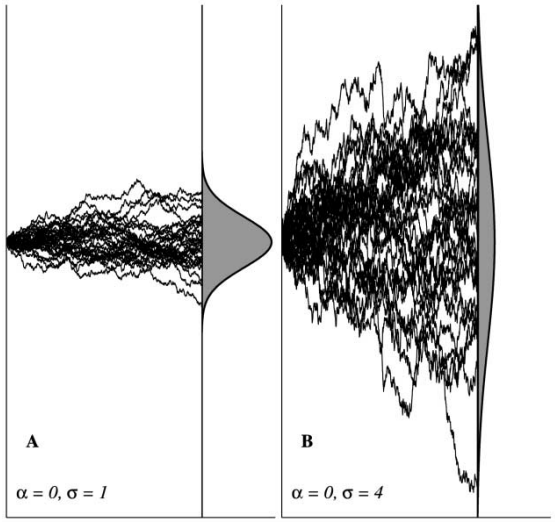
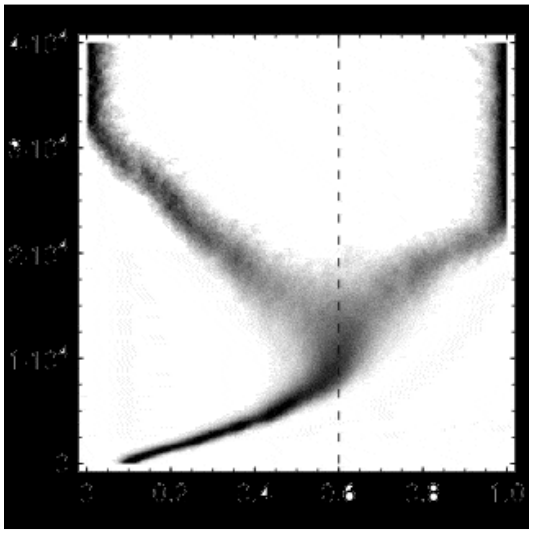
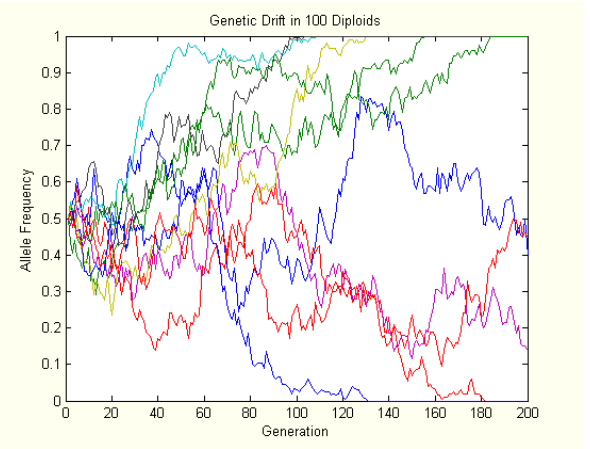
CHANGES IN VARIATION WITHIN AND BETWEEN LINES







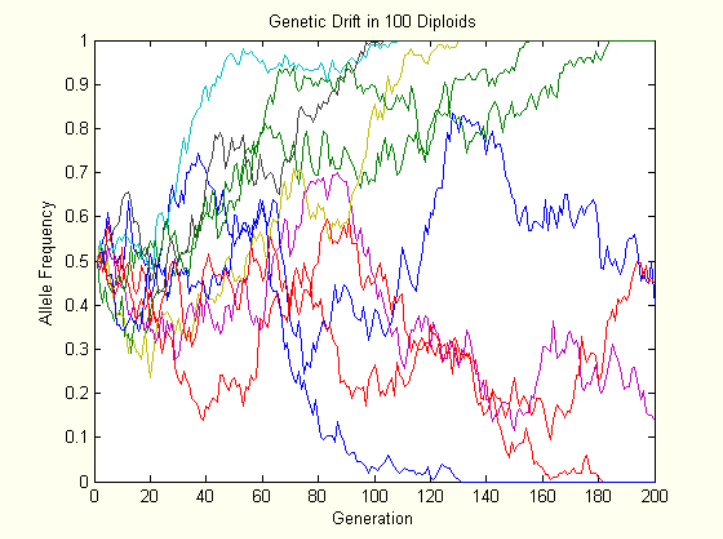
Tom Van Dooren 11/2012
vandoore@biologie.ens.fr

1. Changes in G during short-term selection

2. Interactions between loci

3. Changes by drift - finite populations

4. Variation between populations



1. MULTIVARIATE SELECTION WITH GAUSSIAN SELECTION/FITNESS FUNCTIONS

Fitness function

$$W(\boldsymbol{z}) = \exp(\boldsymbol{\alpha}^{T}\boldsymbol{z} - \tfrac{1}{2}(\boldsymbol{z} - \boldsymbol{\theta})^{T}\boldsymbol{W}(\boldsymbol{z} - \boldsymbol{\theta})) = \sum_{i}\alpha_{i}z_{i} - \frac{1}{2}\sum_{i}\sum_{j}(z_{i} - \theta_{i})(z_{j} - \theta_{j})W_{ij}$$

(1) Directional selection occurs when $\boldsymbol{\alpha} \neq 0$ and/or $\boldsymbol{\mu}_{z} \neq \boldsymbol{\theta}$
(2) Stabilising selection occurs, for example, when $\boldsymbol{W}$ is a diagonal matrix and $W_{ii} > 0$
(3) Disruptive selection occurs when $\boldsymbol{W}$ is a diagonal matrix and $W_{ii} < 0$

These types of selection can occur jointly!

Selection modifies the distribution of phenotypes

$p(z)$ the probability distribution of phenotypes before selection,
$p^*(z)$ the probability distribution of phenotypes after selection

$$p^*(z) = \frac{p(z)W(z)}{\int dz p(z)W(z)} = cp(z)W(z)$$

$$p^*(z) = cp(z)W(z)$$

$$= c(2\pi)^{-n/2}|\boldsymbol{P}|^{-1/2}\exp(-\frac{1}{2}(z-\boldsymbol{\mu}_z)^T\boldsymbol{P}^{-1}(z-\boldsymbol{\mu}_z))\cdot\exp(\boldsymbol{\alpha}^T z-\frac{1}{2}(z-\boldsymbol{\theta})^T\boldsymbol{W}(z-\boldsymbol{\theta}))$$

$$= c(2\pi)^{-n/2}|\boldsymbol{P}|^{-1/2}\exp\left(\boldsymbol{\alpha}^T z-\frac{1}{2}\left((z-\boldsymbol{\mu}_z)^T\boldsymbol{P}^{-1}(z-\boldsymbol{\mu}_z)+(z-\boldsymbol{\theta})^T\boldsymbol{W}(z-\boldsymbol{\theta})\right)\right)$$

If $p^*(z)$ is MVN, then we can write it as

$$p^*(z) = (2\pi)^{-n/2}\left|\boldsymbol{P}^*\right|^{-1/2}\exp\left(-\frac{1}{2}(z-\boldsymbol{\mu}_z^*)^T\boldsymbol{P}^{*-1}(z-\boldsymbol{\mu}_z^*)\right)$$

Given this assumption, we can derive the mean and variance after selection

$$\boldsymbol{\mu}_z^* = \boldsymbol{P}^*(\boldsymbol{W}\boldsymbol{\theta}+\boldsymbol{\alpha}+\boldsymbol{P}^{-1}\boldsymbol{\mu}_z)$$

$$\boldsymbol{P}^* = (\boldsymbol{W}+\boldsymbol{P}^{-1})^{-1} = \boldsymbol{P}(\boldsymbol{P}+\boldsymbol{W}^{-1})^{-1}\boldsymbol{W}^{-1}$$

Mean fitness will be

$$\overline{W}(z) = \sqrt{\frac{\left|\boldsymbol{P}^*\right|}{\left|\boldsymbol{P}\right|}}\exp\left(-\frac{1}{2}(\boldsymbol{\mu}_z^{\,T}\boldsymbol{P}^{-1}\boldsymbol{\mu}_z+\boldsymbol{\theta}^T\boldsymbol{W}\boldsymbol{\theta}-\boldsymbol{\mu}_z^{\,*T}\boldsymbol{P}^{*-1}\boldsymbol{\mu}_z^{\,*})\right)$$

business as usual:

P* is normal g* is normal, ... is normal, ... is normal, ...

The selection differential $\boldsymbol{s} = \boldsymbol{\mu}_z^* - \boldsymbol{\mu}_z$

$$\boldsymbol{s} = \boldsymbol{P}^*(\boldsymbol{W}(\boldsymbol{\theta} - \boldsymbol{\mu}_z) + \boldsymbol{\alpha})$$

and

$$\boldsymbol{G}^* = \boldsymbol{G} - \boldsymbol{G}(\boldsymbol{P} + \boldsymbol{W}^{-1})^{-1}\boldsymbol{G}$$
$$\boldsymbol{G}^* = \boldsymbol{G}(\boldsymbol{P} + \boldsymbol{W}^{-1})^{-1}(\boldsymbol{E} + \boldsymbol{W}^{-1})$$

$$\mu_z(t+1) = \mu_z(t) + \boldsymbol{G}(t)(\boldsymbol{P}(t) + \boldsymbol{W}^{-1})^{-1}\boldsymbol{W}^{-1}(\boldsymbol{W}(\theta - \mu_z(t)) + \alpha)$$

and

$$\boldsymbol{G}(t+1) = \boldsymbol{G}(t)/2 + \boldsymbol{G}(t)(\boldsymbol{P}(t) + \boldsymbol{W}^{-1})^{-1}(\boldsymbol{E} + \boldsymbol{W}^{-1})/2$$

or

$$\boldsymbol{G}(t+1) = 3\boldsymbol{G}(t)/2 - \boldsymbol{G}(t)(\boldsymbol{P}(t) + \boldsymbol{W}^{-1})^{-1}\boldsymbol{G}/2$$

- The genetic variance changes after selection, values in ***W*** determine whether it will become wider or narrower for a trait.

- The genetic variance after selection can increase for one trait and decrease for another

Disruptive selection and then what? (Rueffler et al. 2006)

Table 1. Adaptive responses favored by disruptive selection

Responses	Favorable conditions
Changes in the ecological community	
Immigration by similar species	Availability of species and high migration rates
Destabilization of coevolutionary equilibrium into an evolutionary arms race	High variance of trait(s) in species under disruptive selection
Increase in genetic variation	
Frequencies of alleles at polymorphic loci are equalized	Additive polygenic inheritance
Positive linkage disequilibria	Low recombination rates
Increased effect size of alleles at polymorphic loci	No constraints on evolution of effect sizes
Decreased recombination rates	Linkage disequilibrium present
Dominance modification	Available genetic variation in shape of genotype-phenotype map
Assortative mate choice	Mating in the selective environment
Individual specialization with a genetic basis	(Not known at present)
Increase in phenotypic variation	
Sexual dimorphism	No developmental constraints on phenotypic divergence between sexes
Phenotypic plasticity	Reliable environmental cues, low cost of plasticity, high rates of gene flow
Bet-hedging strategies	No reliable cues and stochastic environments
Increase in individual niche width	Low costs of generalism
Specialization through learning	Behavioral flexibility through high learning and cognition abilities

The other side of the same coin:

Change in environmental variance (Gavrilets and Hastings 1994)

The strength of susceptibility to environmental variation shows genetic variation

$$z = g + \gamma_g \varepsilon$$

$$P = G + E_g$$

$$G = \mathrm{var}(G)$$

$$E_g = (\bar{\gamma}_g^{\,2} + \mathrm{var}(\gamma_g))\sigma_\varepsilon^{\,2}$$

$$z = g + \gamma_g \varepsilon$$

Quadratic fitness function

$$W(z) = 1 - s(z - \theta)^2$$

Expected fitness of a genotype is

$$\int d\varepsilon f(\varepsilon) w(z) = 1 - s\left((g - \theta)^2 + \gamma_g{}^2 \sigma_\varepsilon{}^2\right)$$

There is selection on environmental sensitivity

$s < 0$ it will increase
$s > 0$ it will decrease


DEVELOPMENTAL PLASTICITY AND EVOLUTION
MARY JANE WEST-EBERHARD
Plasticity, Robustness, Development and Evolution
Patrick Bateson and Peter Gluckman
CAMBRIDGE
OXFORD
Andreas Wagner
THE ORIGINS OF EVOLUTIONARY INNOVATIONS

2. LINKAGE – EFFECTS OF COMBINATIONS OF LOCI

The famous infinitesimal model:

Huge number of loci, each with very small effects, negligible changes in allele frequencies due to selection.

With random mating, equal selection on both sexes, traits in parents and offspring multivariate normal distributions

$$\Delta \boldsymbol{G} = \boldsymbol{G}^* - \boldsymbol{G} = \boldsymbol{G}\boldsymbol{P}^{-1}(\boldsymbol{P}^* - \boldsymbol{P})\boldsymbol{P}^{-1}\boldsymbol{G}$$

$$\boldsymbol{G}_o = \boldsymbol{G} + \frac{1}{2}(\boldsymbol{G}^* - \boldsymbol{G})$$

$$\boldsymbol{G}_o - \boldsymbol{G} = \frac{1}{2}(\boldsymbol{G}^* - \boldsymbol{G})$$

Infinitesimal model : **A** does not change! The change in $\boldsymbol{G}$ is due to interactions between loci – effects of gametic-phase linkage disequilibrium

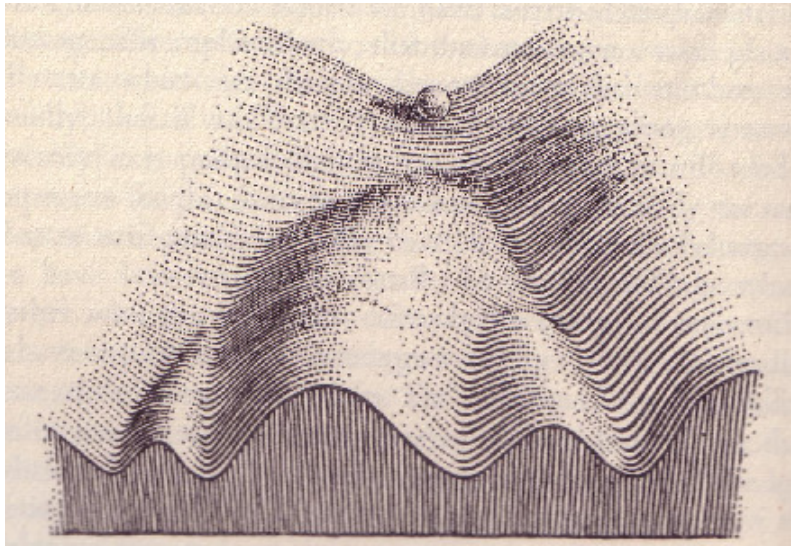
2. LINKAGE – EFFECTS OF COMBINATIONS OF LOCI

Other genetic architectures:

intricate effects expected

'conversions' between variance components

Waddington (1957) The Strategy of the Genes

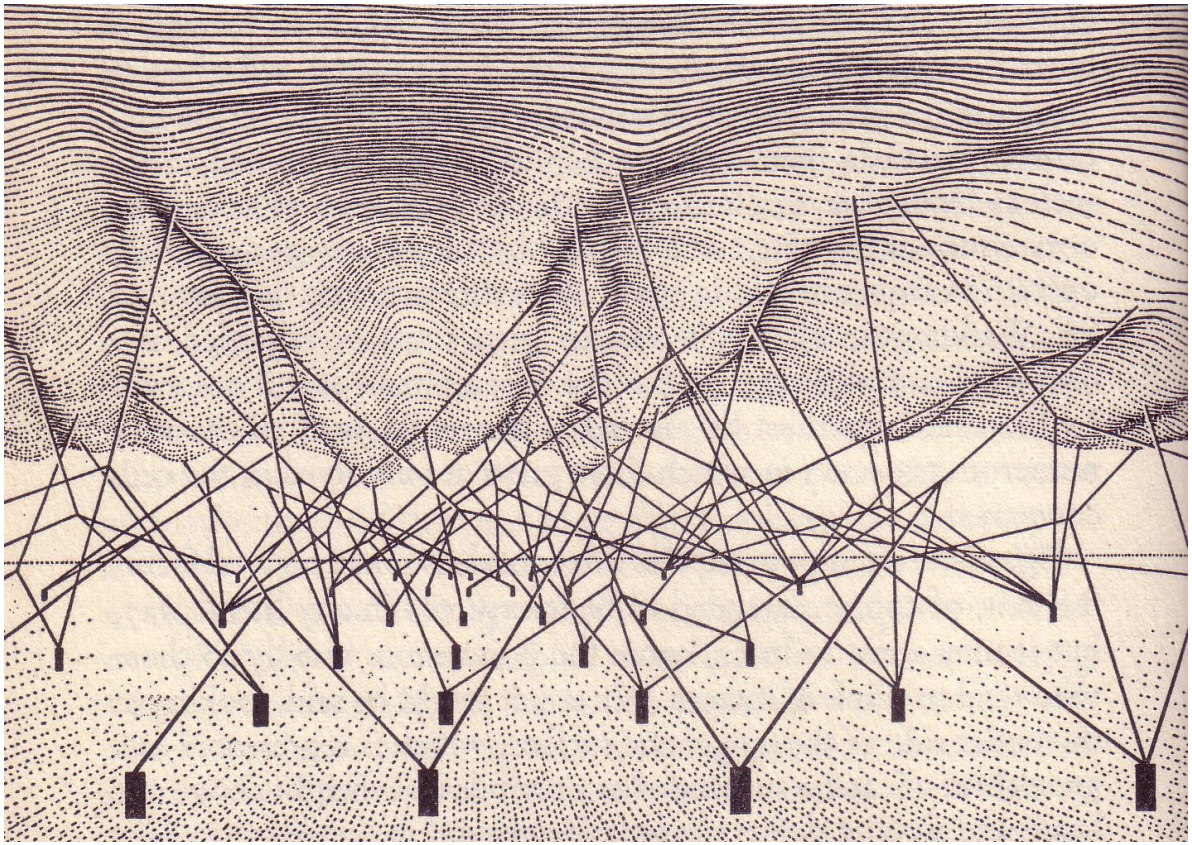


The epigenetic landscape

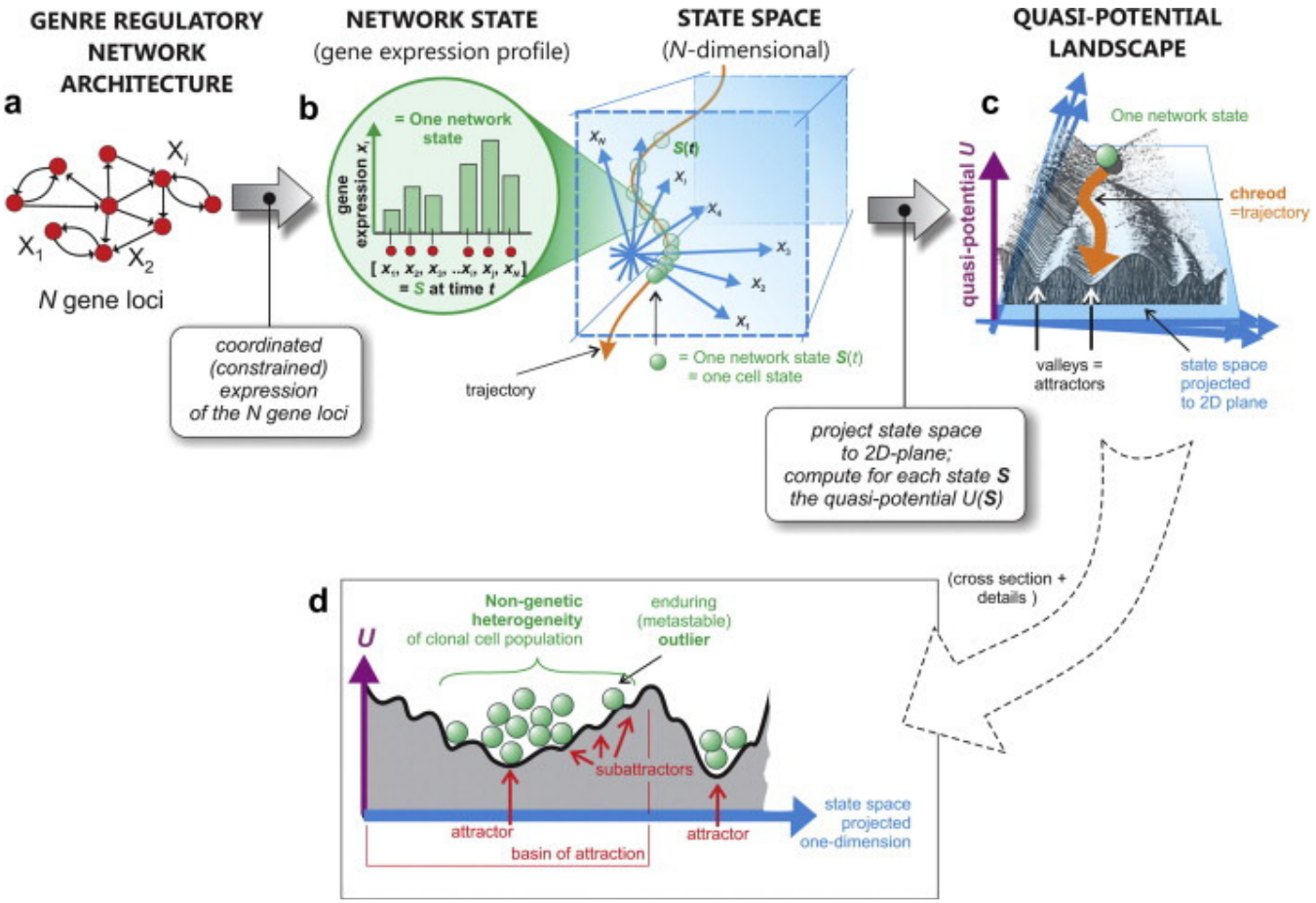
Developmental switches and variation between very similar trajectories

Many dimensions, one for each trait

Waddington (1957) The Strategy of the Genes



- gene (pegs) and gene products determine the epigenetic landscape
- there are many interactions



Huang S. (2012) Tumor progression: Chance and necessity in Darwinian and Lamarckian somatic (mutationless) evolution. doi: 10.1016/j.pbiomolbio.2012.05.001.

3. FINITE POPULATIONS - DRIFT

A concept to preserve an ideal (Wright 1931, 1938):

"the number of breeding individuals in an idealized population that would show the same amount of dispersion of allele frequencies under random genetic drift or the same amount of inbreeding as the population under consideration"

idealized population: monoecious, diploid, random mating

Inbreeding effective population size $N_e^{(i)}$

Variance effective population size $N_e^{(v)}$

Often identical or very similar

Inbreeding coefficient f

panmictic - random mating - selfing allowed- no selection

$$f_t = \frac{1}{2N_e^{(i)}} + \left(1 - \frac{1}{2N_e^{(i)}}\right) f_{t-1}$$

$f_0 = 0$ gives

$$f_t = 1 - \left(1 - \frac{1}{2N_e^{(i)}}\right)^t$$

Variance per locus within populations - no selection

Diallel loci, two alleles with effects 0 (frequency 1-p_i) and a_i (frequency p_i)

(Lynch and Walsh 1997)

$$\sigma_A^{\,2}(t) = 2\sum_i E\left[a_i p_i(t)(1-p_i(t))\right]$$

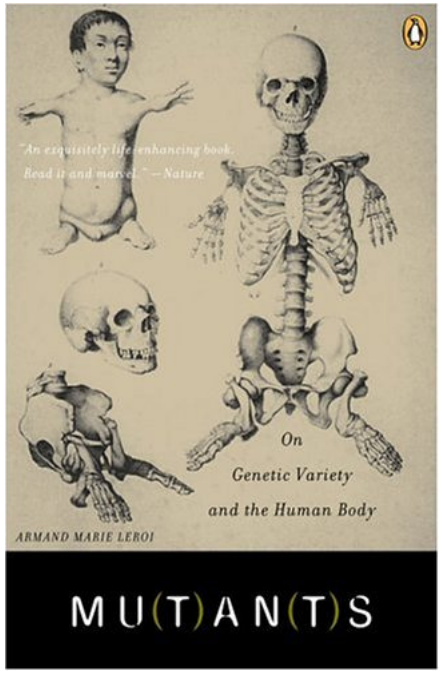
$$\sigma_A^{\,2}(t) = 2\sum_i E\left[a_i p_i(t-1)(1-p_i(t-1))\right]\left(1-\frac{1}{2N_e^{(v)}}\right)$$

The standing genetic variance decreases per generation due to drift

Symmetric mutation

$$\sigma_M{}^2 = 2\sum_i u_i E[a_i{}^2]$$

a squared heterozygotic effect of mutant in locus *i*



Purely additive genetics, no selection

mutation replenishes variation

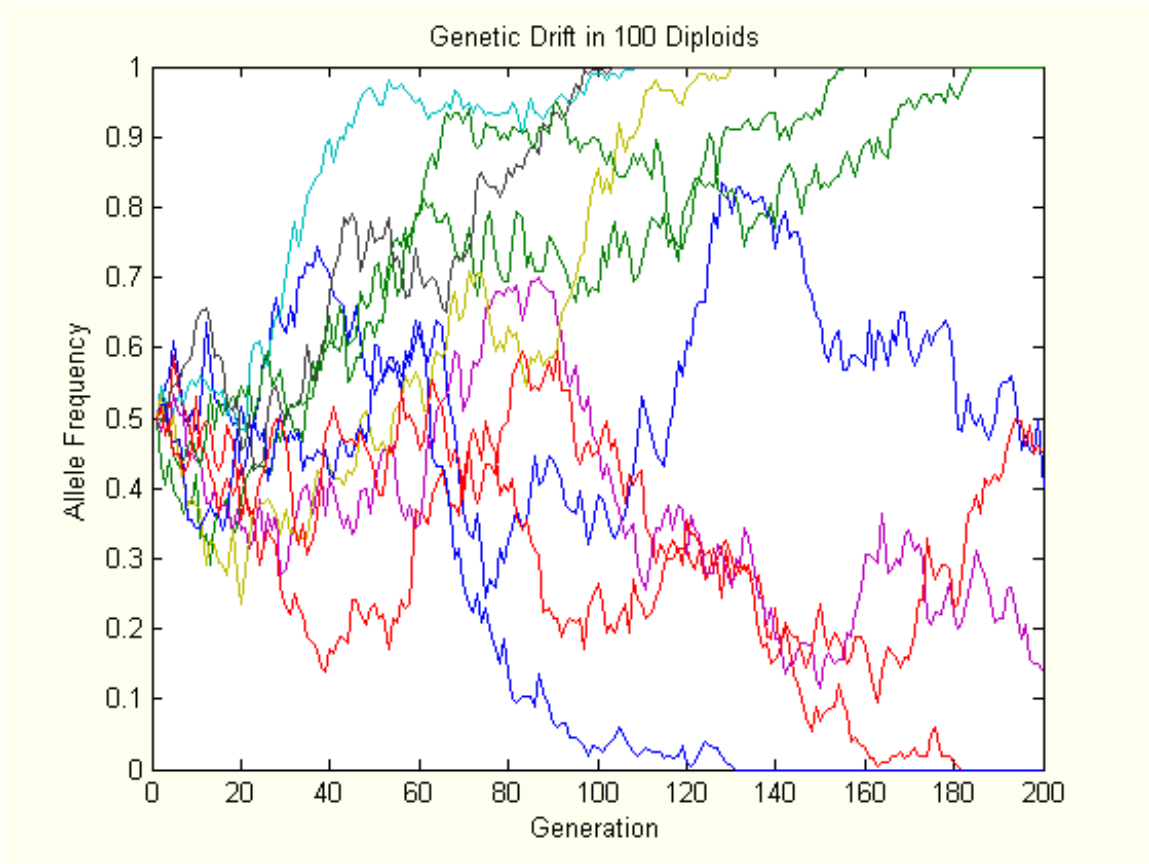
$$\sigma_A{}^2(t+1) = \left(1 - \frac{1}{2N_e{}^{(v)}}\right)\sigma_A{}^2(t) + \sigma_M{}^2$$

The equilibrium genetic variance is

$$\sigma_A{}^2 = 2N_e{}^{(v)}\sigma_M{}^2$$

(substitute it in as a trial solution)

Variation between populations – no mutation – no selection



Additive gene action – random genetic drift

Diallel loci, two alleles with effects 0 (frequency 1-p_i) and a_i (frequency p_i)

Average genotypic value at each locus is $2a_ip_i$

Variance among populations per locus is

$$E\left[(2a_i p_i(t))^2\right] - \left(E\left[2a_i p_i(t)\right]\right)^2$$

$$4a_i^2 E\left[p_i(t)^2\right] - 4a_i^2 \left(E\left[p_i(t)\right]\right)^2 = 4a_i^2 \sigma^2_{p_i}(t)$$

Variance of allele frequencies

$$\sigma_{p_i}^2(t) = E\left[p_i(t)^2\right] - \left(E\left[p_i(t)\right]\right)^2$$
$$= E\left[p_i(t)^2\right] - E\left[p_i(t)\right] - \left(E\left[p_i(t)\right]\right)^2 + E\left[p_i(t)\right]$$
$$= -E\left[p_i(t)\left(1 - p_i(t)\right)\right] + E\left[p_i(t)\right]\left(1 - E\left[p_i(t)\right]\right)$$

$$\sigma_{p_i}^2(t) = p_i(0)\left(1 - p_i(0)\right)\left(1 - \left(1 - \frac{1}{2N_e^{(v)}}\right)^t\right)$$

Total additive genetic variance between populations

$$\sigma_B^2(t) = 4\sum_i a_i{}^2 p_i(0)\big(1-p_i(0)\big)\left(1-\left(1-\frac{1}{2N_e{}^{(v)}}\right)^t\right)$$

$$\sigma_B^2(t) = 2\sigma_A^2(0)\left(1-\left(1-\frac{1}{2N_e{}^{(v)}}\right)^t\right)$$

with mutation

Variation within populations has equilibrium value

$$\sigma_A^{\,2} = 2N_e^{(v)}\sigma_M^{\,2}$$

The mean of a population changes per generation with variance

$$2\sigma_M^{\,2}$$

Brownian motion: the variance of the mean increases linearly with time

$$\sigma_B^{\,2} = 4t\sigma_M^{\,2}$$

- Independent of $N_e^{(v)}$

Conclusions

We can build an intuition based on some idealized models

Additional concepts help us translate the results from these models to a broader range of situations

In reality, things seem more in a messy interaction:

NATURE | Vol 441 | 25 May 2006 NEWS FEATURE

WHAT IS A GENE?

The idea of genes as beads on a DNA string is fast fading. Protein-coding sequences have no clear beginning or end and RNA is a key part of the information package, reports **Helen Pearson**.

'Gene' is not a typical four-letter word. It is not offensive. It is never bleeped out of TV shows. And Laurence Hurst at the University of Bath, UK. "All of that information seriously challenges our conventional definition of a gene," says viously unimagined scope of RNA.

The one gene, one protein idea is coming under particular assault from researchers who

(Pearson 2006)

References

Lynch, M and B Walsh. 1997. Genetics and Analysis of Quantitative Genetic Traits. Sinauer Ass.

Bulmer, MG. The mathematical Theory of Quantitative Genetics. Clarendon Press, Oxford.

Gavrilets, S., and Hastings, A. 1994. A quantitative genetic model for selection on developmental noise. Evolution 48: 1478-1486.

http://nitro.biosci.arizona.edu/zbook/NewVolume_2/newvol2.html