Molecular evolution: traditional tests of neutrality

ppt credit

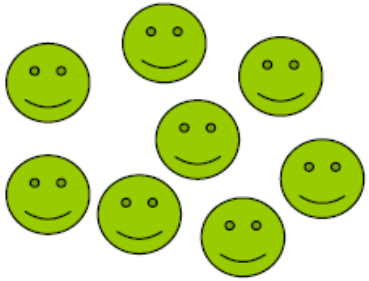
Dr. Daniel Neafsey
Research Scientist, Broad Institute

Mutation+Selection=Evolution

Relative importance of each for maintaining variation in population?

Early Criticism of Darwin

Blending inheritance, ‘gemmules’



Fleeming Jenkin (1867):

$Var[X(t+1)] = \frac{1}{2}\, Var[X(t)]$

Mendelian Inheritance

published 1865-66, rediscovered 1900

Law of Segregation:

- allelic variation
- offspring receive 1 allele from each parent
- dominance/recessivity
- parental alleles ‘segregate’ to form gametes

Law of Independent Assortment

Simple case: no selection

The Hardy-Weinberg Law (1908)

Requires:

- infinite population size
- random mating
- non-overlapping generations
- no selection, mutation, or migration

The Hardy-Weinberg Law

Genotype:	*AA*	*Aa*	*aa*
Frequency at time 0:	u_0	v_0	w_0

$$u_0 + v_0 + w_0 = 1$$

frequency of A (p_0) = $u_0 + v_0/2$

frequency of a (q_0) = $w_0 + v_0/2$

$$p_0 + q_0 = 1$$

The Hardy-Weinberg Law

Genotype:	*AA*	*Aa*	*aa*
Frequency at time 0:	u_0	v_0	w_0

Mating Pair	Frequency	Offspring		
		AA	Aa	aa
AA x AA	u_0^2	1	0	0
AA x Aa	u_0v_0	½	½	0
Aa x AA	u_0v_0	½	½	0
Aa x Aa	v_0^2	¼	½	¼

Frequency of AA in next generation: $u_1 = u_0^2 + u_0v_0 + 1/4\ v_0^2$

$= (u_0 + v_0/2)^2$

$= p_0^2$

The Hardy-Weinberg Law

If assumptions met:

- allele frequencies don't change
- after a single generation of random mating, genotype frequencies are:

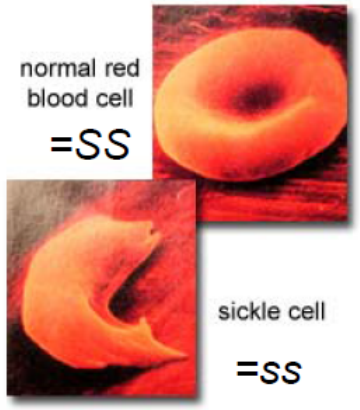
$u = p^2$ $v = 2pq$ $w = q^2$

- entire system characterized by one parameter (p)

Deviation from expectations indicates failure of 1 or more assumptions—selection?

HW application: Sickle cell anemia

	Observed Counts	Expected Counts
SS	834	
Ss	161	$2pq$ *1000= 129
ss	5	



$p = \sqrt{0.834} = 0.91$

$q = \sqrt{0.005} = 0.071$

Approach: Detect selection through comparison to neutral expectation

Kimura: neutral theory

Ewens: sampling formula

Coalescence

Neutral Theory History

- Motoo Kimura (1924-1994)
- 1968: a large proportion of genetic change is not driven by selection
- Adapted diffusion approximations to genetics
- Dealt with finite pops

WIKIPEDIA
The Free Encyclopedia

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia
Wikimedia Shop

Interaction
Help
About Wikipedia
Community portal
Recent changes
Contact page

Tools
What links here
Related changes
Upload file
Special pages
Permanent link
Page information
Wikidata item
Cite this page

Print/export
Create a book
Download as PDF
Printable version

Languages
Català
Deutsch
Español
فارسی
Français
Galego
한국어
Italiano
Қазақша

Motoo Kimura

From Wikipedia, the free encyclopedia

Motoo Kimura (木村 資生 *Kimura Motoo*, November 13, 1924 – November 13, 1994) was a Japanese biologist best known for introducing the neutral theory of molecular evolution in 1968.[2][3] He became one of the most influential theoretical population geneticists. He is remembered in genetics for his innovative use of diffusion equations to calculate the probability of fixation of beneficial, deleterious, or neutral alleles.[4] Combining theoretical population genetics with molecular evolution data, he also developed the neutral theory of molecular evolution in which genetic drift is the main force changing allele frequencies.[5] James F. Crow, himself a renowned population geneticist, considered Kimura to be one of the two greatest evolutionary geneticists, along with Gustave Malécot, after the great trio of the modern synthesis (Haldane, Wright, Fisher).[6]

Contents [hide]
1 Life and work
2 Honors
3 See also
4 References

Motoo Kimura

Born	November 13, 1924 Okazaki, Japan
Died	November 13, 1994
Alma mater	University of Wisconsin
Doctoral advisor	James F. Crow
Other academic advisors	Sewall Wright, Hitoshi Kihara
Known for	development of neutral theory of molecular evolution, and contributions to evolutionary biology, population genetics
Notable awards	Asahi Prize (1986) John J. Carty Award (1987) International Prize for Biology (1988) Darwin Medal (1992) Fellow of the Royal Society[1]

Life and work [edit]

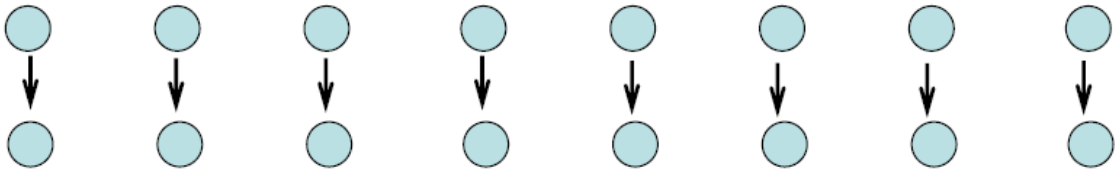
Kimura was born in Okazaki, Aichi Prefecture. From an early age he was very interested in botany, though he also excelled at mathematics (teaching himself geometry and other maths during a lengthy convalescence due to food poisoning). After entering a selective high school in Nagoya, Kimura focused on plant morphology and cytology; he worked in the laboratory of M. Kumazawa studying the chromosome structure of lilies. With Kumazawa, he also discovered how to connect his interests in botany and mathematics: biometry.[1]

Due to World War II, Kimura left high school early to enter Kyoto Imperial University in 1944. On the advice of the prominent geneticist Hitoshi Kihara, Kimura entered the botany program rather than cytology because the former, in the Faculty of Science rather than Agriculture, allowed him to avoid military duty. He joined Kihara's laboratory after the war, where he studied the introduction of foreign chromosomes into plants and learned the foundations of population genetics. In 1949, Kimura joined the National Institute of Genetics in Mishima, Shizuoka. In 1953 he published his first population genetics paper (which would eventually be very influential), describing a "stepping stone" model for population structure that could treat more complex patterns of immigration than Sewall Wright's earlier "island model". After meeting visiting American geneticist Duncan McDonald (part of Atomic Bomb Casualty Commission), Kimura arranged to enter graduate school at Iowa State College in summer 1953 to study with J. L. Lush.[1]

Kimura soon found Iowa State College too restricting; he moved to the University of Wisconsin to work on stochastic models with James F. Crow and join a strong intellectual community of like-minded geneticists, including Newton Morton and most significantly, Sewall Wright. Near the end of his graduate study, Kimura gave a paper at the 1955 Cold Spring Harbor Symposium; though few were able to understand it (both because of mathematical complexity and Kimura's English pronunciation) it received strong praise from Wright and later J.B.S. Haldane. His accomplishments at Wisconsin included a general model for genetic drift, which could accommodate multiple alleles, selection, migration, and mutations, as well as some work based on R.A. Fisher's fundamental theorem of natural selection. He also built on the work of Wright with the Fokker-Planck equation by introducing the Kolmogorov backward equation to population genetics, allowing the calculation of the probability of a gene to become fixed in a population. He received his PhD in 1956, before returning to Japan (where he would remain for the rest of his life, at the National Institute of Genetics).[1]

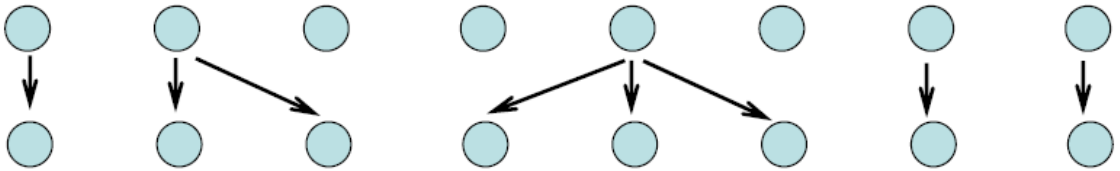
Kimura worked on a wide spectrum of theoretical population genetics problems, many of them in collaboration with Takeo Maruyama. He introduced the "infinite alleles" and "infinite sites" models for the study of genetic drift, both of which would be used widely as the field of molecular evolution grew alongside the number of available peptide and genetic sequences. He also created the "ladder model" that could be applied to electrophoresis studies where

Genetic Drift



no drift

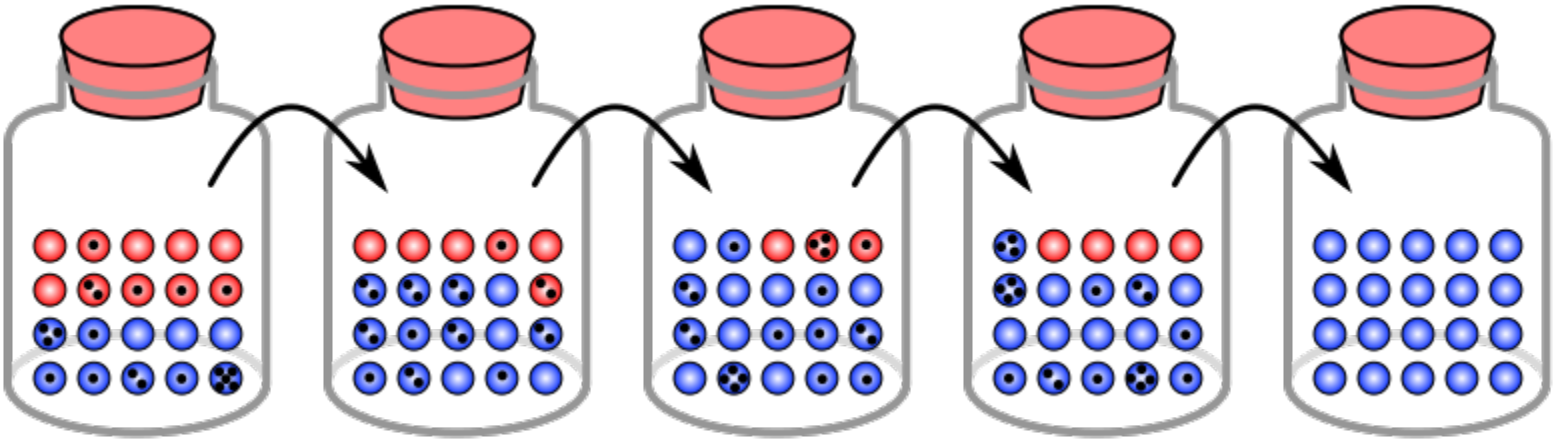
infinite pop

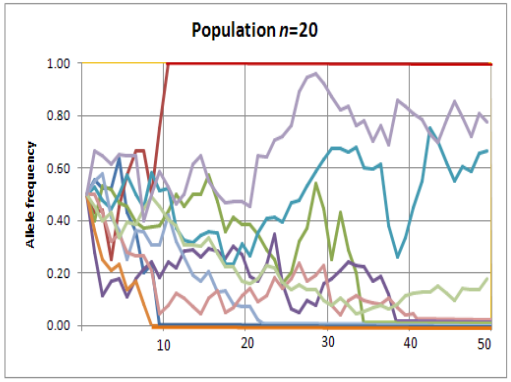


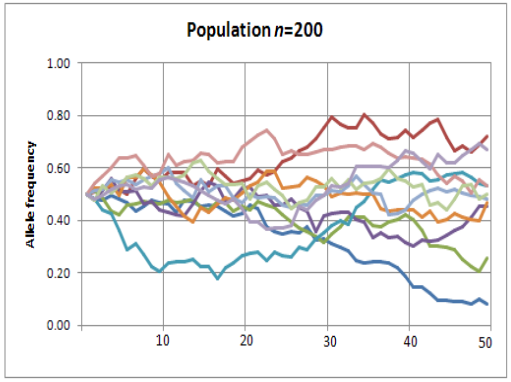
drift

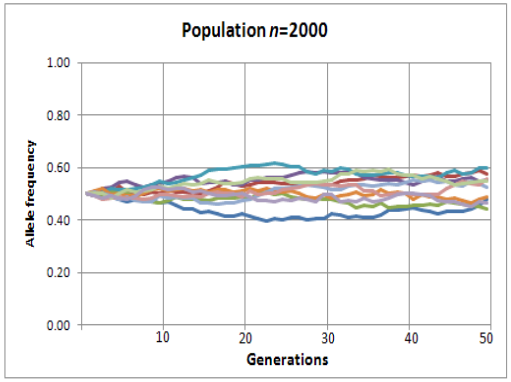
finite pop

Genetic drift




Population n=20
Allele frequency
1.00
0.80
0.60
0.40
0.20
0.00
10
20
30
40
50


Population n=200
Allele frequency
1.00
0.80
0.60
0.40
0.20
0.00
10
20
30
40
50


Population n=2000
Allele frequency
1.00
0.80
0.60
0.40
0.20
0.00
10
20
30
40
50
Generations

Test of the evolution model

Rate-based selection metric: d_N/d_S

d_N = no. nonsynonymous changes/ no. nonsynonymous sites

d_S = no. synonymous changes/ no. synonymous sites

Counting codon 'sites' example: CAT

Histidine is encoded by only one other codon: CAC

full nonsyn sites

fractional site

P(T⇒C) = fractional syn sites

P(T⇒G or A) = fractional nonsyn sites

Rate-based selection metric: d_N/d_S

$d_N/d_S < 1$ purifying selection

$d_N/d_S = 1$ neutral expectation

$d_N/d_S > 1$ positive selection

Rate-based selection metric: d_N/d_S

- Can be calculated using various methods
- Goldman & Yang implementation (PAML):

nucleotide changes modelled as continuous-time Markov chain with state space = 61 codons

$$q_{ij} = \begin{cases} 0: \text{ if the two codons differ at > 1 position} \\ \pi_j: \text{ synonymous transversion} \\ \kappa\pi_j: \text{ synonymous transition} \\ \omega\pi_j: \text{ nonsynoymous transversion} \\ \omega\kappa\pi_j: \text{ nonsynonymous transition} \end{cases}$$