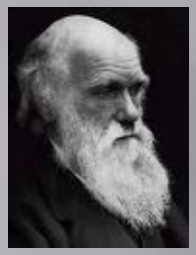
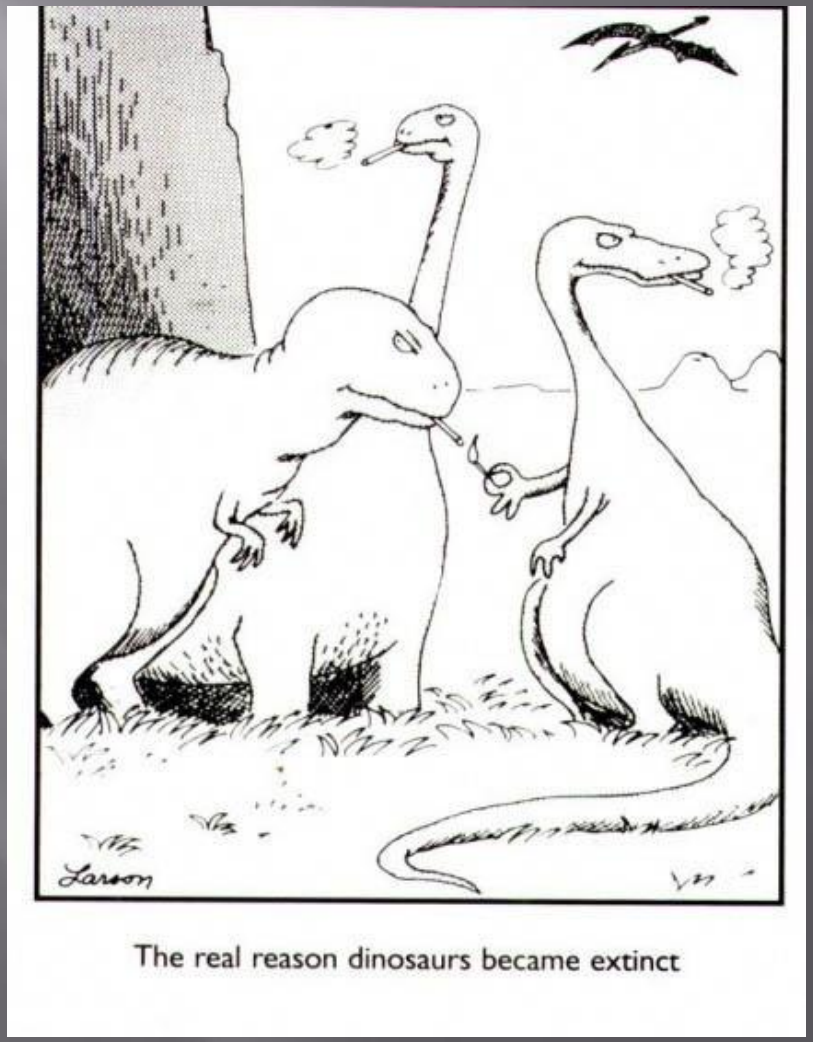
Neutral dynamics and family (type) statistics :





The real reason dinosaurs became extinct

Nadav Shnerb (BIU)

Yosi Maruvka (BIU)

David Kessler (BIU)

Gur Yaari (Yale)

Sorin Solomon (HUJI)

Robert Ricklefs (St. Louis)

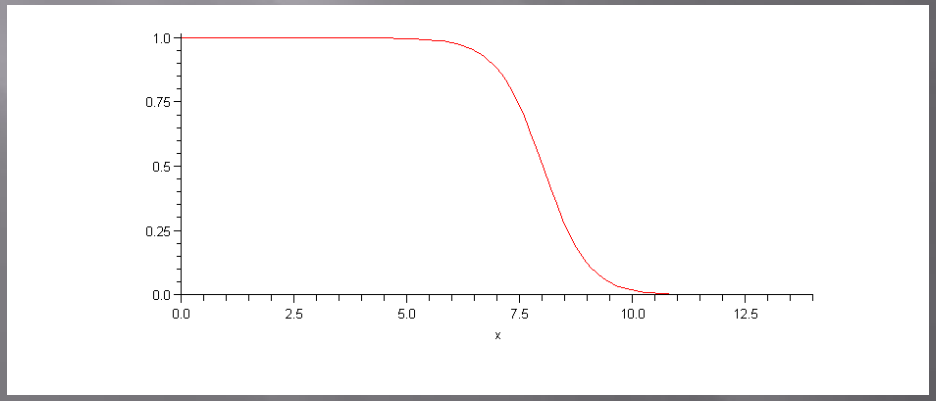
John Wakeley (Harvard)

PLoS computational biology 5(4): e1000359 (2009)

What is the difference between pants and glasses?

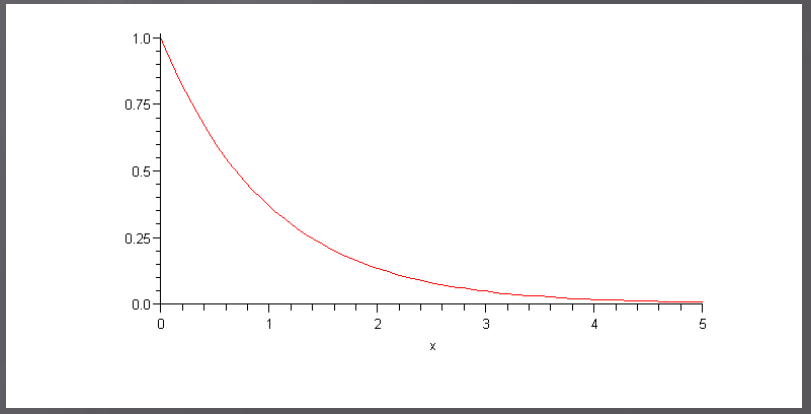
Survival of the fittest

Pants "die" (ripped) due to accumulated wear.



Survival of the luckiest

Glasses "die" when they shattered, usually due to random and uncorrelated events

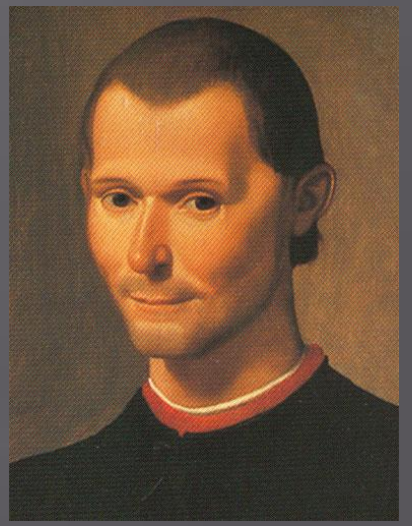


What is the signal to noise ratio in complex system?
Social dynamics:

The prince, Capitolo XXV

[What fortune can effect in human affairs and how to withstand her]

"Fortune is the mistress of one half our actions, and yet leaves the control of the other half, or a little less, to ourselves"



Deterministic (fitness)

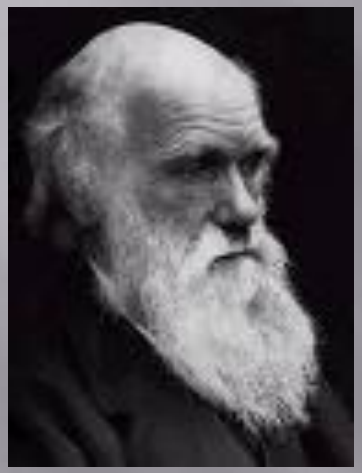
Stochastic (neutral)

it is the ultimate aim of this work, to lay bare the economic law of motion of modern society

Karl Marx, Das Capital (introduction)

Cleopatra's nose, had it been shorter, the whole face of the world would have been changed (Pascal)

Ecology and evolution: Neutral vs. adaptive (niche) theories



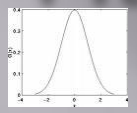
*"When we look at the plants and bushes clothing an entangled bank, we are **tempted** to attribute their proportional numbers and kinds to what we call **chance.** But how **false** a view is this!"*
Darwin On the Origin of Species (1859)

*Hubbell started from a postulate that most consider preposterous: that one tree or one bird is just like any other. His patterns result **solely from random fluctuations** in births, deaths and the arrival of new species.*
Scientific American April 29 2002.

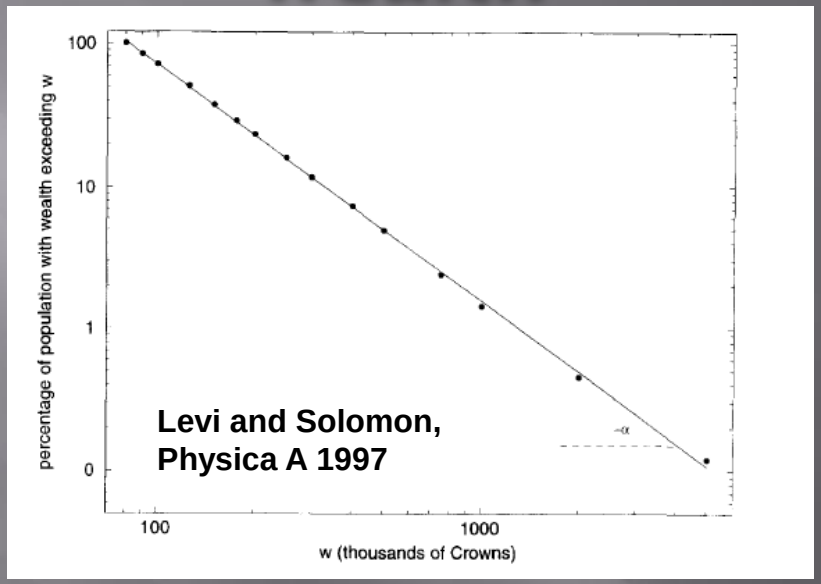
Deterministic (fitness) ⟷ Stochastic (neutral)

Plugging numbers: Finance and wealth

$$P(x) \approx e^{-\frac{x^2}{2\sigma^2}}$$



Neutrality=
Gaussian statistics



Pareto: Saying "everything is governed by chance" is a prediction. The central limit theorem guarantees that the resulting statistics will be Gaussian.

$$x_{t+1} = x_t + \eta_t$$
$$\eta \in [-1,1]$$

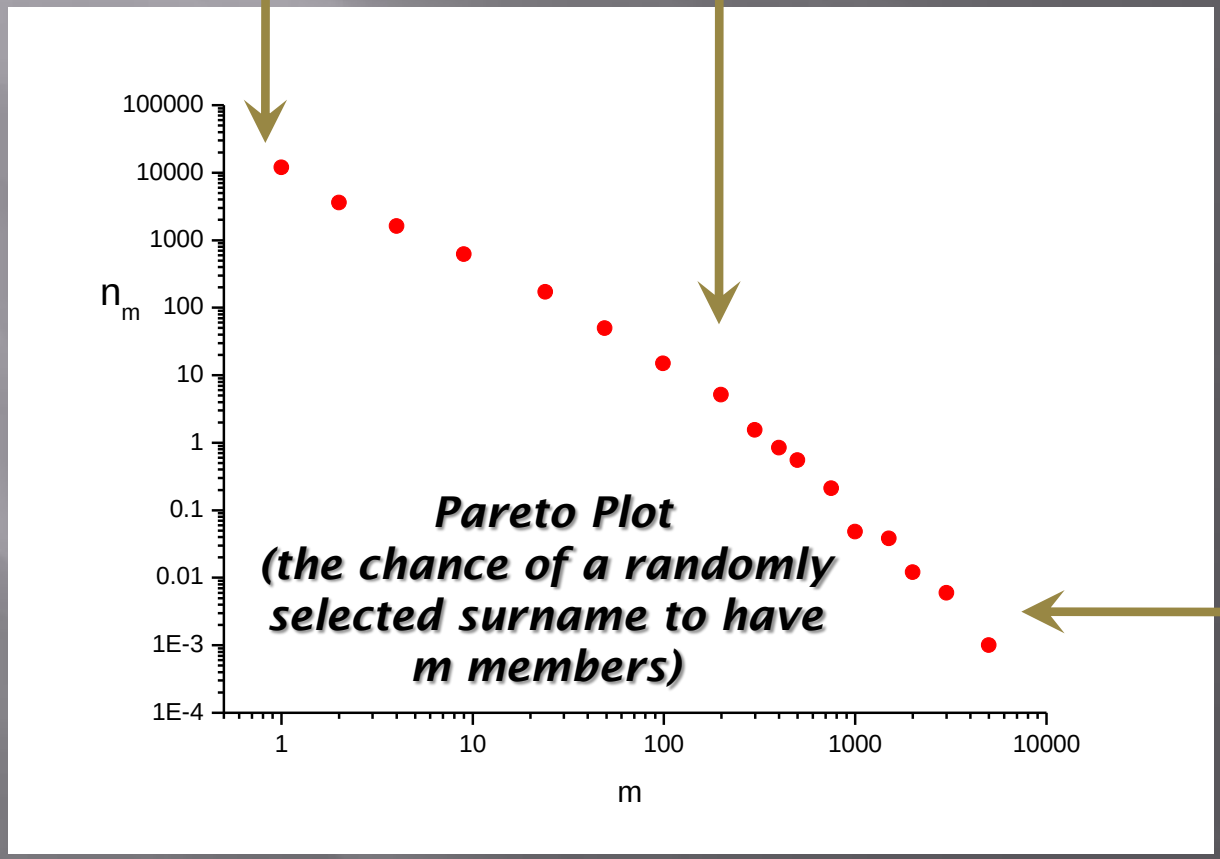
$$P(x) \approx e^{-\frac{x^2}{\sigma^2(t)}}$$

$$x_{t+1} = x_t \eta_t$$
$$\eta \in [1/2, 2]$$
$$\log(x_{t+1}) = \log(x_t) + \log(\eta_t)$$

$$P(x) \approx e^{-\frac{[\log(x)]^2}{\sigma^2(t)}} \approx x^{-\frac{\log(x)}{\sigma^2(t)}}$$

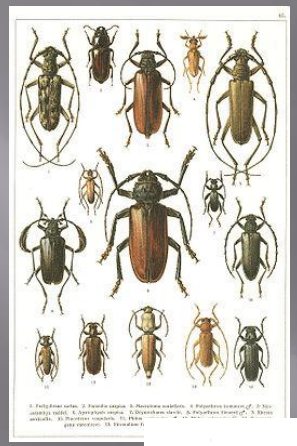
Deterministic (fitness) ⟷ Stochastic (neutral)

example for neutral dynamics: surnames



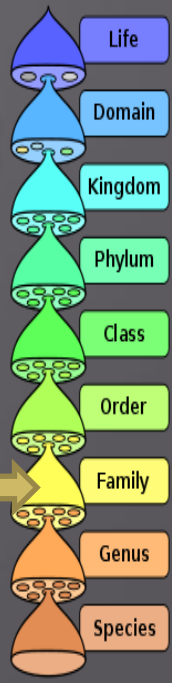
The underlying stochastic process: Yule 1925

a special preference for beetles … here they are:



Genus = surname
Species = first name

# of species in genus varies tremendously



$$n_m \approx m^{-\alpha}$$

$$\log(n_m) \approx -\alpha \log m$$

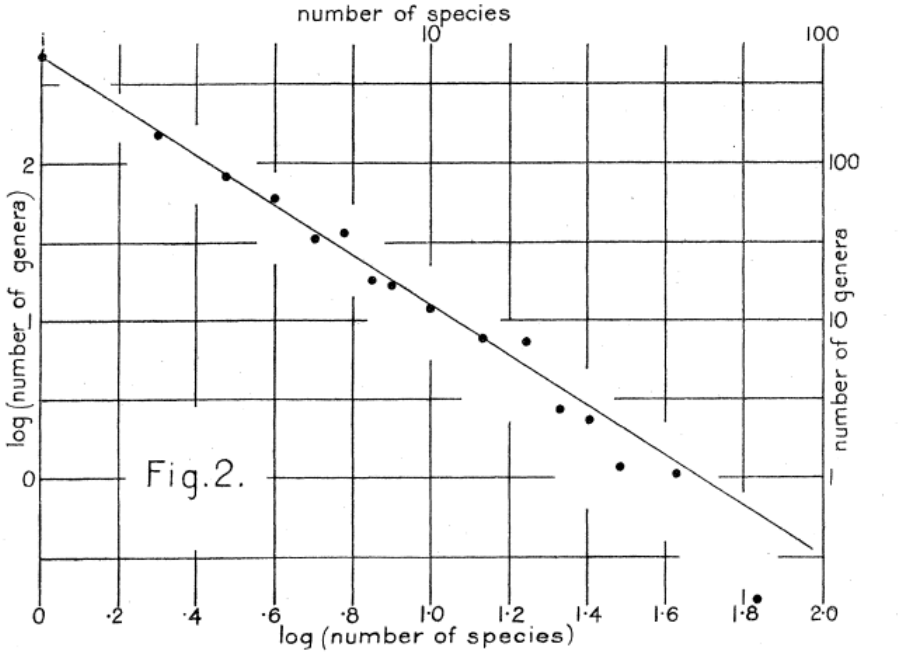


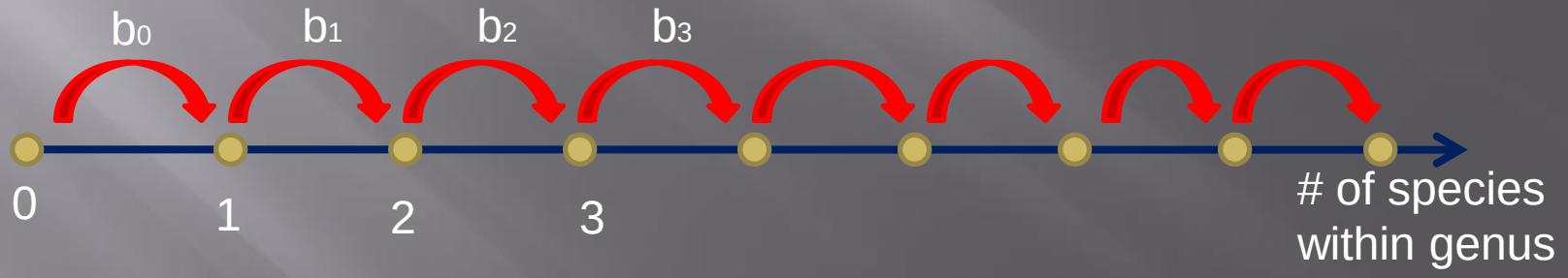
Fig. 2.—Double logarithmic chart for the frequency distribution of sizes of genera in the *Cerambycinæ*. Data in Appendix, Table B.

Modeling neutral processes :
The Yule process (preferential attachment)

No fitness, no selection, everybody are equal

1. ***Birth=speciation:*** *A species is chosen at random to reproduce = to undergo speciation (no fitness, selection, whatever).*

2. ***Mutation = speciation that creates new genus:*** *an offspring is identical to its father with probability $1-\mu$ and is a mutant (new species, surname, genus) with probability μ.*



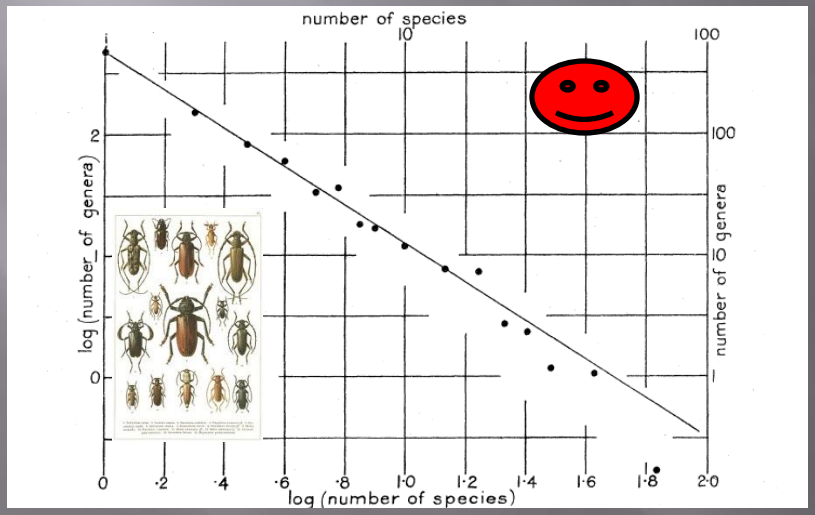
$$b_n = \gamma(1-\mu)n \quad (n \geq 1) \qquad b_0 = \mu\gamma \cdot total\ \ population$$

$$Where\ \ \gamma\ \ is\ \ the\ \ birth\ \ rate$$

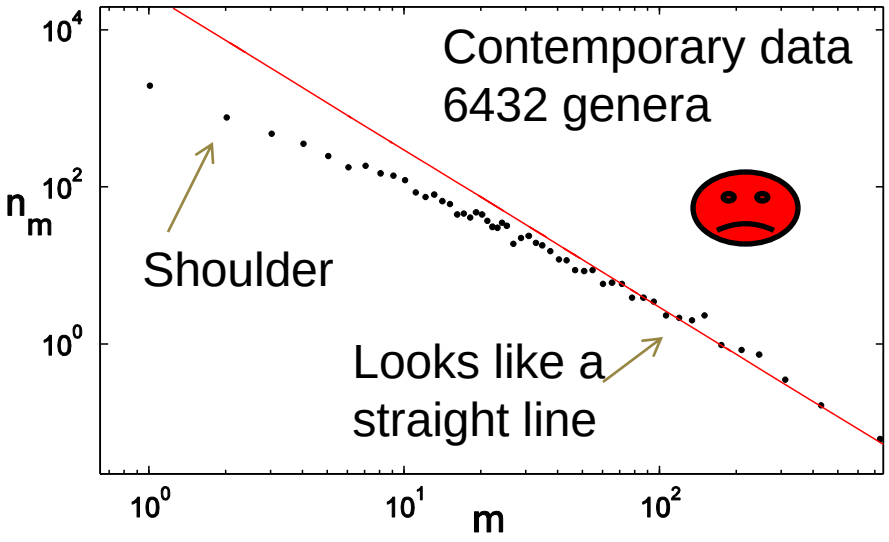
Yule-Simon steady state

$$n_m = \left(1+\frac{1}{\mu}\right)\frac{\Gamma(m)\Gamma(2+1/\mu)}{\Gamma(m+2+1/\mu)} = \left(1+\frac{1}{\mu}\right)B(m,2+1/\mu) \approx m^{-\left(2+\frac{1}{\mu}\right)}$$

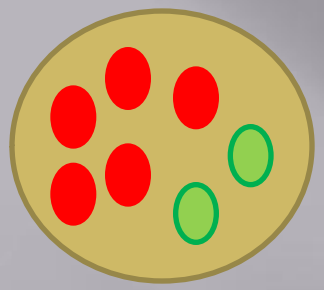
Original Yule 1024 genera



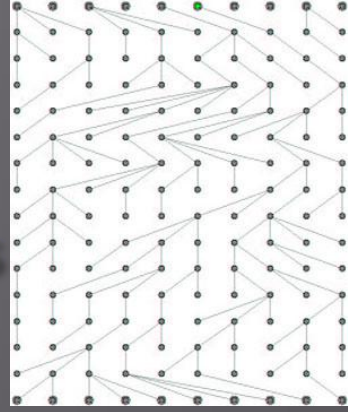
Cerambycidae
Data from catalogue of life



JTB **262** 245 (2010)



Beyond Yule:
The Birth-Death-Mutation (BDM) process



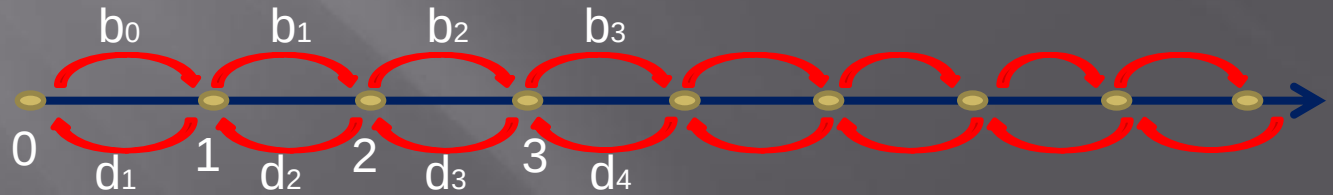
Moran Version
("overlapping generations")

1. ***Birth:*** *choose a random agent to reproduce.*

2. ***Mutation:*** *an offspring is identical to its "mother" with probability $1-\mu$ and is a mutant (new genus, surname, allele) with probability μ.*

3. ***Death:*** *a random agent is chosen to die with probability δ.*

Important parameters are μ and the growth rate $\gamma = 1-\delta$



JTB **262** 245 (2010)

Universality

The results are independent of the microscopic details of the process (overlapping or nonoverlapping generations, the order of events, IF

$$\mu, \gamma < 1.$$

Under these conditions the number of families of size m, n_m, varies only slightly with m so for both Moran and Wright-Fisher process, independent of their details, the family statistics is described by a universal differential equation, the Kummer equation:

$$\frac{\partial n_m}{\partial t} = (\mu - \gamma)\frac{\partial}{\partial m}(m n_m) + \frac{\sigma^2}{2}\frac{\partial^2}{\partial m^2}(m n_m)$$

$$\gamma > \mu \quad n_m = \frac{\nu R_c \Gamma(2+\nu)}{m} U\left(1+\nu, 0, \frac{R_c}{N} m\right)$$

$$\gamma < \mu \quad n_m = \frac{R_c \Gamma(1+\nu)}{m} U\left(\nu, 0, \frac{R_c}{N} m\right) e^{-\frac{R_c}{N} m}$$

Large m asymptotic

$$\gamma > \mu$$

$$n_m \approx m^{-1-[\gamma/(\gamma-\mu)]}$$

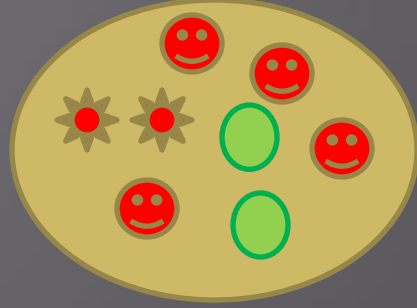
$$\gamma < \mu$$

$$n_m \approx m^{-1-\frac{\gamma}{\mu-\gamma}} e^{-(\mu-\gamma)m}$$

First presented by Manrubia and Zannette, JTB 216 461 (2002), in the context of Moran process

JTB **262** 245 (2010)

Effect of sampling



Let us assume that only R_0 individuals were sampled out of the total population of size N_0.

Hypergeometric distribution

$$n_m^R = \sum_{p \geq m} n_p \frac{\binom{p}{m}\binom{N_o - p}{R_o - m}}{\binom{N_o}{R_o}}$$

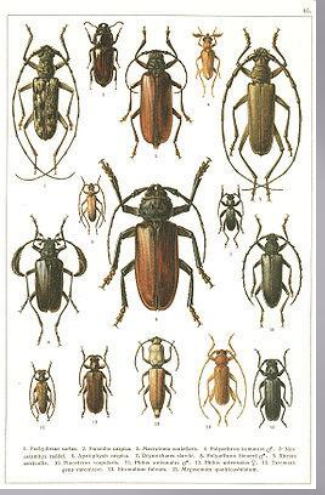
Effective sampling strength

$$\nu \equiv \frac{\mu}{\gamma - \mu}$$

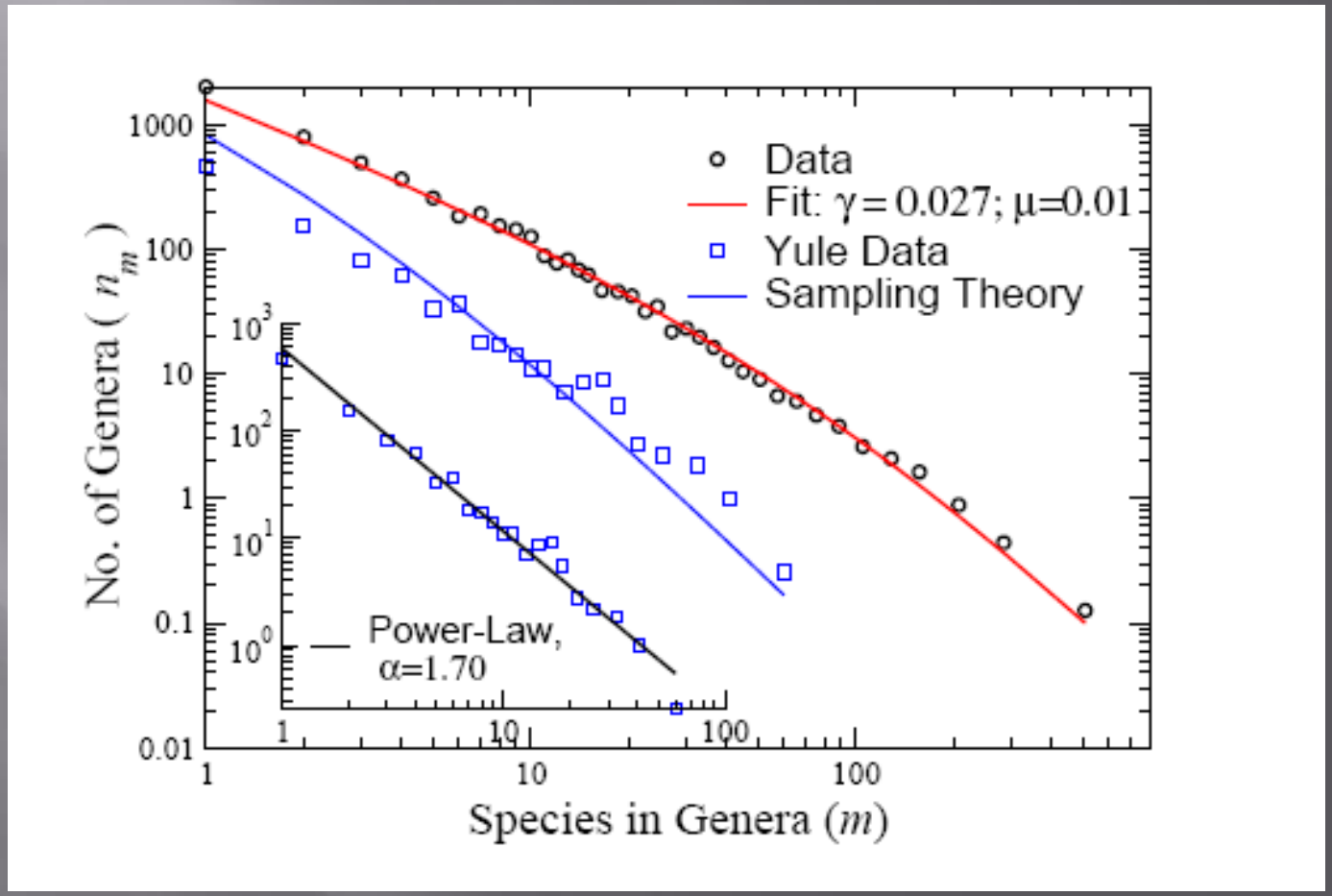
$$s \equiv \frac{\sigma^2(1+\nu)}{2\gamma}\frac{R_0}{N_0}$$

$$n_m^R \approx \nu R_c \mathrm{B}\left(2+\nu, m\right) s^m \, {}_2F_1\left(m, m+1; m+2+\nu; 1-s\right)$$

Weak sampling ⟹ ***back to Yule-Simon***

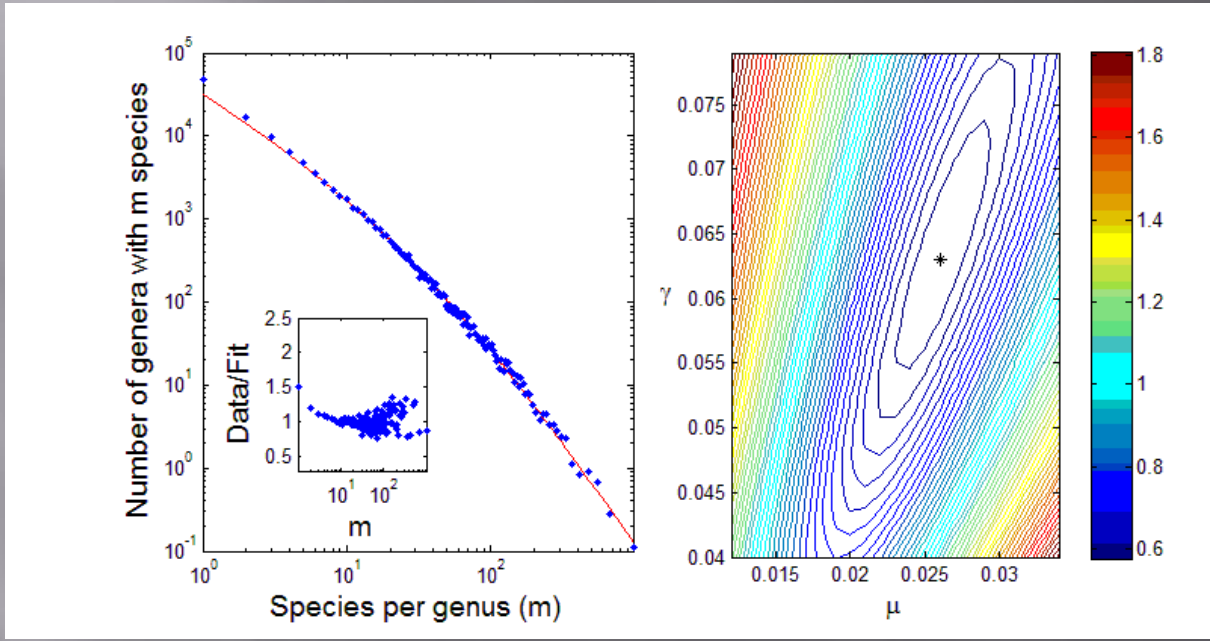


Back to the bugs



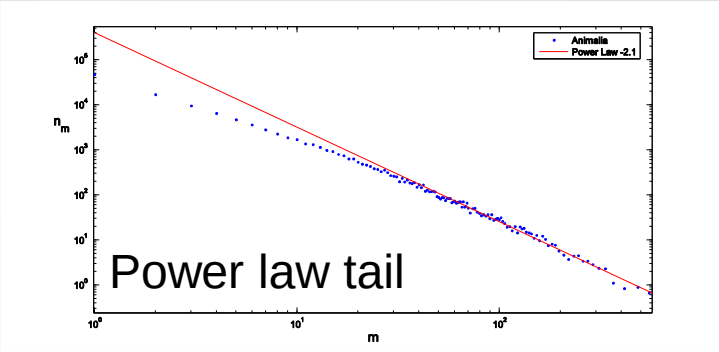
Applications
Species within genera: animalia

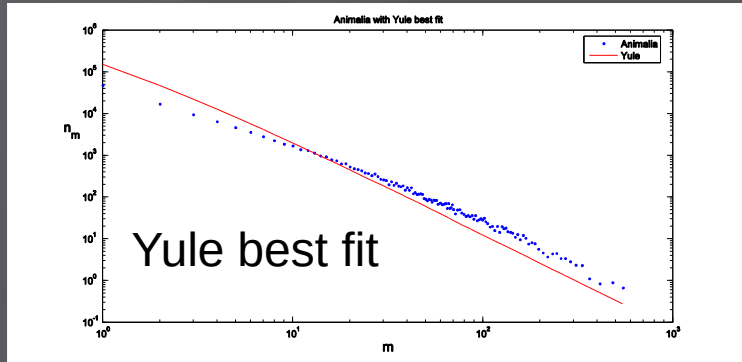
- 2 parameters fit.
- Logarithmic binning.
- PDF (NOT cumulative distribution)

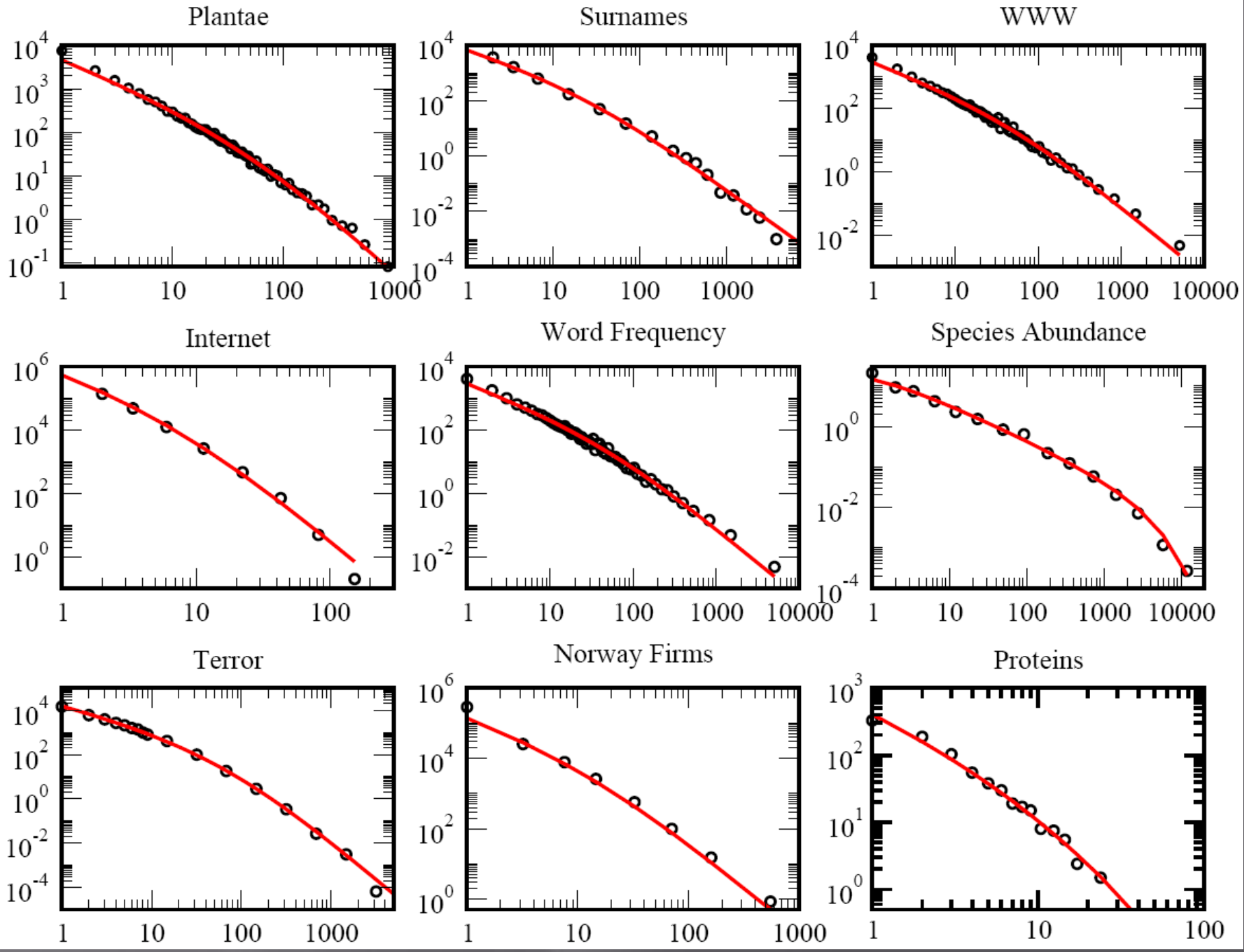


www.catalogueoflife.org

Species 2000 & ITIS Catalogue of Life: 2009 Annual Checklist






Plantae
Surnames
WWW
Internet
Word Frequency
Species Abundance
Terror
Norway Firms
Proteins

Inferring historic demographic parameters from current polymorphism (SNP) data

J. Stat. Phys. **142** 1302 (2011)

Molecular Biology and Evolution **28** 1617 (2011)

The data

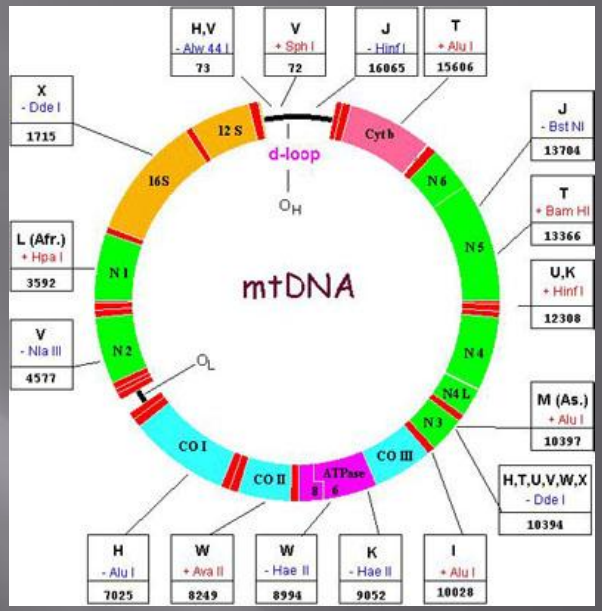
1212 sequences (China)

μ=0.00241/sequence/generation

The challenge:

Infer the growth rate γ

and the effective population size N_0



Haplotype (family) statistics:

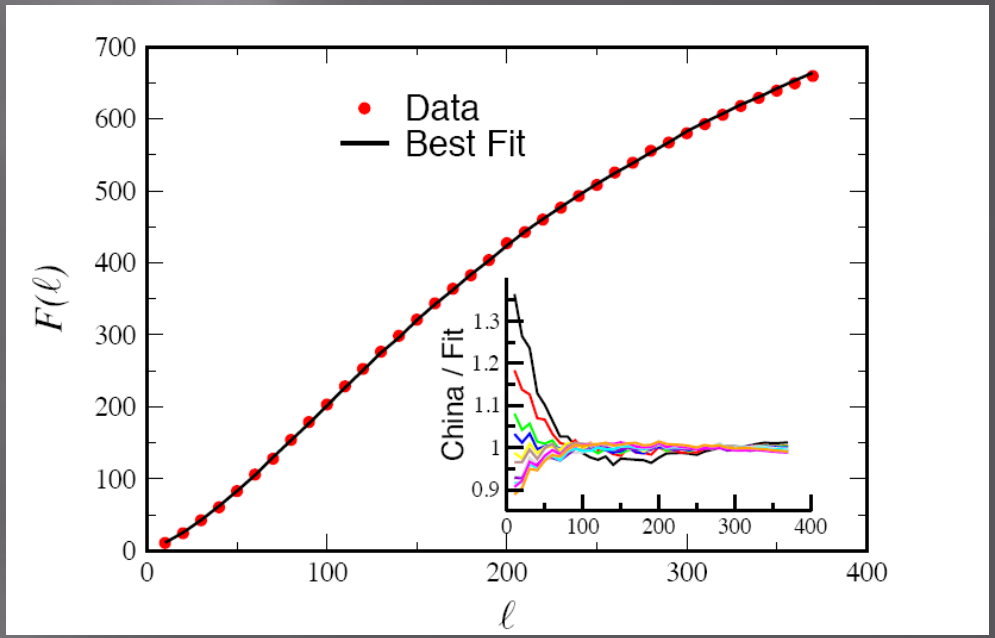
Any sequence inherited along maternal lineages is like a surname.

These surnames are subject to a birth-death-mutation process.

Ends up with haplotype (surname, family) statistics: the number of occurrences of singletons, doubletons and so one.

This statistics depends on the demographic parameters, thus in principle we can retrieve this information from it.

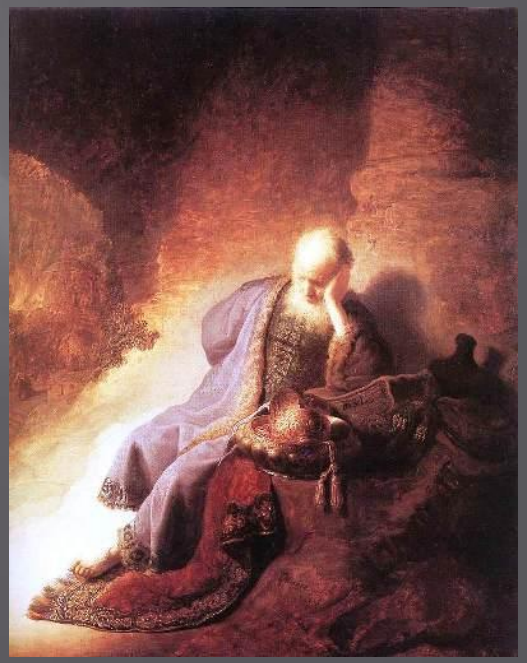
Individual	Loci 1	2	3	4	5	6	7
α	A	G	C	T	A	G	C
β	A	A	C	T	A	G	C
γ	A	G	C	A	A	G	C
δ	A	G	T	T	A	G	C
ϵ	G	G	T	T	A	G	T



Conclusions

- *The birth-death-mutation process yields the universal distribution function . It deviates strongly from Yule-Simon for small families since families cannot die in Yule.*
- *Yule is the weak sampling limit of BDM.*
- *One can use the BDM distribution in order to retrieve demographic parameters like growth rate in the prehistoric times.*

Neutral theory, King Solomon Version



Ecclesiastes (Qoheleth), Chapter 2:

14: The wise man's eyes are in his head; but the fool walketh in darkness: and I myself perceived also that <u>one event happeneth to them all</u>.

*15: Then said I in my heart, As it happeneth to the fool, so it happeneth even to me; and **<u>why</u>** was I then more wise?*

יד הֶחָכָם עֵינָיו בְּרֹאשׁוֹ, וְהַכְּסִיל בַּחֹשֶׁךְ הוֹלֵךְ; וְיָדַעְתִּי גַּם-אָנִי, שֶׁמִּקְרֶה אֶחָד יִקְרֶה אֶת-כֻּלָּם.

טו וְאָמַרְתִּי אֲנִי בְּלִבִּי, כְּמִקְרֵה הַכְּסִיל גַּם-אֲנִי יִקְרֵנִי, וְלָמָּה חָכַמְתִּי אֲנִי, אָז יֹתֵר; וְדִבַּרְתִּי בְלִבִּי, שֶׁגַּם-זֶה הָבֶל.


Blagodarya
感谢侬
Multumesc
Dziekuje
Grazie mille
NAGYON KÖSZÖNÖM
Gracias
Thank You
ありがとう
Spasibo
dunke schoen
Tika hoki
Благодаря
תודה רבה
Shukriya
Спасибо
merci beaucoup
значения
شكرا