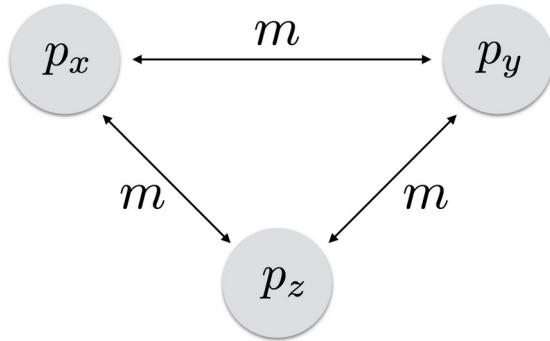




Πληθυσμιακή και
Εξελικτική Γενετική

Ροή γονιδίων και πληθυσμιακή δομή

Αριστοτέλης Παπαγεωργίου, Τμ. ΜΒΓ ΔΠΘ, apapage@mbg.duth.gr

Ευρωπαίοι άποικοι στην Κολομβία

Η Ισπανία αποίκισε τη βορειοδυτική Κολομβία (Antioquia) από τον 16ο αιώνα. Σήμερα τα άτομα σε αυτήν την περιοχή θεωρούν ότι έχουν ευρωπαϊκή καταγωγή.

Η εξέταση του mtDNA δείχνει ότι σχεδόν το σύνολο (90%) της μητρικής γραμμής καταγωγής σε αυτόν τον πληθυσμό είναι από ιθαγενείς Αμερικανίδες.

Η ανάλυση του χρωμοσώματος Y δείχνει ότι σχεδόν το σύνολο (94%) της πατρικής γραμμής καταγωγής είναι ευρωπαϊκή.

Στον παρακάτω πίνακα φαίνονται τα αποτελέσματα των γενετικών αναλύσεων (observed) και γίνεται ο έλεγχος του σεναρίου ότι οι σημερινοί κάτοικοι είναι απόγονοι ευρωπαίων ανδρών και γυναικών αμερικανικής καταγωγής (expected).

TABLE 7.1 The expected (assuming the population was founded by Native American women and European men) and observed mtDNA, Y-chromosome, autosome, and X-chromosome Native American and European ancestry in an isolated northwestern Colombian population. Data from G. Bedoya, et al., 2006.

	<i>mtDNA</i>	<i>Y chromosome</i>	<i>Autosome</i>	<i>X chromosome</i>
Expected				
Native American	1.0	0.0	0.5	0.67
European	0.0	1.0	0.5	0.33
Observed				
Native American	0.90	0.01	0.16	0.25
European	0.02	0.94	0.79	0.69

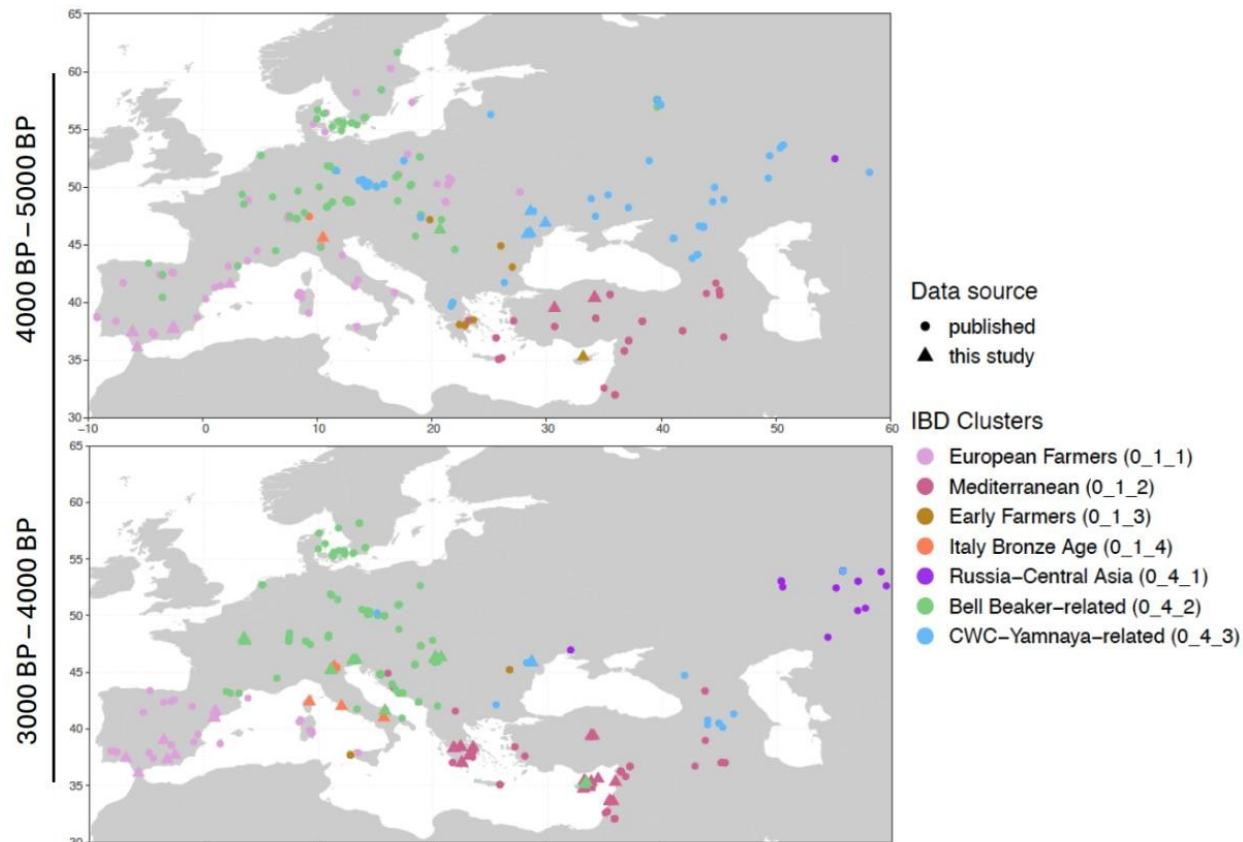
Αρχαίο DNA και οι λαοί της στέπας

Η εξάπλωση των ινδοευρωπαϊκών γλωσσών στην Ευρώπη από την 5η χιλιετία BP σχετίζεται με τις γραμμές γενετικής καταγωγής από τη στέπα κατά την έναρξη της Εποχής του Χαλκού. Έγινε ανάλυση αρχαίων γονιδιωμάτων από τη Μεσόγειο (5.200 ως 2.100 BP).

Η άφιξη της στεπικής γραμμής καταγωγής στην Ισπανία, τη Γαλλία και την Ιταλία μετριάστηκε από την παρουσία των Bell Beaker (BB) της Δυτικής Ευρώπης, συμβάλλοντας πιθανότατα στην εμφάνιση της πλάγιας και κελτικής γλώσσας.

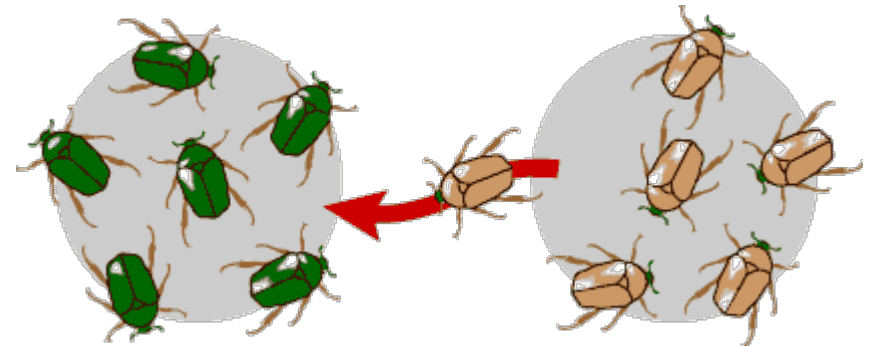
Αντίθετα, οι αρμενικοί και οι ελληνικοί πληθυσμοί δέχτηκαν την γραμμή καταγωγής της στέπας απευθείας από ομάδες Yamnaya της Ανατολικής Ευρώπης.

Τα ευρήματά ευθυγραμμίζονται με μοντέλα γλωσσικής απόκλισης για την ινδοευρωπαϊκή γλωσσική οικογένεια.



Γονιδιακή ροή (gene flow)

- Λέγεται και μετανάστευση
- Μετακίνηση ατόμων ή γαμετών μεταξύ πληθυσμών (ομάδων) που καταλήγουν σε ανταλλαγή γενετικού υλικού
- Η μετανάστευση αλλάζει τις συχνότητες των αλληλομόρφων στους φυσικούς πληθυσμούς και καθορίζει σε μεγάλη έκταση τη γενετική τους διαφοροποίηση
- Η γονιδιακή ροή δρα αντίθετα στη γενετική εκτροπή
 - Αυξάνει το λειτουργικό μέγεθος του πληθυσμού
 - Αυξάνει τη γενετική ποικιλότητα μέσα στους πληθυσμούς
 - Μειώνει τη διαφοροποίηση μεταξύ πληθυσμών

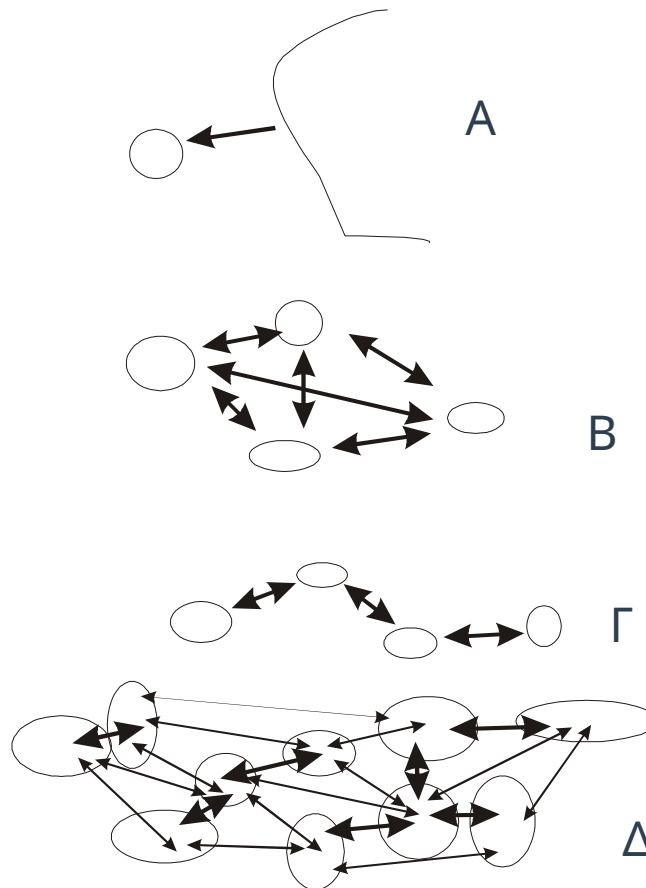


Ένταση ροής γονιδίων

- Η ροή γονιδίων μπορεί να έχει μεγάλη ή μικρή ένταση
 - Ανάλογα με την αλλαγή στις συχνότητες των αλληλομόρφων που προκαλείται εξ αιτίας της
- Η ένταση της γονιδιακής ροής εξαρτάται από:
 - τη διαφορά των συχνοτήτων των αλληλομόρφων μεταξύ των πληθυσμών που γίνεται η μετακίνηση και
 - το ποσοστό των ατόμων που μεταναστεύουν σε κάθε γενιά
- Όταν τα επίπεδα ροής γονιδίων μεταξύ των διαφορετικών (υπο-)πληθυσμών είναι μεγάλα, τότε υπάρχει ομογενοποίηση της γενετικής σύστασης του συνόλου
- Όταν αυτά είναι μικρά, τότε αυξάνει η διαφοροποίηση μεταξύ των πληθυσμών λόγω επιλογής, εκτροπής και μετάλλαξης

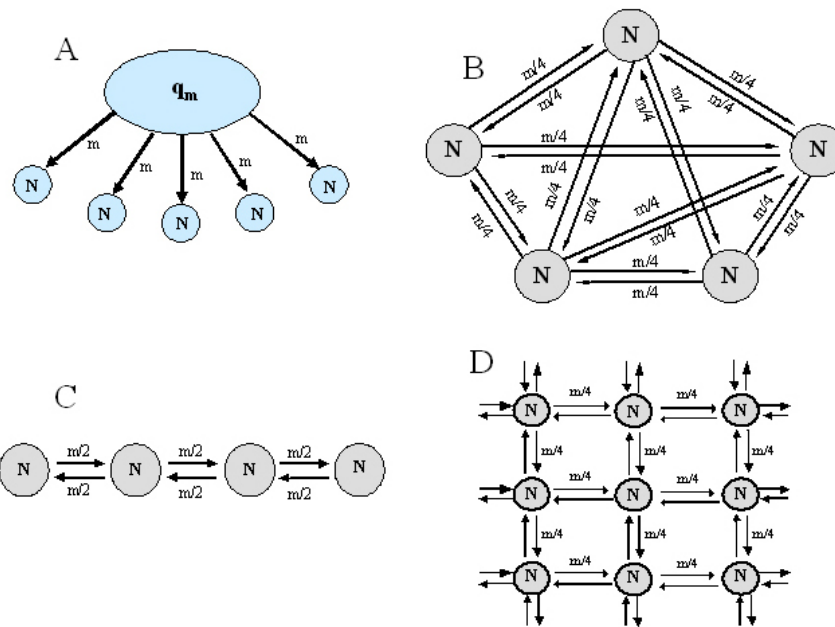
Μοντέλα γονιδιακής ροής

- [Α] Μοντέλο «ήπειρος-νησί» από μεγάλο σε μικρό πληθυσμό
- [Β] Μοντέλο «των νησιών» μεταξύ μικρών πληθυσμών κατά τυχαίο τρόπο
- [Γ] Μοντέλο «του βηματισμού» (stepping stone) μόνο μεταξύ γειτονικών πληθυσμών στη διάρκεια μιας γενιάς
- [Δ] Μοντέλο «απομόνωσης λόγω απόστασης» (isolation-by-distance) με εξάρτηση από την απόσταση μεταξύ πληθυσμών που αποτελούν μια συνέχεια

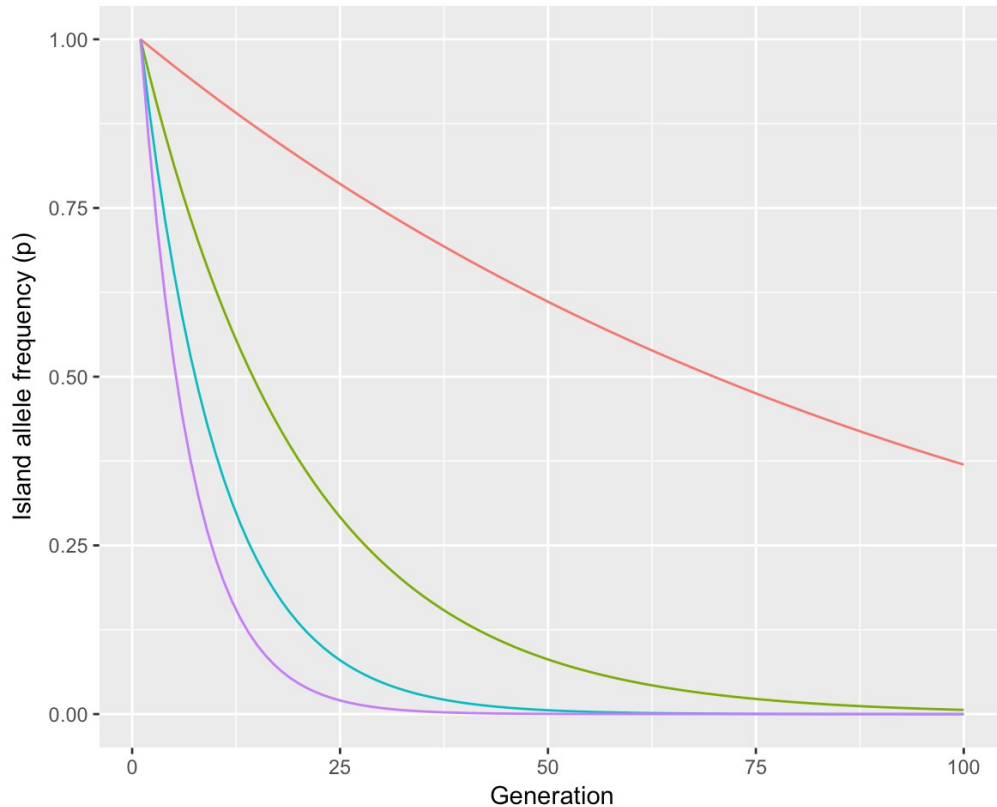


Μοντέλα γονιδιακής ροής

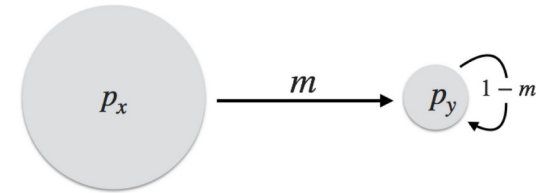
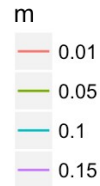
- Όλα τα μοντέλα γονιδιακής ροής έχουν το κοινό γνώρισμα της τάσης ομογενοποίησης των πληθυσμών
- Αν αυτή η τάση δεν εξισορροπηθεί, θα οδηγήσει στη σύγκλιση των συχνοτήτων των αλληλομόρφων προς κάποια μέση τιμή



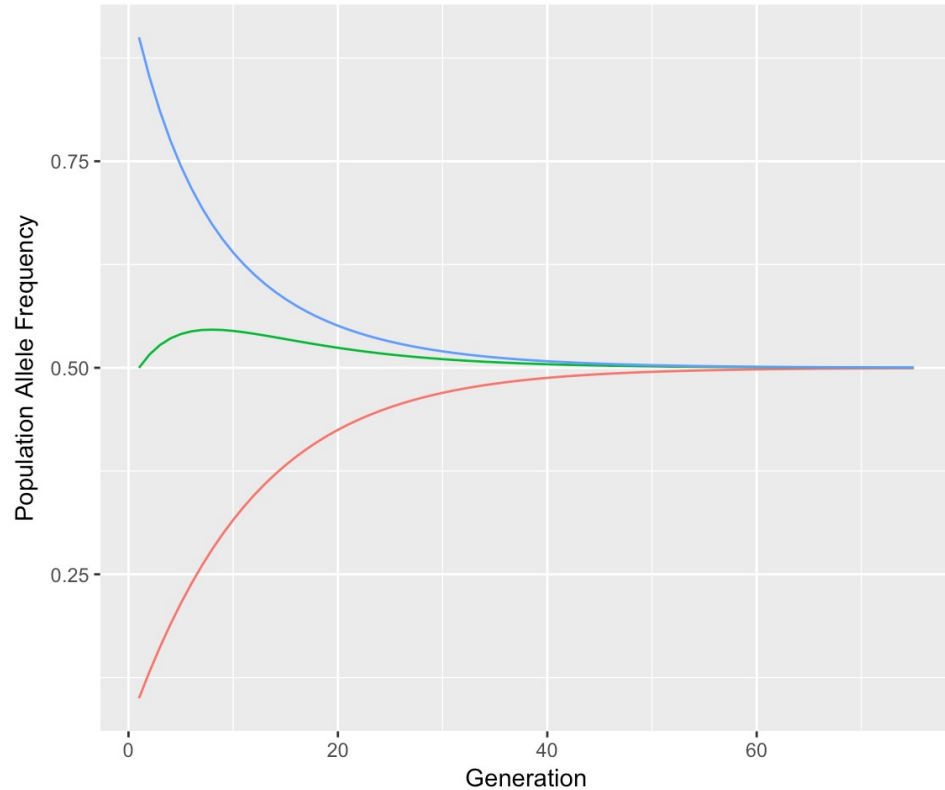
Αλλαγή συχνοτήτων αλληλομόρφων στο μοντέλο νησιού-ηπείρου



Συχνότητες αλληλόμορφων για έναν νησιωτικό πληθυσμό που είναι παγιωμένος για ένα αλληλόμορφο και δέχεται μετανάστες από πληθυσμό της ηπειρωτικής χώρας με διάφορους ρυθμούς (m), που δεν έχει το αλληλόμορφο που υπάρχει στο νησί.



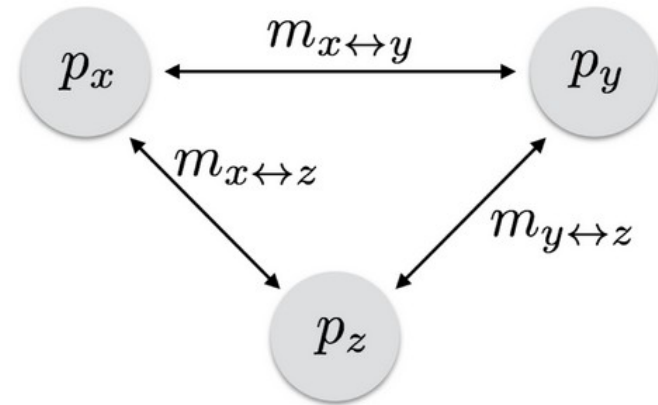
Αλλαγή συχνοτήτων αλληλομόρφων στο γενικό μοντέλο



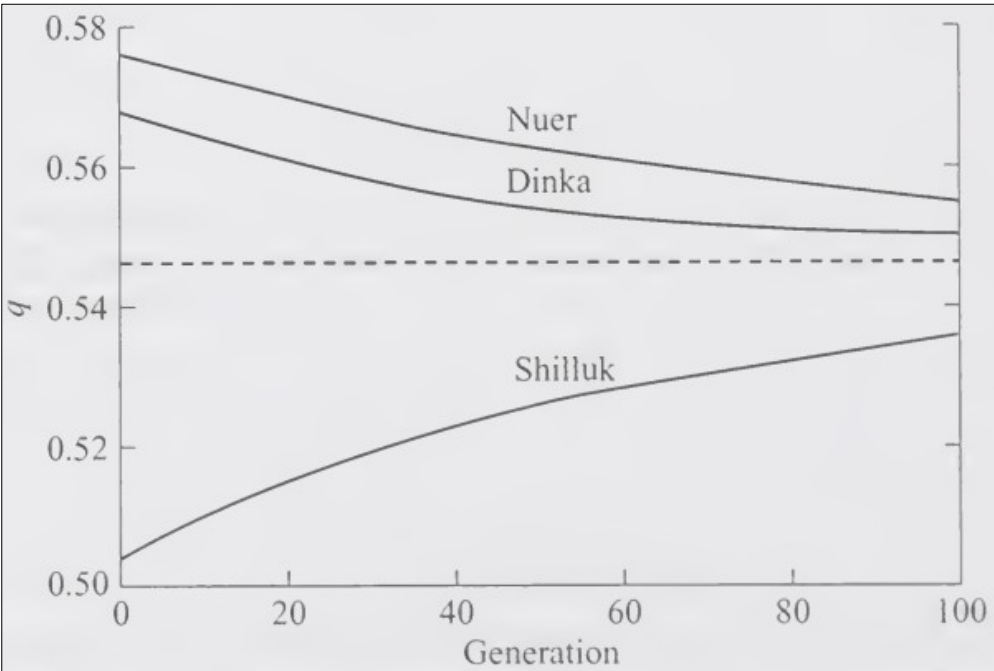
Αλλαγές στις συχνότητες των αλληλομόρφων στον χρόνο για τρεις πληθυσμούς που συνδέονται μεταξύ τους με διαφορετικούς ρυθμούς ροής γονιδίων: ($m_{XY}=0.04$, $m_{XZ}=0.02$, $m_{YZ}=0.08$)

Population

X
Y
Z



Αλλαγή συχνοτήτων αλληλομόρφων στο γενικό μοντέλο



Αλλαγές στις συχνότητες για το αλληλόμορφο M των ομάδων αίματος MN σε για τρεις σουδανικούς πληθυσμούς.

Ο Πίνακας 7.3 δίνει την αναλογία της γονιδιακής ροής μεταξύ αυτών των ομάδων.

Αν και μεγάλο ποσοστό αυτών των ατόμων δεν μεταναστεύουν, τελικά η ροή γονιδίων τείνει να εξισώσει τις συχνότητες του αλληλομόρφου M στον χρόνο.

TABLE 7.3 The proportion of gene flow among three Sudanese populations. Data from D. F. Roberts and R. W. Hiorns, 1962.

Recipient population	Source population		
	(1) Nuer	(2) Dinka	(3) Shilluk
(1) Nuer	0.9850	0.0125	0.0025
(2) Dinka	0.0138	0.9775	0.0087
(3) Shilluk	0.0000	0.0098	0.9902

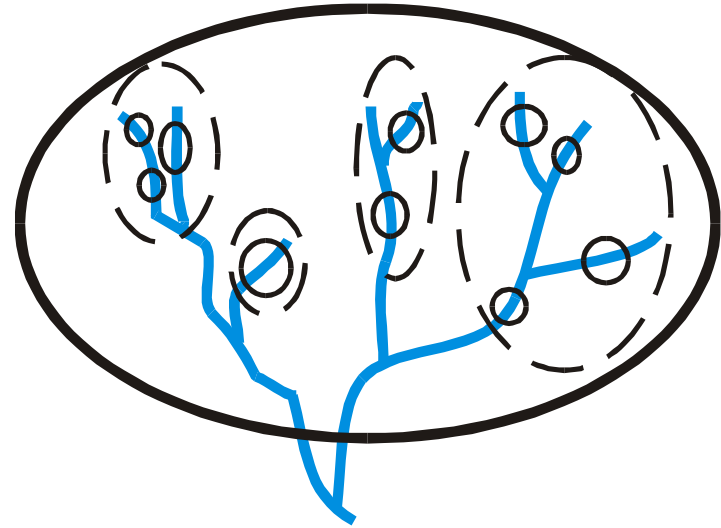
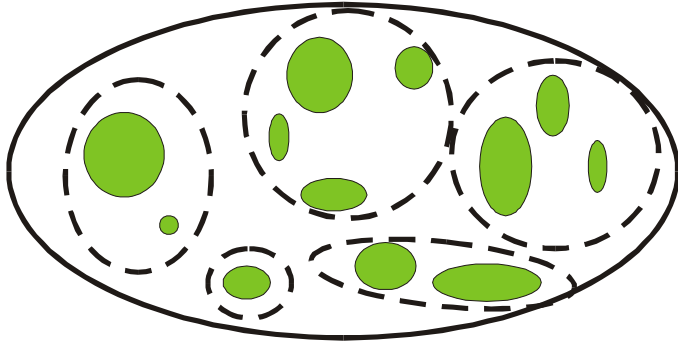
Γενετική διαφοροποίηση

- Γενετική ποικιλότητα ανάμεσα σε διαφορετικούς πληθυσμούς του ίδιου είδους
 - Ή διαφορετικά είδη
 - Ή διαφορετικούς υποπληθυσμούς
- Γενετική απόσταση ανάμεσα σε δύο πληθυσμούς
- Δείκτες γενετικής διαφοροποίησης ανάμεσα σε ένα σύνολο πληθυσμών / υποπληθυσμών
 - F statistics
 - AMOVA
 - PCA based
 - Model based
- Περιγραφή της **πληθυσμιακής δομής** ενός πληθυσμού ή ενός συνόλου πληθυσμών

Γενετική Δομή Πληθυσμών - Πληθυσμιακή Υποδιαίρεση

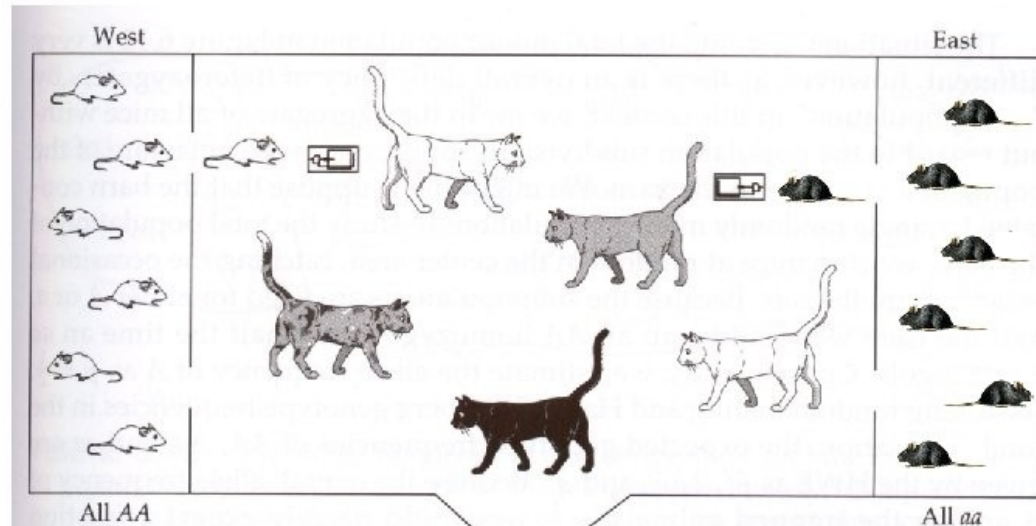
- Τα είδη συνήθως σχηματίζουν υποπληθυσμούς
 - πληθυσμιακή υποδιαίρεση
 - γενετική διαφοροποίηση (διαφορετικές γονιδιακές συχνότητες μεταξύ των υποπληθυσμών)
- Ιεραρχική δομή πληθυσμών: οι υποπληθυσμοί μπορούν να ομαδοποιηθούν σε ομάδες και αυτές με τη σειρά τους σε ευρύτερες ομάδες
 - π.χ. πληθυσμοί ψαριών στη λεκάνη απορροής ποταμού

Γενετική Δομή Πληθυσμών - Πληθυσμιακή Υποδιαίρεση



Γενετική Δομή Πληθυσμών - Πληθυσμιακή Υποδιαίρεση

Wahlund Effect

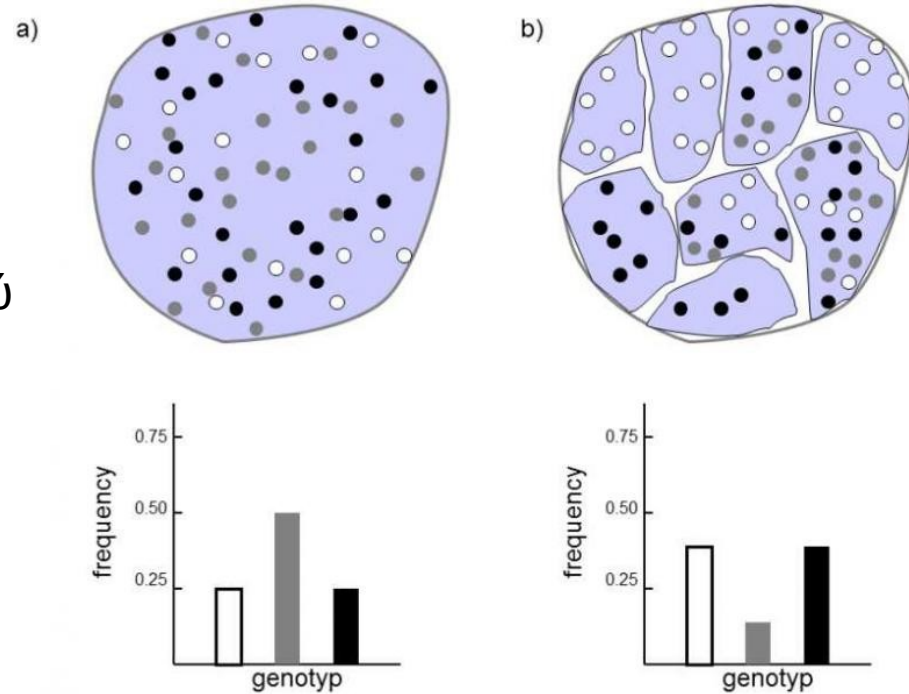


Hartl and Clark 1997

Trapped mice will always be homozygous even though $H_E = 0.5$

Φαινόμενο Wahlund

- Πολλές φορές η υποδιαίρεση ενός πληθυσμού σε υποπληθυσμούς δεν είναι εμφανής
 - Πληθυσμός δέντρων με συγκέντρωση των απογόνων (από σπέρματα) κοντά στους θηλυκούς γονείς
 - Πληθυσμός ανθρώπων με πολιτιστικού ή θρησκευτικούς φραγμούς στην αναπαραγωγή
- Ακόμα και να υπάρχει ισορροπία HW στους υποπληθυσμούς, στο σύνολο έχουμε έλλειμμα ετερόζυγων
 - Sten Wahlund (1928)



Φαινόμενο Wahlund

$$\overline{Q} = \overline{q}^2 + V_q$$

$$\overline{H} = 2\overline{p}\overline{q} - 2V_q$$

$$\overline{P} = \overline{p}^2 + V_q$$

- Λόγω της υποδιαίρεσης του πληθυσμού, αυξάνονται οι ομόζυγοι κατά V_q ο καθένας και κατά $2V_q$ μειώνονται οι ετερόζυγοι
 - Όπου V_q η διακύμανση της συχνότητας του αλληλομόρφου a
- Όσο πιο διαφορετικοί γενετικά είναι μεταξύ τους οι υποπληθυσμοί, τόσο πιο μεγάλη είναι η απόκλιση του συνόλου από την ισορροπία
 - Μειώνονται οι ετερόζυγοι
 - Αυξάνονται οι ομόζυγοι

Δείκτης F

- Δείκτης που υπολογίζει το έλλειμμα ετεροζυγωτίας (ετερόζυγων γονότυπων) σε έναν πληθυσμό
- Η σύγκριση μεταξύ της παρατηρούμενης και αναμενόμενης ετεροζυγωτίας
 - Όταν $F > 0$ τότε υπάρχει περίσσια ομόζυγων γονοτύπων στον πληθυσμό,
 - ενώ όταν $F < 0$ περίσσια ετερόζυγων
- Σε έναν πληθυσμό ισορροπίας ισχύει $F = 0$
 - Επειδή $H_e = H_o$

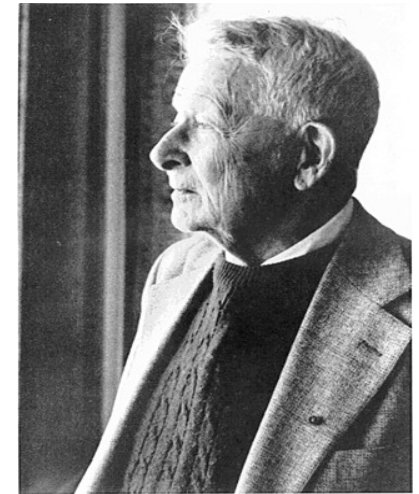
$$F = \frac{H_E - H_O}{H_E} = \frac{2pq - H_O}{2pq}$$

Πληθυσμιακή υποδιαίρεση: συντελεστές F

- Όταν υπάρχει πληθυσμιακή δομή, τότε ο συντελεστής F μπορεί να υπολογιστεί
 - Για τα άτομα ενός πληθυσμού F_{IT}
 - Για τα άτομα που ανήκουν στον ίδιο υποπληθυσμό F_{IS}
 - Για τους υποπληθυσμούς ενός πληθυσμού F_{ST}

$$F_{IT} = F_{IS} + (1 - F_{IS})F_{ST}$$

Wright's F statistics



Sewall Wright

Δείκτης διαφοροποίησης F_{ST}

$$F_{ST} = \frac{F_{IT} - F_{IS}}{1 - F_{IS}}$$

- Είναι μέτρηση της διαφοροποίησης μεταξύ των υποπληθυσμών και είναι πάντα θετική
 - Ευρέως χρησιμοποιούμενη σε πληθυσμιακές μελέτες φυσικών ειδών
 - Τα F_{IS} και F_{IT} μπορούν να μετρηθούν με την απόκλιση από την ισορροπία HW
 - Το F_{ST} μετράει ουσιαστικά τη μείωση της ετεροζυγωτίας λόγω πληθυσμιακής υποδιαίρεσης

Δείκτης διαφοροποίησης G_{ST}

- Για πολλαπλά γονίδια, είναι H_0 η μέση παρατηρούμενη ετεροζυγωτία σε έναν υποπληθυσμό, H_S η μέση αναμενόμενη ετεροζυγωτία σε έναν υποπληθυσμό και H_T η μέση συνολική αναμενόμενη ετεροζυγωτία ενός πληθυσμού
- Μπορούμε τελικά να υπολογίσουμε τη διαφοροποίηση ανάμεσα στους υποπληθυσμούς σαν συνάρτηση της αναμενόμενης ετεροζυγωτίας στον πληθυσμό και στους υποπληθυσμούς

$$\bar{F}_{IS} = \frac{\bar{H}_S - \bar{H}_0}{\bar{H}_S}$$

$$\bar{F}_{IT} = \frac{\bar{H}_T - \bar{H}_0}{\bar{H}_T}$$

$$\bar{F}_{ST} = \frac{\bar{H}_T - \bar{H}_S}{\bar{H}_T}$$

$$G_{ST} = \frac{H_T - H_S}{H_T}$$



M. Nei

Ιδιότητες του F_{ST}

- $F_{ST} = 0$ (min)
 - Καμία διαφοροποίηση: οι συχνότητες των αλληλομόρφων είναι ίδιες ανάμεσα στους (υπο)πληθυσμούς που μελετάμε
- $F_{ST} = 1$ (max)
 - Σε κάθε (υπο)πληθυσμός είναι παγιωμένο (fixed) διαφορετικό αλληλόμορφο

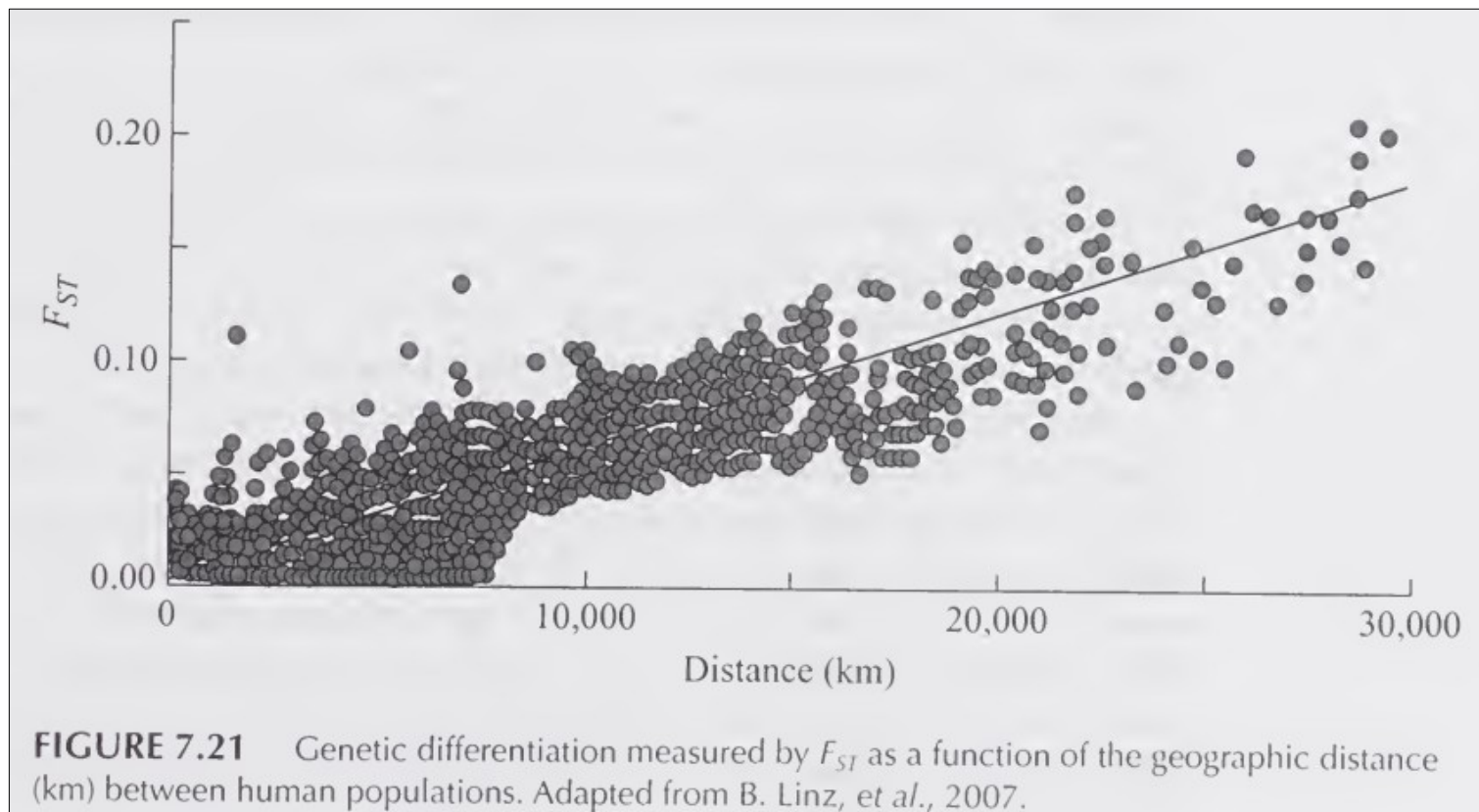
F_{ST}	Γενετική διαφοροποίηση
0 – 0.05	μικρή
0.05 – 0.15	μεσαία
0.15 – 0.25	μεγάλη
> 0.25	πολύ μεγάλη

Χρήση του F_{ST}

Οργανισμός	Πληθυσμοί	Γενετικοί τόποι	F_{ST}
Άνθρωπος (τρεις ήπειροι)	3	35	0,069
Άνθρωπος -Yanomama (Αμαζόνιος)	37	15	0,077
Ποντίκι (<i>Mus musculus</i>)	4	40	0,113
<i>Drosophila equinoxialis</i>	5	27	0,109
Φυτό (<i>Lycopodium lucidulum</i>)	4	13	0,282

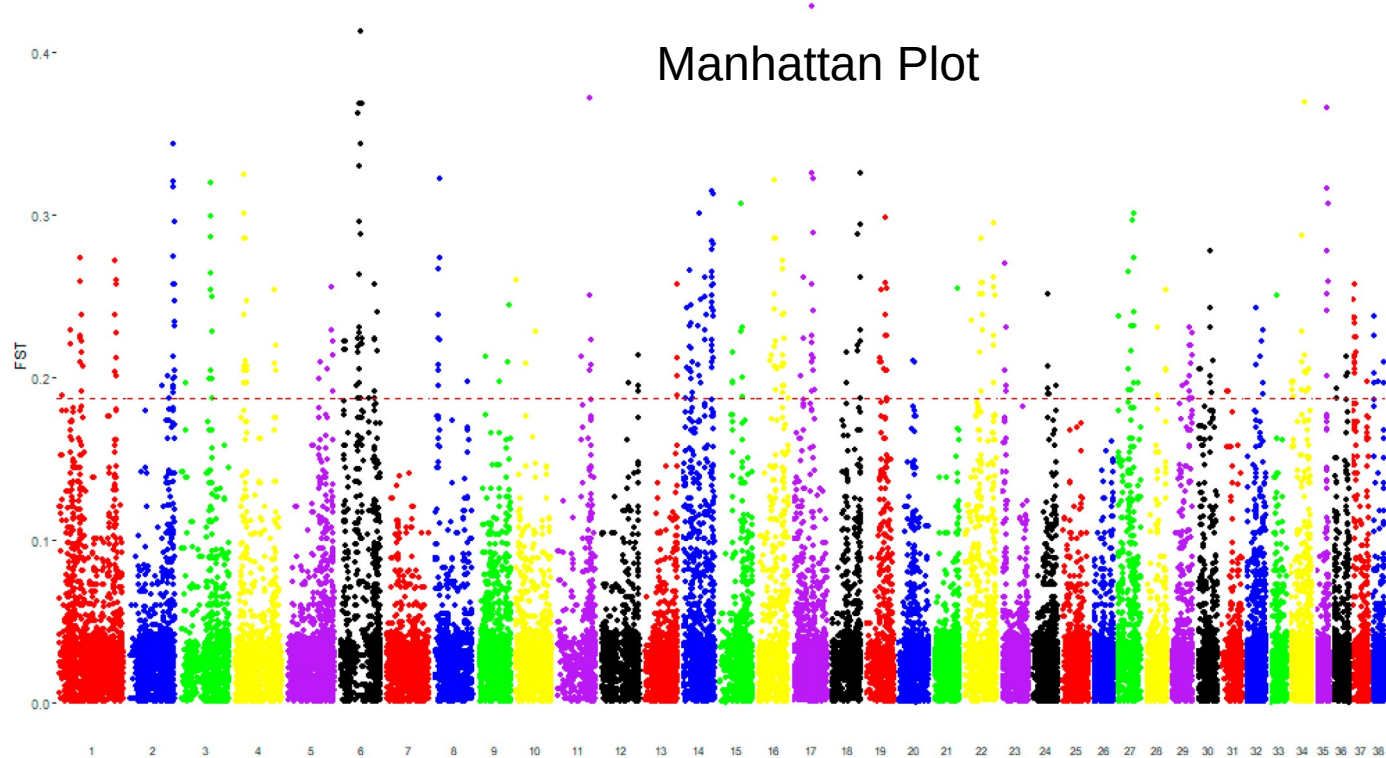


Διαφοροποίηση και το μοντέλο stepping stone



Philip W. Hedrick.
Genetics of
Populations, 4th
Edition (2010)

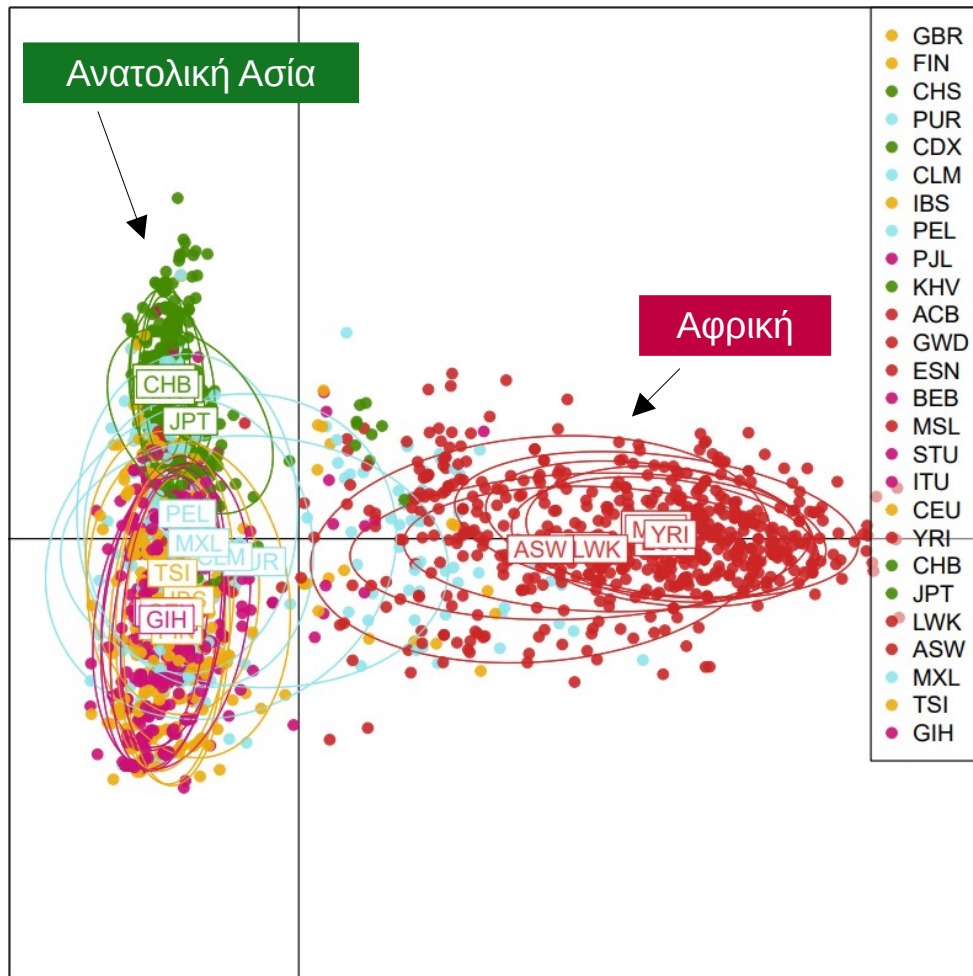
Μελέτες συσχέτισης GWAS



Bionda, A.; Cortellari, M.; Bagardi, M.; Frattini, S.; Negro, A.; Locatelli, C.; Brambilla, P.G.; Crepaldi, P.
A Genomic Study of Myxomatous Mitral Valve Disease in Cavalier King Charles Spaniels. *Animals* **2020**, *10*, 1895.

Κατάταξη σε γενετικές ομάδες

- Συχνά δεν γνωρίζουμε από πριν αν ένας πληθυσμός έχει γενετική δομή ή όχι
- Με μια σειρά αναλύσεων μπορούμε να διακρίνουμε την ύπαρξη γενετικών ομάδων (clusters) μέσα σε ένα σύνολο δεδομένων
 - PCA based
 - Model based

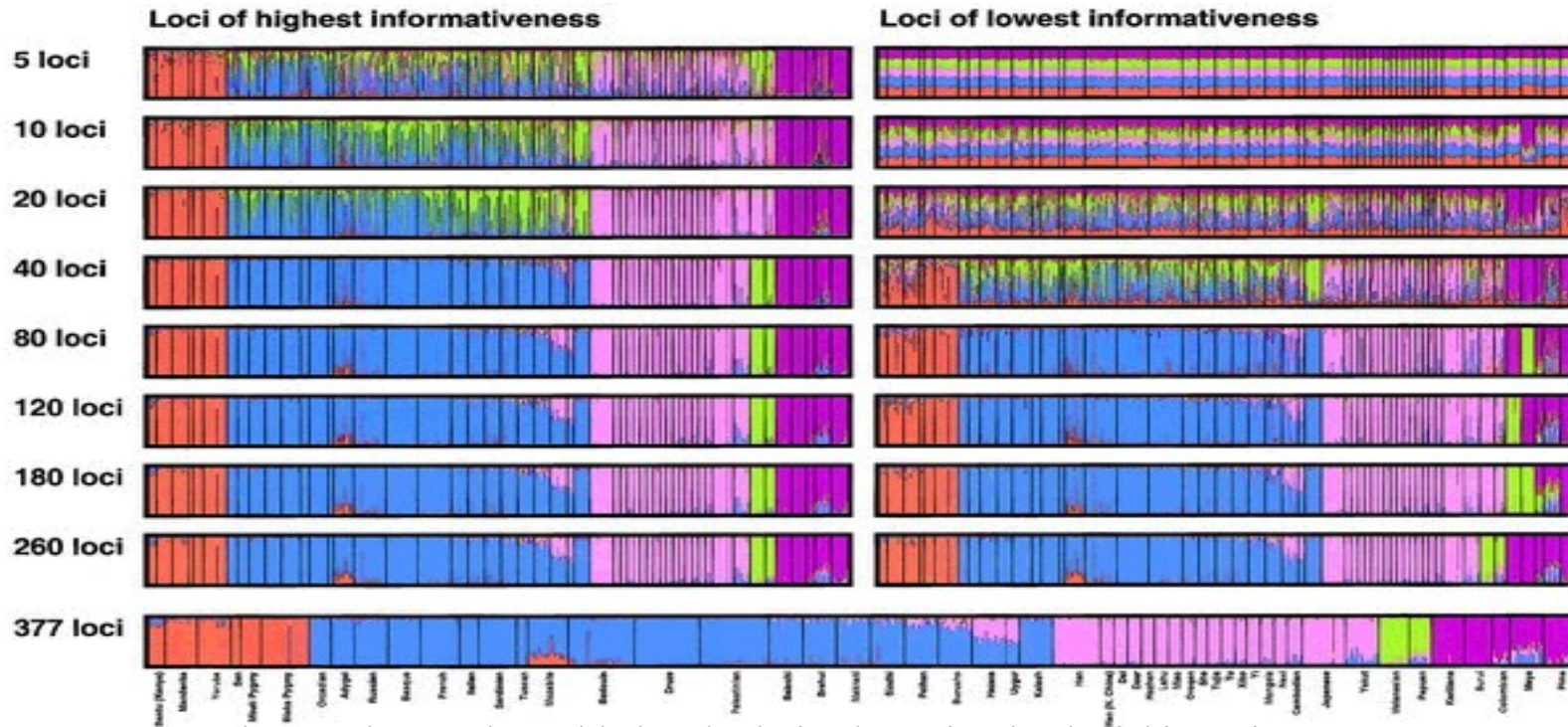


PCA - based

Ανάλυση διακρίτοτητας βασισμένη σε PCA ανάμεσα σε 25 ανθρώπινους πληθυσμούς για τα γονίδια LZTFL1 και CCR9, που είναι σημαντικά για τη λειτουργία του αναπνευστικού συστήματος και σχετίζονται με την απόκριση απέναντι σε προσβολή από COVID-19.

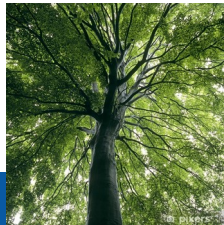
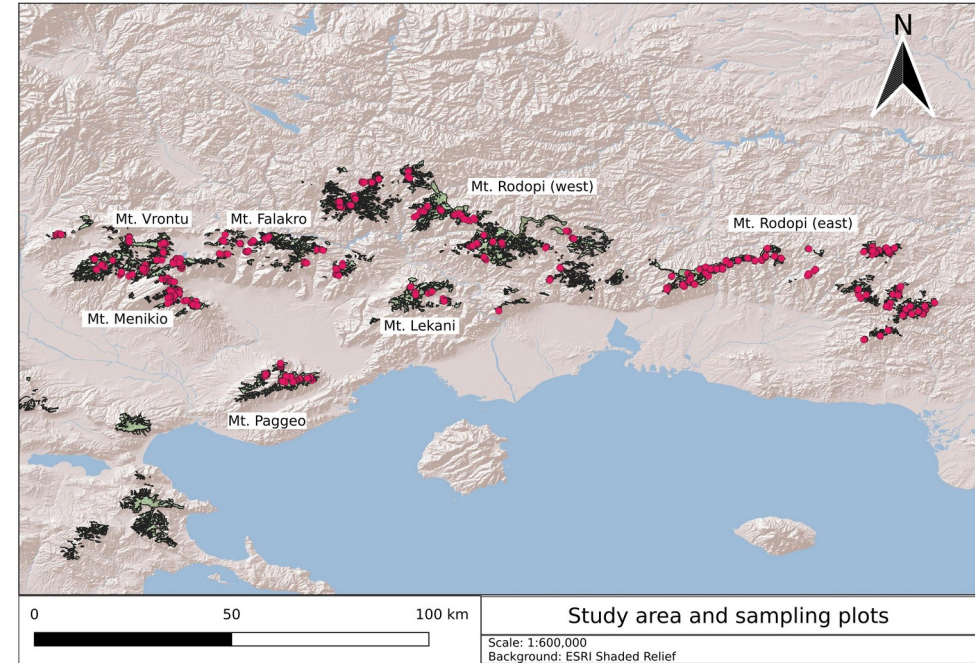
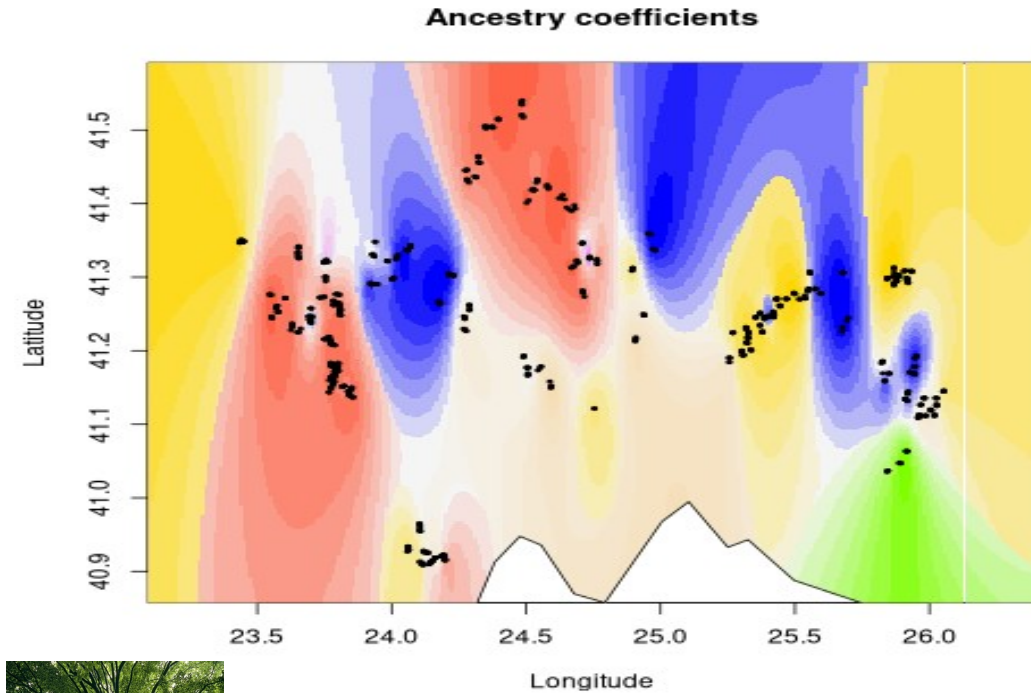
Διακρίνονται καθαρά οι πληθυσμοί αφρικανικής καταγωγής από τους υπόλοιπους ενώ ο δεύτερος άξονας διαχωρίζει τους πληθυσμούς της ανατολικής Ασίας.

Εκτίμηση διαφοροποίησης με Bayesian μοντέλα



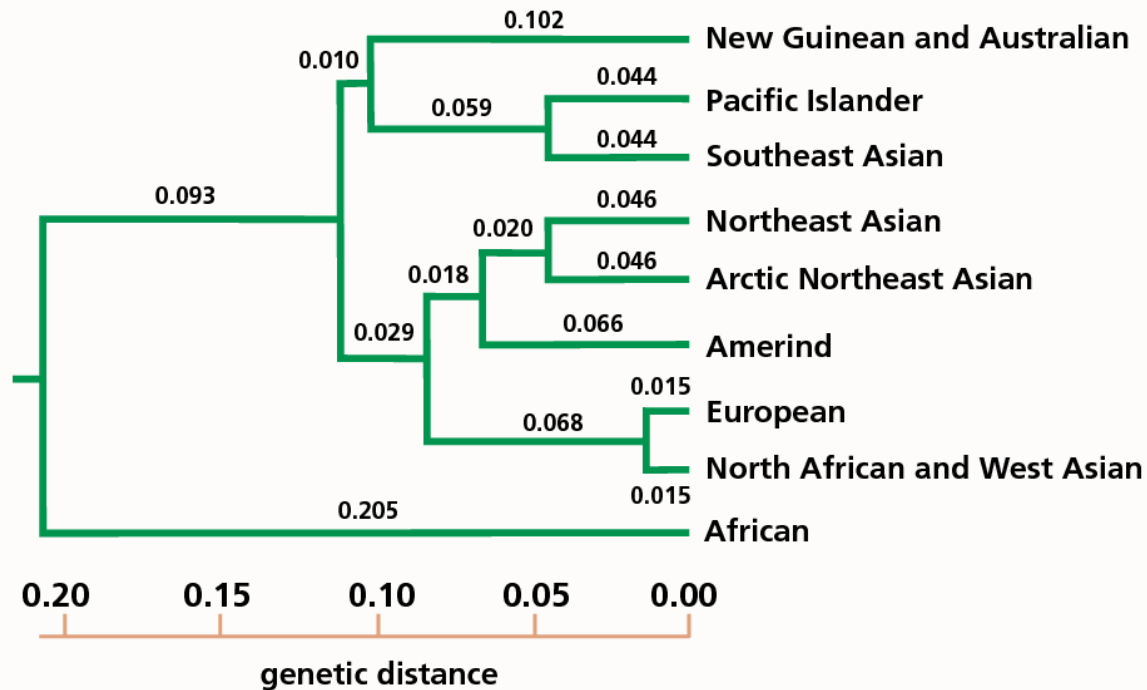
Rosenberg, N. A., Li, L. M., Ward, R., & Pritchard, J. K. (2003). Informativeness of genetic markers for inference of ancestry. *The American Journal of Human Genetics*, 73(6), 1402-1422.

Χωρική εκτίμηση διαφοροποίησης με μοντέλα



Manolis et al. 2021. Geographic distribution of individual beech trees, clustered in six ancestral populations derived from a model-based algorithm.

Απεικόνιση διαφοροποίησης με γενετικές αποστάσεις



- Μέσα από τις γενετικές αποστάσεις για κάθε ζευγάρι πληθυσμών, μπορεί να προκύψει φυλογενετικό δέντρο

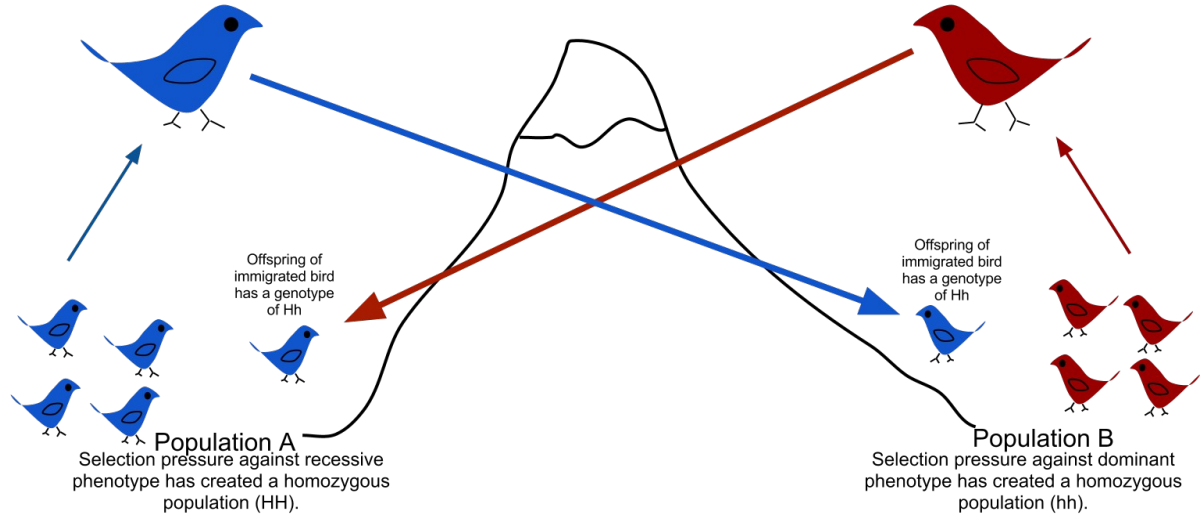
The application of molecular genetic approaches to the study of human evolution
L. Cavalli-Sforza, M. Feldman
2003, Nature Genetics

Γονιδιακή ροή και φυσική επιλογή

- Έστω ότι ένα αλληλόμορφο είναι επιλεκτικά υπολειπόμενο στον πληθυσμό και διατηρείται σε χαμηλή συχνότητα
 - Κατευθύνουσα επιλογή
- Το αλληλόμορφο αυτό είναι δυνατό να φτάσει στον πληθυσμό μέσω μετανάστευσης (γονιδιακής ροής) από άλλους πληθυσμούς, όπου το αλληλόμορφο βρίσκεται σε υψηλή συχνότητα
- Η ροή γονιδίων **καθυστερεί** την προσαρμογή στην περίπτωση αυτή
 - Όμως εμπλουτίζει συνεχώς τη γενετική ποικιλότητα του πληθυσμού

Ισορροπία μεταξύ γονιδιακής ροής & επιλογής

- Διαφορετικά αλληλόμορφα ευνοούνται σε διαφορετικούς πληθυσμούς και περιβάλλοντα
- Η ποικιλότητα διατηρείται χάρη στη γονιδιακή ροή



Κλιμακωτή διάταξη / κλινές

- Σε πολλούς γεωγραφικά διαφοροποιημένους πληθυσμούς παρατηρείται κλιμακωτή γενετική δομή ή διάταξη
- Συμβαίνει κυρίως όταν το περιβάλλον μεταβάλλεται βαθμηδόν κατά μήκος της γεωγραφικής κατανομής του είδους και η επιλογή επιβάλλει ένα συνεχές φάσμα γονιδιακών συχνοτήτων

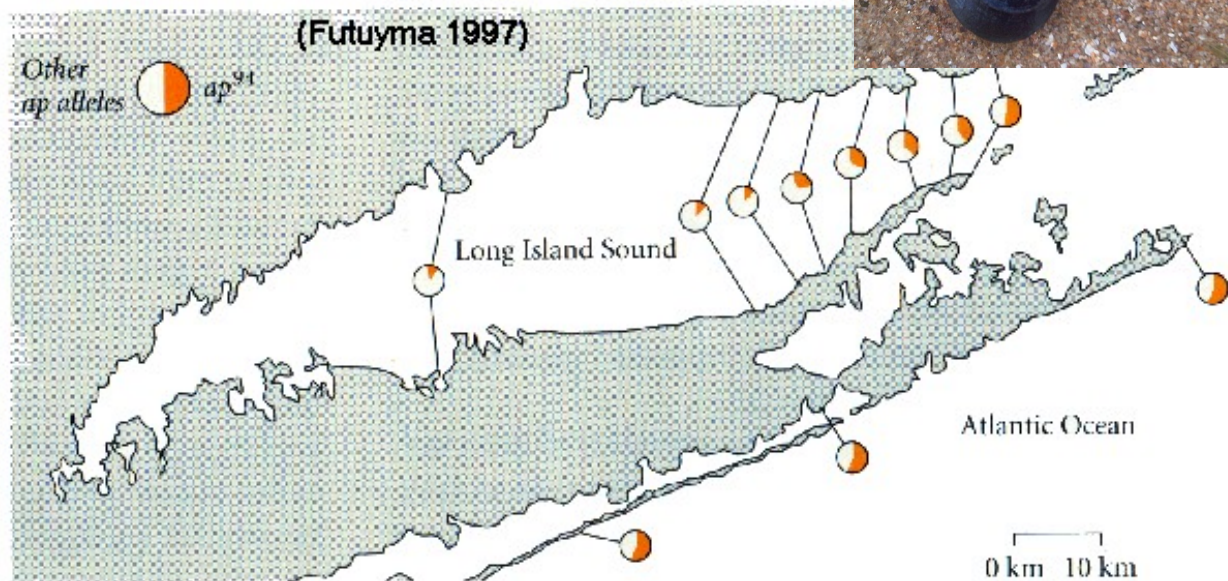


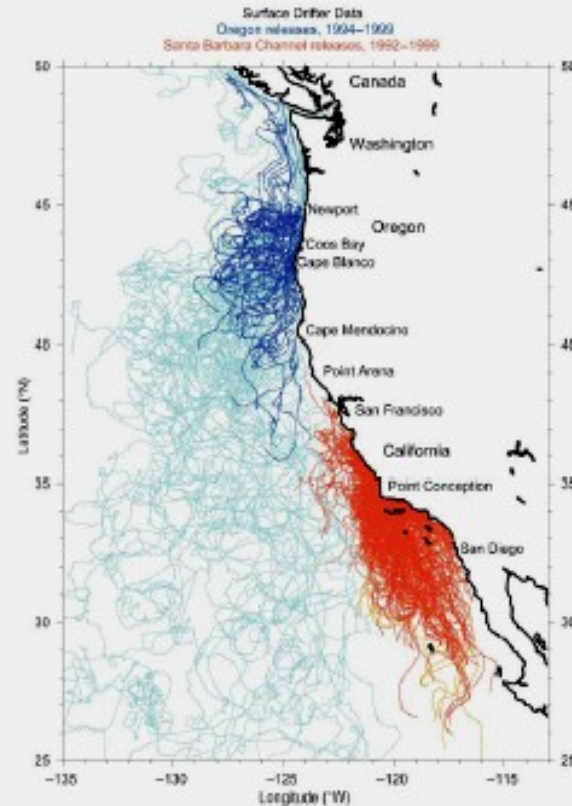
FIGURE 13.20 The frequency of the ap^{94} allele, indicated by the dark portion of each circle, in samples of the mussel *Mytilus edulis* in Long Island Sound and nearby sites. The

frequency of the allele drops rapidly over a 30-kilometer distance in the Sound, where salinity decreases from east to west, despite high gene flow. (After Koehn and Hilbish 1987.)



Κλιμακωτή διάταξη / κλινές

Fig.2 Frequencies of haplotypes of COI and EF1 α for populations from Vancouver Island to central California. Haplotype groups are described in the text and Fig. 1. See Table 1 for key to abbreviations of locations and sample sizes.



*Balanus
grandula*

Γονιδιακή ροή και γενετική εκτροπή

- Ισορροπία γονιδιακής ροής - γενετικής εκτροπής
 - Η γενετική εκτροπή τείνει να επιβάλει μονομορφισμό σε κάποιους πληθυσμούς, αλλά για διαφορετικά αλληλόμορφα σε διαφορετικούς πληθυσμούς
 - Τυχαία δύναμη
 - Η γονιδιακή ροή εισάγει νέα αλληλόμορφα σε έναν πληθυσμό από άλλους πληθυσμούς
- Η γονιδιακή ροή δρα **αντίθετα** στη γενετική εκτροπή
 - Διατηρεί τη γενετική ποικιλότητα μέσα στους πληθυσμούς
 - Αποτρέπει τη διαφοροποίηση ανάμεσα τους πληθυσμούς

Ισορροπία ροής γονιδίων & εκτροπής

- Πότε επέρχεται ισορροπία μεταξύ ροής γονιδίων και γενετικής εκτροπής στο νησιωτικό μοντέλο;
 - Αυτό συμβαίνει στην περίπτωση που η διαφοροποίηση μεταξύ των υποπληθυσμών είναι:
- Και επειδή οι ποσότητες m^2 και $2m$ είναι πολύ μικρές μπορούμε να πούμε ότι:

$$F_{ST} = \frac{(1 - m)^2}{2N - (2N - 1)(1 - m)^2}$$

$$F_{ST} \approx \frac{1}{4Nm + 1}$$

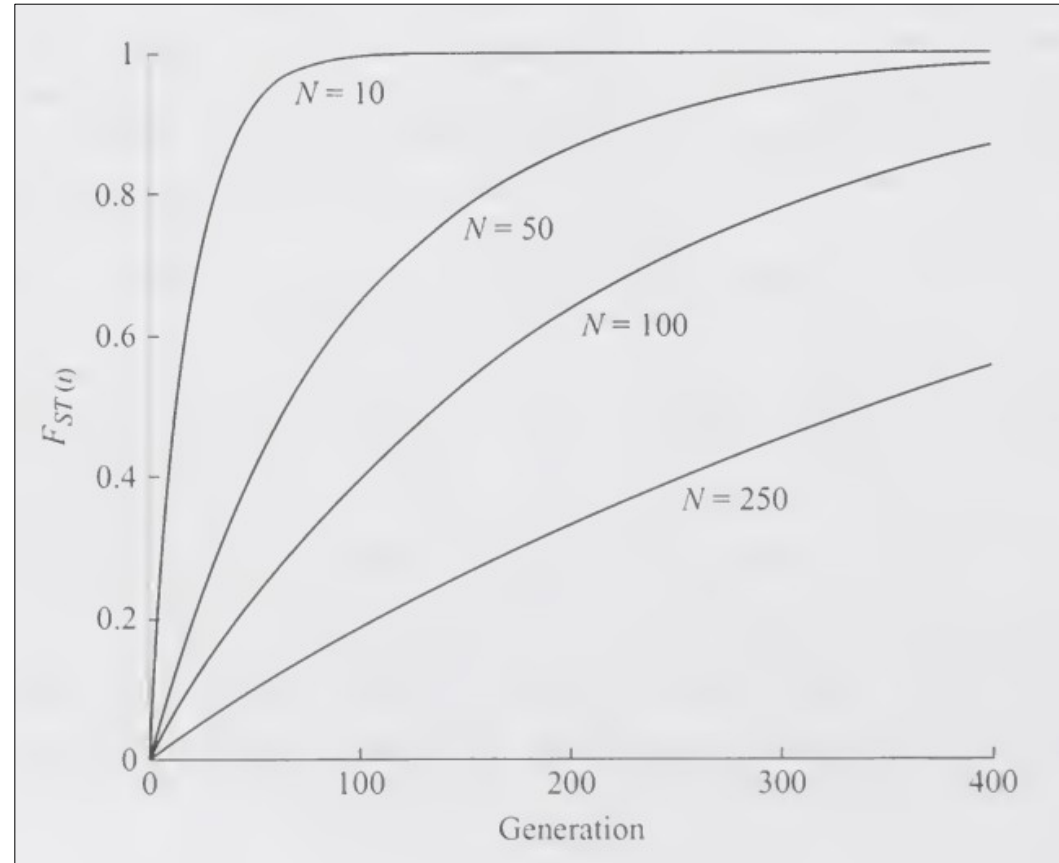
Ισορροπία ροής γονιδίων & εκτροπής

- Από την προηγούμενη σχέση μπορούμε να υπολογίσουμε τον αριθμό των ατόμων που μετανάστευσαν ανά γενιά:
- Ο συγκεκριμένος τύπος χρησιμοποιείται συχνά για τον υπολογισμό της μετανάστευσης στα ζώα
 - Προσοχή: Το μοντέλο είναι θεωρητικό!

$$Nm \cong \left(\frac{1 - F_{st}}{4F_{st}} \right)$$

Γενετική εκτροπή και διαφοροποίηση

- Όταν απουσιάζει η ροή γονιδίων, σε ένα σύνολο πληθυσμών πεπερασμένου μεγέθους, η γενετική εκτροπή οδηγεί σε αύξηση της διαφοροποίησης
- Στην εικόνα, ένα σύνολο πληθυσμών ξεκινά με $F_{ST}=0$ και μηδενική ροή γονιδίων
- Ανάλογα με το μέγεθος των πληθυσμών, η διαφοροποίηση αυξάνει στον χρόνο



Γενετική εκτροπή και διαφοροποίηση

- Η συνύπαρξη ροής γονιδίων και εκτροπής οδηγεί σε ισορροπία
- Οι δύο δυνάμεις δρουν αντίθετα
- Στην εικόνα βλέπουμε τρεις περιπτώσεις ροής γονιδίων (m) σε πληθυσμούς διαφορετικών μεγεθών (N)
- Οι κύκλοι συμβολίζουν τον μέσο χρόνο επίτευξης της ισορροπίας

