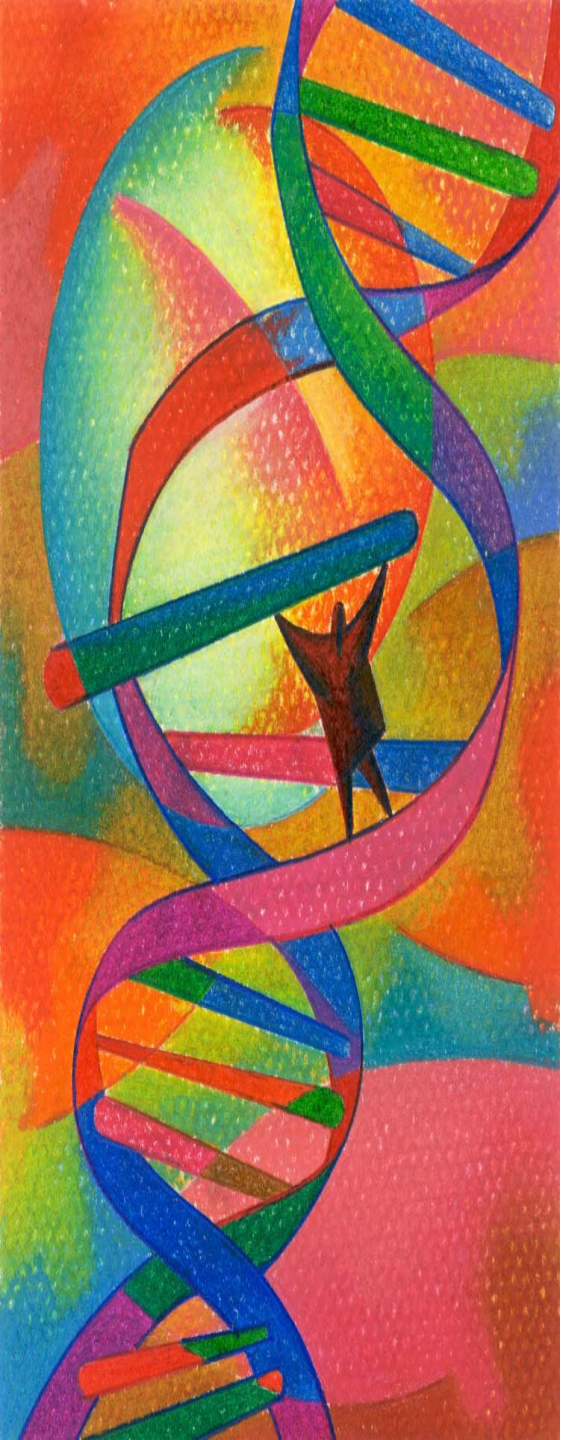
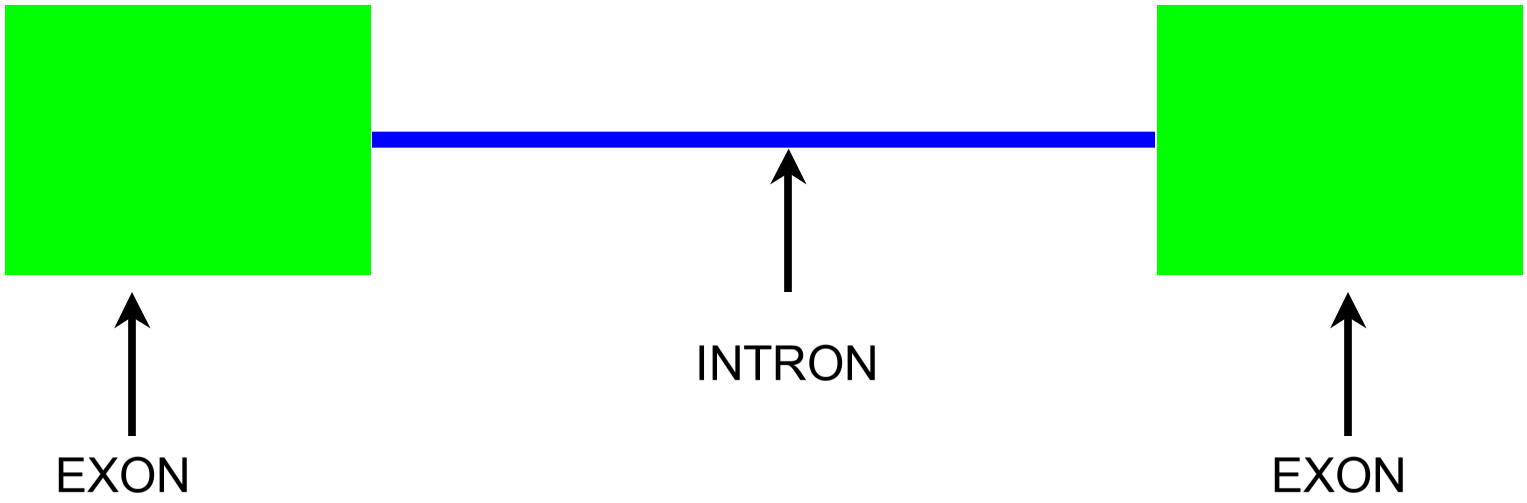


Natural Selection on Introns in *Cryptococcus neoformans*

Stephanie Hughes
Jackson State University
Summer 2006

What is an intron?

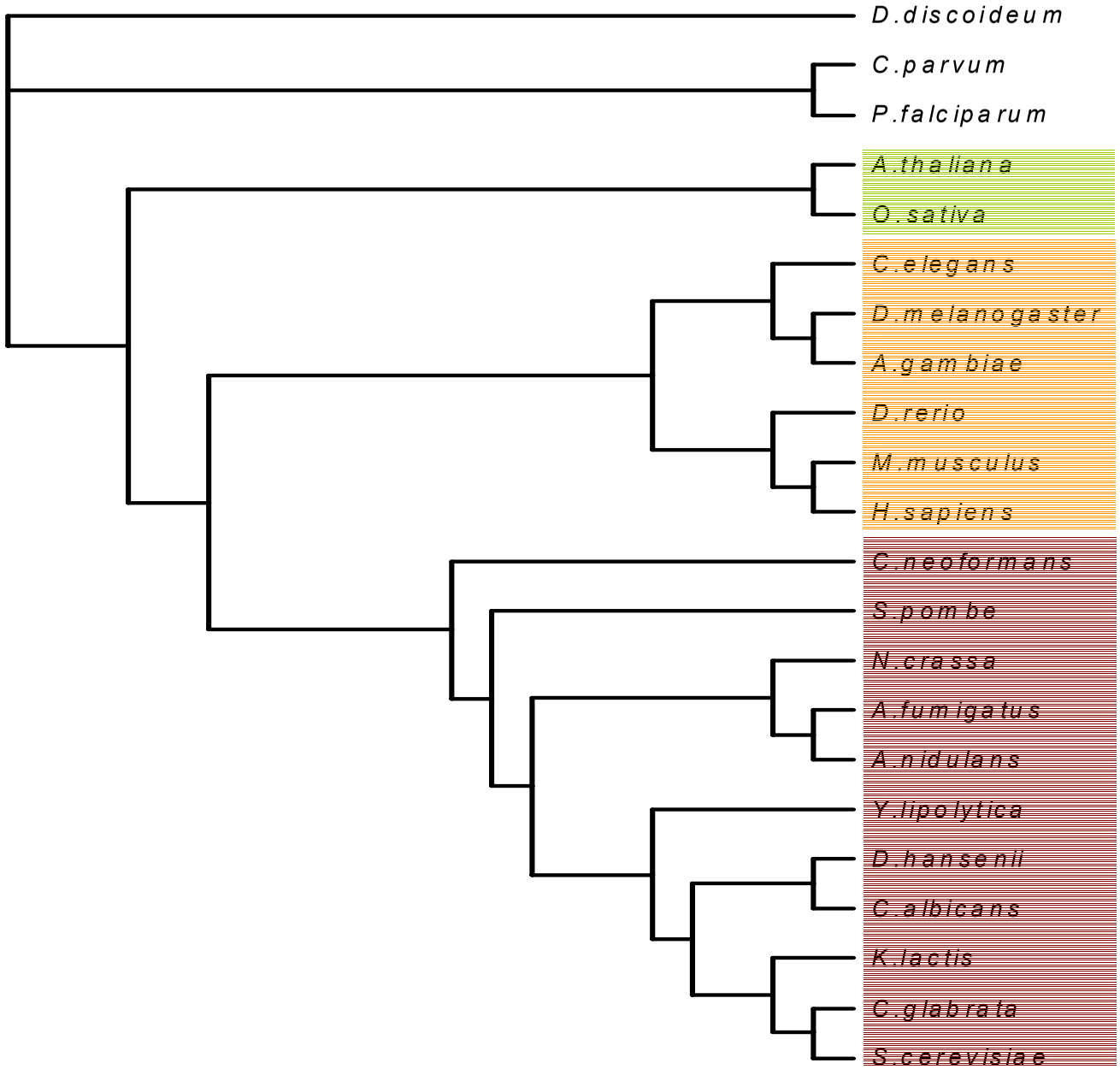


An intron is a noncoding region that intervenes exons.

Introduction

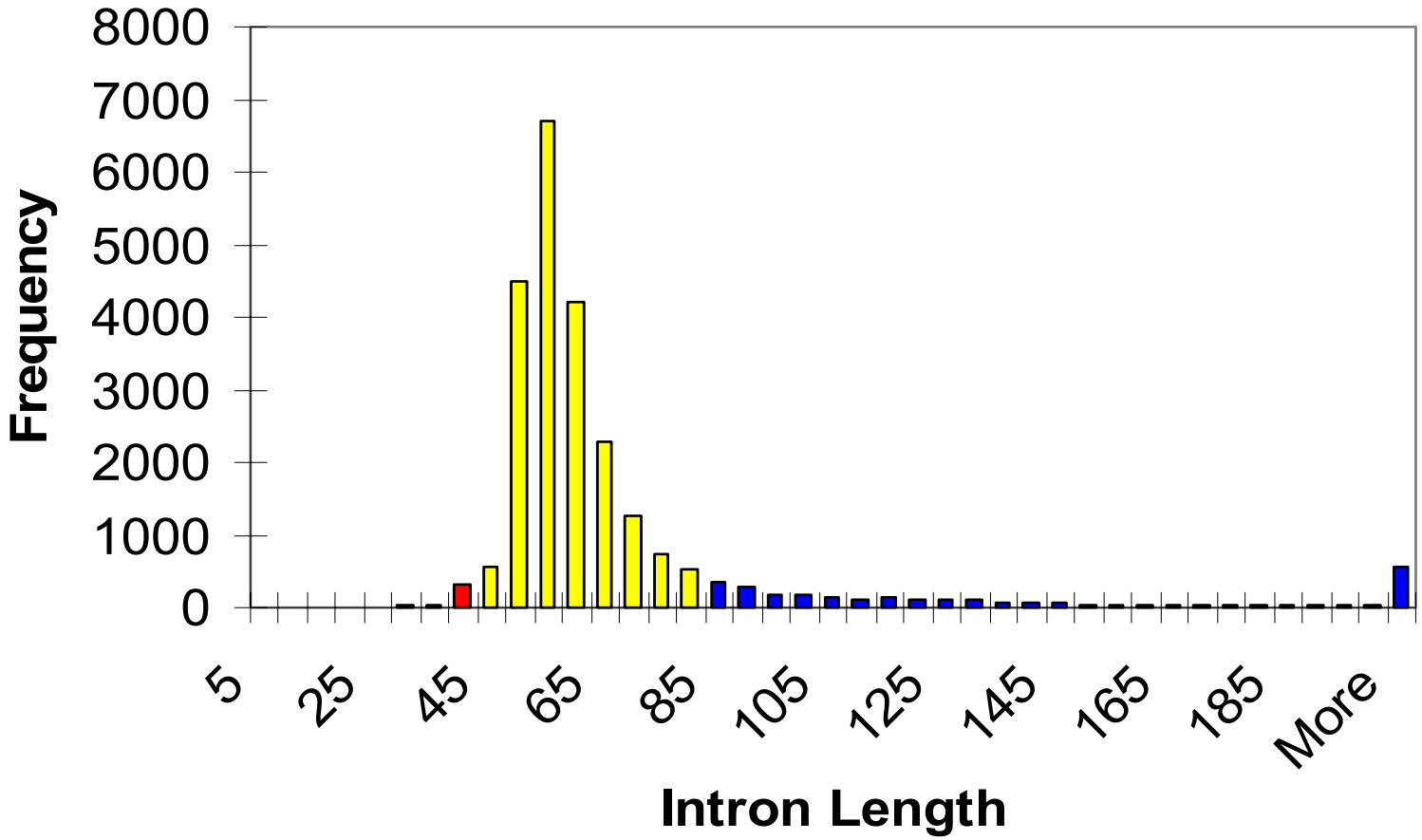
- Function of introns is unclear
- Natural selection is involved with introns:
  - Size correlates with areas of high recombination in Drosophila
  - Size correlation in highly expressed genes
  - Size correlation and intron rank
- Role in protein evolution and gene expression

Intron Profiles Across The Eukaryotes



	genes w/ introns	Avg. intron length	Avg # introns/gene
D.discoideum	70%	150 bp	1.90
C.parvum	1%	93 bp	1.29
P.falciparum	51%	179 bp	2.61
A.thaliana	78%	160 bp	5.63
O.sativa	77%	374 bp	4.04
C.elegans	98%	353 bp	5.36
D.melanogaster	84%	781 bp	4.32
A.gambiae	86%	978 bp	3.23
D.rerio	95%	2661 bp	7.90
M.musculus	74%	4229 bp	8.57
H.sapiens	77%	4281 bp	7.55
C.neoformans	95%	69 bp	5.48
S.pombe	43%	84 bp	2.07
N.crassa	75%	122 bp	2.07
A.fumigatus	77%	102 bp	2.36
A.nidulans	86%	104 bp	3.03
Y.lipolytica	10%	271 bp	1.1
D.hansenii	5%	178 bp	1.06
C.albicans	3%	127 bp	1.18
K.lactis	2%	345 bp	1.02
C.glabrata	2%	454 bp	1.02
S.cerevisiae	5%	109 bp	1.03

Intron length has a broad range

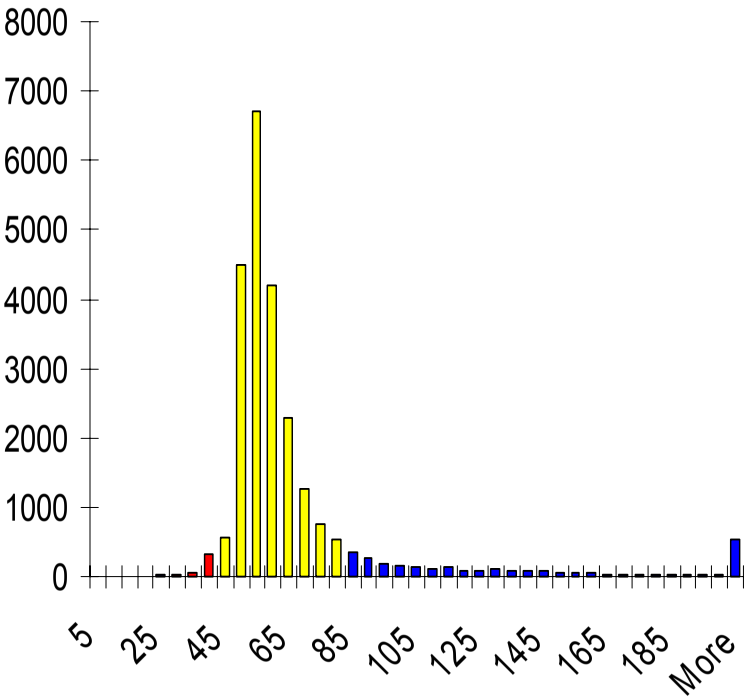


*Mode = 52 bp

Why *C. neoformans*?

- Four strains sequenced
- (Two at the Broad)
- Strains diverged but remain similar (0.5 subs/site)

Hypotheses



Assuming that typical introns are in mutation-selection equilibrium :

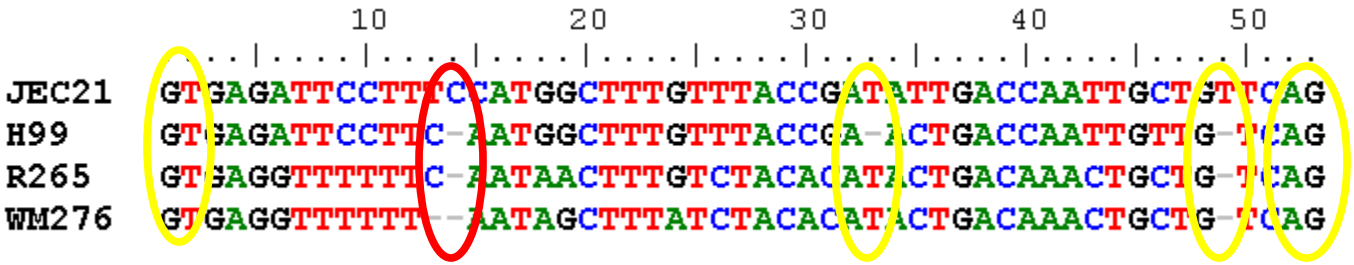
A) Long introns are temporarily in disequilibrium.

-OR-

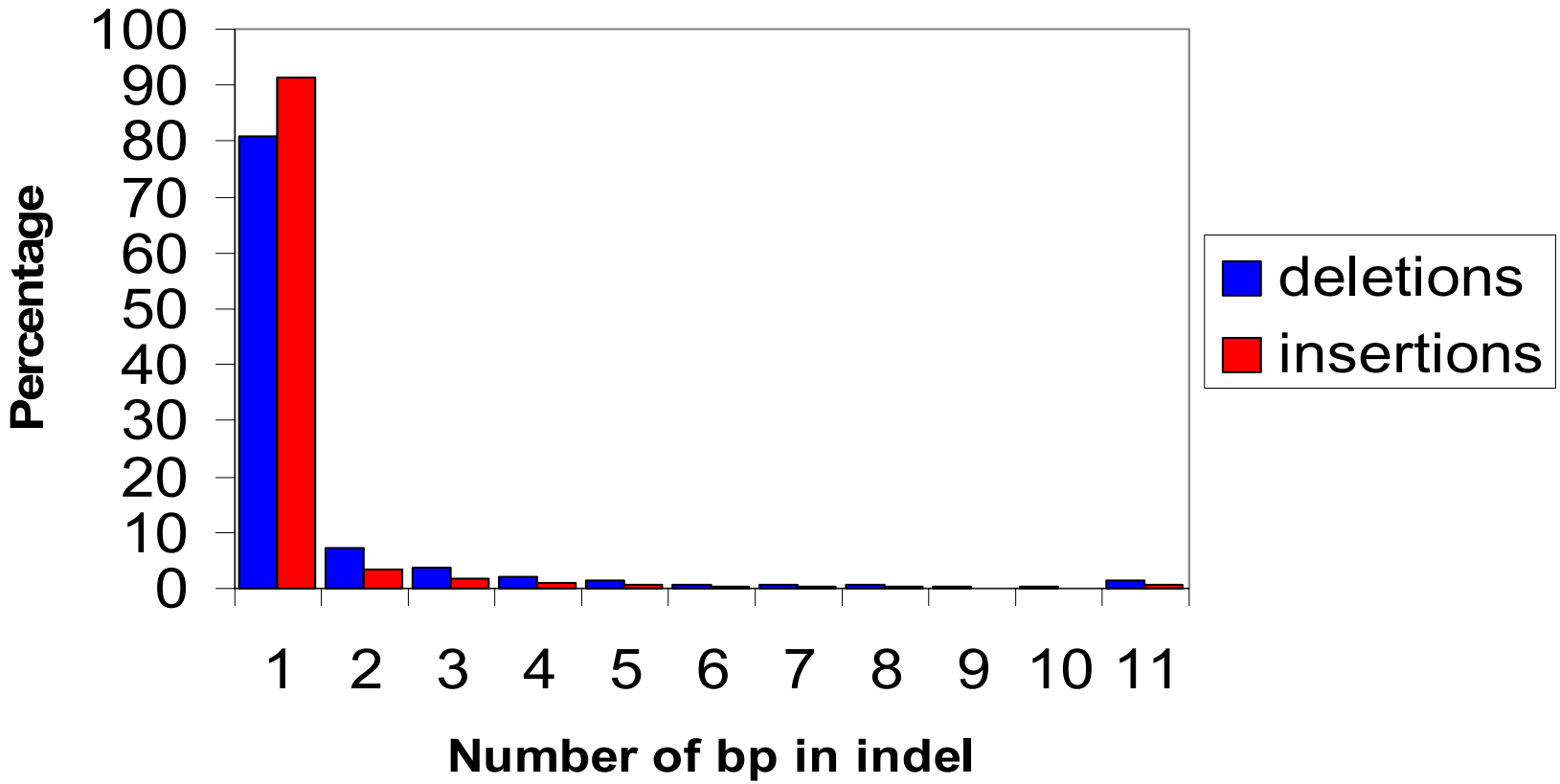
B) Long introns are maintained by natural selection.

Methods

- Analyze introns by doing the following:
  - Test to confirm introns
  - Distinguish intron size
  - Analyze indels
  - Choose only rootable indels
  - Count number of bp in each indel
  - Calculate percentage of conservation



Most indels are 1 bp



Total bp deleted = 17604

Total bp inserted = 11920

Total deletions = 12604

Total insertions = 5894

Indel Rates (indel bp/intronlength)

Intron Size	Deletion Rate	Insertion Rate
Long (>86bp)	0.0110 [.0110-.0122]	0.0069 [.0056-.0070]
Typical (47-86bp)	0.0120 [.0118-.0122]	0.0078 [.00757-.00816]
	p=0.003	p=0.045

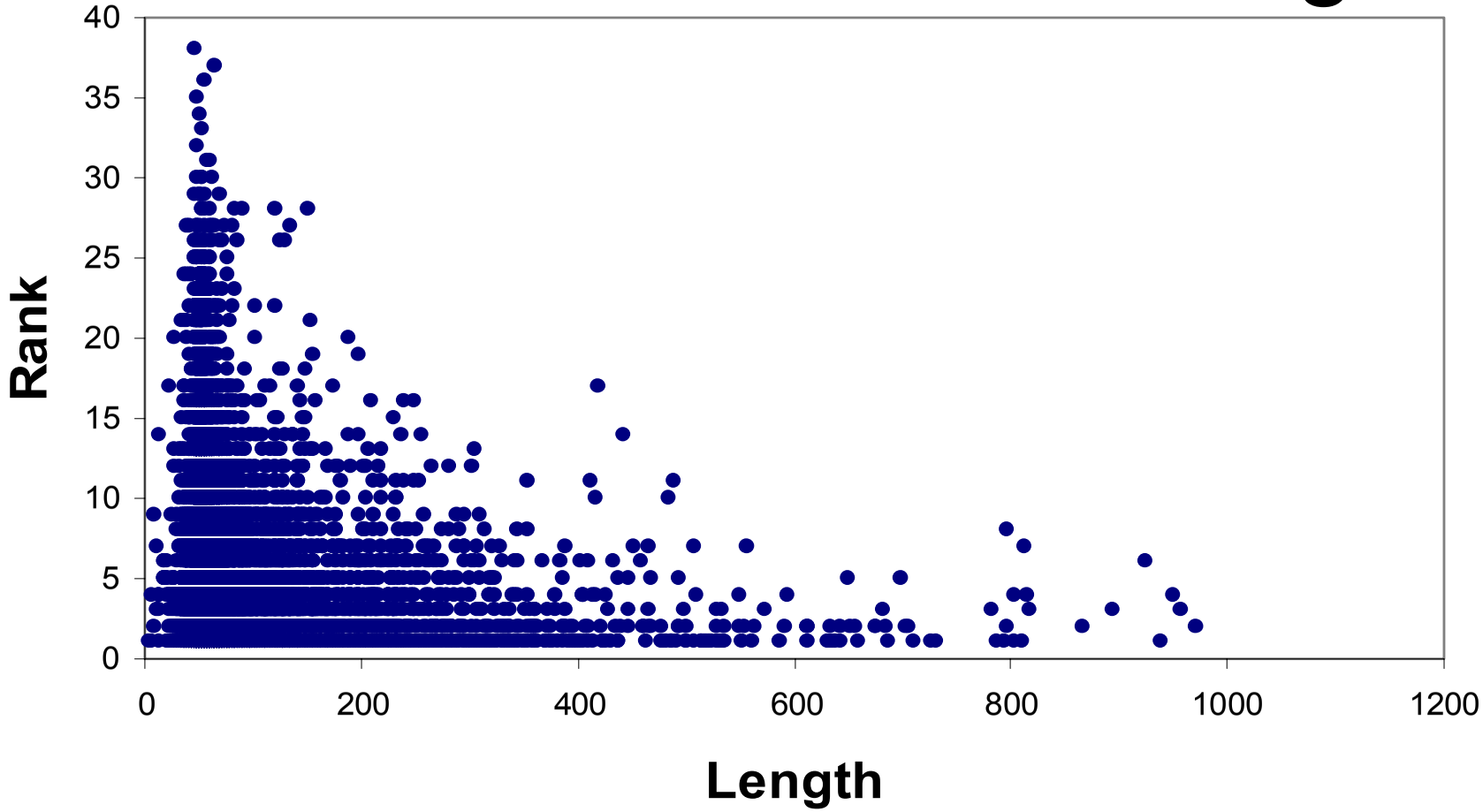
* Magnitude of deletion bias = 0.3% length change over 140 million years

Conclusions

- Long introns show a different indel rate as typical introns suggesting they are not under mutation-selection equilibrium.

- Are some long introns maintained while others are in disequilibrium?

First Rank Introns are Longer



Avg. length of first rank introns = 79.7 bp

Avg. length of other introns = 64.8 bp

Spearman's rho = - 0.089 p = 0.001 p=1.14x10-85 (t-test)

Future Studies

Effect of intron length and gene expression

Effect of intron size and rate of evolution

Acknowledgments

- The Broad Institute
- Eric Lander
- Daniel Neafsey
- Bruce Birren
- Angela Brunache