Supplementary Figures

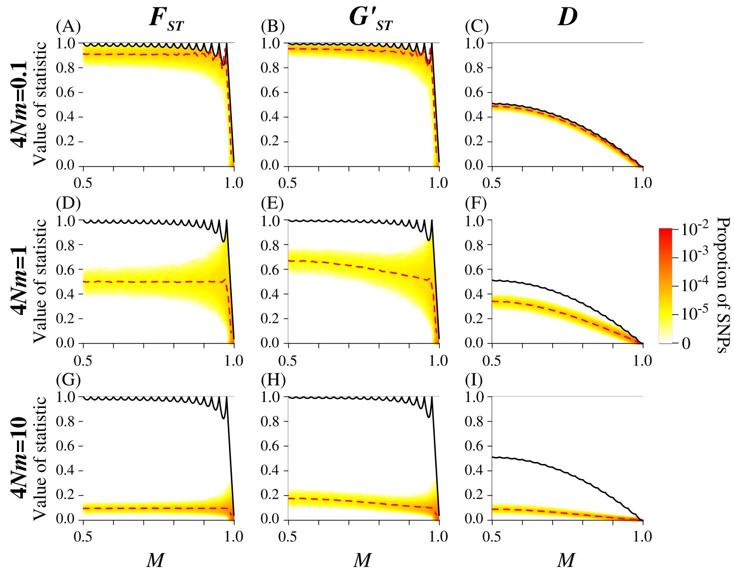

Figure S1 Joint density of the frequency M of the most frequent allele and statistics F_{ST}, G'_{ST}, and D, for different scaled migration rates $4Nm$, considering $K = 40$ subpopulations. The simulation procedure and figure design follow Figure 3.

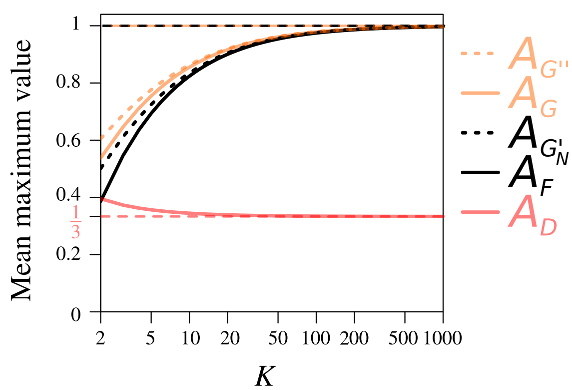

Figure S2 The means A_F, A_G, A_D, $A_{G'_N}$, and $A_{G''}$ of the maximal values of F_{ST}, G'_{ST}, D, $G'_{ST,\text{Nei}}$, and G''_{ST} respectively, over the interval $M \in [\frac{1}{2}, 1)$, as functions of the number of subpopulations K. $A_F(K)$, $A_G(K)$, and $A_D(K)$ are copied from Figure 2. $A_{G'_N}$ and $A_{G''}$ are computed numerically from eqs. S4.24 and S4.27. The x-axis is plotted on a logarithmic scale. The figure design follows Figure 2.

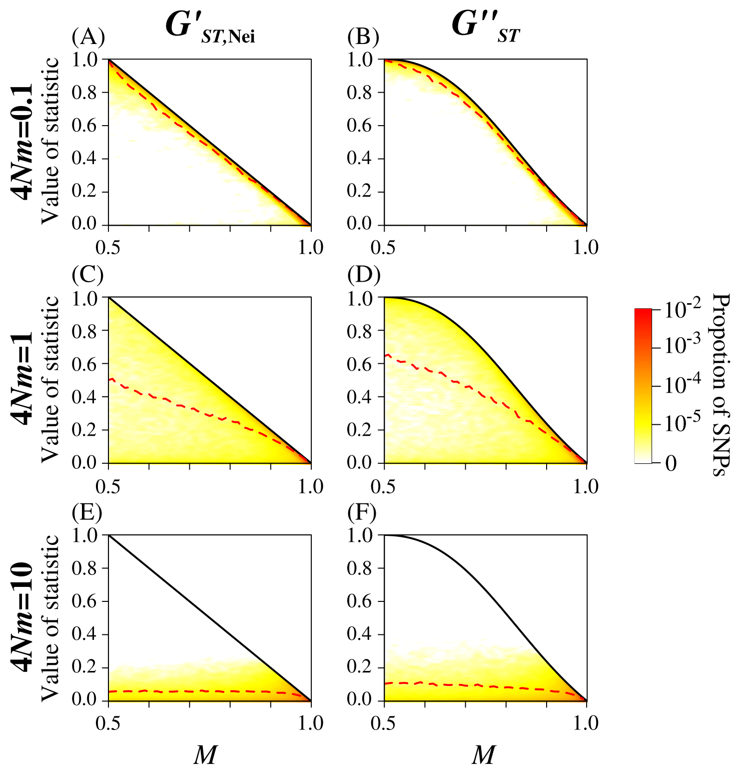

Figure S3 Joint density of the frequency M of the most frequent allele and statistics $G'_{ST,\text{Nei}}$ and G''_{ST}, for different scaled migration rates $4Nm$, considering $K = 2$ subpopulations. The black solid line represents the maximum value of $G'_{ST,\text{Nei}}$ and G''_{ST} in terms of M (eqs. S4.24 and S4.27); the red dashed line represents the mean $G'_{ST,\text{Nei}}$, and G''_{ST} in sliding windows of M of size 0.02 (plotted from 0.51 to 0.99). Colors represent the density of loci, estimated using a Gaussian kernel density estimate with a bandwidth of 0.007, with density set to 0 outside the minimum and maximum values. Loci are simulated using coalescent software `MS`, assuming an island model of migration and conditioning on 1 segregating site. Each panel considers 100,000 replicate simulations, with 100 lineages sampled per subpopulation. The figure design follows Figure 3.

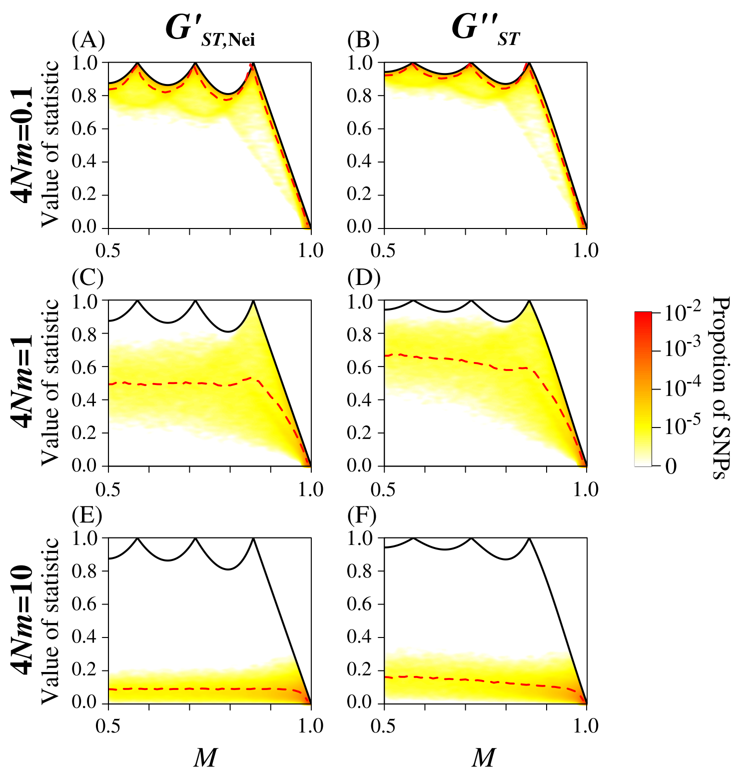

Figure S4 Joint density of the frequency M of the most frequent allele and statistics $G'_{ST,\text{Nei}}$ and G''_{ST}, for different scaled migration rates $4Nm$, considering $K = 7$ subpopulations. The simulation procedure follows Figure S3. The figure design follows Figures 4 and S3.

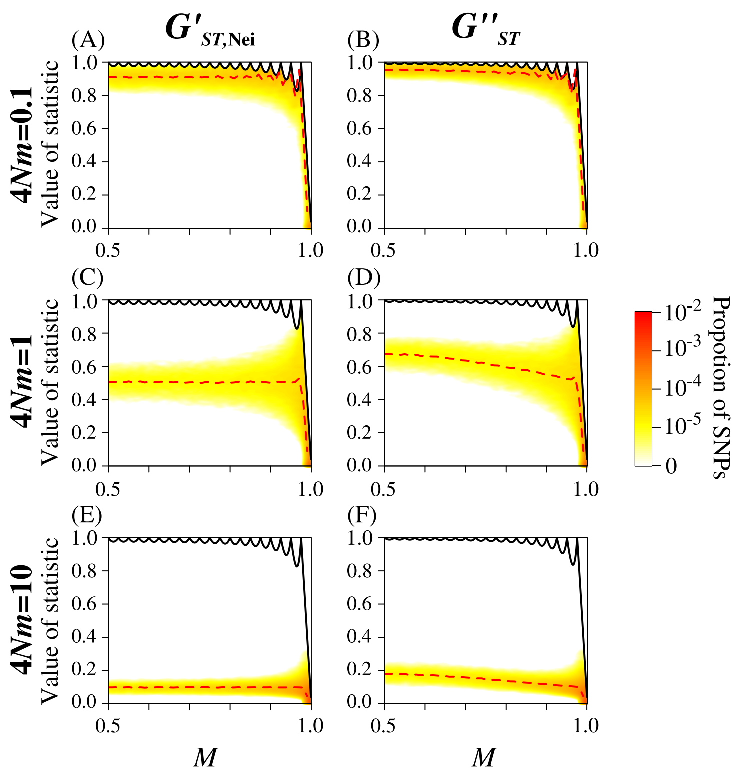

Figure S5 Joint density of the frequency M of the most frequent allele and statistics $G'_{ST,\mathrm{Nei}}$ and G''_{ST}, for different scaled migration rates $4Nm$, considering $K = 40$ subpopulations. The simulation procedure follows Figure S3. The figure design follows Figures S1 and S3.

Supplementary File S1. PROPERTIES OF THE MAXIMAL VALUES OF G'_{ST} AND D AS FUNCTIONS OF M

In this file, we derive the conditions under which the peaks (local maxima) of the maximal values of G'_{ST} (eq. 10) and D (eq. 11) in terms of M are reached, we derive their values, and we show the non-differentiability of the maximal G'_{ST} and D at the peaks.

1. Position and value of the peaks in the maximum value of G'_{ST} as a function of M

From eq. 2, $G'_{ST} = 1$ if and only if

$$\frac{(H_T - H_S)(K - 1 + H_S)}{H_T(K-1)(1-H_S)} = 1.$$

Solving for H_S, $G'_{ST} = 1$ if and only if $H_S = 0$ or $1 - H_S = K(1 - H_T)$. $H_S = 0$ leads to $S = M$, or $\frac{1}{K}\sum_{k=1}^{K} p_k^2 = \frac{1}{K}\sum_{k=1}^{K} p_k$. This equation is in turn equivalent to $\sum_{k=1}^{K} p_k(1-p_k) = 0$. Thus, $H_S = 0$ if and only if each p_k is equal either to 0 or to 1.

Because for all $\frac{1}{2} \leqslant M < 1$, $0 < H_T \leqslant \frac{1}{2}$, and it follows that $\frac{K}{2} \leqslant K(1 - H_T) < K$. In addition, because $H_S \geqslant 0$, $1 - H_S \leqslant 1$. Thus, $1 - H_S = K(1 - H_T)$ requires that $K = 2$, $H_T = \frac{1}{2}$, and $H_S = 0$, which is equivalent to having $p_1 = 1$ and $p_2 = 0$, or $p_1 = 0$ and $p_2 = 1$. We conclude $G'_{ST} = 1$ if and only if all p_k are either equal to 0 or equal to 1. This condition is the same as the condition for $F_{ST} = 1$ derived by Alcala & Rosenberg (2017, p. 1583), and thus leads to local maxima in the maximal value of G'_{ST} as a function of M at the same positions as the peaks in the maximum of F_{ST}: at $M = \frac{i}{K}$, with $i = \lceil \frac{K}{2} \rceil, \lceil \frac{K}{2} \rceil + 1, \ldots, K-1$.

Because the maximum value of G'_{ST} as a function of M (eq. 10) is continuous, and because it is bounded above by 1 and is equal to 1 only at the peaks, it follows that the maximum value of G'_{ST} is strictly below 1 between the peaks.

2. Position of the peaks in the maximum value of D as a function of M

From eq. 3, $D = 1$ if and only if

$$\frac{K(H_T - H_S)}{(K-1)(1-H_S)} = 1.$$

Solving for H_S, $D = 1$ if and only if $1 - H_S = K(1 - H_T)$. As shown in Supplementary File S1.1, this condition is met if and only if $K = 2$ and $M = \frac{1}{2}$, with $p_1 = 1$ and $p_2 = 0$, or $p_1 = 0$ and $p_2 = 1$. Thus, D values are only unconstrained in the unit interval in one specific case.

For $i = \lfloor \frac{K}{2} \rfloor, \lfloor \frac{K}{2} \rfloor + 1, \ldots, K-1$, we define the interval I_i by $[\frac{1}{2}, \frac{i+1}{K})$ for $i = \lfloor \frac{K}{2} \rfloor$ in the case that K is odd, and by $[\frac{i}{K}, \frac{i+1}{K})$ for all other (i, K). For M in I_i, $\lfloor KM \rfloor = i$ is constant. We denote $x = \{KM\} = KM - i$, so that $M = \frac{i+x}{K}$. Denote by $Q_i^D(x)$ the function of x that gives the maximum value of D on interval I_i for M,

$$Q_i^D(x) = \frac{2K(i + x^2) - 2(i+x)^2}{(K-1)[K - 2x(1-x)]}, \tag{S1.1}$$

where x ranges in $[0, 1)$ (or $[\frac{1}{2}, 1)$ in the case of odd K and $i = \lfloor \frac{K}{2} \rfloor$) and integers K and i satisfy $K \geqslant 2$ and $\lfloor \frac{K}{2} \rfloor \leqslant i \leqslant K - 1$.

$D^*(M)$ is continuous at each $M = \frac{i}{K}$, with $\lim_{x \to 1} Q_i^D(x) = Q_{i+1}^D(0)$ for each i with $\lfloor \frac{K}{2} \rfloor \leqslant i \leqslant K-2$. The derivative of $Q_i^D(x)$ is

$$\frac{dQ_i^D(x)}{dx} = 4\frac{(2i - K + 1)x^2 + [(K-i-1)^2 + (i-1)^2 + (K-2)]x - i^2}{(K-1)[K - 2x(1-x)]^2}. \tag{S1.2}$$

For $x \geqslant 0$, $\lim_{x \to 0^+} dQ_i^D(x)/dx = -4i^2/[K^2(K-1)]$, a strictly negative quantity for all $K \geqslant 2$ and $\lfloor \frac{K}{2} \rfloor \leqslant i \leqslant K-1$. In addition, $\lim_{x \to 1^-} dQ_i^D(x)/dx = 4(K-i-1)^2/[K^2(K-1)]$, a strictly positive quantity for all $K \geqslant 2$ and $\lfloor \frac{K}{2} \rfloor \leqslant i \leqslant K-2$. Hence, changing variables back from x to M, we see that for each point $M = \frac{i}{K}$ where $\lfloor \frac{K}{2} \rfloor \leqslant i \leqslant K-1$, the continuous function $D^*(M)$ has a positive derivative when approaching from the left and a negative derivative when approaching from the right. Thus, $D^*(M)$ has a local maximum at each $\frac{i}{K}$.

3. No other peaks in the maximum value of D as a function of M

We show that for each $K \geqslant 2$, the only peaks in $D^*(M)$ occur at $M = \frac{i}{K}$ for $\lfloor \frac{K}{2} \rfloor \leqslant i \leqslant K-1$.

For each i, we have shown that $\lim_{x\to 0^+} dQ_i^D(x)/dx < 0$ and $\lim_{x \to 1^-} dQ_i^D(x)/dx \geqslant 0$, with equality in the latter equation if and only if $i = K-1$. As a smooth function on $[0,1]$ with the property that its derivative changes from negative to nonnegative on $[0,1]$, $Q_i^D(x)$ has at least one critical point on $[0,1]$ that represents a local minimum. We show that $Q_i^D(x)$ has no more than one critical point in $[0,1]$; because it has a local minimum, it can have no local maxima interior to the interval $[0,1]$, so that $D^*(M)$ can only have local maxima at points $M = \frac{i}{K}$.

The denominator of $dQ_i^D(x)/dx$ is positive in $[0,1]$. We find the roots of the numerator of $dQ_i^D(x)/dx$ to obtain the critical points of $Q_i^D(x)$. Excluding the case of odd K and $i = \lfloor \frac{K}{2} \rfloor$, we have

$$x = \frac{-[(K-i-1)^2+(i-1)^2+K-2] \pm \sqrt{[(K-i-1)^2+(i-1)^2+K-2]^2+4i^2(2i-K+1)}}{2(2i-K+1)}. \tag{S1.3}$$

The negative root is negative for $K \geqslant 2$, leaving only a single critical point in the interval $[0,1]$.

For the case of odd K and $i = \lfloor \frac{K}{2} \rfloor$, the numerator of $dQ_i^D(x)/dx$ is linear in x, with root $x = \frac{1}{2}$. Hence, noting that for odd K and $i = \lfloor \frac{K}{2} \rfloor$, $Q_i^G(x)$ approaches its local maximum on $[\frac{1}{2}, 1)$ as $x \to 1$, on the interval $[\frac{1}{2}, \frac{K+1}{2K})$, a local minimum occurs at $M = \frac{1}{2}$.

4. Value of the peaks in the maximum value of D as a function of M

For $M = \frac{i}{K}$, with integers $K \geqslant 2$ and $\lfloor \frac{K}{2} \rfloor \leqslant i \leqslant K-1$, the maximal D from eq. 11 becomes

$$D^*(M) = \frac{2KM(1-M)}{K-1} = \frac{KH_T}{K-1}. \tag{S1.4}$$

The function $2KM(1-M)/(K-1)$ serves as an upper bound for D at all values of M, as $D \leqslant \frac{KH_T}{K-1}$ for all $H_S \geqslant 0$, with equality if and only if $H_S = 0$, $D^*(M)$ touches the curve $2KM(1-M)/(K-1)$ only at values M for which H_S can equal 0, or $M = \frac{i}{K}$ for $\lfloor \frac{K}{2} \rfloor \leqslant i \leqslant K-1$.

5. Non-differentiability of the maximal values of G'_{ST} and D at the peaks

Because $\lfloor KM \rfloor$ and $\{KM\}$ are non-differentiable for $M = \frac{i}{K}$ with $i = \lceil \frac{K}{2} \rceil, \lceil \frac{K}{2} \rceil + 1, \ldots, K-1$, the numerators and denominators of the maximum values of G'_{ST} and D (eqs. 10 and 11) are also non-differentiable at these points, and thus, the maximal values of G'_{ST} and D are also non-differentiable at these points.

6. Limit of the maximal value of G'_{ST}

From eq. 10, for fixed M in $[\frac{1}{2}, 1)$, because $0 \leqslant \{KM\} < 1$ and $\lfloor KM \rfloor/(KM) \to 1$ when $K \to \infty$,

$$\lim_{K\to\infty} \frac{[K(K-1)+2\{KM\}(1-\{KM\})](\lfloor KM \rfloor + \{KM\}^2 - KM^2)}{K(K-1)[K-2\{KM\}(1-\{KM\})]M(1-M)} = 1.$$

7. Limit of the maximal value of D

Similarly, applying eqs. 5 and 11, for fixed M in $[\frac{1}{2}, 1)$, because $0 \leqslant \{KM\} < 1$ and and $\lfloor KM \rfloor/(KM) \to 1$ when $K \to \infty$,

$$\lim_{K\to\infty} \frac{2K\left(\lfloor KM \rfloor + \{KM\}^2 - KM^2\right)}{(K-1)\left(K-2\{KM\}(1-\{KM\})\right)} = 2M(1-M) = H_T.$$

Supplementary File S2. THE SIZE OF THE PERMISSIBLE RANGE FOR G'_{ST} AND D

This file provides the computation of the integrals $A_G(K)$ (eq. 13) and $A_D(K)$ (eq. 14).

1. Computing $A_G(K)$ (eq. 13)

$A_G(K)$ is the integral of the maximum value of G'_{ST} in terms of M (eq. 10), divided by the size range of possible M values, $\frac{1}{2}$:

$$A_G(K) = 2\int_{\frac{1}{2}}^{1}\left[\frac{[K(K-1)+2\{KM\}(1-\{KM\})](\lfloor KM\rfloor+\{KM\}^2-KM^2)}{K(K-1)[K-2\{KM\}(1-\{KM\})]M(1-M)}\right]dM. \tag{S2.1}$$

For each interval I_i, the maximum value of G'_{ST} is a smooth function

$$Q_i^G(x) = \frac{[K(K-1)+2x(1-x)]\left[K(i+x^2)-(i+x)^2\right]}{(K-1)\left[K-2x(1-x)\right](i+x)(K-i-x)}, \tag{S2.2}$$

where x lies in $[0,1)$, i is an integer that lies in $[\lfloor\frac{K}{2}\rfloor, \lfloor\frac{K}{2}\rfloor+1, \ldots, K-1]$, and K is an integer greater than or equal to 2. Using the fact that $x = KM - i$, we obtain $dx = K\,dM$. We can break integral $A_G(K)$ into a sum of integrals of $Q_i^G(x)$ over intervals I_i,

$$A_G(K) = \begin{cases} \dfrac{2}{K}\displaystyle\sum_{i=\frac{K}{2}}^{K-1}\int_0^1 Q_i^G(x)\,dx \;\; \text{for even } K, \\ \dfrac{2}{K}\left[\displaystyle\int_{\frac{1}{2}}^{1} Q_{\frac{K-1}{2}}^G(x)\,dx + \sum_{i=\frac{K+1}{2}}^{K-1}\int_0^1 Q_i^G(x)\,dx\right] \text{for odd } K. \end{cases} \tag{S2.3}$$

Because $Q_i^G(x)$ is a rational function, we use partial fraction decomposition to compute its integral. $Q_i^G(x)$ can be written

$$Q_i^G(x) = 1 - \frac{Kh_2(K,i)}{2(x+i)} - \frac{Kh_2(K,K-i-1)}{2(K-x-i)} - \frac{K\sqrt{2K-1}\,h_1(K,i)}{(2x-1)^2+(2K-1)} - \frac{2K(2x-1)h_0(K,i)}{(2x-1)^2+(2K-1)}. \tag{S2.4}$$

In this expression, $h_1(K,i)$, $h_2(K,i)$, and $h_3(K,i)$ are functions that are independent of x:

$$h_0(K,i) = \frac{K^3(2i-K+1)}{(K-1)[i^2+(i+1)^2+(K-1)][(K-i-1)^2+(K-i)^2+(K-1)]} \tag{S2.5}$$

$$h_1(K,i) = \frac{4Ki(K-i-1)[(K-i)^2+(i+1)^2-1]}{(K-1)[i^2+(i+1)^2+(K-1)][(K-i-1)^2+(K-i)^2+(K-1)]\sqrt{2K-1}} \tag{S2.6}$$

$$h_2(K,i) = \frac{2i(i+1)}{K(K-1)}\left[1-\frac{K^2}{i^2+(i+1)^2+(K-1)}\right]. \tag{S2.7}$$

Letting $y = 2x-1$, we integrate by noting $\int \frac{y}{y^2+c}\,dy = \frac{1}{2}\log(y^2+c)$ and $\int \frac{1}{y^2+c}\,dy = \frac{1}{\sqrt{c}}\arctan(y/\sqrt{c})$, where c is a positive constant not dependent on y. For $\lceil\frac{K}{2}\rceil \leqslant i \leqslant K-1$,

$$\frac{2}{K}\int_0^1 Q_i^G(x)\,dx = \frac{2}{K} - 2h_1(K,i)\arctan\left(\frac{1}{\sqrt{2K-1}}\right) + h_2(K,i)\log\left(\frac{i}{i+1}\right) + h_2(K,K-i-1)\log\left(\frac{K-i-1}{K-i}\right), \tag{S2.8}$$

and for $i = \frac{K-1}{2}$,

$$\frac{2}{K}\int_{\frac{1}{2}}^{1} Q_{\frac{K-1}{2}}^G(x)\,dx = \frac{1}{K} - h_1\left(K,\frac{K-1}{2}\right)\arctan\left(\frac{1}{\sqrt{2K-1}}\right) + h_2\left(K,\frac{K-1}{2}\right)\log\left(\frac{K-1}{K+1}\right). \tag{S2.9}$$

From eq. S2.6, $h_1(K,i) = h_1(K, K-i-1)$. Consequently, for even K, we simplify the expression for $A_G(K)$ by noting that $\sum_{i=K/2}^{K-1} h_1(K,i) = \sum_{i=0}^{K/2-1} h_1(K,i)$, and $2\sum_{i=K/2}^{K-1} h_1(K,i) = \sum_{i=0}^{K-1} h_1(K,i)$. We can similarly simplify the expression for $A_G(K)$ when K is odd, because $\sum_{i=(K+1)/2}^{K-1} h_1(K,i) = \sum_{i=0}^{(K-3)/2} h_1(K,i)$, and thus, $[2\sum_{i=(K+1)/2}^{K-1} h_1(K,i)] + h_1[K, \frac{K-1}{2}] = \sum_{i=0}^{K-1} h_1(K,i)$.

Because $\sum_{i=K/2}^{K-1} h_2(K, K-i-1)\log[(K-i-1)/(K-i)] = \sum_{i=0}^{K/2-1} h_2(K,i)\log[i/(i+1)]$, we can group terms involving h_2 in the expression for $A_G(K)$ when K is even (eq. S2.3) into a single sum $\sum_{i=0}^{K-1} h_2(K,i)\log[i/(i+1)]$. Similarly, because $\sum_{i=(K+1)/2}^{K-1} h_2(K, K-i-1)\log[(K-i-1)/(K-i)] = \sum_{i=0}^{(K-3)/2} h_2(K,i)\log[i/(i+1)]$, we can group the terms involving h_2 in the expression for $A_G(K)$ when K is odd into a sum $\sum_{i=0}^{K-1} h_2(K,i)\log[i/(i+1)]$.

Substituting eqs. S2.8 and S2.9 into eq. S2.3, grouping the expressions with h_1 and h_2, taking $0\log 0 = 0$, and simplifying, the expressions for $A_G(K)$ for even and odd K equalize and we obtain eq. 13.

2. Increase of $A_G(K)$ as a function of K

We must show that $\Delta_G(K) = A_G(K+1) - A_G(K) \geqslant 0$. We numerically computed $A_G(K)$ (eq. 13) and $\Delta_G(K)$ for K ranging from 2 to 10,000; we found that $\Delta_G(K) > 0$ for all K in that range.

Although this numerical result does not formally prove that $\Delta_G(K) > 0$ for all K, we note that because $1 \geqslant G'_{ST} \geqslant F_{ST}$ owing to the normalization in the definition of G'_{ST}, $1 \geqslant A_G(K) \geqslant A_F(K)$. Hence, because $\lim_{K\to\infty} A_F(K) = 1$, we also have $\lim_{K\to\infty} A_G(K) = 1$.

3. Computing $A_D(K)$ (eq. 14)

$A_D(K)$ is the integral of the maximum value of D in terms of M (eq. 11), divided by the size range of possible M values, $\frac{1}{2}$:

$$A_D(K) = 2\int_{\frac{1}{2}}^{1} \frac{2K}{K-1} \frac{\lfloor KM \rfloor + \{KM\}^2 - KM^2}{K - 2\{KM\}(1-\{KM\})} dM. \tag{S2.10}$$

Using $Q_i^D(x)$ (eq. S1.1), we break $A_D(K)$ into a sum of integrals over domains I_i,

$$A_D(K) = \begin{cases} \dfrac{2}{K} \displaystyle\sum_{i=\frac{K}{2}}^{K-1} \int_0^1 Q_i^D(x)\, dx \text{ for even } K, \\ \dfrac{2}{K}\left[\displaystyle\int_{\frac{1}{2}}^{1} Q_{\frac{K-1}{2}}^D(x)\, dx + \sum_{i=\frac{K+1}{2}}^{K-1} \int_0^1 Q_i^D(x)\, dx \right] \text{ for odd } K. \end{cases} \tag{S2.11}$$

We use a partial fraction decomposition of the rational function $Q_i^D(x)$:

$$Q_i^D(x) = 1 - \frac{2}{K-1}\left[\frac{(2x-1)f_1(K,i)}{(2x-1)^2 + (2K-1)} + \frac{f_2(K,i)}{(2x-1)^2 + (2K-1)} \right], \tag{S2.12}$$

where $f_1(K,i) = 2i - K + 1$ and $f_2(K,i) = i^2 + (K-i-1)^2$ are functions that do not depend on x.

Letting $y = 2x - 1$, $Q_i^D(x)$ can be integrated by again applying $\int \frac{y}{y^2+c}\, dy = \frac{1}{2}\log(y^2 + c)$ and $\int \frac{1}{y^2+c}\, dy = \frac{1}{\sqrt{c}}\arctan(y/\sqrt{c})$, where c is a positive constant that does not depend on y. For $\lceil \frac{K}{2} \rceil \leqslant i \leqslant K-1$,

$$\frac{2}{K}\int_0^1 Q_i^D(x)\, dx = \frac{2}{K} - \frac{4[i^2 + (K-1-i)^2]}{K(K-1)\sqrt{2K-1}} \arctan\left(\frac{1}{\sqrt{2K-1}}\right), \tag{S2.13}$$

and for $i = \frac{K-1}{2}$,

$$\frac{2}{K}\int_{\frac{1}{2}}^{1} Q_{\frac{K-1}{2}}^D(x)\, dx = \frac{1}{K} - \frac{(K-1)}{K\sqrt{2K-1}} \arctan\left(\frac{1}{\sqrt{2K-1}}\right). \tag{S2.14}$$

We can eliminate terms in i from eq. S2.11 when K is even by noting

$$\sum_{i=\frac{K}{2}}^{K-1}[i^2+(K-1-i)^2]=\sum_{i=\frac{K}{2}}^{K-1}i^2+\sum_{i=\frac{K}{2}}^{K-1}(K-1-i)^2=\sum_{i=0}^{K-1}i^2. \tag{S2.15}$$

Similarly, we can eliminate terms in i from the expression for $A_D(K)$ when K is odd:

$$\sum_{i=\frac{K+1}{2}}^{K-1}[i^2+(K-1-i)^2]=\left(\sum_{i=0}^{K-1}i^2\right)-\left(\frac{K-1}{2}\right)^2. \tag{S2.16}$$

Because $\sum_{i=0}^{K-1}i^2 = K(K-1)(2K-1)/6$, we substitute eqs. S2.15 and S2.16 into eq. S2.11 and simplify the sums. The expressions for $A_D(K)$ when K is even and odd equalize, and we obtain eq. 14.

4. Decrease of $A_D(K)$ as a function of K

To show that $A_D(K)$ is decreasing in K, we must show that $dA_D/dK < 0$ for all $K \geqslant 2$. From the expression for $A_D(K)$ in eq. 14,

$$\frac{dA_D}{dK}=\frac{1}{3}\left[\frac{1}{K}-\frac{2\arctan\left(\frac{1}{\sqrt{2K-1}}\right)}{\sqrt{2K-1}}\right]. \tag{S2.17}$$

Let $f(x)=\arctan(x)-(x-x^3/3)$. Because $f(0)=0$ and $f'(x)=x^4/(1+x^2)>0$, $f(x)>0$ for positive x. Hence,

$$\arctan\left(\frac{1}{\sqrt{2K-1}}\right)\geqslant\frac{1}{\sqrt{2K-1}}-\frac{1}{3(2K-1)^{3/2}}. \tag{S2.18}$$

Applying inequality S2.18 in eq. S2.17, we obtain

$$\frac{dA_D}{dK}\leqslant-\frac{4K-3}{9K(2K-1)^2}, \tag{S2.19}$$

which is strictly negative for all $K \geqslant 2$. We conclude that $dA_D/dK < 0$ and hence that $A_D(K)$ decreases monotonically as a function of K.

For the limit of $A_D(K)$ as $K \to \infty$, we use l'Hôpital's rule to find $\lim_{K\to\infty}\arctan(\frac{1}{\sqrt{2K-1}})/(\frac{1}{\sqrt{2K-1}}) = 1$, so that $\lim_{K\to\infty}A_D(K)=\frac{1}{3}$.

Supplementary File S3. MS COMMANDS TO SIMULATE NUCLEOTIDE SEQUENCES UNDER AN EQUILIBRIUM ISLAND MODEL WITH K SUBPOPULATIONS AND SCALED MIGRATION RATE $4NM$

$K = 2$

```
./ms 200 100000 -s 1 -I 2 100 100 0.1
./ms 200 100000 -s 1 -I 2 100 100 1
./ms 200 100000 -s 1 -I 2 100 100 10
```

For $K = 7$

```
./ms 700 100000 -s 1 -I 7 100 100 100 100 100 100 100 0.1
./ms 700 100000 -s 1 -I 7 100 100 100 100 100 100 100 1
./ms 700 100000 -s 1 -I 7 100 100 100 100 100 100 100 10
```

For $K = 40$

```
./ms 4000 100000 -s 1 -I 40 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100
    100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 0.1
./ms 4000 100000 -s 1 -I 40 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100
    100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 1
./ms 4000 100000 -s 1 -I 40 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100
    100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 100 10
```

Supplementary File S4. PROPERTIES OF NEI'S G'_{ST} AND MEIRMANS AND HEDRICK'S G''_{ST}

This supplementary information file provides results regarding alternative formulations of F_{ST} and G'_{ST}—Nei's G'_{ST} and Meirmans and Hedrick's G''_{ST}—that include a multiplicative term based on the number of sampled populations K.

Mathematical constraints on G'_{ST} and G''_{ST}

Using H_T, H_S, and $D_{ST} = H_T - H_S$, Nei (1987, pp. 188-191) defined a measure $G'_{ST,\text{Nei}}$:

$$\begin{aligned} D'_{ST} &= \frac{K}{K-1}(H_T - H_S) \\ H'_T &= H_S + D'_{ST} \\ G'_{ST,\text{Nei}} &= \frac{D'_{ST}}{H'_T} = \frac{K(H_T - H_S)}{KH_T - H_S}. \end{aligned}$$

From eqs. 4 and 5, substituting $H_S = 2(M - S)$ and $H_T = 2M(1 - M)$ for the biallelic case, $G'_{ST,\text{Nei}}$ becomes

$$G'_{ST,\text{Nei}} = \frac{K(S - M^2)}{M(K - 1 - KM) + S}. \tag{S4.20}$$

Meirmans & Hedrick (2011, eq. 4) defined a second quantity G''_{ST} by

$$G''_{ST} = \frac{G'_{ST,\text{Nei}}}{G'_{ST,\text{Nei,max}}} = \frac{K(H_T - H_S)}{(KH_T - H_S)(1 - H_S)}. \tag{S4.21}$$

From eqs. 4 and 5, substituting H_S and H_T by their values as functions of M and S, eq. S4.21 becomes:

$$G''_{ST} = \frac{K(S - M^2)}{[M(K - 1 - KM) + S](1 - 2M + 2S)}. \tag{S4.22}$$

Maximal values of G'_{ST} and G''_{ST}

We first show that if M is fixed, $G'_{ST,\text{Nei}}$ is increasing when treated as a function of S. By Theorem 1 of Alcala & Rosenberg (2017), for fixed M and K, S is positive, satisfying $M^2 \leqslant S \leqslant [\lfloor KM \rfloor + \{KM\}^2]/K$. In particular, because $\frac{1}{2} \leqslant M < 1$, we have $\frac{1}{4} \leqslant S \leqslant 1$. The derivative of $G'_{ST,\text{Nei}}$ with respect to S is

$$\frac{dG'_{ST,\text{Nei}}}{dS} = \frac{K(K-1)M(1-M)}{[M(K - 1 - KM) + S]^2}. \tag{S4.23}$$

The numerator is positive, as $\frac{1}{2} \leqslant M < 1$ and $K \geqslant 2$. Noting $H_T = 2M(1 - M)$ and $H_S = 2(M - S)$, the denominator equals $\frac{1}{4}(KH_T - H_S)^2$, a quantity that is also strictly positive, as $H_T \geqslant H_S$ by the Wahlund principle, $H_T > 0$, and $K \geqslant 2$. $G'_{ST,\text{Nei}}$ is therefore an increasing function of S, so that its maximum as a function of M occurs when S lies at its largest permissible value given M. For fixed M, $\frac{1}{2} \leqslant M < 1$, Theorem 1 of Alcala & Rosenberg (2017) gives the maximum for S as a function of M. Inserting this maximum, we have:

$$G'_{ST,\text{Nei}} \leqslant \frac{\lfloor KM \rfloor + \{KM\}^2 - KM^2}{KM(1-M) - M + \lfloor KM \rfloor / K + \{KM\}^2/K}. \tag{S4.24}$$

Similarly, we show that G''_{ST} is an increasing function of S for M in $[\frac{1}{2}, 1)$ and integers $K \geqslant 2$. For fixed M, the derivative of G''_{ST} with respect to S is

$$\begin{aligned} \frac{dG''_{ST}}{dS} &= \frac{K\left[-2(S - M^2)^2 + (K-1)M(1-M)[1 - 2M(1-M)]\right]}{[M(K - 1 - KM) + S]^2(1 - 2M + 2S)^2}, \\ &= \frac{K[(K-1)H_T(1 - H_T) - (H_T - H_S)^2]}{2[M(K - 1 - KM) + S]^2(1 - H_S)^2}. \end{aligned} \tag{S4.25}$$

The denominator in eq. S4.25 is positive, as a product of the positive $\frac{1}{4}(KH_T - H_S)^2$ and $(1 - H_S)^2$ for $0 \leqslant H_S \leqslant \frac{1}{2}$. Hence, the sign of dG''_{ST}/dS is determined by the sign of its numerator. We find the roots of the numerator as a function of H_T, denoted $H_{T,1}$ and $H_{T,2}$:

$$\begin{aligned} H_{T,1} &= \frac{K - 1 + 2H_S + \sqrt{(K-1)[K-1+4H_S(1-H_S)]}}{2K}, \\ H_{T,2} &= \frac{K - 1 + 2H_S - \sqrt{(K-1)[K-1+4H_S(1-H_S)]}}{2K}. \end{aligned} \tag{S4.26}$$

To show that G''_{ST} is increasing in S for permissible values of S, we examine the roots. For biallelic markers, $0 \leqslant H_S \leqslant \frac{1}{2}$, and $H_S(1 - H_S)$ is an increasing function of H_S. It then follows that $H_{T,1}$ is also an increasing function of H_S. Consequently, the minimum value of $H_{T,1}$, treated as a function of H_S, is reached at $H_S = 0$, yielding $H_{T,1} \geqslant (K-1)/K \geqslant \frac{1}{2}$ for all $K \geqslant 2$. Because for biallelic markers $H_T \leqslant \frac{1}{2}$, $H_{T,1} \geqslant H_T$. For $H_{T,2}$, because $0 \leqslant H_S \leqslant \frac{1}{2}$, $H_S(1 - H_S) \geqslant 0$, and $H_{T,2} \leqslant [K - 1 + 2H_S - \sqrt{(K-1)^2}]/(2K) = H_S/K < H_S$ for all $K \geqslant 2$. Because the Wahlund principle ensures that $H_T \geqslant H_S$, $H_{T,2} < H_T$.

Because the numerator of dG''_{ST}/dS is a quadratic expression in H_T with a negative leading term $-KH_T^2$, it is positive between its roots. We have shown that the permissible values of H_T satisfy $H_{T,2} < H_T \leqslant H_{T,1}$. Hence, the numerator of dG''_{ST}/dS is non-negative for all $0 \leqslant H_S \leqslant \frac{1}{2}$ and $H_S \leqslant H_T \leqslant \frac{1}{2}$, with equality requiring $H_T = H_{T,1} = \frac{1}{2}$, $K = 2$, $H_S = 0$, and $M = \frac{1}{2}$.

This argument demonstrates that for each $M \neq \frac{1}{2}$, G''_{ST} is an increasing function of S on the permissible interval for S. Its maximum as a function of M occurs when S lies at its largest permissible value given M. Hence, for fixed $K \geqslant 2$ and fixed M, $\frac{1}{2} \leqslant M < 1$, using Theorem 1 from Alcala & Rosenberg (2017) to specify this value of S given M, we substitute the maximum of S into eq. S4.20:

$$G''_{ST} \leqslant \frac{\lfloor KM \rfloor + \{KM\}^2 - KM^2}{[KM(1-M) - M + \lfloor KM \rfloor/K + \{KM\}^2/K](1 - 2M + 2(\lfloor KM \rfloor + \{KM\}^2)/K)}. \tag{S4.27}$$

Note that for $M = \frac{1}{2}$, eq. S4.27 evaluates to 1, the largest possible value for G''_{ST}, so that eq. S4.27 is also valid for $M = \frac{1}{2}$.

Comparison of the ranges of possible values of F_{ST}, G'_{ST}, D, $G'_{ST,\mathrm{Nei}}$, and G''_{ST}

We computed the ranges of possible values of $G'_{ST,\mathrm{Nei}}$ and G''_{ST}, denoted $A_{G'_N}$ and $A_{G''}$, numerically, for K ranging from 2 to 10,000, using the same procedure as for G'_{ST}. Results appear in Figure S2.

Simulation-based distributions of $G'_{ST,\mathrm{Nei}}$ and G''_{ST}

We performed simulations using the same procedure as that used to produce Figures 3, 4, and S1. Results appear in Figures S3–S5.

Literature Cited

Alcala N, Rosenberg NA (2017) Mathematical constraints on F_{ST}: biallelic markers in arbitrarily many populations. *Genetics*, **206**, 1581–1600.

Meirmans PG, Hedrick PW (2011) Assessing population structure: F_{ST} and related measures. *Molecular Ecology Resources*, **11**, 5–18.

Nei M (1987) *Molecular evolutionary genetics*. Columbia university press.