

Theories of speciation

Today:

Mathematical theories of speciation:
what do they tell us?

1

Mathematical theory in evolution,
including speciation

- Identify key parameters
- Attempts to simplify – “caricature” processes
- Almost inevitably over-simplifies, because speciation is complex, multilocus

Pairs of coexisting closely related sexual species tend to be:

- Ecologically different (some sort of divergent, disruptive selection).
- Tend not to mate together (some sort of assortative mating).

How might that happen, and could it occur with gene flow?

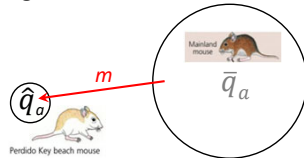
3/24/2025

2

2

Simplest key model: Haldane's model of
migration/selection balanceAdditive divergent
selection*
against a alleles, s

AA Aa aa
* Fitnesses 1 1- s 1-2 s

If $m \gg s$ then equilibrium $\hat{q} \approx \bar{q}$ -- “swamping!”If $m \ll s$ then equilibrium $\hat{q} \approx \frac{m}{s} \bar{q}$ -- migration-selection
equilibrium!

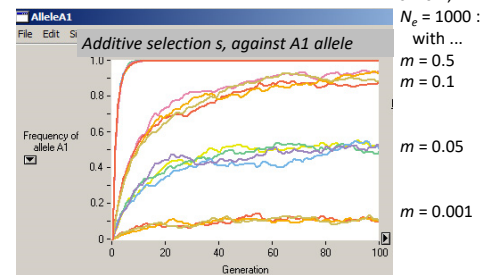
Too much gene flow, then no divergence

3/24/2025

3

3

Selection vs. migration

If $m \ll s$ then selection “wins,” maintains divergence, $p \approx m/s$ If $m \gg s$ then migration “wins,” pops. are “swamped”

3/24/2025

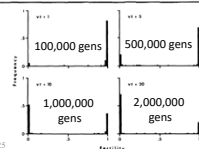
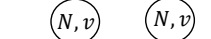
4

4

Nei et al. 1983:

Allopatric speciation. Accumulation of incompatibility via genetic
drift: stepwise mutation, heterozygote disadvantageTABLE 1
Fertilities or viabilities for various genotypes in the stepwise mutation model.
Here only five alleles are shown

	A_{-2}	A_{-1}	A_0	A_1	A_2
A_{-2}	1	1	0	0	0
A_{-1}	1	1	1	0	0
A_0	0	1	1	1	0
A_1	0	0	1	1	1
A_2	0	0	0	1	1

 v is the mutation rateFIGURE 4.—Relationships between the average fertility over all replications (\bar{f}) and evolutionary time (vt) for various values of $2Nv$, where $v = 10^{-5}$ is used. The stepwise mutation model is used. The number of replications used is 100. The symbols (□, ▲, ○, ●, and ■) represent the cases of $N = 5 \times 10^4, 2.5 \times 10^5, 5 \times 10^6$ and 900, respectively. The solid line shows the expected lowest fertility.

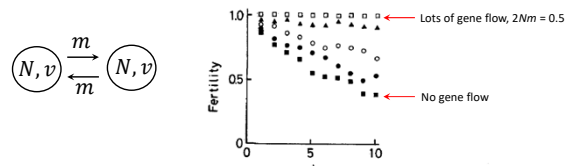
Nei et al. 1983

3/24/2025

5

5

Effect of gene flow

FIGURE 8.—Relationships between the average fertility over all replications (\bar{f}) and evolutionary time in the presence of migration. $N = 5000$ and $v = 10^{-5}$ were used. The number of replications is 100. The symbols (□, ▲, ○, ●, and ■) represent the cases of $2Nm = 0.5, 0.2, 0.1, 0.05$ and 0, respectively.

Conclusions

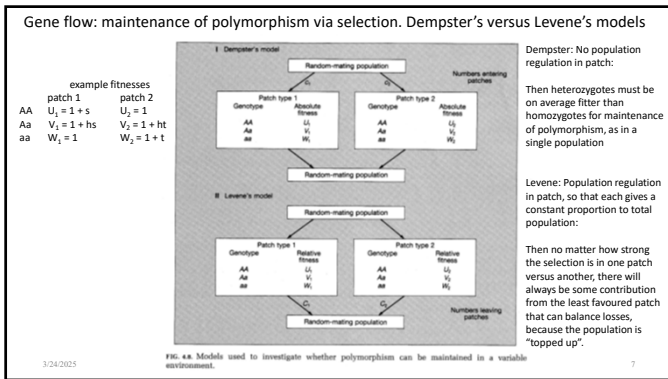
- Allopatric speciation via drift will be slow!
- Faster in smaller population sizes
- Even moderate gene flow (e.g. parapatry) stops speciation

Nei et al. 1983

3/24/2025

6

6



7

Felsenstein, 1981. Skepticism towards Santa Rosalia, or why are there so few kinds of animals*? Evolution 35: 124-138

HOMAGE TO SANTA ROSALIA
or
WHY ARE THERE SO MANY KINDS OF ANIMALS?*

G. E. HUTCHINSON
Department of Zoology, Yale University, New Haven, Connecticut

Reference to a famous paper in ecology by George E. Hutchinson (1959) American Naturalist 93: 145-159.

When you did me the honor of asking me to fill your presidential chair, I accepted perhaps without duly considering the duties of the president of a society, founded largely to further the study of evolution, at the close of the year that marks the centenary of Darwin and Wallace's initial presentation of the theory of natural selection. It seemed to me that most of the significant aspects of modern evolutionary theory have come either from geneticists, or from those heroic museum workers who suffering through years of neglect, were able to establish about 20 years ago what has come to be called the "new systematics." You had, however, chosen an ecologist as your president and one of that school at times supposed to study the environment without any relation to the organism.

A few months later I happened to be in Sicily. An early interest in zoogeography and in aquatic insects led me to attempt to collect near Palermo, certain species of water-bugs, of the genus *Corixa*, described a century ago

8

A concluding sentence from Hutchinson:

"... the reason why there are so many species of animals is at least partly because a complex trophic organization of a community is more stable than a simple one, but that limits are set by the tendency of food chains to shorten or become blurred, by unfavorable physical factors, by space, by the fineness of possible subdivision of niches, and by those characters of the environmental mosaic which permit a greater diversity of small than of large allied species."

Mosaicism of niches: maybe small size of organisms and "fineness of possible subdivision of niches" is related to gene flow among niches?

9

Felsenstein (1981), Skepticism towards Santa Rosalia, or why are there so few kinds of animals? :

Introduction p. 124

Speciation in current mathematical models is too easy! Maynard Smith 1966, Dickinson & Antonovics 1973, Caisse & Antonovics 1978, Balkau & Feldman 1973. All show sympatric speciation is possible: speciation is easy!

"...One would expect to find nearly infinite numbers of species. ... There would be a balance between speciation and the extinction of small species. ...

It is my impression that the number of species in nature is far smaller, and their size [i.e. nos. of individuals] far larger, than such a model would predict."

Maybe there's a genetic constraint on speciation?

10

Felsenstein's model p. 125 et seq.

Haploid, deterministic, $N \rightarrow \infty$

Levene-type model of population structure (soft selection, density-regulated)

Two loci, B,C experience divergent selection, s , in two subpopulations (niches), which maintains B,C polymorphisms overall, as follows:

Population regulation: absolute

Within each population, random mating every generation for recombination ($m=0.5$).

Linkage Disequilibrium (LD) forms between B and C in such a Levene model.

Unrelated assortative mating locus, A, that causes preferential mating with parameter d , as follows:

"Progress towards speciation" = stable linkage disequilibrium between A and B+C.

Genotype	Subpopulation I	Subpopulation II
BC	$(1 + s)^2$	1
Bc	$1 + s$	$1 + s$
bC	$1 + s$	$1 + s$
bc	1	$(1 + s)^2$

	A	a	Totals
A	$p^2(1 - d) + pd$	$p(1 - p)(1 - d)$	p
a	$p(1 - p)(1 - d)$	$(1 - p)^2(1 - d) + (1 - pd)$	$1 - p$
Totals	p	$1 - p$	1

11

What is linkage disequilibrium (LD)

When two genes deviate from the expected two locus equilibrium, the genes are said to be in **linkage disequilibrium** (or **gametic disequilibrium**).

The strength of this deviation is measured by the linkage disequilibrium (LD) coefficient, D . Suppose we have two diallelic loci, A/a and B/b

Observed gametic frequencies = random two locus expectation + deviation

Genotype	Observed gametic frequencies	random two locus expectation	+ deviation
P_{AB}	P_{AB}	$P_A P_B$	$+ D$
P_{Ab}	P_{Ab}	$P_A (1 - p_B)$	$- D$
P_{aB}	P_{aB}	$(1 - p_A) P_B$	$+ D$
P_{ab}	P_{ab}	$(1 - p_A)(1 - p_B)$	$- D$

$\Sigma = 1$

D can vary between a maximum of $+0.25$ and a minimum of -0.25 , but the range is often smaller if the frequency of alleles is not exactly 0.5. (Because $p_i \geq 0$ for all i, j).

12

Factors that affect LD

1. Decay

Disequilibrium declines by a fraction given by the recombination rate every generation

If c = recombination rate between

2 genes, then: $D_t = D_{t-1}(1 - c)$

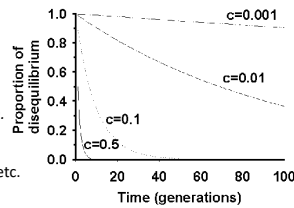
So, after many generations (t):

$$D_t = D_0(1 - c)^t$$

D can therefore decline by at most 50% in each generation.

2. Build-up

- Epistatic selection ($AB \uparrow$, $Ab \downarrow$ etc.
- Genetic drift
- Gene flow between divergent populations



3/24/2025

13

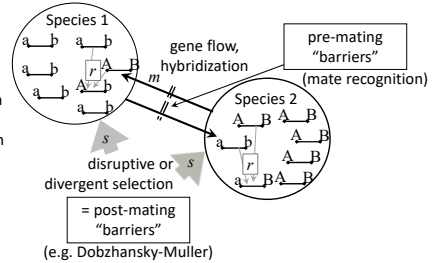
13

Genotypic clusters: in stable linkage disequilibrium

Disequilibrium decay is balanced by build-up due to selection

Felsenstein's criterion for speciation is not reproductive isolation alone

It's that there's enough reproductive isolation to maintain stable disequilibrium



3/24/2025

14

14

Alterations to the model: migration, recombination p. 128

TABLE 1. Threshold values of the amount of assortative mating in various cases involving different amounts of migration and linkage. The threshold value of d lies between the two values given in the appropriate column. SMR and SRM refer to the two orderings of events (Selection, Migration, and Recombination) in the life cycle.

	AB	BC		d threshold	
	s	r_{12}	r_{13}	m	
1.0	.5	.5	.5	.82-.83	.81-.82
1.0	.5	.5	.1	.59-.60	.57-.58
1.0	.5	.5	.01	.51-.55	.53-.54
1.0	.1	.5	.1	.28-.29	.27-.28
1.0	.5	.1	.1	.86-.87	.86-.87

Migration, m , is acting as a weak inverse kind of assortative mating given divergence of B, C loci.

Speciation more likely if A and B tightly linked: Association increases fitness due to assortative mating affecting one of the ecological loci

Speciation less likely if B & C are tightly linked: fewer deleterious recombinants, so lowered selection for assortative mating

Order of processes: Selection, gene flow, recombination or Selection, recombination, migration

3/24/2025

15

15

Evolutionary forces, p. 129

How does this work?

Remember, no direct selection on A locus.

Supposing all in perfect linkage disequilibrium, then ABC and abc only.

If rare recombinant Abc were to occur, would mate with ABC, and AbC and AbC offspring would result. Bc and bC are less fit than BC and bc overall, so there is indirect selection against A-(B+C) recombinants.

This is why reduction in recombination rate B-C makes speciation more difficult. ... And why A-B recombination rate reduction makes speciation easier in Table 1.

3/24/2025

16

16

Numerical results p. 127

"A computer program has been written to iterate genotype frequencies, using the language PASCAL on an [Intel] 8080 (SOL-20) microcomputer. This iteration is entirely deterministic."

"The Sol-20 was the FIRST ... microcomputer with a built-in keyboard and television output, ... later known as a home computer." (Wikipedia)

The SOL-20 microcomputer was invented by Joe's younger brother Lee Felsenstein. Lee also wrote the PASCAL compiler program for the SOL-20! Up to 8kb RAM!

Joe Felsenstein wrote his program in PASCAL and also the first versions of his phylogenetics program PHYLIP for this system, and for another microcomputer designed by Lee Felsenstein, the Osborne.

3/24/2025

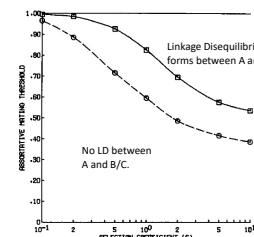
17

17

Numerical results p. 127

"While some special cases can be treated analytically, this is quite tedious."

... So numerical results were largely used for this 3-locus model.



"The immediate conclusion which we can draw from these results is that it is possible to construct a sympatric speciation model which sometimes does not speciate."

[Heh, heh! This was definitely the agenda all along!]

3/24/2025

18

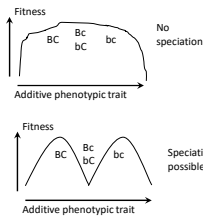
18

Mean fitness of BC and bc vs. Bc and bC generalists, p. 136

"When the mean fitnesses of BC and bc, averaged across the two subpopulations, did not exceed the fitnesses of Bc and bC, then sympatric speciation became impossible no matter how strong the assortative mating. In this case the generalists Bc and bC do not have lower average fitness than the specialists.

This suggests that we may be able to relate these fitness patterns to ... niche overlap between the forms adapted to the two subpopulations, and that when this overlap is too great, speciation will not occur."

Niches must be "discrete" relative to quantitative variation so that recombinants/intermediates survive poorly.



3/24/2025

19

19

"One-allele" vs. "two-allele" models, pp. 133-135

In Felsenstein's model, sympatric speciation requires very strong selection, and strong assortative mating. Otherwise gene flow with $m = 0.5$ or even $m = 0.1$ swamps recombination and prevents speciation.

Why is sympatric speciation "nearly impossible" in this model, when some previous models show it is likely?

The current model is a two-allele model of assortative mating, in that A mates with A, and a mates with a. Progress towards speciation only if one allele, say A, becomes associated with BC, and the other, say a, with bc.

One-allele models, in contrast, favor assortative mating by population or by phenotype, with say A giving assortative mating within populations, and a not. So fixation of A is favoured because it enhances fitness in both populations, & does not require linkage disequilibrium between A and BC/bc, making it easier to attain. Previous models of sympatric speciation have been of this type.

"Allopatry is a situation favorable to either..." one- or two-allele models.

3/24/2025

20

20

Twenty-one years later: Mark Kirkpatrick & Virginie Ravigné 2002

3/24/2025

21

21

Mark Kirkpatrick & Virginie Ravigné

Why do we "still lack a coherent understanding of speciation in terms of population genetic principles"? Three main reasons:

- 1) Speciation takes a long time.
- 2) Speciation is complex, many parameters, "spatial structure, non-random mating, epistasis, etc."
- 3) "Theoreticians have balkanized the subject of speciation." Too many models of many highly specific scenarios.

Surveyed ~ 100 models of speciation via natural selection.
Review is restricted to evolution of prezygotic isolation only.

Kirkpatrick & Ravigné 2002

3/24/2025

22

22

Mark Kirkpatrick & Virginie Ravigné

Five common elements:

- 1) A source of disruptive selection (i.e. one that generates linkage disequilibrium)
- 2) A prezygotic isolating mechanism – assortative mating, or mating preference
- 3) A way to transmit the disruptive selection to the isolating mechanism: direct or indirect selection
- 4) A genetic basis for increased isolation
- 5) An initial condition. Start near panmixia, or start with strong isolation?

Kirkpatrick & Ravigné 2002

3/24/2025

23

23

By 2002:

Many
mathematical
models of
speciation!

Table 1: Models of speciation classified by five major elements

	Reference ^a
I. A form of disruptive selection:	
A. Fitnesses vary in space	1, 3-7, 9, 12, 13, 21, 22, 25, 26, 34, 38, 49, 51-53, 55
B. Frequency-dependent natural selection:	
1. Two niches with independent density regulation	10, 12, 14, 15, 20, 25-30, 32, 37, 44, 45, 60
2. Tension zone	23, 25, 26, 37, 48
3. Competition within a niche	2, 8, 11, 12, 14-16, 25-30, 32, 39, 41, 43, 45, 46, 50, 60, 62
C. Sexual selection	2, 11, 17, 18, 24-26, 31, 33, 35, 38-40, 42, 48, 54, 56-59, 62
II. A prezygotic isolating mechanism:	
A. Mating preferences	16-18, 23, 24, 26, 30, 31, 33, 34-36, 40, 42, 43, 48-52, 54, 56-59, 61, 62
B. Assortment traits (including geographical isolation)	1, 3-9, 11-15, 20-22, 25, 27-29, 31, 32, 37-39, 41, 43, 47, 48, 50, 51, 53, 55, 60
C. A geographical setting:	
1. Allopatry	18, 24, 33, 35, 42, 49
2. Parapatry	1, 3-7, 12, 13, 21-23, 25, 26, 34, 36, 38, 40, 47, 51-53, 55
3. Sympatry	2, 6, 8-12, 14-17, 20, 25-32, 37-39, 41, 43-46, 48, 50, 54, 56-62
III. Transmitting the force of selection to the isolating mechanism:	
A. Direct selection	5, 6, 11, 20, 25, 27, 28, 31, 32, 37-39, 45, 48-50, 53-57
B. Indirect selection	1, 3-10, 12-18, 20-23, 25, 26, 28-30, 33, 34-37, 40, 42, 43, 47, 50-53, 55, 58-62
IV. A genetic basis for increased isolation:	
A. One-allele mechanisms	1, 3, 6, 8, 10, 12-15, 21-23, 29, 37, 47, 50, 51, 4, 5, 7, 9, 11, 13, 16-19, 21, 22, 24-33, 34-40, 42, 43, 45, 48-62
B. Two-allele mechanisms	
V. An initial condition:	
A. Divergence initially low	2, 8, 10-12, 14-17, 20, 28-32, 37, 39, 41, 43, 45, 50, 58-62
B. Divergence initially large (including geographical isolation)	1, 3-7, 9, 12, 13, 18, 21-27, 29, 33, 34, 36, 40, 42, 43, 47, 49, 51-57, 61

3/24/2025

24

Kirkpatrick & Ravigné

Haploid “toy model” of speciation. Studied the effect of “postzygotic isolation” and “assortative mating”

Two haploid loci, each have + and – alleles that affect a quantitative trait
Allele frequencies are set at 50% (“some form of [disruptive] selection maintains this polymorphism”)

Selection is such that fitnesses of ++, +–, and –– are 1, 1- S , and 1
 $S > 0$ is disruptive, $S < 0$ is stabilizing

Fixed assortative mating parameter A : $A = 0$, is random, $A = 1$ complete assortment, based on number of + alleles

Kirkpatrick & Ravigné 2002

Kirkpatrick & Ravigné

Using this greatly simplified, haploid two locus “toy model” of speciation, K & R and looked at the effect of disruptive selection and “assortative mating.” Like Felsenstein, K & R are interested in the maintenance of linkage disequilibrium. Even this “toy model” is quite complicated – here’s an analytical result for linkage disequilibrium:

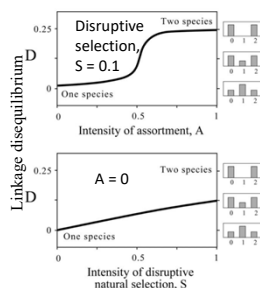
After mating, haploid offspring are produced following free recombination.
The dynamics of the linkage disequilibrium are given by

$$\Delta D = \frac{-64AS^2 - 4A^3 + A^3 - 4AS + 25S^2D^2 - 16(2A^2 - 8A^2 + 12A^2 - 8A - 3S - 2)SD^2 - 4(4 + 6A^2 - 4A^2 + A^3 + 4A(S - 2) - 4S(D - SS - 2))}{4(6A^2 - 4A^2 + A^3 + 4A(S - 2) + 2S - 2S^2 + 32(-4A + 6A^2 - 4A^2 + A^3 + 4S - 25S^2)D + 64AS^2 - 4A^3 - 4AS + 25S^2D^2)}$$

Here, A is the degree of assortative mating, S is the selection pressure, and D is linkage disequilibrium.

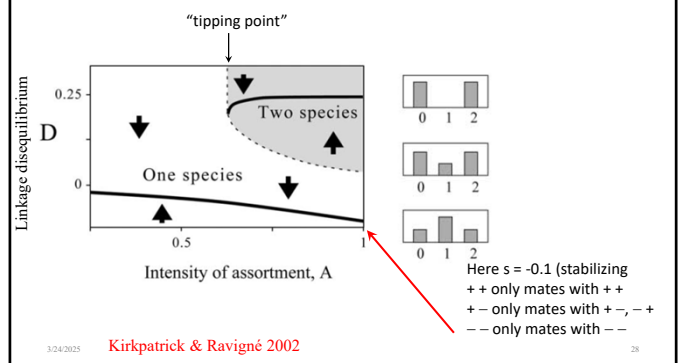
Kirkpatrick & Ravigné 2002

Kirkpatrick & Ravigné, toy model. Result 1:



Assortative mating has a stronger effect on linkage disequilibrium than simple disruptive selection

Kirkpatrick & Ravigné 2002



Here $s = -0.1$ (stabilizing)
++ only mates with ++
+- only mates with +-
-- only mates with --

Kirkpatrick & Ravigné 2002

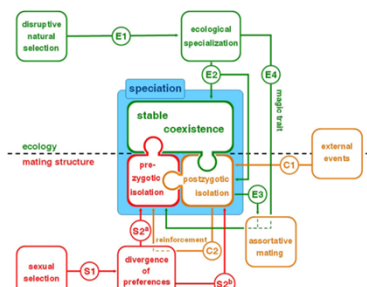
“Adaptive dynamics” models: back to Darwin

“Classical” speciation models (C, orange): RI + genetics only, ignore stable coexistence

Adaptive dynamics models (E, green): put ecology back into model. Need for coexistence.

Sexual selection may also be involved (S, red): affects pre-zygotic isolation. Can improve possibility of speciation

Need to put all three together



Weissing, Edelaar & van Doorn 2011

Adaptive dynamics

Putting ecology into speciation (Darwin’s original insight from Malthus)
Utilization of resources and population growth leads to depletion and competition.
The idea of evolutionary branching.

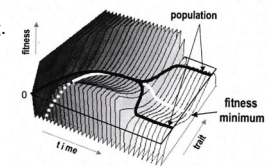
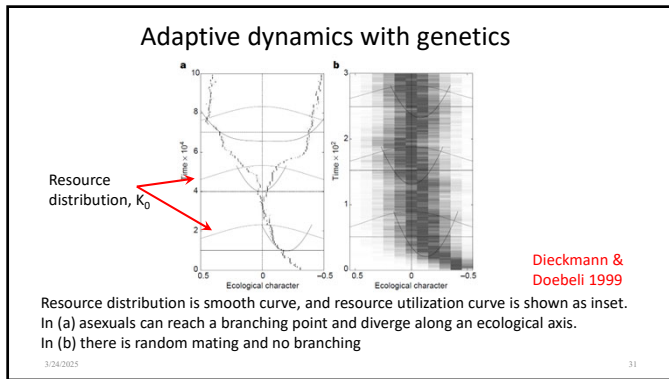
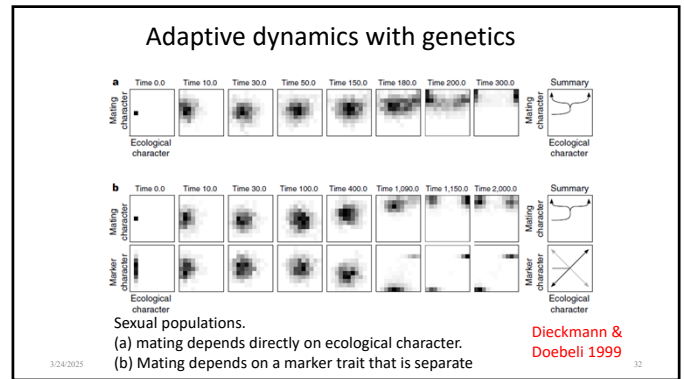


FIGURE 2.10. The development of the fitness landscape during a branching event.

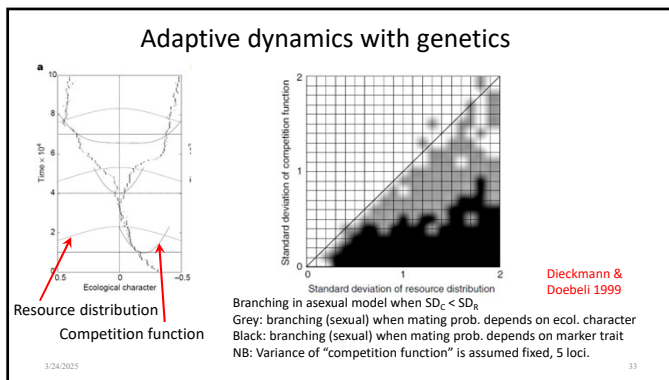
Metz 2011



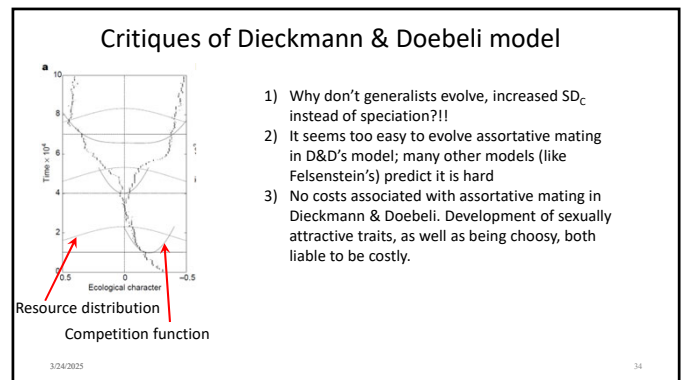
31



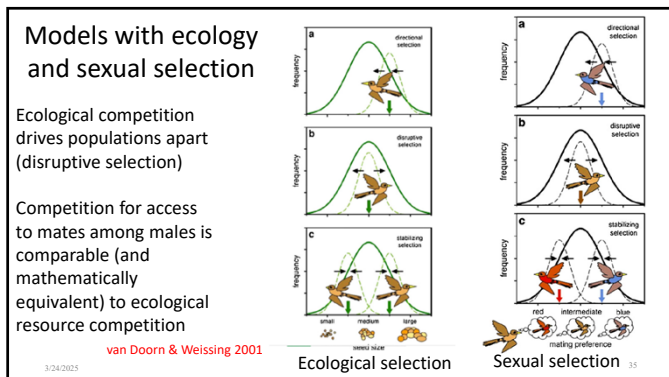
32



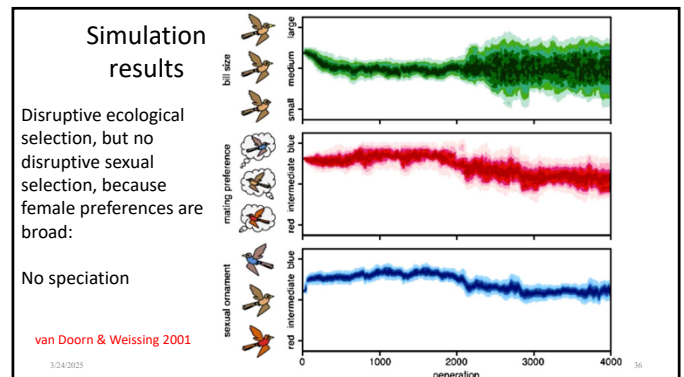
33



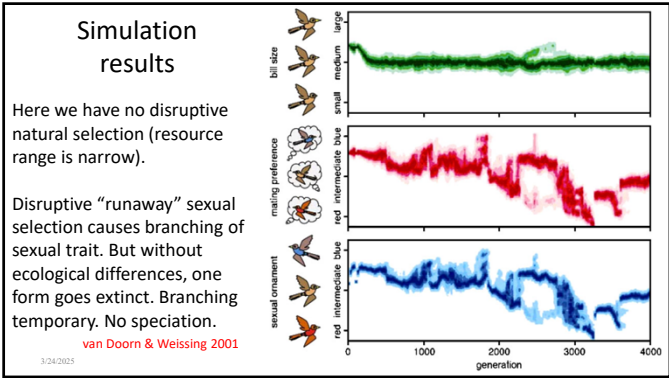
34



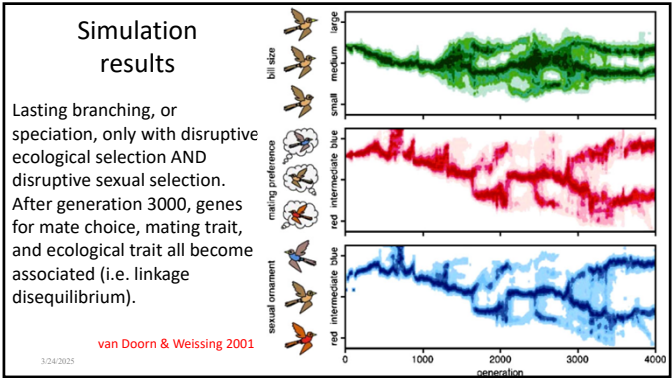
35



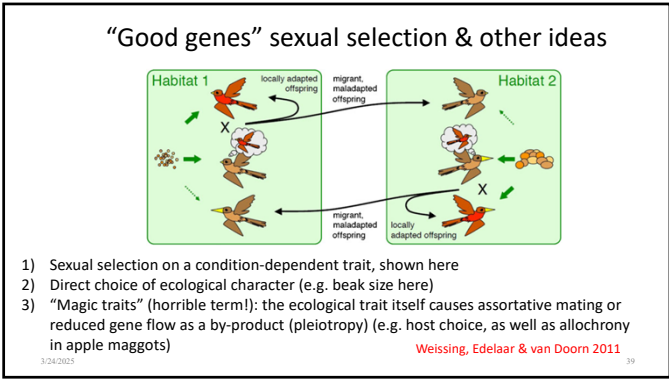
36



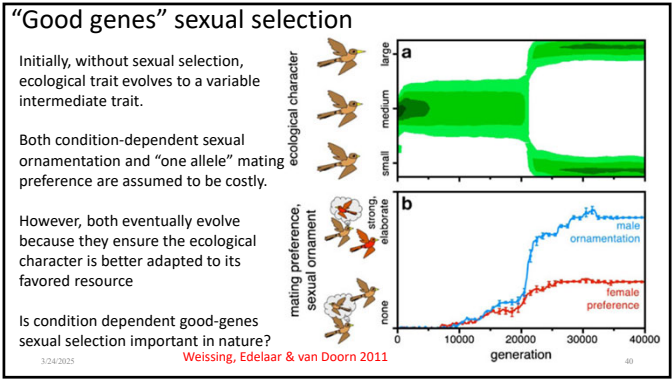
37



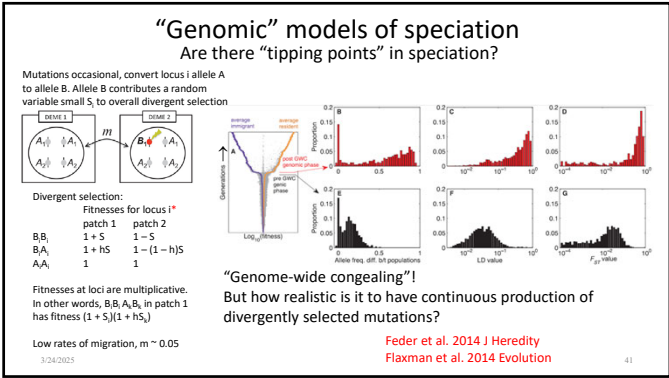
38



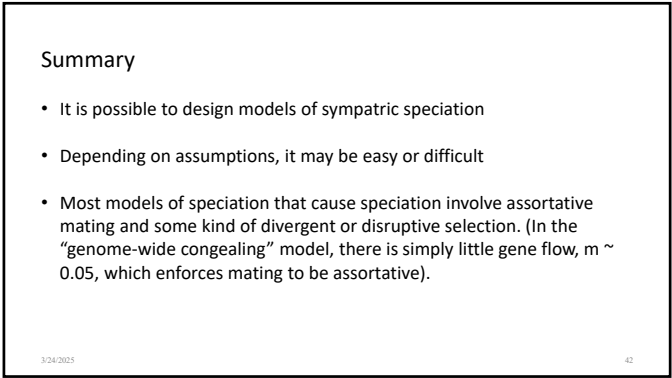
39



40



41



42

Some references (part I):

Bolnick, D.I., & Fitzpatrick, B.M. 2007. Sympatric speciation: models and empirical evidence. *Annual Review of Ecology, Evolution, and Systematics* 38:459-487. <https://doi.org/10.1146/annurev.ecolsys.38.091206.095804>

*Felsenstein, J. 1981. Skepticism towards Santa Rosalia, or why are there so few kinds of animals? *Evolution* 35:124-138. <https://doi.org/10.1111/j.1558-5646.1981.tb04864.x>

Gavrilets, S. 2014. Models of speciation: where are we now? *Journal of Heredity* 105:743-755. <https://doi.org/10.1093/hered/esu045>

*Kirkpatrick, M., & Ravigné, V. 2002. Speciation by natural and sexual selection. *American Naturalist* 159:522-535. <http://www.jstor.org/stable/10.1086/338370>

Kopp, M., ..., & van Doorn, G.S. 2018. Mechanisms of assortative mating in speciation with gene flow: connecting theory and empirical research. *The American Naturalist* 191:1-20. <http://www.journals.uchicago.edu/doi/abs/10.1086/694889>

(*)Maynard Smith, J. 1966. Sympatric speciation. *American Naturalist* 100:637-650

* = discussed in lecture

3/24/2025

43

43

Some more references (part II):

Nei, M. 1976. Mathematical models of speciation and genetic distance, Pages 723-765 in S. Karlin & E. Nevo, eds. *Population Genetics and Ecology*.

*Nei, M., Maruyama, T., & Wu, C.-I. 1983. Models of evolution of reproductive isolation. *Genetics* 103:557-579. <https://doi.org/10.1093/genetics/103.3.557>

*Nosil, P., Feder, J.L., Flaxman, S.M., & Gompert, Z. 2017. Tipping points in the dynamics of speciation. *Nature Ecology & Evolution* 1:0001. <http://dx.doi.org/10.1038/s41559-016-0001>

Ortiz-Barrientos, D., & Noor, M.A.F. 2005. Evidence for a one-allele assortative mating locus. *Science* 310:1467-. <https://doi.org/10.1126/science.1121260>

Turelli, M., Barton, N.H., & Coyne, J.A. 2001. Theory and speciation. *Trends in Ecology and Evolution* 16:330-343. [https://doi.org/10.1016/S0169-5347\(01\)02177](https://doi.org/10.1016/S0169-5347(01)02177)

*Weissing, F., Edelaar, P., & van Doorn, G. 2011. Adaptive speciation theory: a conceptual review. *Behavioural Ecology and Sociobiology* 65:461-480. <https://doi.org/10.1007/s00265-010-1125-7>

van Doorn, G.S., Dieckmann, U., & Weissing, F.J. 2004. Sympatric speciation by sexual selection: a critical reevaluation. *American Naturalist* 163:709-725

van Doorn, G.S., Edelaar, P., & Weissing, F.J. 2009. On the origin of species by natural and sexual selection. *Science* 326:1704-1707. <https://doi.org/10.1126/science.1181661>

3/24/2025

* = discussed in lecture

44

44