Population Genetics: Its Mathematical Principles and Some Recent Trends

Tomoko Ohta, Ph.D.
National Institute of Genetics
Mishima, Japan

Population Genetics

- Darwin + Mendel
- Population consequences of Mendelian Inheritance worked out in 1930s

R. A. Fisher, J. B. S. Haldane and S. Wright

Change of gene frequencies in populations by natural selection, random genetic drift, mutation and migration

Change in gene frequency

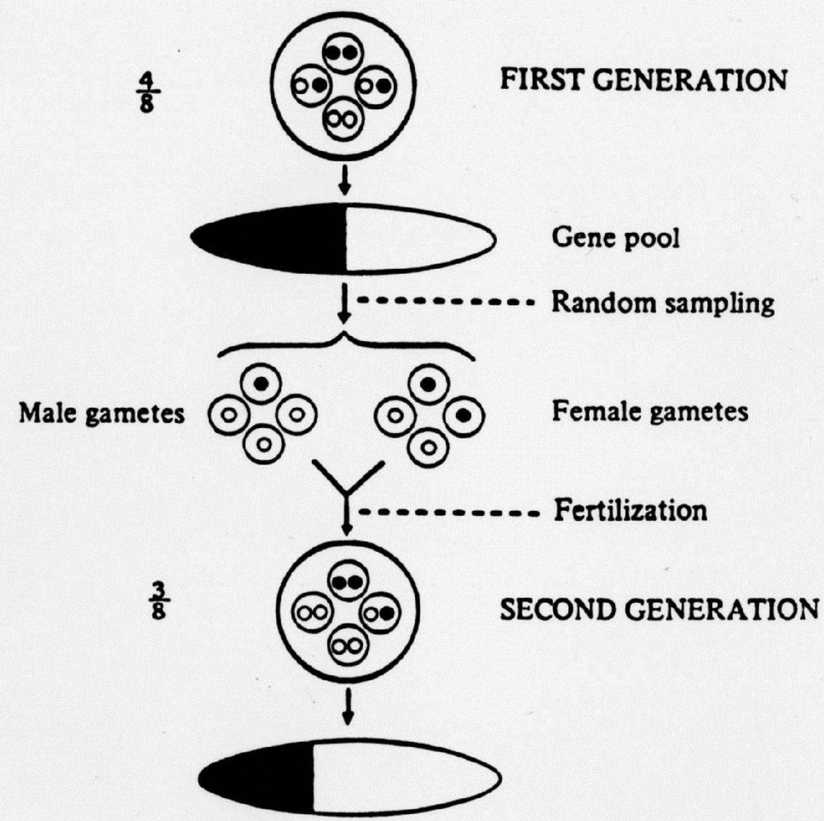
Natural selection

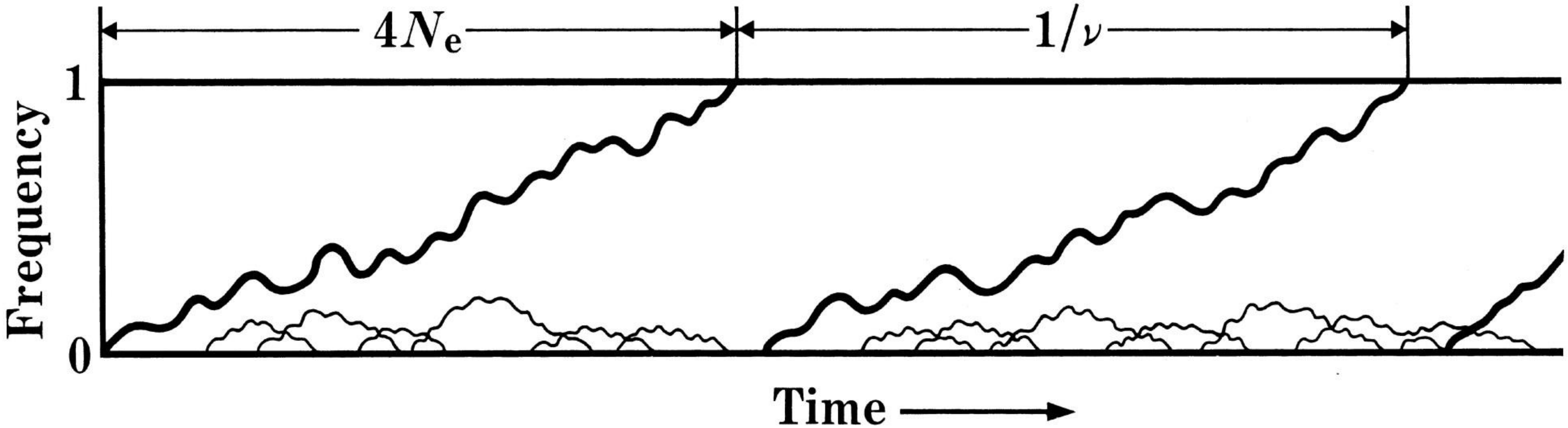
- Deterministic: Necessity
  - selection coefficient: s

Genetic drift

- Erratic movement of gene frequency
- Stochastic: Chance
  - Mainly by random sampling of gametes at reproduction
  - Effective population size: N

Random sampling




4N_e
1/ν
1
0
Frequency
Time

Beginning of Population Genetics + Molecular Evolution

- 1960s
- Application of data on genetic variation and evolution to population genetic models

- Stochastic processes becoming more applicable.

1960s

Kimura: Population genetics + Molecular biology

Zuckerkandl-Pauling: Molecular clock (1965)

Kimura: Neutral theory (1968)

King-Jukes: Non-Darwinian evolution (1969)

Ohno: Evolution by gene duplication (1970)

Protein polymorphisms by electrophoresis

Lewontin-Hubby 1966

Kimura-Crow model 1964

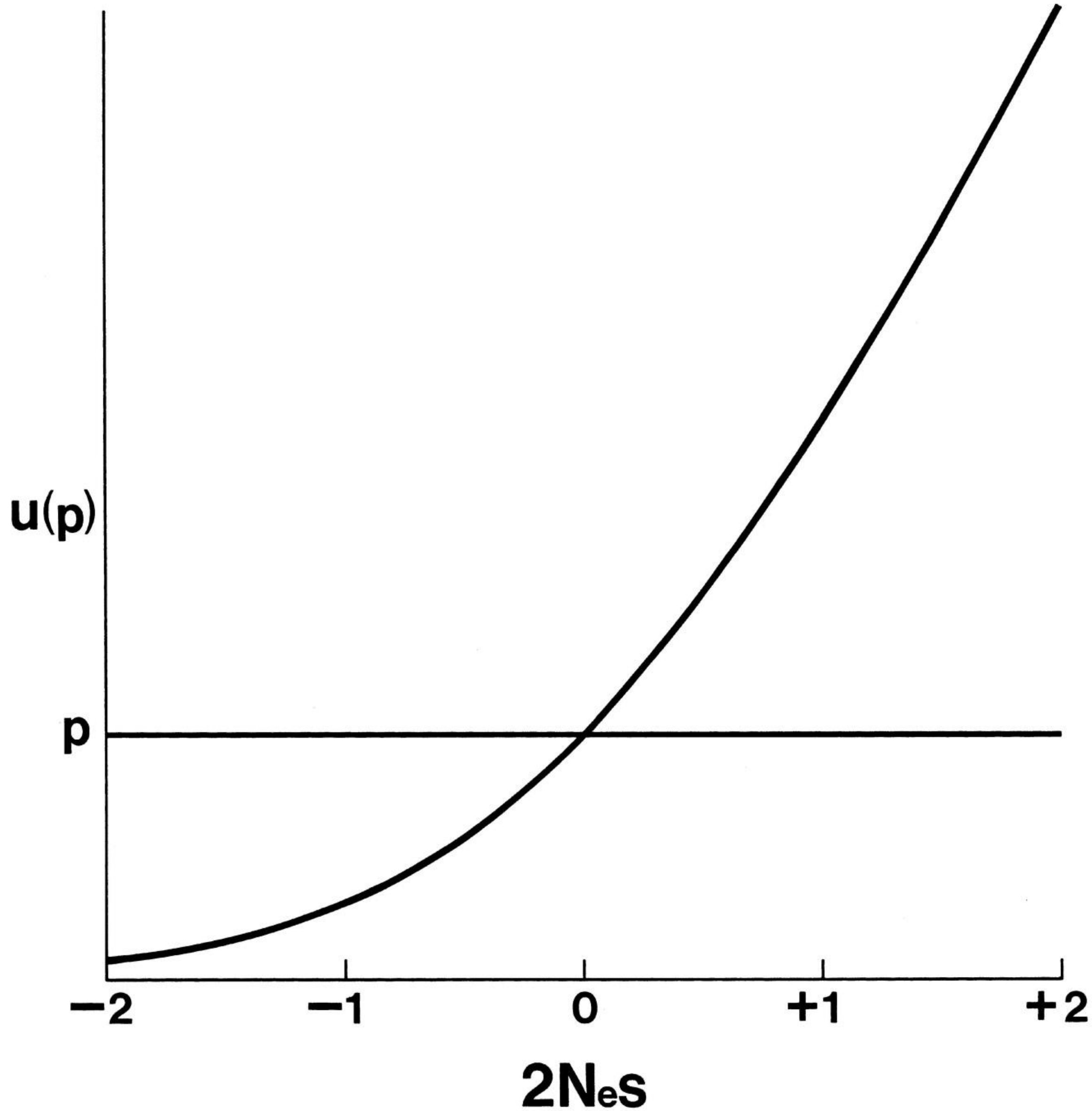
Infinite-site model 1969

Step-mutation model 1973

Three Problems

- What are borderline mutations between the selected and the neutral?
- Molecular clock: year or generation-time dependent?
- Narrow range of heterozygosity?

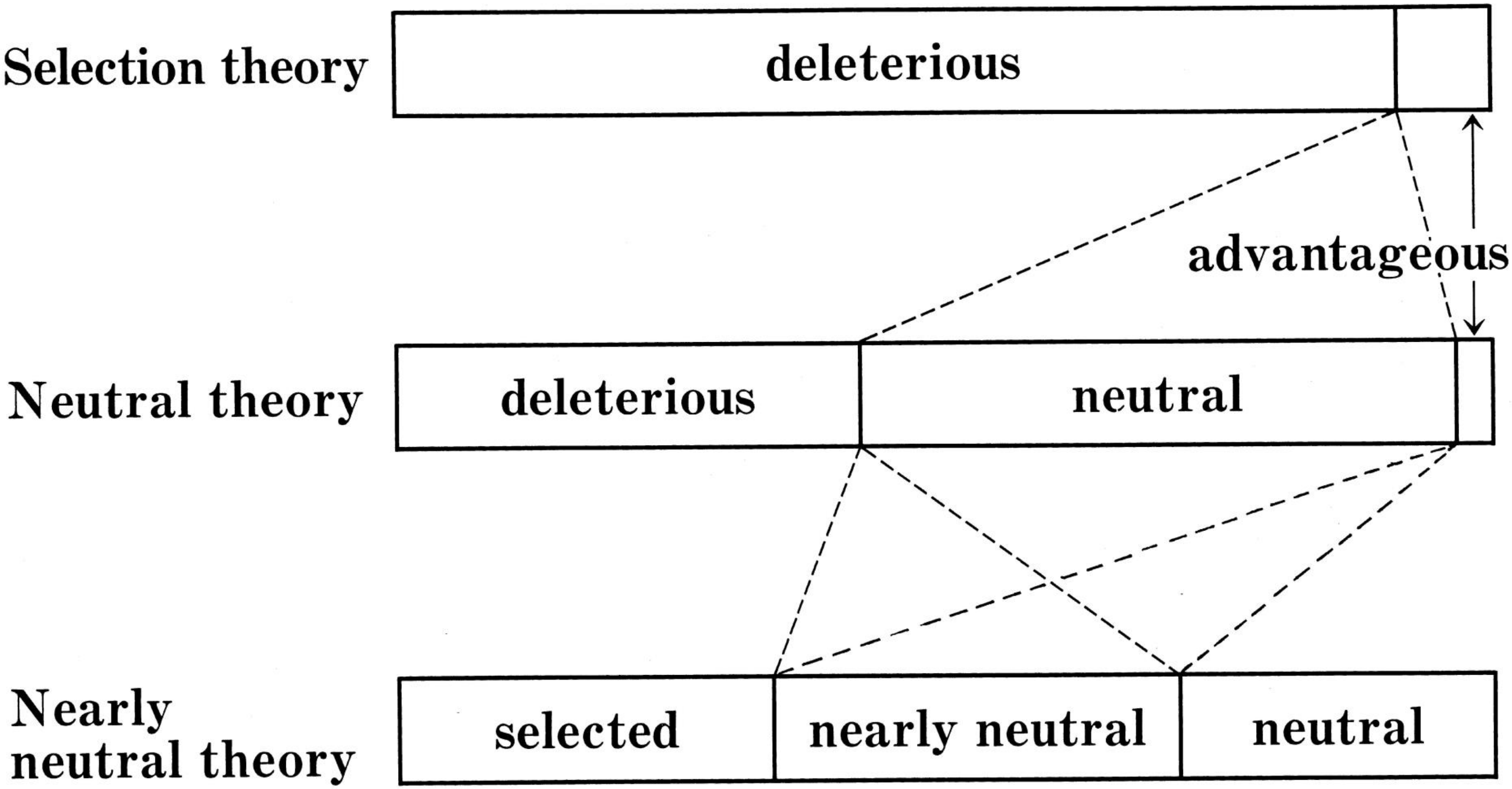
Ohta: Slightly deleterious mutation theory 1973



u(p)
p
−2
−1
0
+1
+2
2NeS

Step-mutation model

- Rare Common Rare
- Slightly deleterious mutation
- Used in probability theory by JFC Kingman


Selection theory
deleterious
advantageous
Neutral theory
deleterious
neutral
Nearly
neutral theory
selected
nearly neutral
neutral

- Prediction of the Neutral Theory

Evolutionary rate = Neutral mutation rate

- Prediction of the Nearly Neutral Theory

Negative correlation between the population size and evolutionary rate

Early support of rapid evolution of small populations

- Endosymbiotic bacteria

  Moran 1996

- Island species of birds

  Johnson and Seger 2001

  Woolfit and Bronham 2005

dN/dS

Human > Pig > Mouse

Jorgensen et al. 2005

Human > Dog > Mouse

Lindblad-Toh et al. 2005

Hominid > Rodent

The Chimpanzee Sequencing Consortium 2005

Yeast laboratory strain > Yeast wild strain

Gu et al. 2005

Polymorphism

Human SNPs : Abundant rare alleles

Bacteria 84 species Tajima's D

Dnon 68% negative
Dsyn 70% positive

Hughes 2005

Protein Folding and Proteostasis

- Aberrant folding and aggregation - Toxic

  Peak and valley of free-energy landscape

- Chaperones - Dynamics

  F U Hartl et al. 2011

Intrinsically Disordered Protein

- Signaling and regulatory functions:
  tightly regulated but provide versatility
  via post-translational modification
  Babu et al. 2011

Examples:

Histone tail for chromatin function

Tissue specific splicing disordered segment
for protein interaction network

Evolution of Complex Gene Regulatory Systems

Phenotype, especially morphology, changes:

mainly coming from gene regulatory changes

Gene expression: Time and Space

Expression quantity

Robustness of gene regulatory network

Gene redundancy

Gene degeneracy

J. M. Whitacre

Distributed robustness: flexible regulatory network

A. Wagner

Robustness→ Increase of nearly neutral mutations

- Ratio of gene expression divergence between species to gene expression diversity within species
  - About equal in the brain, heart, kidney, liver but three fold higher in the testes

Brain: Ratio of the change of the human lineage to that of chimpanzee is larger than the same ratio in the liver or heart

Khaitovich et al. 2006

Weak stabilizing selection on gene expression

- Divergence of gene expression among Drosophila species initially increases with time, but eventually saturates.
- The strength of stabilizing selection is estimated to be very weak, i. e., nearly neutral.

Bedford- Hartl, 2009

The ENCODE project

- Many transcripts with unknown function and/or without constraint exist in human cells

- Neutral or nearly neutral

- Provide opportunities for new regulatory systems

Genome Accessibility

- Nucleosome packaging and positioning

    DNA wrapped around a octamer: H3,H4,H2A,H2B

    Linker histone between nucleosomes: H1

- DNA sequence and chromatin remodeling

O. Bell et al. 2011

Linker Histone Variability

- H1 family: subtypes
- differential expression, mobile,

  non-random distribution on chromatin regions

  Polymorphic: physiological traits

  Intrinsically disordered region for signaling system

AS. Kowalsky and J. Palyga   22012

Chromatin Modifying Enzymes

readers, writers, erasers

histone tails (intrinsically disordered)

acetylation, methylaton, phosphorilaton etc

histone mimick sequences of nuclear proteins

Digital to analog transformation

A. I. Badeaux and Y. Shi 2013

Protein kinases

Phosphorylation sites in orthologous proteins of different species, not conservative

human vs mouse: 12% of phosphoserine

15% of phosphothreonine

Lienhard, 2008

~500 protein kinases in human

10%: pseudokinases

Taylor and Kornev 2010

Near-neutrality
Genotype

Epigenetics
Chromatin structure

DNA methylation and others

Stochastic variation of gene expression → ← Modification of gene network

↑
environment

Robustness
Extra-activity of genome

Cooption in gene regulation

Phenotype
Darwinian selection

Environment -> Gene Expression

Ligand { endogenous / exogenous } Receptor

steroid — nuclear-receptor superfamily

retinoid

vitamin D — flexible

thyroid hormone

hormone response elements: numerous and often weak-binding

Ligand-dependent transcription

Evans and Mangelsdorf 2014

Attractors and Democratic Dynamics in Gene Regulation

- Autocratic: master gene
- Intermediate: common in real regulation
- Democratic: all genes interact

attractors, self-organization

Bar-Yam t al. 2009

Evolution of complex systems

- Epigenetics
- Genetic Drift
- Natural selection

Work together