversity: Molecular Evolution Meets Functional Genom

Improvement in Phylogenetic Reconstruction Using the Right Model

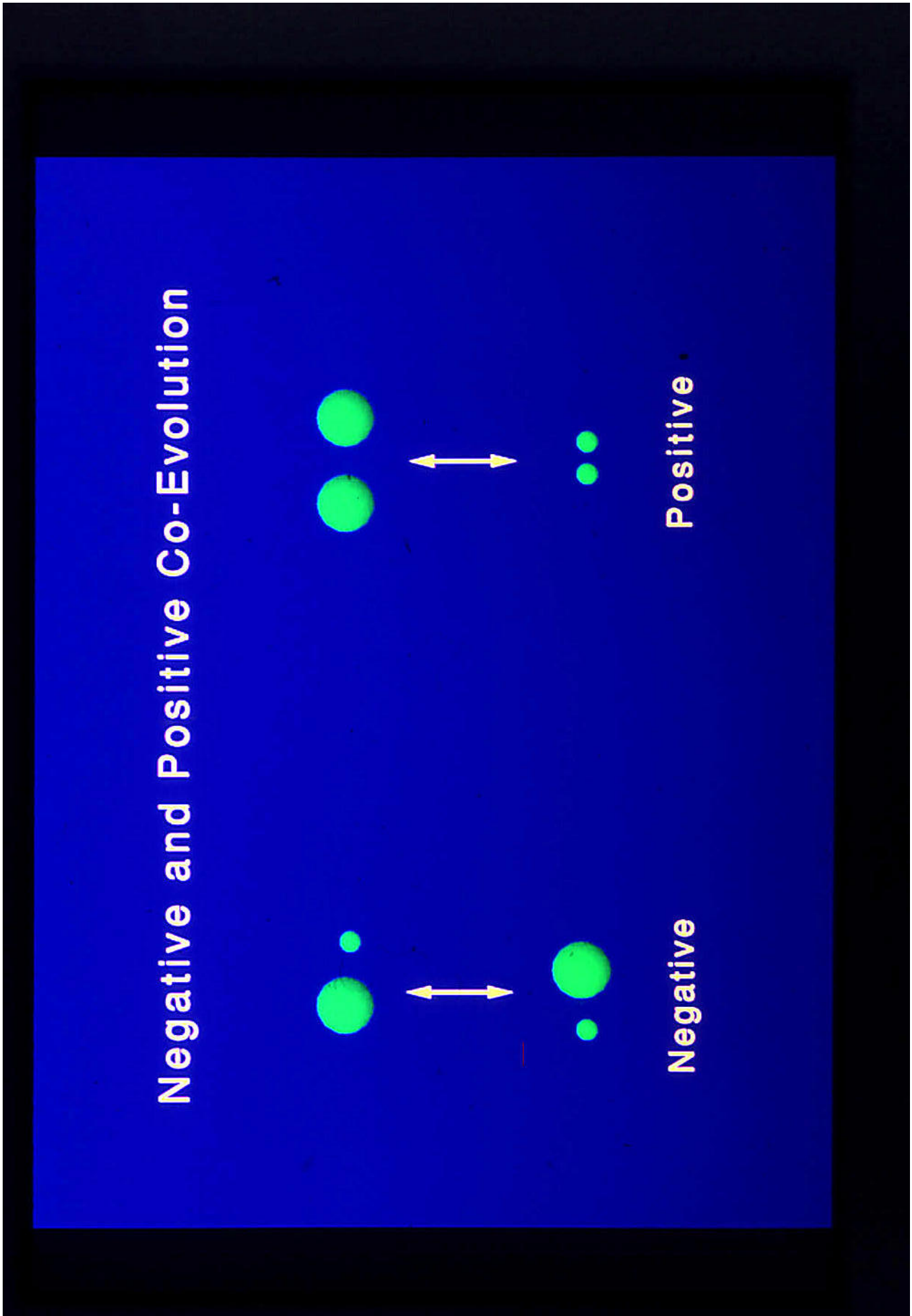
Adding Information about the Model
at Each Site: Doppelgänger Trees

Learning the Right Model by Adding Taxa: the Doppelgänger Effect

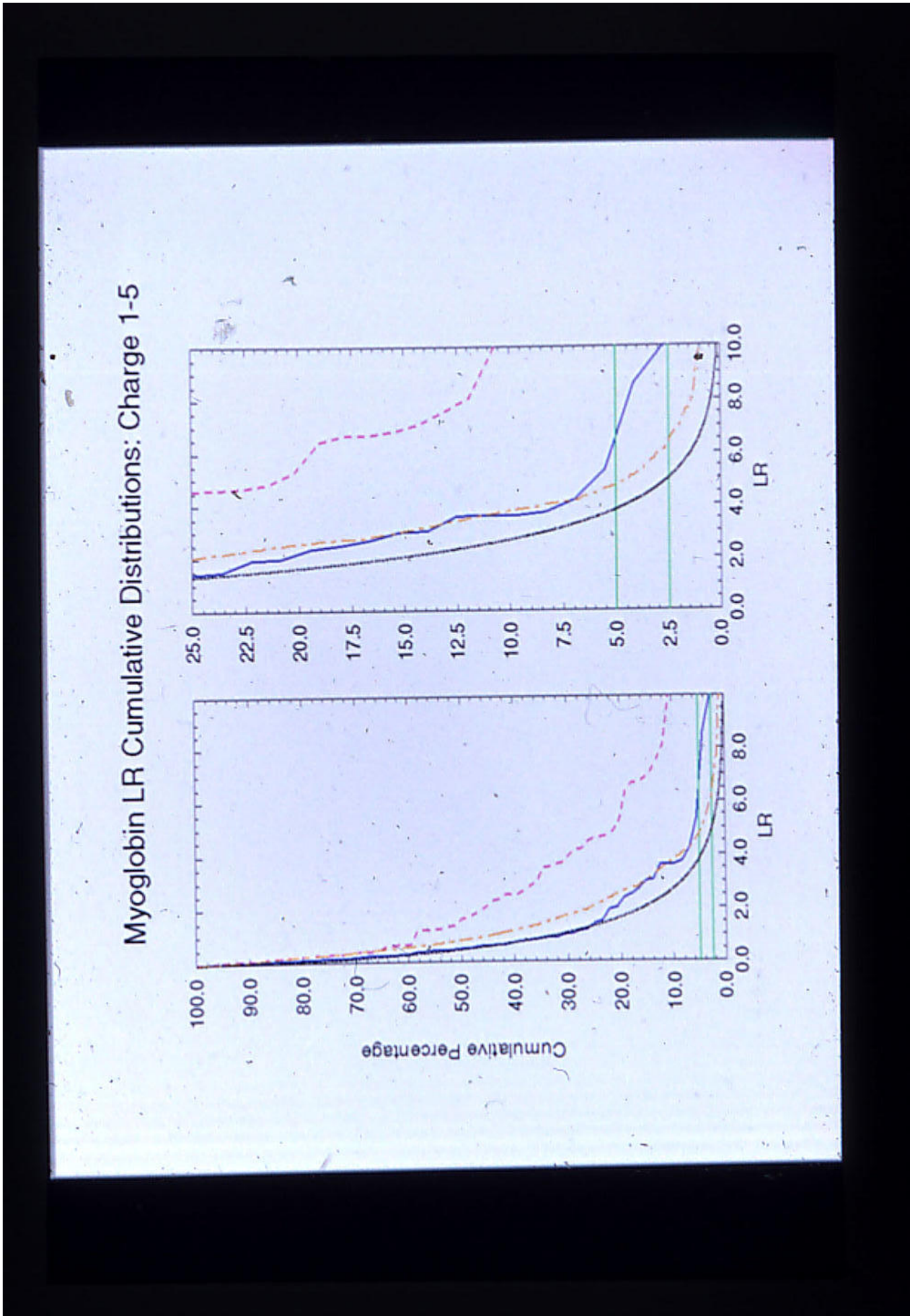


A zone of inefficiency for Parsimony with Increasing Numbers of Taxa

[16] Dr. D. Pollock, LSU (ITP 5/11/01) Genomic Biodiversity: Molecular Evolution Meets Functional Genom



[17] Dr. D. Pollock, LSU (ITP 5/11/01) Genomic Biodiversity: Molecular Evolution Meets Functional Genom

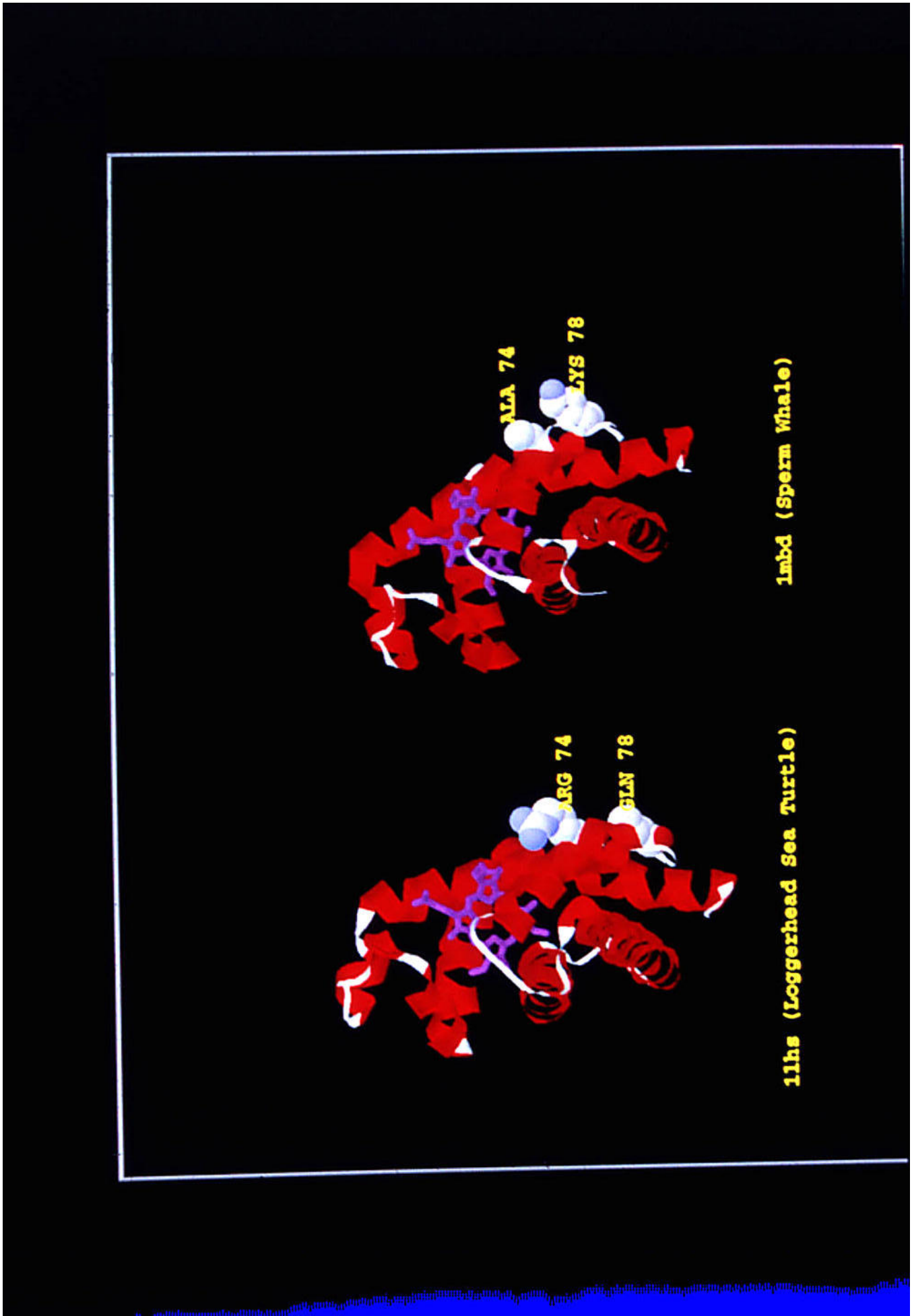


Negative Co-evolution of Adjacent Charge-Segregated Pairs is Biased towards Separation Distances of 3 and 4 along the Sequence

	Separation Distance					
	1	2	3	4	5	
P(LR)<0.05	2	2	8	8	4	24(7.8)
Total	31	27	35	36	27	156

Note: 3.6 residues per turn of α-helix

[19] Dr. D. Pollock, LSU (ITP 5/11/01) Genomic Biodiversity: Molecular Evolution Meets Functional Genom



Evolutionary, Structural, and Functional Genomics

- Better Knowledge of Evolutionary Process
- How does Structure & Function affect Rates and Probabilities of Substitution?
- Changes in Evolutionary Process with Changes in Function (Rapid)
- Change in Evolutionary Process with Change in Structure (Usually Slow)
- Predict Structure & Function Features

Lysozyme Structural Divergence
Chicken
Turkey
Human
Phage T4

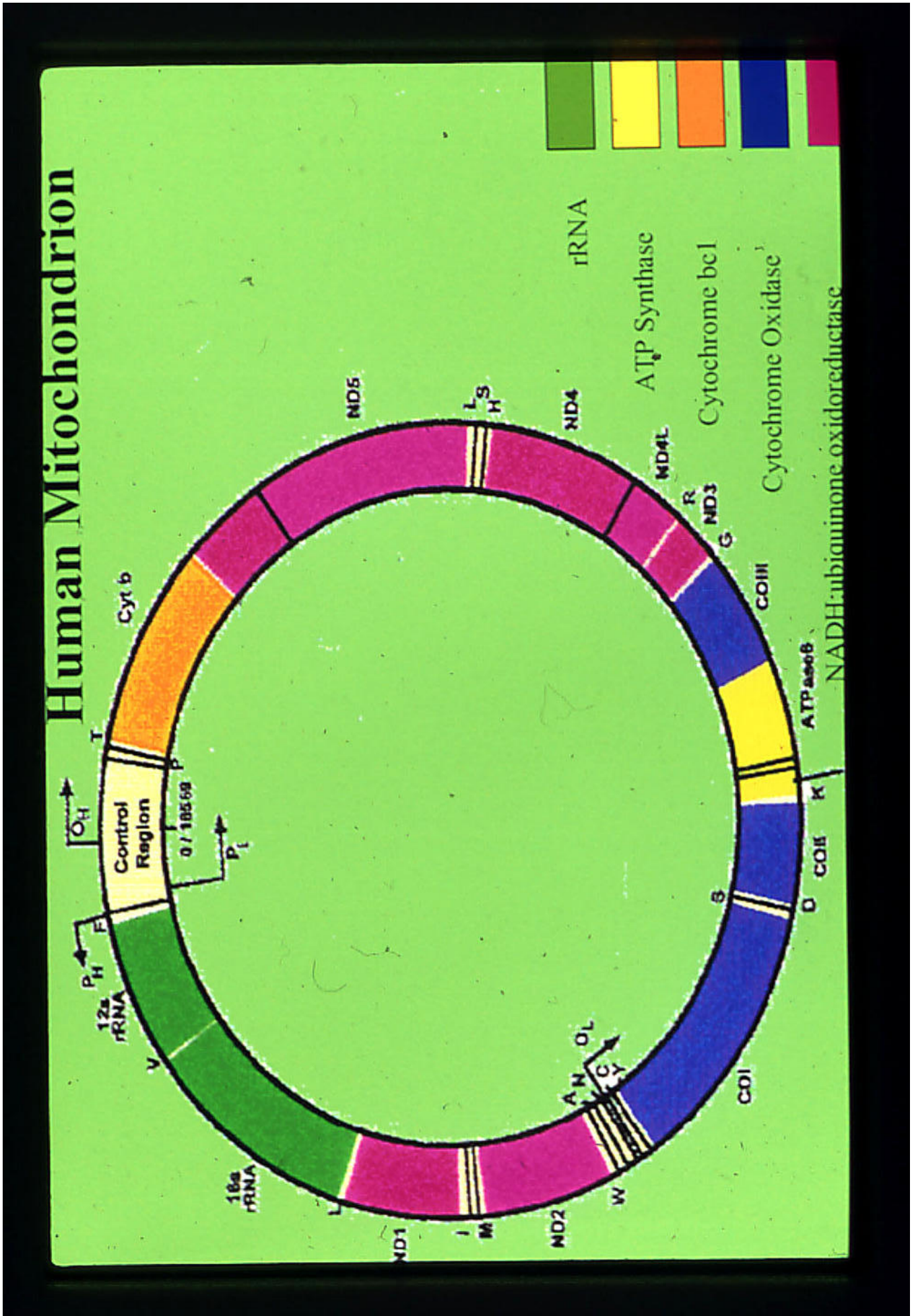
[22] Dr. D. Pollock, LSU (ITP 5/11/01) Genomic Biodiversity: Molecular Evolution Meets Functional Genom

[23] Dr. D. Pollock, LSU (ITP 5/11/01) Genomic Biodiversity: Molecular Evolution Meets Functional Genom

The Near-Human Evolutionary Environment

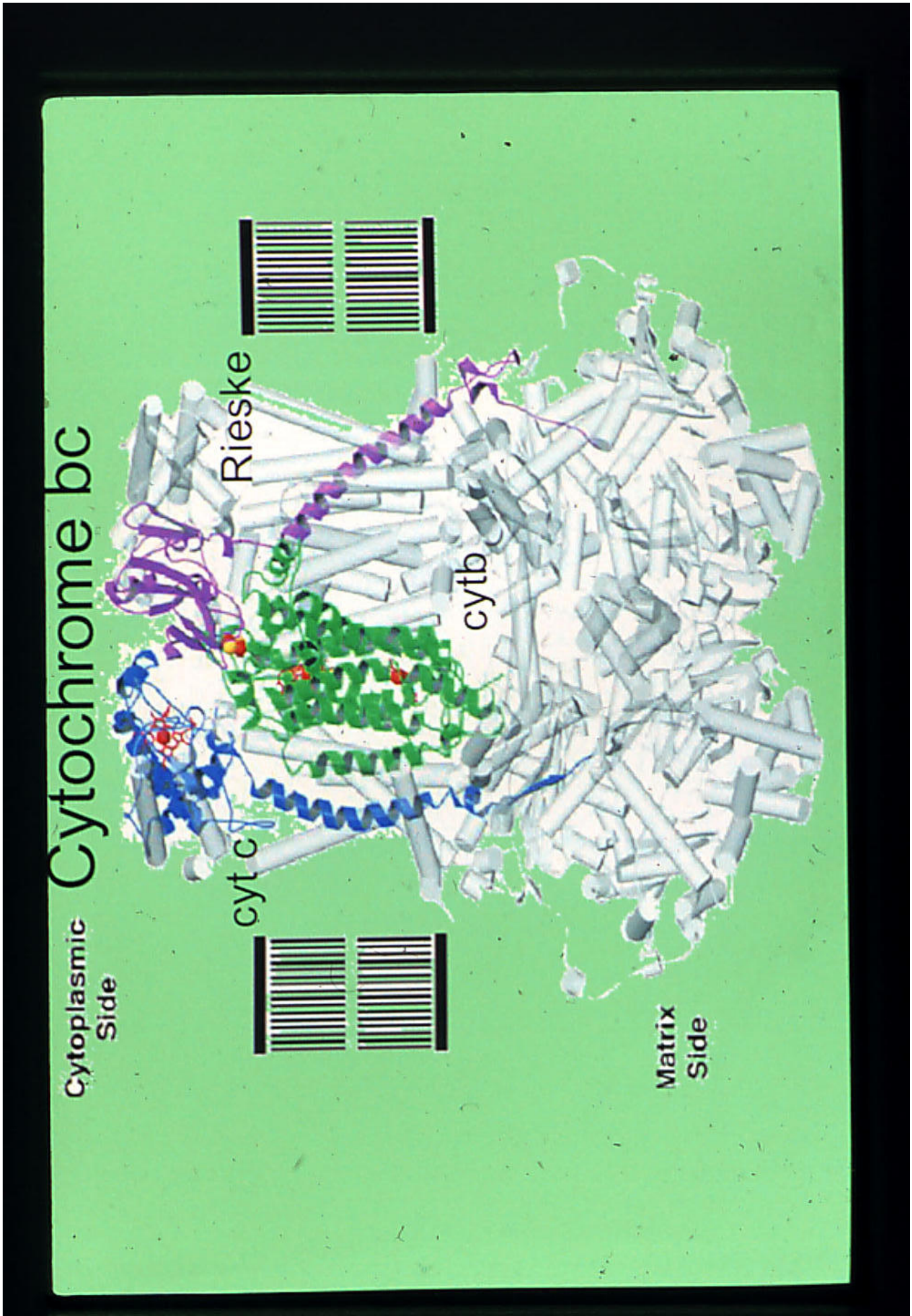
The Vertebrate Mitochondria Project

- Simultaneous Automated Shotgun Sequencing
- 10-20 Genomes per Experiment
- 2002 Genomes by end of 2002
- Density of Sampling
  - Some groups will be heavily sampled (e.g., Mammals. Also Primates, Passerines, Rodents)

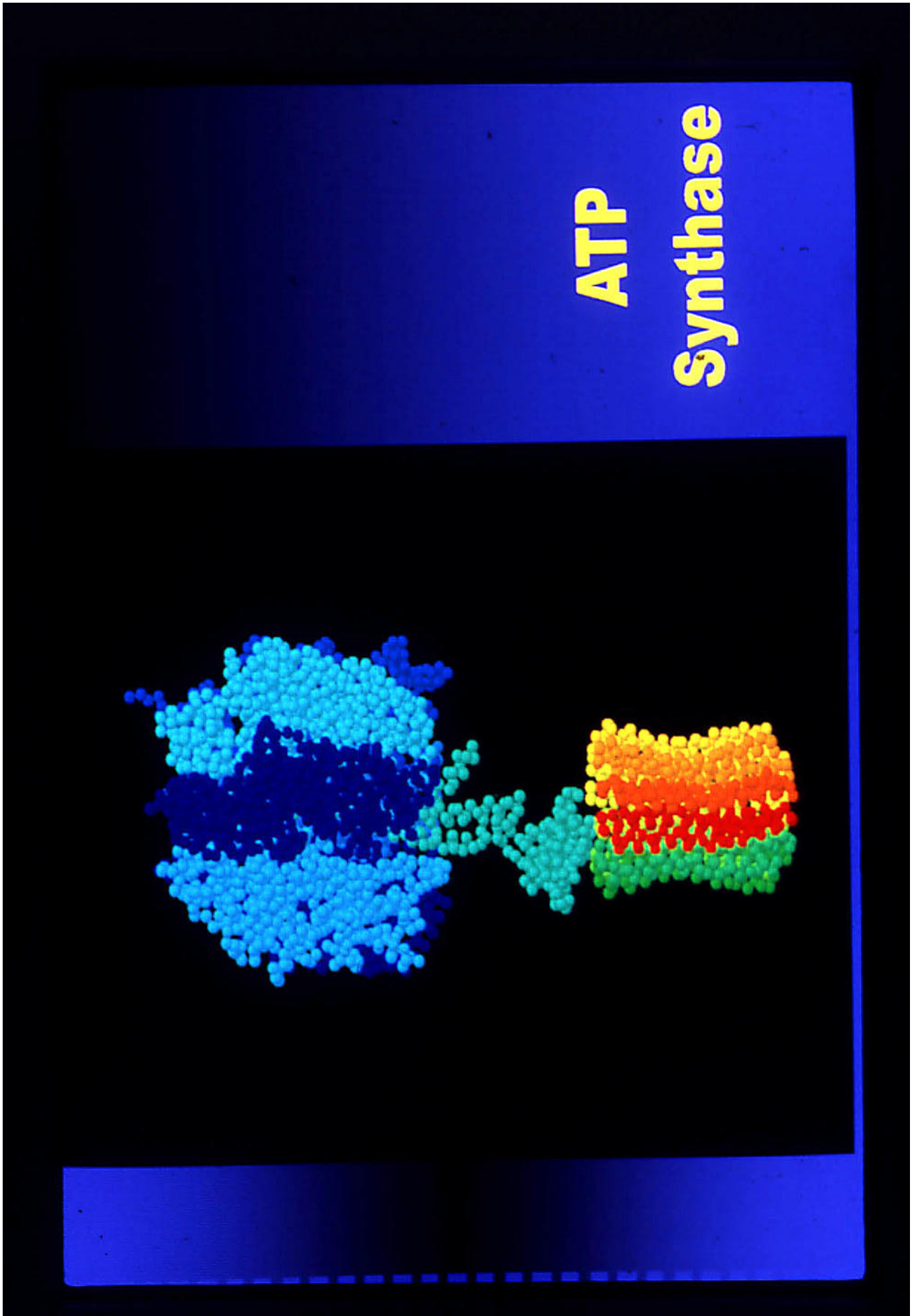

Human Mitochondrion
rRNA
ATP Synthase
Cytochrome bc1
Cytochrome Oxidase
NADH:ubiquinone oxidoreductase
Control Region
Cyt b
ND5
ND4
ND4L
ND3
COIII
ATPase6
COII
COI
ND2
ND1
12s rRNA
16s rRNA
F
V
L
I
M
Q
A
N
C
Y
W
S
D
K
G
R
H
S
L
T
P

[27] Dr. D. Pollock, LSU (ITP 5/11/01) Genomic Biodiversity: Molecular Evolution Meets Functional Genom

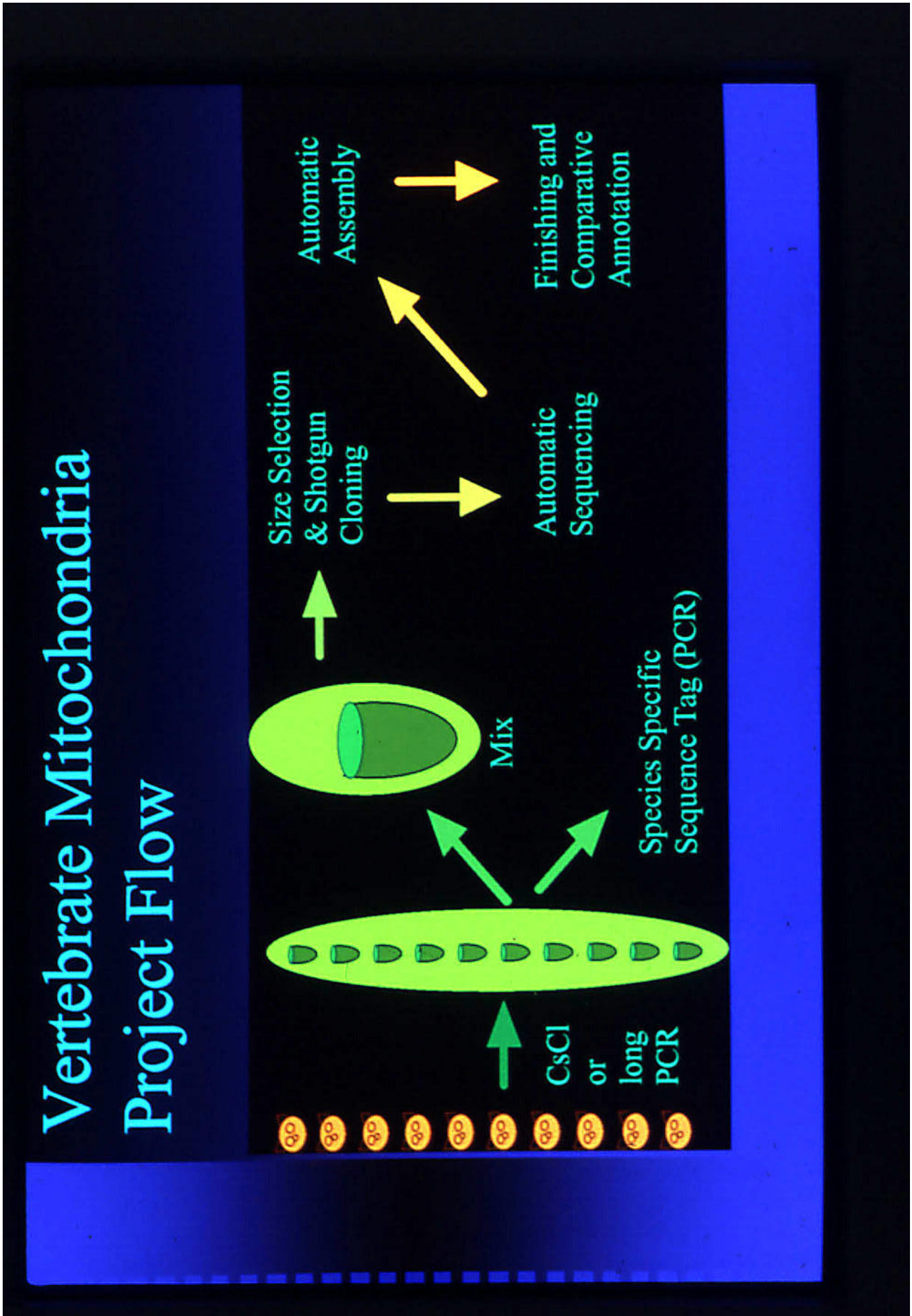
[28] Dr. D. Pollock, LSU (ITP 5/11/01) Genomic Biodiversity: Molecular Evolution Meets Functional Genom



[29] Dr. D. Pollock, LSU (ITP 5/11/01) Genomic Biodiversity: Molecular Evolution Meets Functional Genom

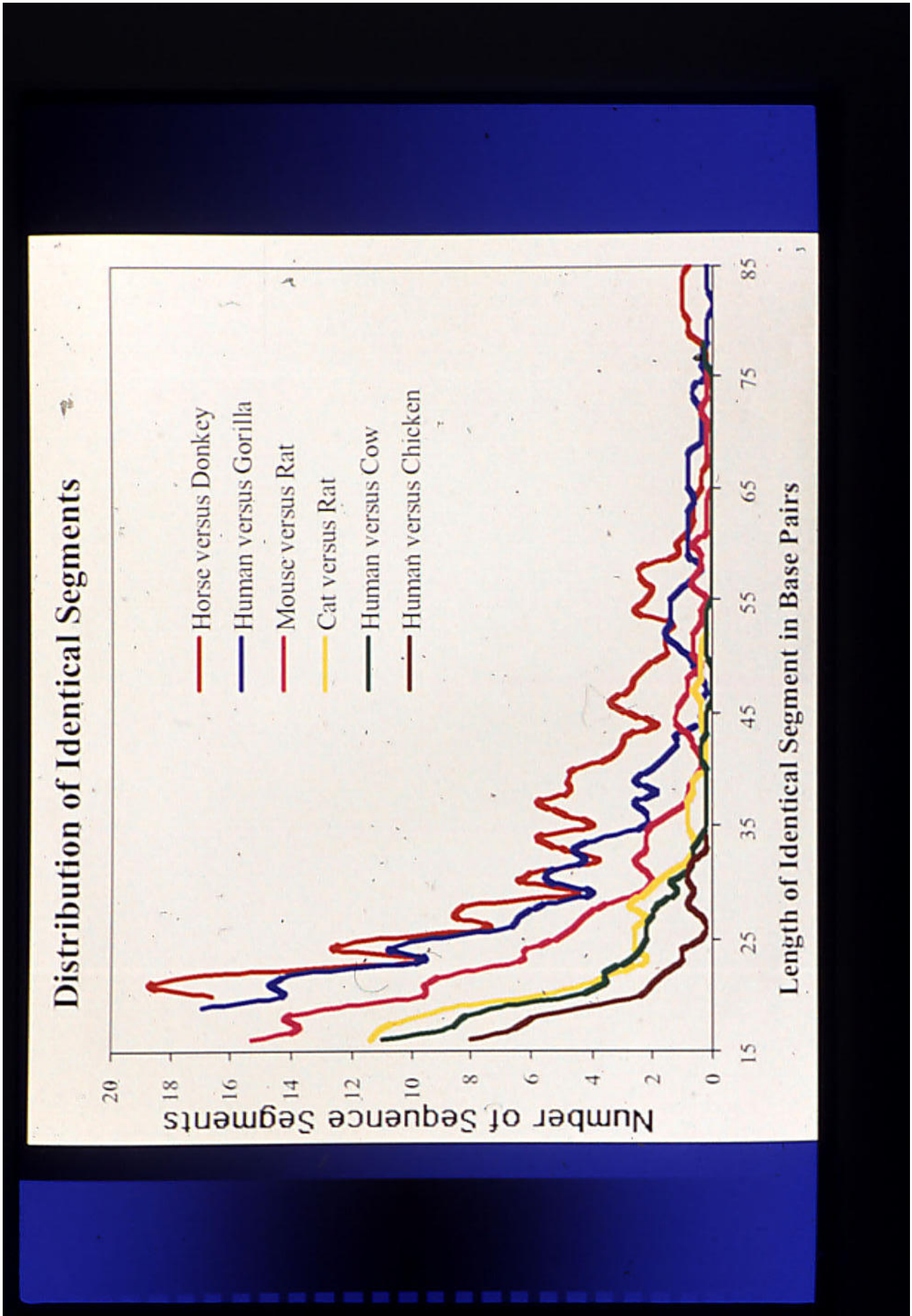


[30] Dr. D. Pollock, LSU (ITP 5/11/01) Genomic Biodiversity: Molecular Evolution Meets Functional Genom



Potential Problems and Solutions

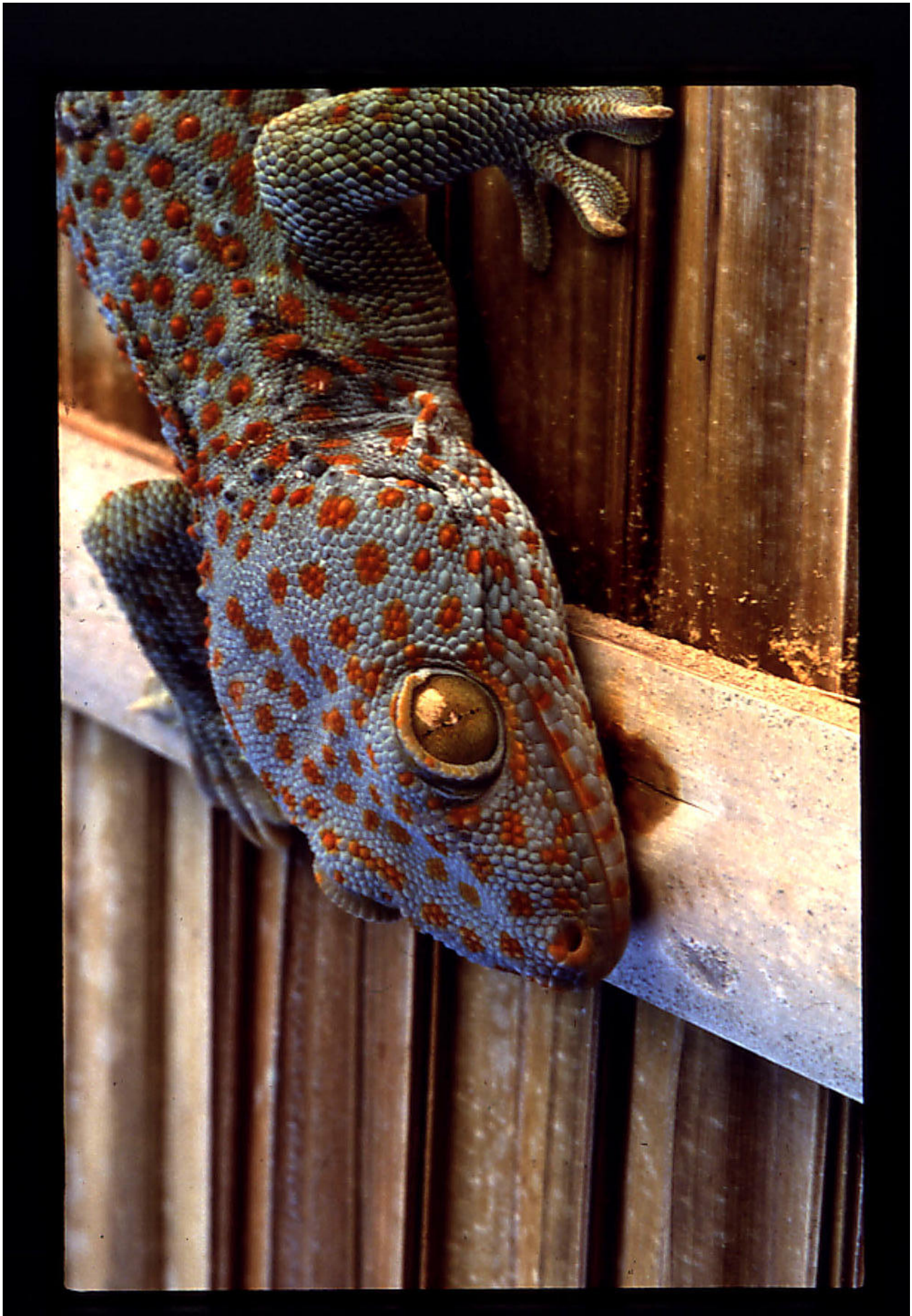
- Inaccurate Taxonomic Identification
  - Short PCR & Sequence on Original
- Errors from PCR and Sequencing
  - Redundancy: Errors will be Known
  - Sequence Quality Values
- Nuclear Transfer of Mitochondrial DNA
  - Mitochondrial Purification & Long PCR
  - Analysis of Polymorphism


Distribution of Identical Segments
Horse versus Donkey
Human versus Gorilla
Mouse versus Rat
Cat versus Rat
Human versus Cow
Human versus Chicken
Number of Sequence Segments
0
2
4
6
8
10
12
14
16
18
20
15
25
35
45
55
65
75
85
Length of Identical Segment in Base Pairs

Evolutionary Questions

- Vertebrate Mitochondrial Topology
- Resolution of Deep (Ordinal) and Shallow (Family and Generic) Relationships
- Test Accuracy of Reconstruction Techniques
- Obtain accurate knowledge of evolutionary process, including
  - changes in the process, functional divergence
  - context dependence
  - interaction between subunits, domains and secondary structures

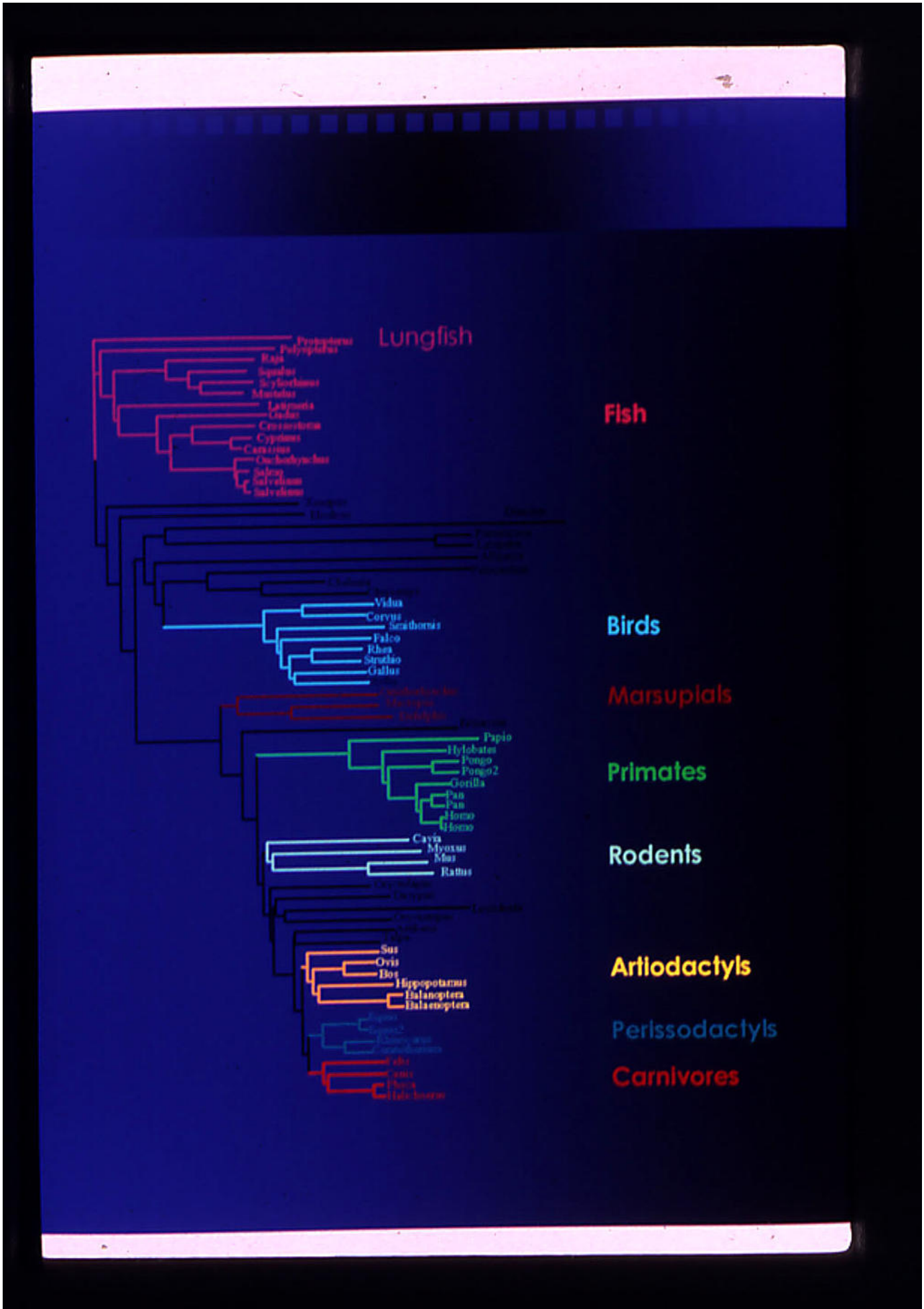
[34] Dr. D. Pollock, LSU (ITP 5/11/01) Genomic Biodiversity: Molecular Evolution Meets Functional Genom



[35] Dr. D. Pollock, LSU (ITP 5/11/01) Genomic Biodiversity: Molecular Evolution Meets Functional Genom

[36] Dr. D. Pollock, LSU (ITP 5/11/01) Genomic Biodiversity: Molecular Evolution Meets Functional Genom

[37] Dr. D. Pollock, LSU (ITP 5/11/01) Genomic Biodiversity: Molecular Evolution Meets Functional Genom



Mitochondrial Genome Sequencing Rates

- Genomes per ABI 3700 Week: 23
  Months per 2000 Genomes: 20
- Weeks required to complete 2000 Genomes by DOE's Joint Genome Institute: 1
- Days required to complete 2000 Genomes running at full capacity,
  Celera : ~2
  Human Genome Project: ~2

Next Steps

- Feasibility and Scale-up
  - 5 Genomes (Currently Underway)
  - 20 Genomes (Samples Being Processed)
  - 100 Genomes (Samples Being Acquired)
  - 2000 Genomes (1-3 Years)
- Automation
- Database and Access
- Automated Large-Scale Analysis
- Further Collaboration with Systematists and Genome Centers

Collaborators: DNA Samples and Taxonomic Analysis

- Robert Zink, U. Minnesota, Bell MNH
- Caro-Beth Stewart, SUNY, Albany
- Jim McGuire, Frank Burbrink, LSU MNH
- Mark Hafner, Fred Sheldon, LSU MNH
- David Penny, Massey U., New Zealand
- David Mindell, U. Michigan
- Maryellen Ruvulo, Randy Collura, Harvard
- Rodney Honeycutt, Texas A&M

[41] Dr. D. Pollock, LSU (ITP 5/11/01) Genomic Biodiversity: Molecular Evolution Meets Functional Genom