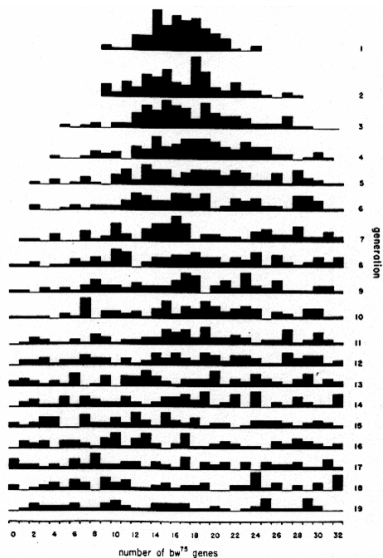
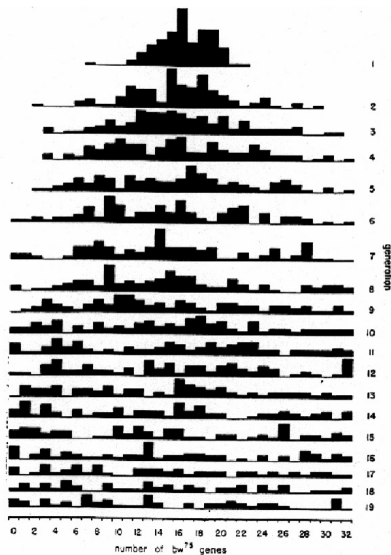


Geographic Population Structure

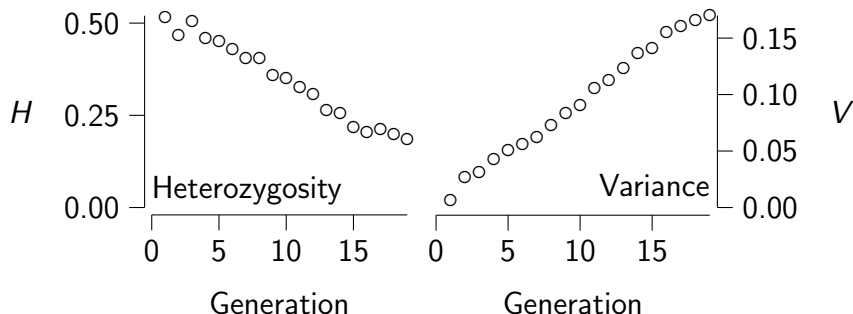
Alan R. Rogers

March 14, 2024

Buri's (1956) data

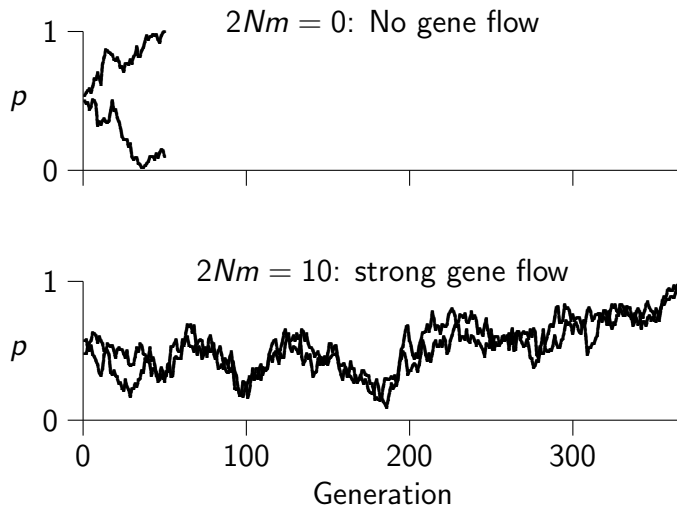


Drift reduces heterozygosity w/i groups while increasing variance among groups

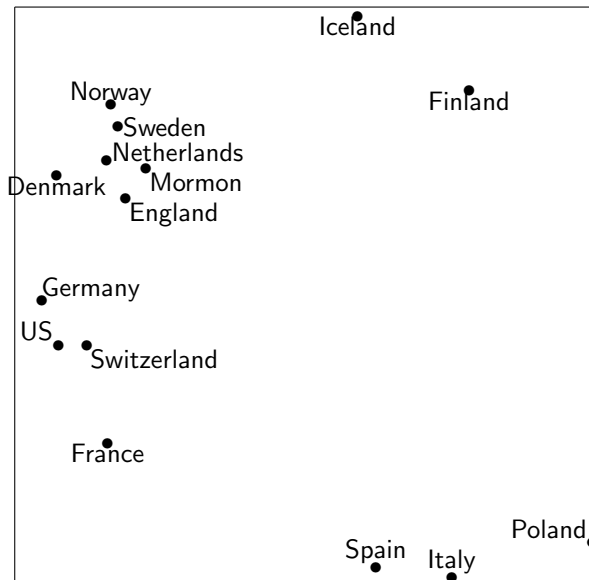


Data: Buri (1956)

Gene flow (migration) reduces population differences

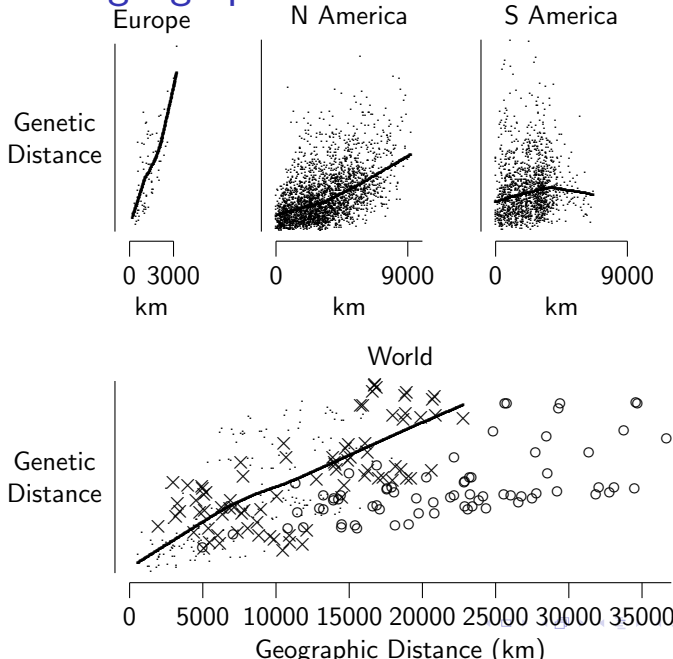


PC map of European genetic distances



(McLellan et al
1984)

Genetic and geographic distance



Wahlund principle: subdivision reduces heterozygosity

	A_1	A_1A_1	A_1A_2	A_2A_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

	$p_i - \bar{p}$	$(p_i - \bar{p})^2$
Pop 1	-4/16	1/16
Pop 2	4/16	1/16
Variance:		1/16

$$\mathcal{H}_S = \mathcal{H}_T - 2V \quad (\text{Wahlund 1928})$$

Wahlund principle: subdivision reduces heterozygosity

	A_1	A_1A_1	A_1A_2	A_2A_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

	$p_i - \bar{p}$	$(p_i - \bar{p})^2$
Pop 1	-4/16	1/16
Pop 2	4/16	1/16
Variance:		1/16

$$\mathcal{H}_S = \mathcal{H}_T - 2V \quad (\text{Wahlund 1928})$$

Wahlund principle: subdivision reduces heterozygosity

	A_1	A_1A_1	A_1A_2	A_2A_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

	$p_i - \bar{p}$	$(p_i - \bar{p})^2$
Pop 1	-4/16	1/16
Pop 2	4/16	1/16
Variance:		1/16

$$\mathcal{H}_S = \mathcal{H}_T - 2V \quad (\text{Wahlund 1928})$$

Wahlund principle: subdivision reduces heterozygosity

	A_1	A_1A_1	A_1A_2	A_2A_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

	$p_i - \bar{p}$	$(p_i - \bar{p})^2$
Pop 1	-4/16	1/16
Pop 2	4/16	1/16
Variance:		1/16

$$\mathcal{H}_S = \mathcal{H}_T - 2V \quad (\text{Wahlund 1928})$$

- ▶ 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

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

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

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 p_0 .

What we already know about heterozygosity

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

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

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 p_0 .

What we already know about heterozygosity

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

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

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 p_0 .

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.

Rearranging gives the variance among groups

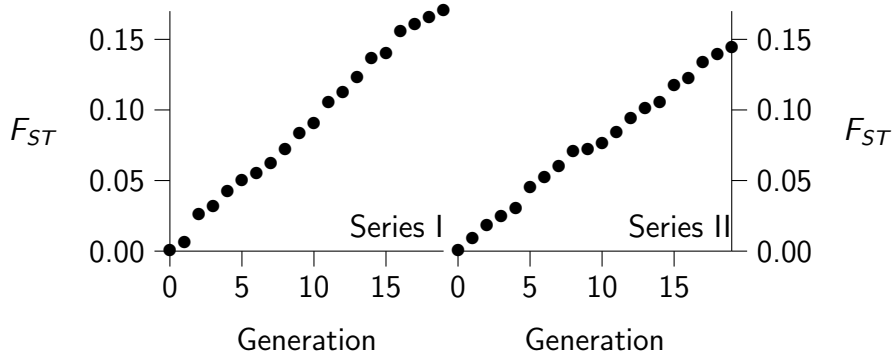
$$V = p_0 q_0 \left[1 - \left(1 - \frac{1}{2N} \right)^t \right]$$

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

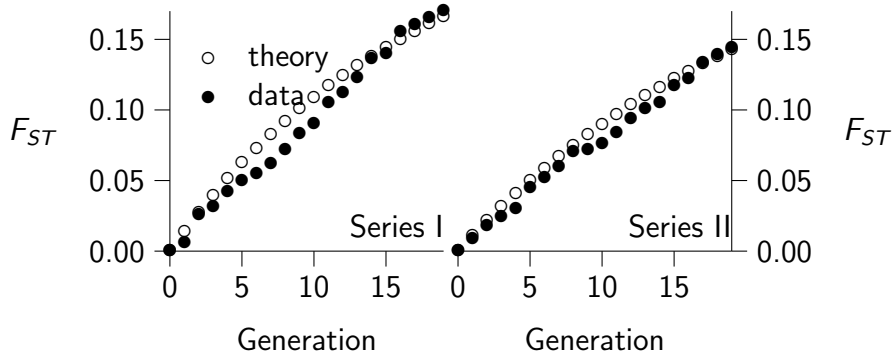
$$F_{ST} = \frac{V}{p_0 q_0} = 1 - \left(1 - \frac{1}{2N} \right)^t$$

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

Observed F_{ST} in Buri's experiments



Observed and expected F_{ST} in Buri's experiments



Application to differences among human races

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

For major human populations, $F_{ST} \approx 0.1$. Modern humans have been in Europe and E Asia for 50,000 years (2000 generations). Plug this into

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

and solve for N_e .

Answer: $N_e \approx 10,000$: agrees with genetic diversity within populations.

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 $N_e m \approx 2$.

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 $N_e m \approx 2$.

Summary

- ▶ Drift increases variance among groups and reduces that within them.
- ▶ Wahlund's principle: $\mathcal{H}_S = \mathcal{H}_T - 2V$
- ▶ F_{ST} measures between-group variance relative to total variance.
- ▶ If major human populations have been isolated for 50,000 y, then $N_e \approx 10,000$.
- ▶ If they are at migration-drift equilibrium, then pairs of populations exchange ~ 2 migrants per generation.