

Geographic Population Structure

Alan R. Rogers

December 4, 2013

Drift Inflates Variance among Populations

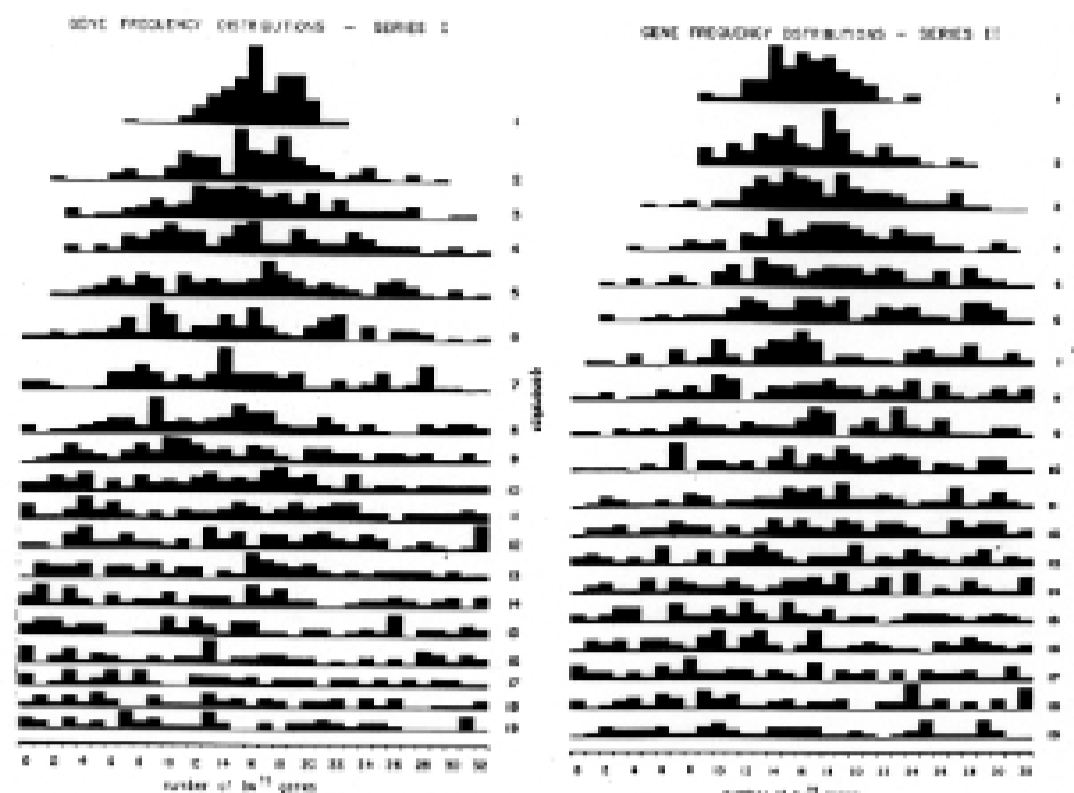


Fig. 6. Gene frequency distributions by generation in series I. Graphical presentation of the data of table 15. Fig. 7. Gene frequency distributions by generation in series II. Graphical presentation of the data of table 16.

Variance among groups increases across generations (Buri 1956)

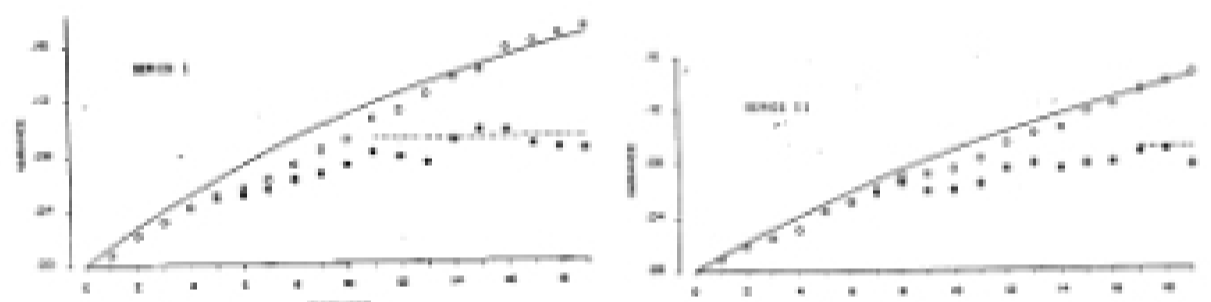


Fig. 12. Theoretical variance of the total frequency distribution by generation including fixed classes and based on a constant variance of 25% = 0.25 for series I. Fig. 13. Theoretical variance of the total frequency distribution by generation including fixed classes and based on a constant variance of 25% = 0.25 for series II.

Wahlund principle: subdivision reduces heterozygosity

	A ₁	A ₁ A ₁	A ₁ A ₂	A ₂ A ₂
Pop 1	4/16	1/16	6/16	9/16
Pop 2	12/16	9/16	6/16	1/16
Species	8/16	5/16	6/16	5/16
Hardy-Weinberg	8/16	4/16	8/16	4/16

Amount of reduction: 2/16

	p _i - p̄	(p _i - p̄) ²
Pop 1	-4/16	1/16
Pop 2	4/16	1/16
Variance:		1/16

$$H_S = H_T - 2V \quad (\text{Wahlund 1928})$$

- ▶ Wahlund showed that heterozygosity is reduced by group differences in allele frequencies.
- ▶ Buri's experiment illustrates that
 - ▶ Drift reduces heterozygosity
 - ▶ Drift increases group differences
- ▶ We need a theory to connect these facts.
- ▶ Let us build one on top of what we have already.

What we already know about heterozygosity

$$E[H_t | p_t] = 2p_t q_t \quad (\text{Hardy-Weinberg})$$

$$E[H_t | p_0] = 2p_0 q_0 \left(1 - \frac{1}{2N}\right)^t \quad (\text{Ch. 2})$$

How can these both be true?
It must be true that

$$E[p_t q_t | p_0] = p_0 q_0 \left(1 - \frac{1}{2N}\right)^t$$

On the other hand, it is also true that

$$E[p_t q_t | p_0] = p_0 q_0 - V_t \quad (\text{Wahlund})$$

where V_t is the variance of p_t about p₀.

Rearranging gives the variance among groups

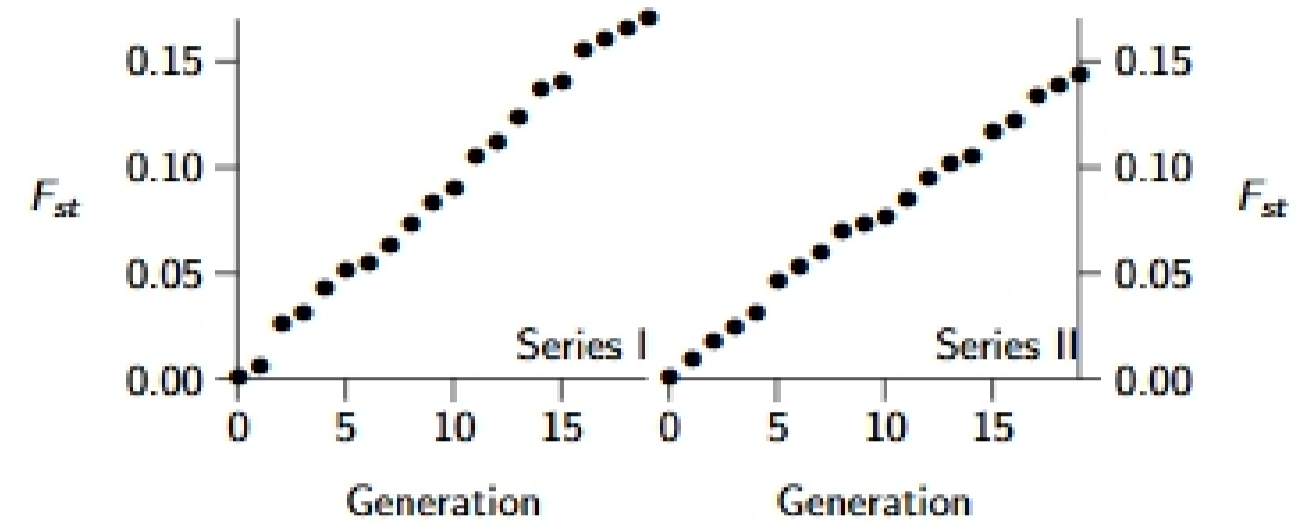
$$V = p_0 q_0 \left[1 - \left(1 - \frac{1}{2N} \right)^t \right]$$

We usually normalize this expression by dividing both sides by $p_0 q_0$. The result is called F_{ST} :

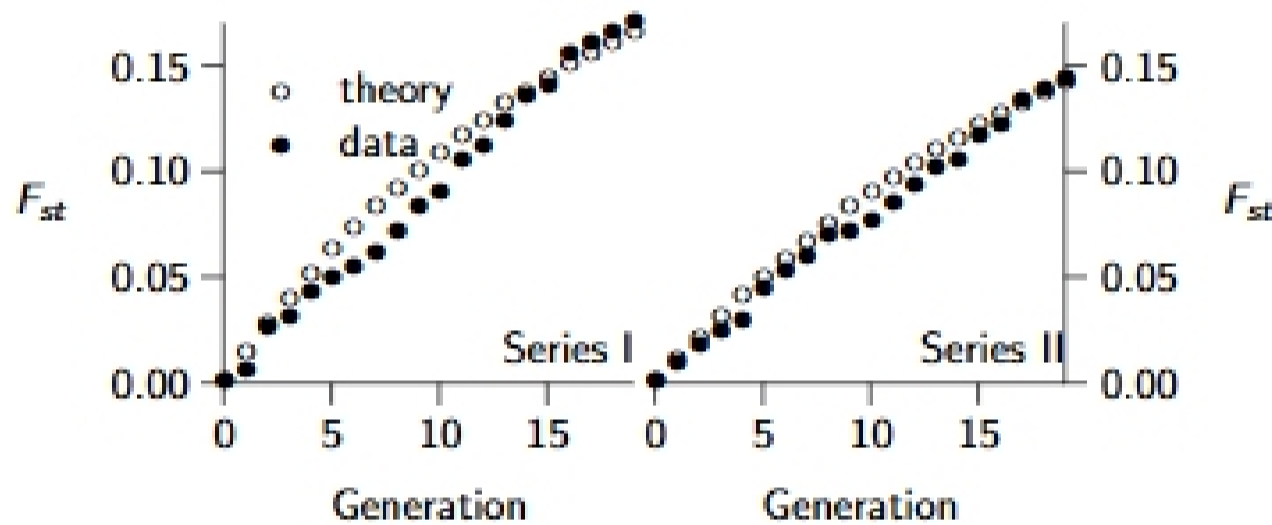
$$F_{ST} = \frac{V}{p_0 q_0} = 1 - \left(1 - \frac{1}{2N} \right)^t$$

In data analysis, we take $p_0 \approx \bar{p}$, the current population mean.

Data from Buri (1956)



Data from Buri (1956)



Application to differences among human races

Theory assumes *no* migration. Seems unlikely, but let's see where it leads.

For major human populations, $F_{ST} \approx 0.1$.

Two major hypotheses about the age human races.

- ▶ Replacement model: about 50,000 years (2000 generations)
- ▶ Multiregional model: at least 1,000,000 years (40,000 generations)

Plug these into

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

and solve for N_e .

What do these imply about N_e of major continental populations?

- ▶ Replacement: $N_e \approx 10,000$
- ▶ Multiregional: $N_e \approx 200,000$

The first figure is more plausible: agrees with genetic diversity within populations.

If human races had been separated for a million years without gene flow, race differences should be larger.

But what about gene flow?

The role of migration

- ▶ Drift increases group differences.
- ▶ Migration (m) reduces them
- ▶ Eventually, an equilibrium is reached.

$$F_{ST} = \frac{1}{4N_e m + 1}$$

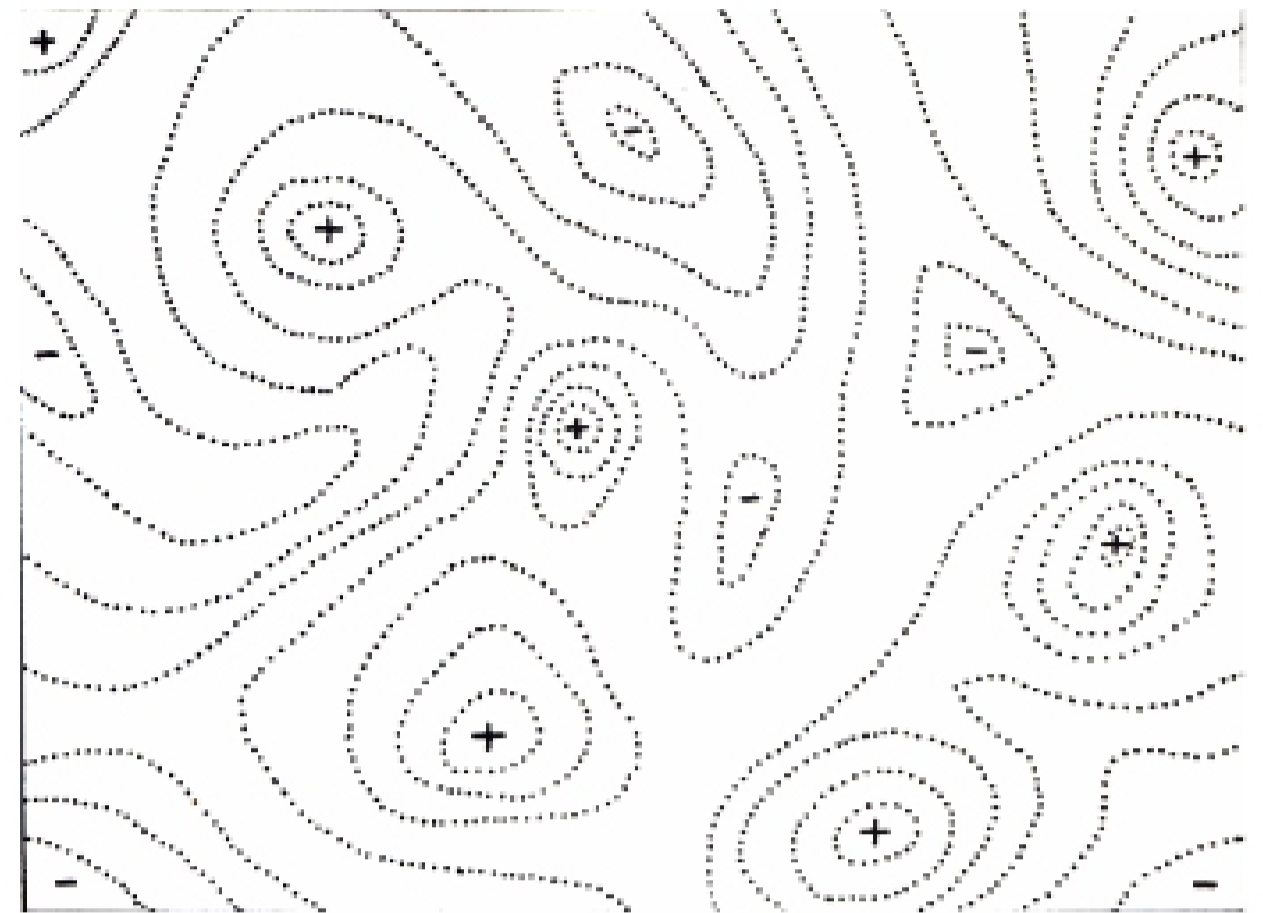
Depends only on $N_e m$. Small if $N_e m > 1$.

Human value ($\approx 1/9$) implies $N_e m \approx 2$.

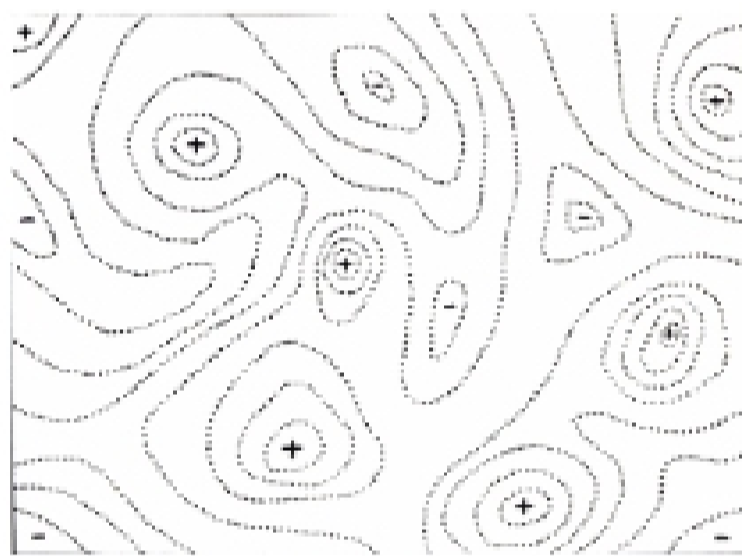
Makes the multiregional model more plausible.

What role does population structure play in adaptive evolution?

Wright's adaptive landscape



Wright's adaptive landscape



Wright's (1932) adaptive landscape. Each (x, y) point represents a multilocus genotype. Contour lines show fitness. Populations tend to move uphill.

Multiple peaks exist because genes do not act additively.

Wright's shifting balance theory

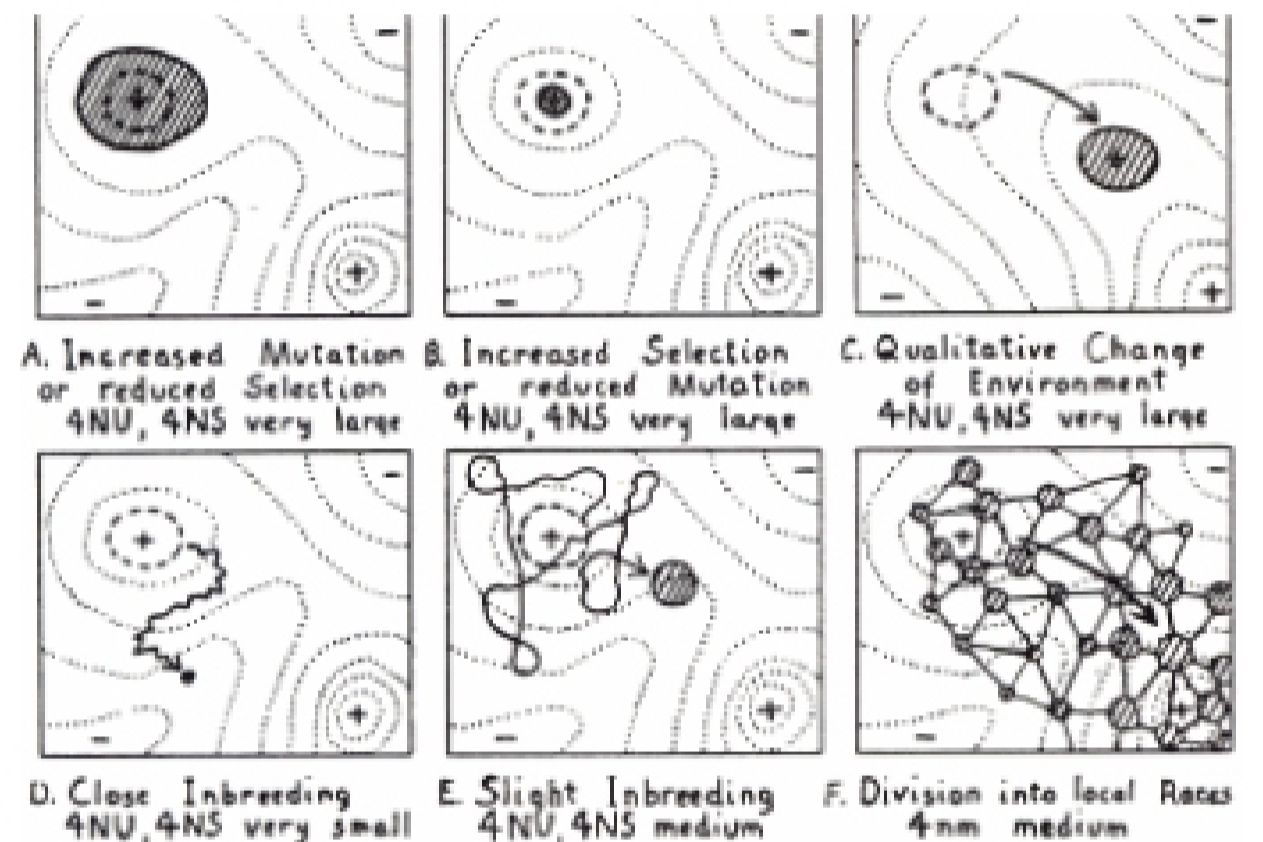


Fig. 4. Field of gene combinations occupied by a population within the general field of

Wright's shifting balance theory

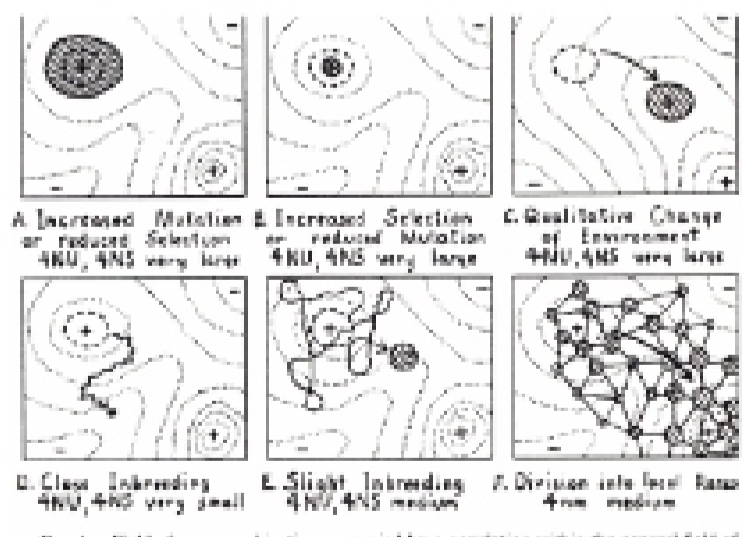


Fig. 4. Field of gene combinations occupied by a population within the general field of

Adaptive evolution is fastest when the population is divided into moderately small local groups that exchange migrants. Only then can selection choose among co-adapted gene complexes.

Extinction and recolonization

