Phenotype, Selection, and Adaptation

or

Everything You Wanted To Know About Evolutionary Theory But Were Afraid To Ask

Amoolya Singh

PhD candidate
UC Berkeley

17 Aug 2005

A Little Etymology

evolve: Latin *volv-re* to roll out, unroll; from *e-* out + *volvre* to roll

7. To develop by natural processes from a more rudimentary to a more highly organized condition; to originate (animal or vegetable species) by gradual modification from earlier forms. (OED 2nd ed. 1989)

1832 LYELL Princ. Geol. II. i. 14 *The orang-outang, having been evolved out of a monad, is made slowly to attain the attributes and dignity of man*.

- the **fact** of evolution: organisms are related by common descent
- the **history** of evolution: the details of when lineages split from one another and of the changes that occurred in each lineage
- **mechanisms** or processes by which evolutionary change occurs

Natural Selection

27 Dec **1831**: Charles Darwin joined the HMS Beagle as a naturalist on a five-year expedition around the world

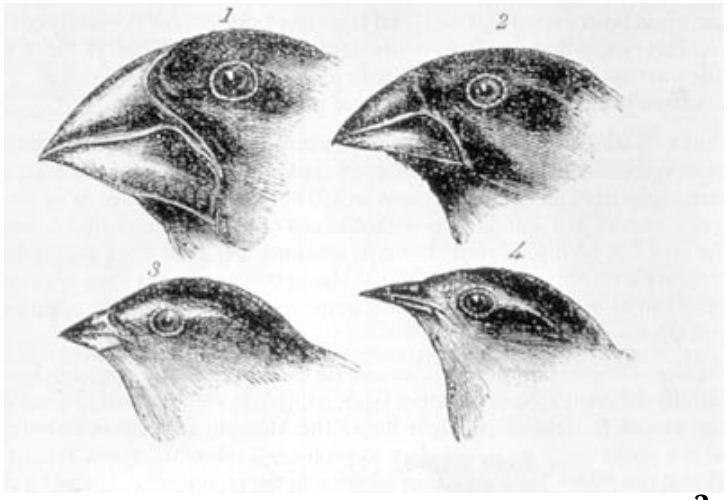
Cabin reading: Lyell (geology, fossils)

Finches and nuts on Galapagos islands: how did the variation come about?

Read Malthus on population, food

Heredity, variation among offspring, and limited food => evolution

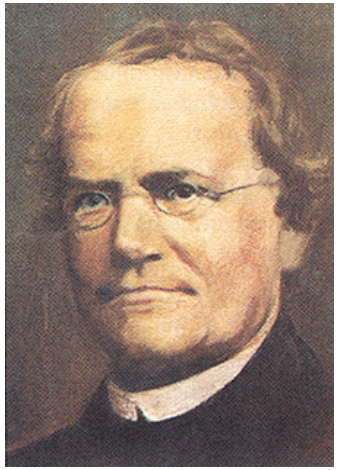
1859 *On the Origin of Species by Means of Natural Selection*



Meanwhile, back in Brno…

1866 Gregor Mendel, *Experiments in Plant Hybridisation*

- inheritance in pea plants
- hypothesized a **factor** that conveys traits from parent to offspring
- independent assortment
- distinction between dominant and recessive traits
- distinction between heterozygote and homozygote
- difference between what would later be described as genotype and phenotype



Fell into obscurity until **1901**, when William Bateson and others revived it

Proposed as a mechanism for evolution with discrete & heritable characters

Mechanism for Evolution?

Even with acceptance of concept of evolution, no good theory of inheritance that could offer mechanism:

- neo-Darwinism (gradual changes)
- neo-Lamarckism (envt acts directly on organs)
- orthogenesis (better and better forms)
- Mendelism (discrete variation)
- biometric approach (continuous variation)
- mutationism ("ordinary" & mutational variation)

Bishop Samuel Wilberforce and TH Huxley, who traded verbal blows debating Biblical genesis vs. Darwinian selection in 1860

Eventually, boiled down to two schools of thought

biometricians headed by Karl Pearson
minute adjustments or micromutations cause adaptation

Mendelians led by William Bateson
discrete and heritable characters lead to phenotypic changes

But how could continuous variation be explained with discrete characters?
What happens in a population, at equilibrium?

Math to the Rescue

1930 Ronald A. Fisher, *The Genetical Theory of Natural Selection*

synthesized biometric and Mendelian views of mechanism

geometric model of adaptation

- each character of an organism = axis in Cartesian coord. system
- optimal combination of trait values = origin
- population no longer at optimum due to recent environmental change

populations must adapt by using mutations that are **random** with respect to the needs of organisms

- random in phenotypic direction, pointing away from the optimum at least as often as towards

mutations have different phenotypic **sizes**

- some mutations are vectors of large magnitude and others are vectors of small magnitude

populations must adapt in the face of **pleiotropy**

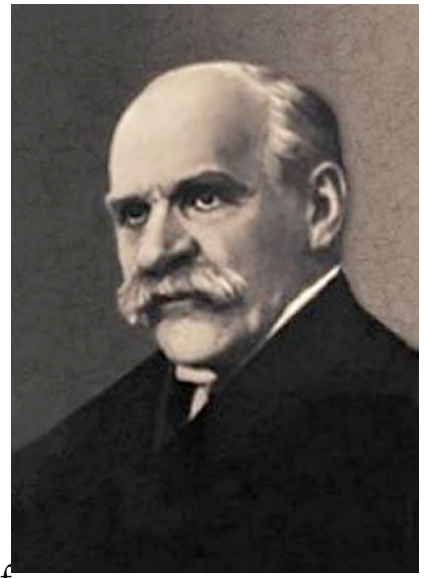
- mutation affects many characters and, although improving one character, might worsen many others

… and population genetics was born

Fisher, along with JBS Haldane and Sewall Wright, architected the modern evolutionary synthesis

Pop gen: study of the **distribution** of & **change** in allele frequencies under the influence of

natural selection, genetic drift, mutation, migration, non-random mating



1932 JBS Haldane, *The Causes of Evolution*

- reestablished natural selection as the premier mechanism of evolution by explaining it in terms of mathematical consequences of Mendelian genetics
- popular science writer (wonderful 1928 essay "On Being the Right Size"); friend of Aldous Huxley; coined the word "clone"

1931 Sewall Wright, *Evolution in Mendelian populations*

- theory of genetic drift (aka "Sewall Wright effect")
- inbreeding coefficient and experiments with guinea pigs
- adaptive surfaces (fitness landscapes)

Dynamics of genetic change

What happens to allele frequencies in a **population**, and at **equilibrium**?

Number theorist GH Hardy played cricket with geneticist R Punnett

1908 Hardy-Weinberg principle

After one generation of random mating, genotype frequencies at a single gene locus will become fixed at a particular equilibrium value

Aa pair of Mendelian characters, A dominant in any given generation

AA:Aa:aa = **p**:2**q**:**r**

p,**q**,**r** are fairly large (random mating)

sexes even distributed among phenotypes

all are equally fertile

Then in next generation, frequencies will be

$(\mathbf{p}+\mathbf{q})^2 : 2(\mathbf{p}+\mathbf{q})(\mathbf{q}+\mathbf{r}) : (\mathbf{q}+\mathbf{r})^2$, or $\mathbf{p'}:2\mathbf{q'}:\mathbf{r'}$

This distribution same as that in the generation before if $\mathbf{q}^2 = \mathbf{pr}$.

But $\mathbf{q'}^2 = \mathbf{p'r'}$… so whatever the values of **p**, **q**, and **r** distribution continues unchanged after the second generation

Even back then, biologists didn't like math

Hardy, Fisher, Haldane, Wright's work had limited impact among biologists

Formulated in very mathematical terms

Until re-interpreted with empirical evidence

1937 T. Dobzhansky *Genetics and the Origin of Species*

student of famous fruit-fly geneticist TH Morgan

zoologists Ernst Mayr, Julian Huxley (brother of Aldous, grandson of TH)

paleontologist George Simpson

botanist G. Ledyard Stebbins

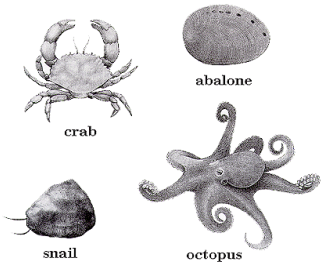
Evolution interdisciplinary

paleontology

large-scale evolution, succession of new dominant types

variations of evolutionary tempo

relationship of evolutionary change to evolutionary opportunity



taxonomy

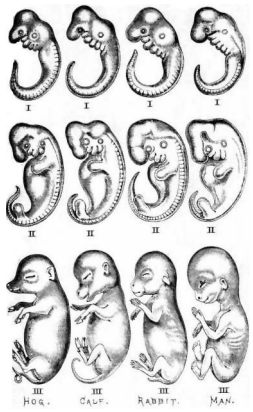
process of species-formation

differences in groups of plants & animals

ecology & comparative physiology, field studies

adaptation, interactions between species & environment

comparative fitness



embryology

ontogeny really does recapitulate phylogeny

evolution must take into account how processes of development are canalized

To this mix was added genetics

1910 TH Morgan: X-linked genes coding for fruit fly eye color

genes reside on specific chromosomes, and occupy specific locations on the chromosome

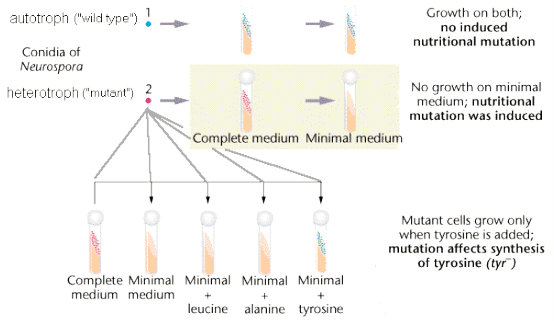
1928 Frederick Griffith: genes could be transferred

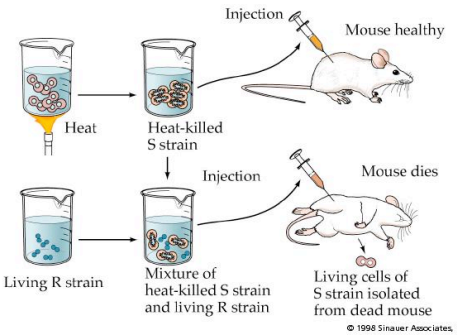
1941 GW Beadle and EL Tatum: mutations in genes caused errors in certain steps in metabolic pathways

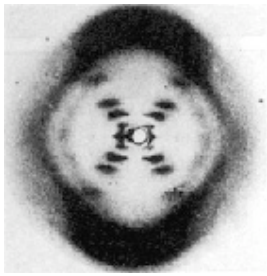
specific genes code for specific proteins => "one gene, one enzyme"

1944 O. Avery, C. Macleod, and M. McCarty: DNA holds the gene's information

1953 JD Watson and F. Crick, using data from M. Wilkins and R. Franklin: molecular structure of DNA







The Molecular Revolution

In studying phenotypes, switch from morphology to sequences

1952 Fred Sanger (2 Nobels)

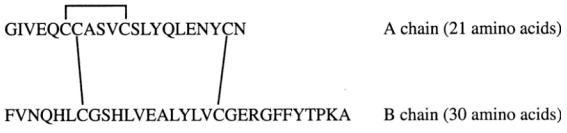
sequenced and compared insulin from cattle, sheep, horse, pig, sperm whale

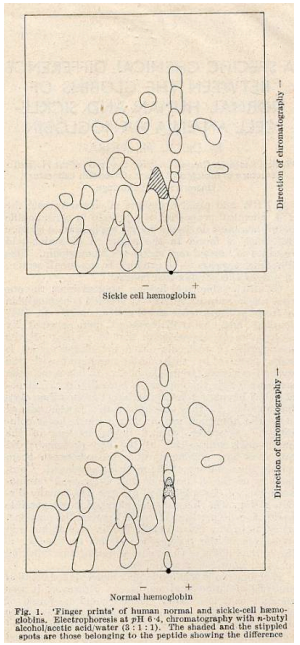
1961 Vernon Ingram

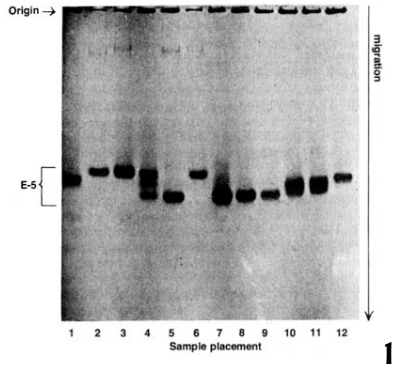
molecular geneology of vertebrate hemoglobin gene family: successive duplications and deletions of single globin gene

1966 Richard Lewontin (student of Dobzhansky) with JL Hubby

electrophoretic variants of Drosophila in enzymes







A closer look at selection

In **1930**, Fisher had asked

what is the probability that a random mutation of a given phenotypic size will be beneficial?

and concluded

very small mutations are the genetic basis of adaptation

influential, but flawed

- micromutationism not really plausible
- cannot study adaptation via inifinitesimally-based quantitative genetics

1968 Motoo Kimura, *Neutral Theory of Molecular Evolution*

- mutations must be more than beneficial!
- must also escape accidental loss when rare
- mutations of larger effect are more likely to escape such loss
- mutations of *intermediate* size are most likely to contribute to adaptation
- size distribution of mutations that are substituted over entire bouts of adaptation is *nearly exponential*
- few mutations of large effect; many mutations of little (neutral) effect

And molecular evolution was born

DNA as "evolutionary document" [Zuckerkandl & Pauling, 1965]

DNA as "molecular clock" [Kimura, 1968]

- mutations occur at constant rates over evolutionary time
- most are "neutral", i.e. don't affect function
- enable us to infer evolutionary history [Woese, 1987]

Find homologous genes [BLAST]

Align DNA/protein sequences [ClustalW]

Construct tree [Phylip]

distance-based methods

- Fitch & Margolias [FITCH, KITSCH]
- neighbor joining [NEIGHBOR]

optimization-based methods

- parsimony [DNAPARS, PROTPARS]
- maximum likelihood [DNAML, PROTML]

Plot tree [DRAWTREE, DRAWGRAM]

Tree Construction: Example

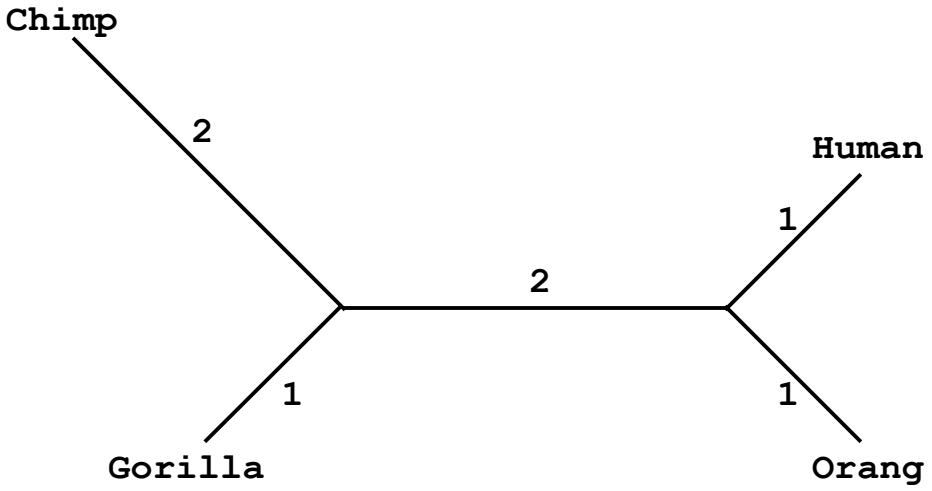
Aligned DNA Sequences

Chimp	A A T T T A G
Gorilla	A A A A A T G
Human	T T A T T A G
Orang	A A A A A A C

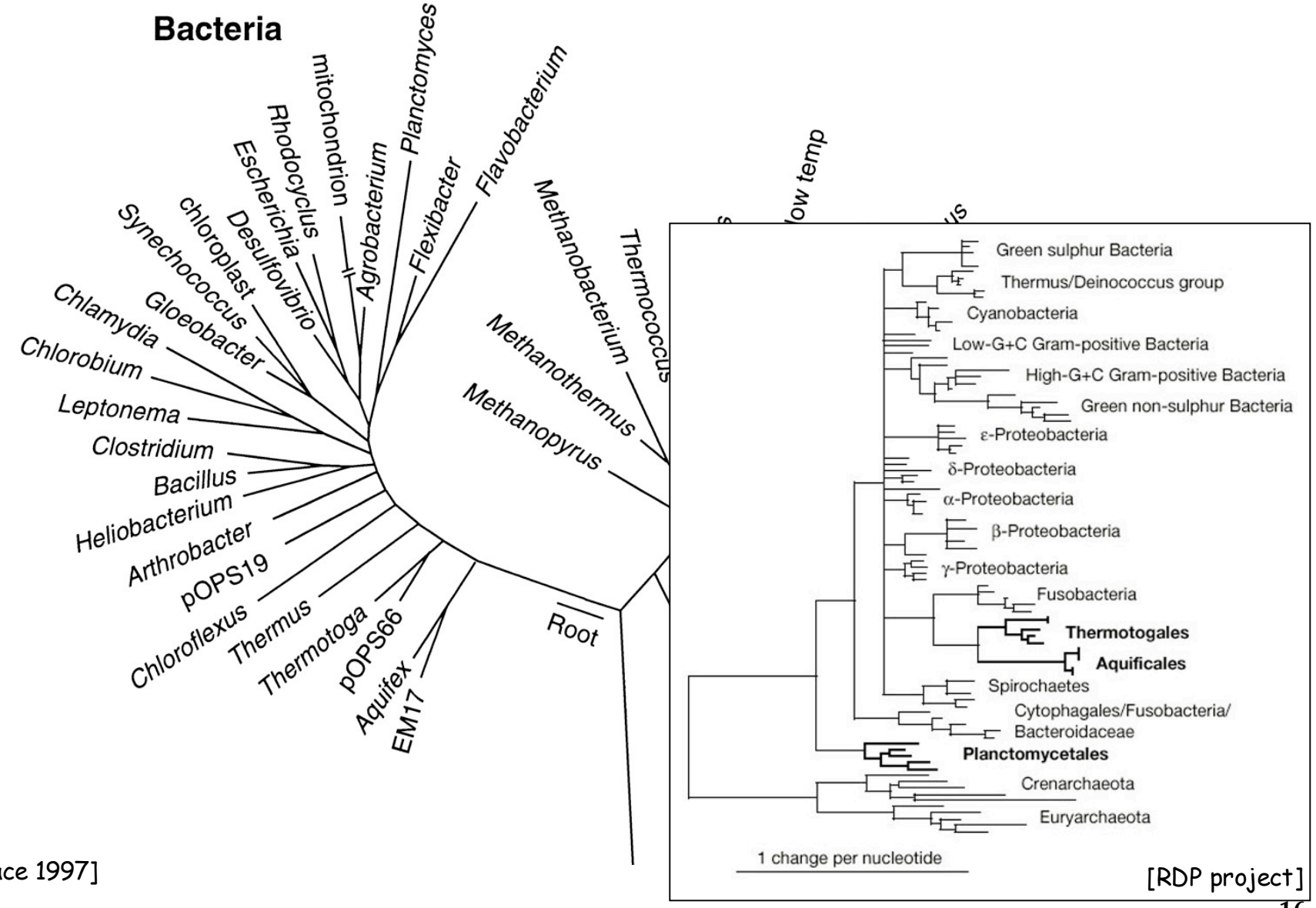
Pairwise Distances

C	0			
G	3	0		
H	5	4	0	
O	5	4	2	0
	C	**G**	**H**	**O**

Tree



rRNA Phylogeny



[Pace 1997]

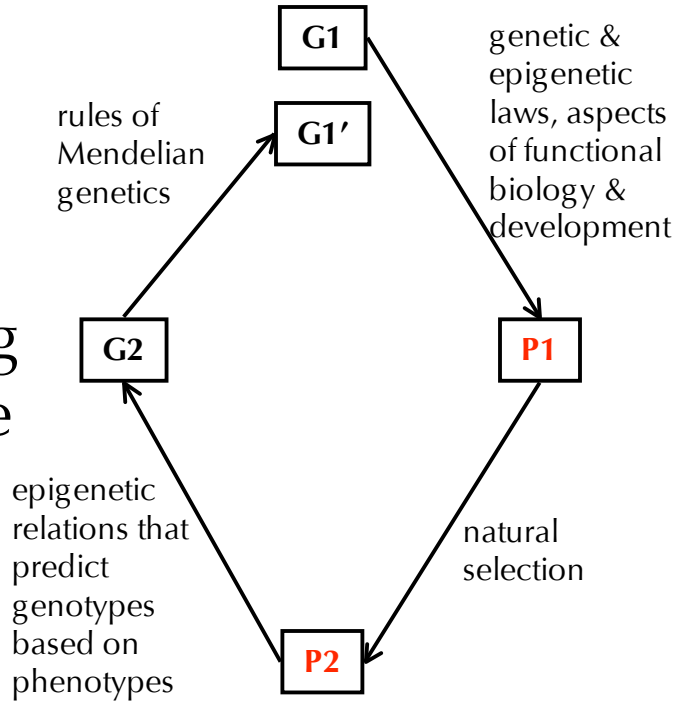
[RDP project]

Phenotype - genotype maps

Lewontin (1974) outlined the theoretical task for population genetics: to provide

a set of laws that predictably map a population of genotypes (G1) to a phenotype space (P1), where selection takes place

another set of laws that map the resulting population (P2) back to genotype space (G2) where Mendelian genetics can predict the next generation of genotypes, thus completing the cycle



Tests of selection and neutrality
dN/dS test

Silent (synonymous) substitution: "steady background noise," random

```
UUU CAU CGU
UUU CAC CGU
Phe His Arg
```

Coding (non-synonymous) substitution: may affect phenotype

```
UUU CAU CGU
UUU CAG CGU
Phe His Arg
    Gln
```

dN/dS: proportion of coding to silent subsitutions - test of how selection is acting
takes into account of transition/transversion rate bias and codon usage bias

<1 more silent than coding mutations, "positive selection"
gene is under strong selective constraints not to evolve too rapidly
mutations very deleterious to function
e.g. major histocompatibility complex in humans (Hughes and Nei 1988)

=1 roughly equal numbers of silent and coding mutations
"neutral evolution"
most mutations have no effect on fitness

>1 more coding than silent mutations
"negative selection" or "purifying selection"
gene evolving rapidly, possibly diverged in function in the gene family under study
mutation does not necessarily reduce fitness

Tests of selection and neutrality (cont'd)

McDonald Kreitman test

two samples of sequences in populations of different species, measure polymorphism between sequences within species and between species

	w/in	b/w
Syn	**a**	**b**
Non-syn	**c**	**d**

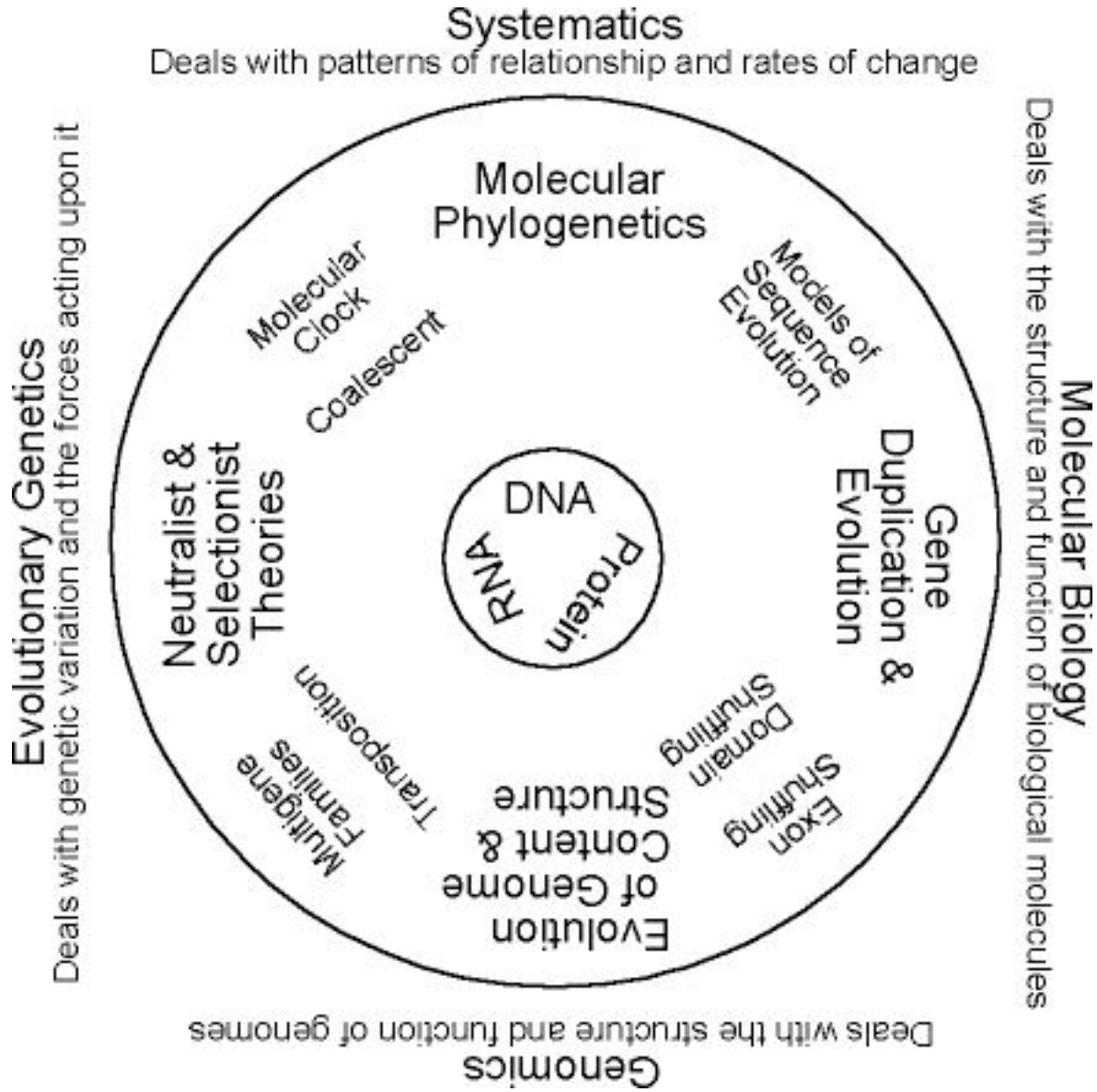
$$X^2 = n(ad-bc)^2 /[(a+b)(a+c)(b+d)(c+d)]$$

where $n=a+b+c+d$ is the total number of polymorphic sites

When n is not small, X^2 follows approx Chi-square distribution with one degree of freedom

e.g. alcohol dehydrogenase Adh gene from *D. simulans* and *D. yakuba*: more synonymous substitutions between species than expected (McDonald and Kreitman 1991)

Recap


Systematics
Deals with patterns of relationship and rates of change
Molecular Biology
Deals with the structure and function of biological molecules
Genomics
Deals with the structure and function of genomes
Evolutionary Genetics
Deals with genetic variation and the forces acting upon it
Molecular Phylogenetics
Molecular Clock
Coalescent
Models of Sequence Evolution
Gene Duplication & Evolution
Domain Shuffling
Exon Shuffling
Evolution of Genome Content & Structure
Transposition
Multigene Families
Neutralist & Selectionist Theories
DNA
RNA
Protein

Thank you!