Lecture 03:
Q_{ST} vs F_{ST}

UNE course:
The search for selection
3 -- 7 Feb 2020

Bruce Walsh (University of Arizona)
jbwalsh@email.arizona.edu

Basic idea

- Is the within- vs. between-population partitioning of the genetic variance for a candidate trait (Q_{ST}) different from the partitioning for random markers (i.e., neutral molecular variation) over the same set of populations (F_{ST})?
- Departures may suggest that the trait is under selection

Table 12.2 Interpretation of Q_{ST} versus F_{ST} comparisons.

Observation	Interpretation
$Q_{ST} > F_{ST}$	**Divergent selection**: spatial variation in trait values in excess of neutral expectation.
$Q_{ST} = F_{ST}$	Consistent with divergence expected under drift. Does not rule out selection, but does not support it either.
$Q_{ST} < F_{ST}$	**Convergent selection**: spatial variation in trait values less than neutral expectation. Similar trait values are favored over populations.

Outline

- Quantifying allele-frequency divergence among subpopulations, F_{ST}
- Q_{ST}
  - Distribution
  - Power
- P_{ST}
- F_{STQ} and the role of LD

F_{ST}, a measure of population structure

- One measure of population structure is given by Wright's F_{ST} statistic (also called the fixation index)

- Essentially, this is the fraction of genetic variation due to between-population differences in allele frequencies

- Changes in allele frequencies can be caused by evolutionary forces such as genetic drift, migration, and local adaptation

- Consider a biallelic locus (A, a). If p denotes overall population frequency of allele A,
  - then the overall population variance is p(1-p)
  - $Var(p_i)$ = variance in p over subpopulations
  - $F_{ST} = Var(p_i)/[p(1-p)]$

Example of F_{ST} estimation

Population	Freq(A)
1	0.1
2	0.6
3	0.2
4	0.7

Assume all subpopulations contribute equally to the overall metapopulation

Overall freq(A) = p = (0.1 + 0.6 + 0.2 + 0.7)/4 = 0.4

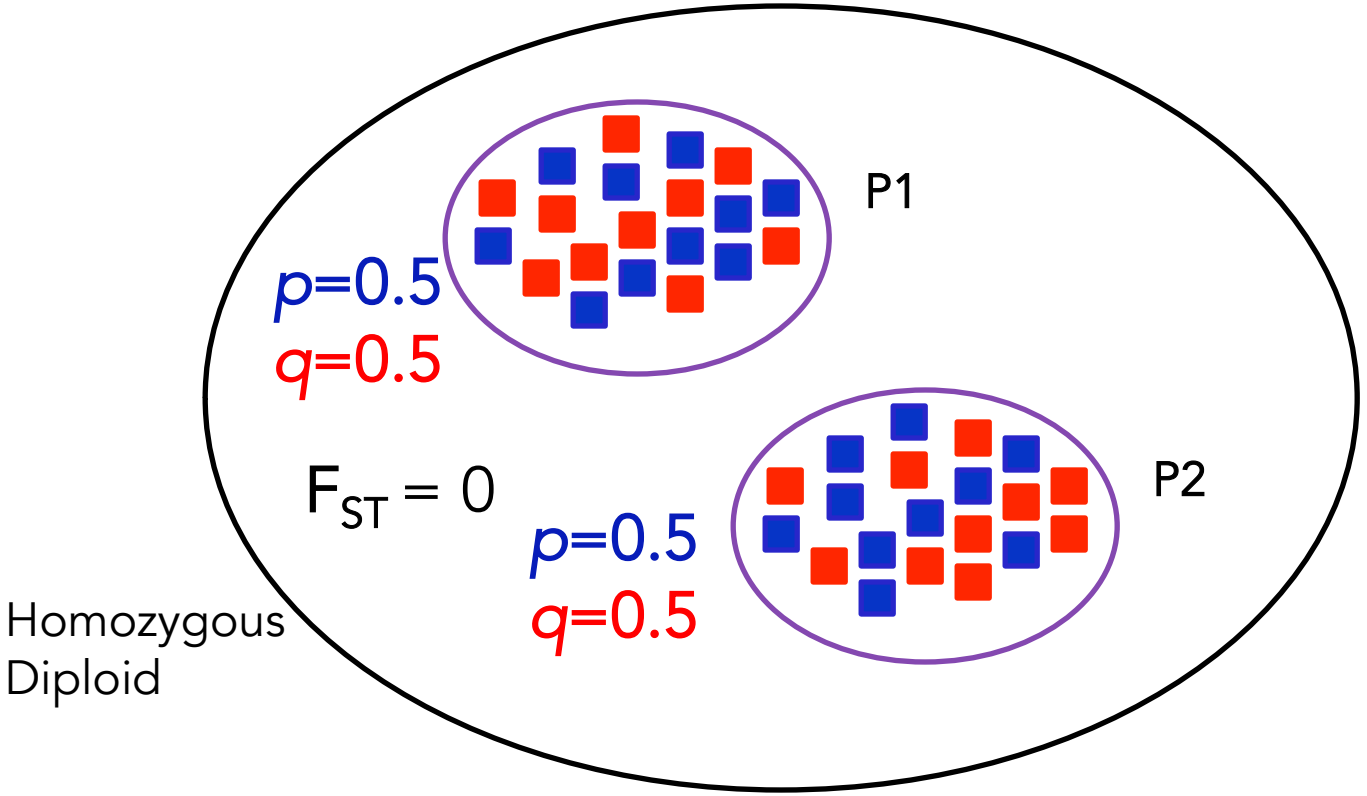
$Var(p_i) = E(p_i^2) - [E(p_i)]^2 = E(p_i^2) - p^2$

$Var(p_i) = [(0.1^2 + 0.6^2 + 0.2^2 + 0.7^2)/4] - 0.4^2 = 0.065$

Total population variance = p(1-p) = 0.4(1-0.4) = 0.24

Hence, $F_{ST} = Var(p_i) /[p(1-p)] = 0.065/0.24 = 0.27$

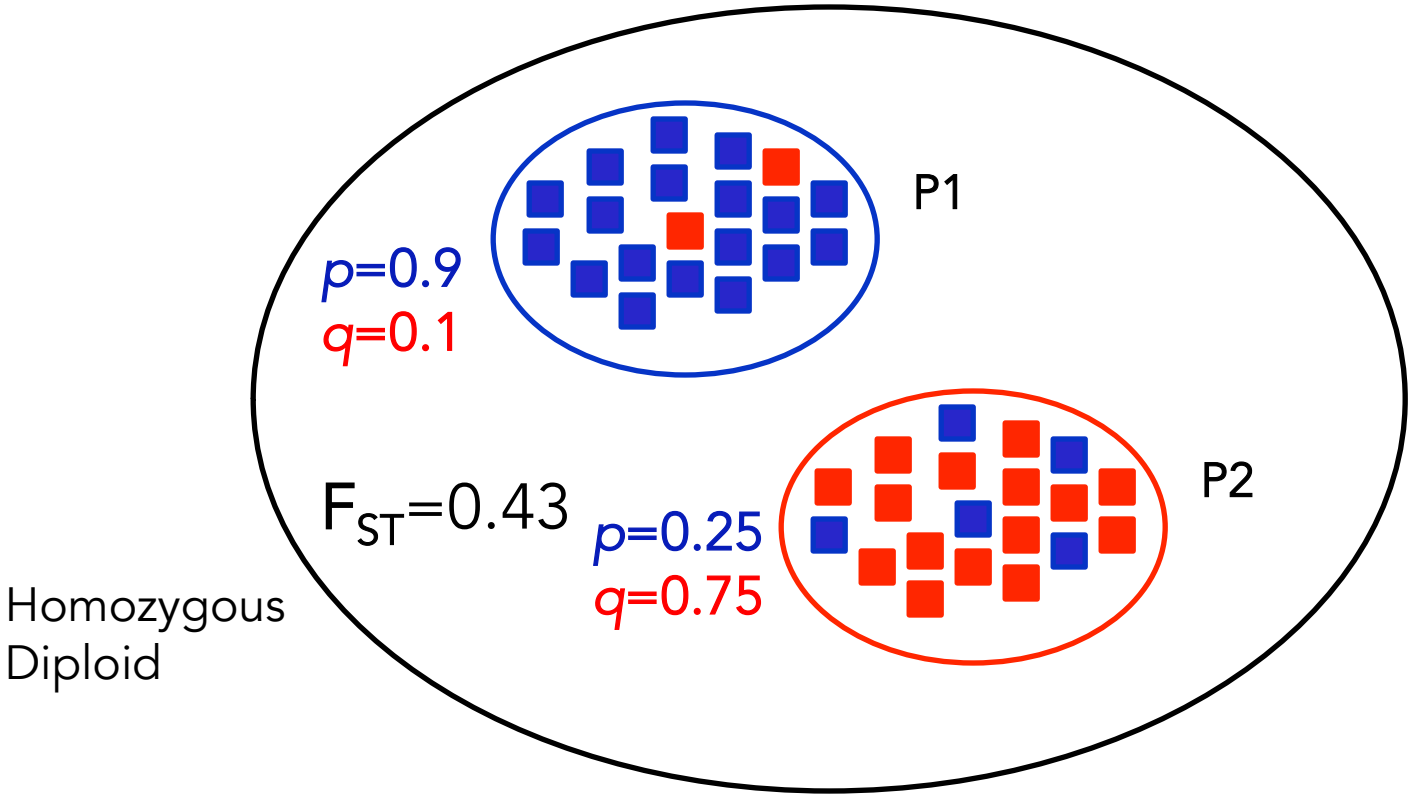
Graphical example of F_{ST}



No population differentiation

Modified from Escalante et al. 2004. Trends Parasitol. 20:388-395

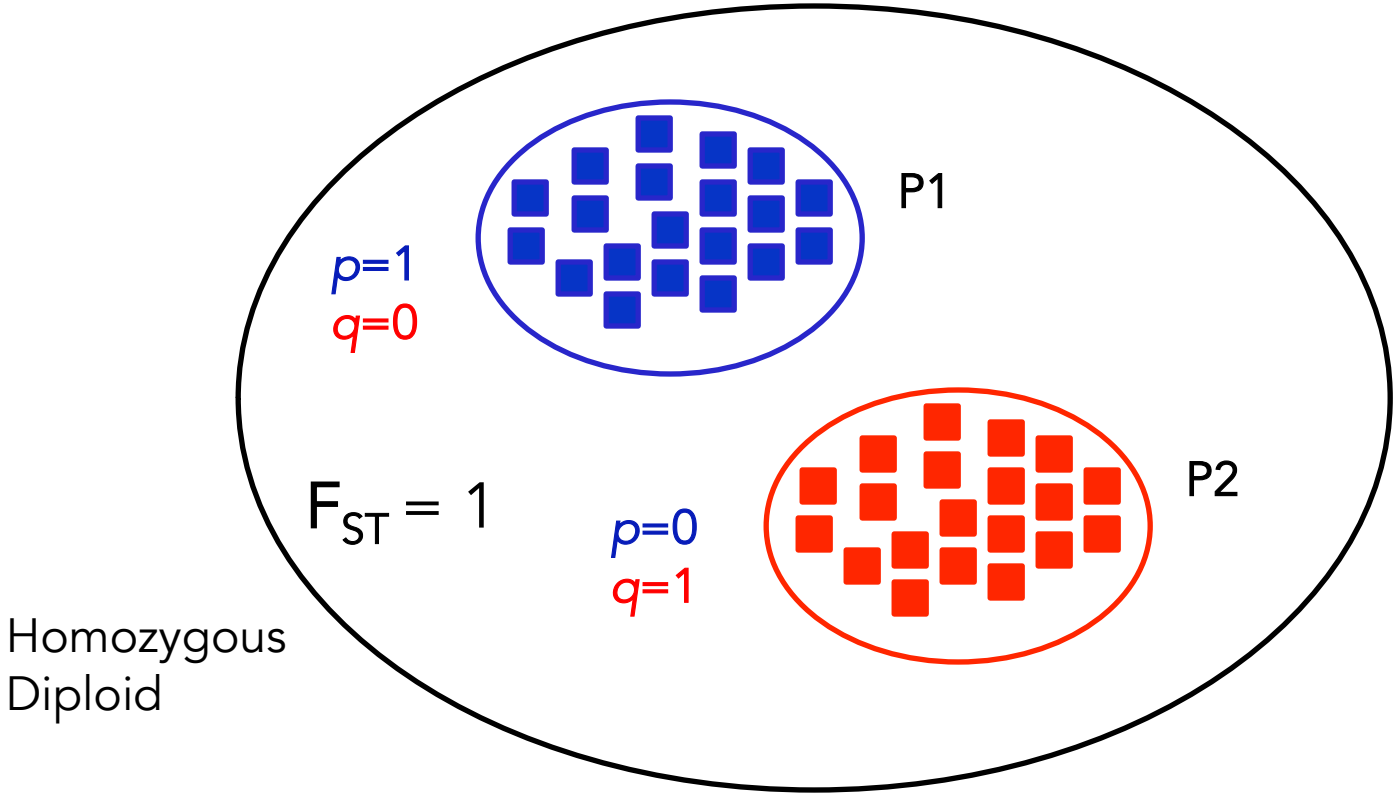
Graphical example of F_{ST}



Strong population differentiation

Modified from Escalante et al. 2004. Trends Parasitol. 20:388-395

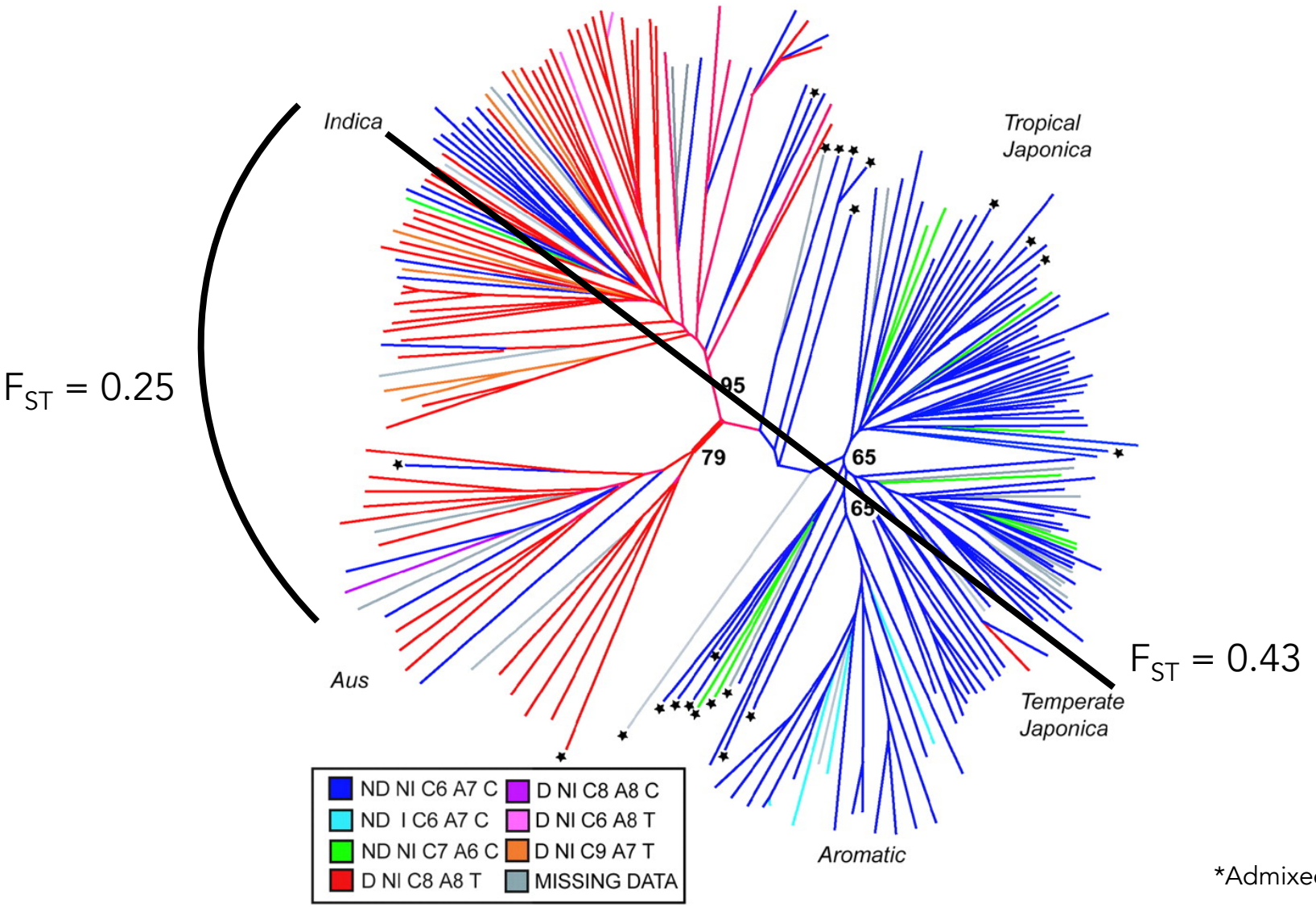
Graphical example of F_{ST}



Complete population differentiation

Modified from Escalante et al. 2004. Trends Parasitol. 20:388-395

Rice population structure

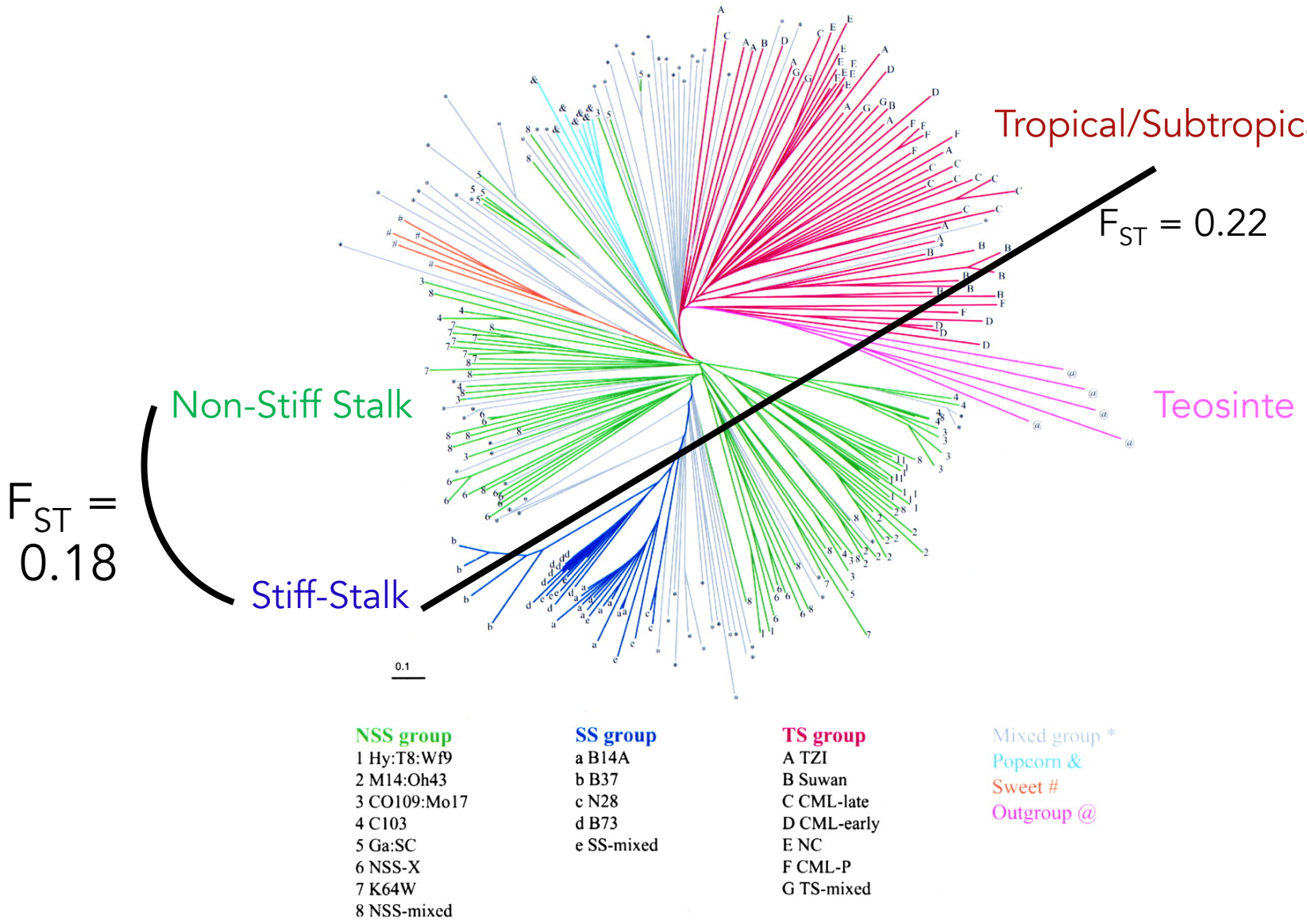


*Admixed individuals

Unrooted neighbor-joining tree based on C.S. Chord (Cavalli-Sforza and Edwards 1967) based on 169 nuclear SSRs. The key relates the color of the line to the chloroplast haplotype based on ORF100 and PS-ID sequences.

Garris et al. 2005. Genetics 169:1631-1638

Maize population structure



Phylogenetic tree for 260 inbred lines using the log-transformed proportion of shared alleles distance

Liu et al. 2003. Genetics 165:2117-2128

Flint-Garcia et al. 2005. Plant J. 144:1054-1064

Q_{ST}: Partitioning Additive Variance Over Populations

Consider a quantitative trait in a diploid with a purely additive-genetic basis, and denote its genetic variance over the entire metapopulation by σ^2_G. From Table 11.3 (setting $f = Q_{ST}$), the within- and among-population components of variance can be represented as $\sigma^2_{GW} = (1 - Q_{ST})\sigma^2_G$ and $\sigma^2_{GB} = 2Q_{ST}\sigma^2_G$, respectively, for a total variance in a structured population of $(1 + Q_{ST})\sigma^2_G$. Rearranging yields

$$Q_{ST} = \frac{\sigma^2_{GB}}{\sigma^2_{GB} + 2\sigma^2_{GW}} \tag{12.26a}$$

While the term Q_{ST} was introduced by Spitze (1993), this metric was proposed earlier by Prout and Barker (1989, 1993) and Lande (1992), and strongly hinted at by Rogers and

For inbred populations,

$$Q_{ST} = \frac{(1+f)\sigma^2_{GB}}{(1+f)\sigma^2_{GB} + 2\sigma^2_{GW}}$$

Distribution of Q_{ST} values

An important advance was the observation by Whitlock (2008) that the distribution of realized Q_{ST} values (ignoring, for now, the additional error introduced by using the sample estimate, $\widehat{Q}_{ST}$, for the true value of the realization for a particular trait) can often be approximated using the Lewontin-Krakauer distribution for F_{ST} values (Equation 9.10a). Simulations by Whitlock confirmed the suggestion by Rogers and Harpending (1983) that, provided F_{ST} is small, the amount of information on population structure derived from the variance components of a quantitative trait is equivalent to that from a single-marker F_{ST}. Provided that the average F_{ST} is small, then under the null that $Q_{ST} = F_{ST}$, to a very good approximation, we have

$$\frac{n_d - 1}{\overline{F}_{ST}} Q_{ST} \sim \chi^2_{n_d-1}, \quad \text{implying} \quad Q_{ST} \sim \frac{\overline{F}_{ST}}{n_d - 1} \chi^2_{n_d-1} \tag{12.28a}$$

where $\overline{F}_{ST}$ is the average F_{ST} over the scored molecular marker loci, and n_d is the number of demes. This expression assumes that Q_{ST} is estimated without error, a point addressed shortly.

Michael Whitlock

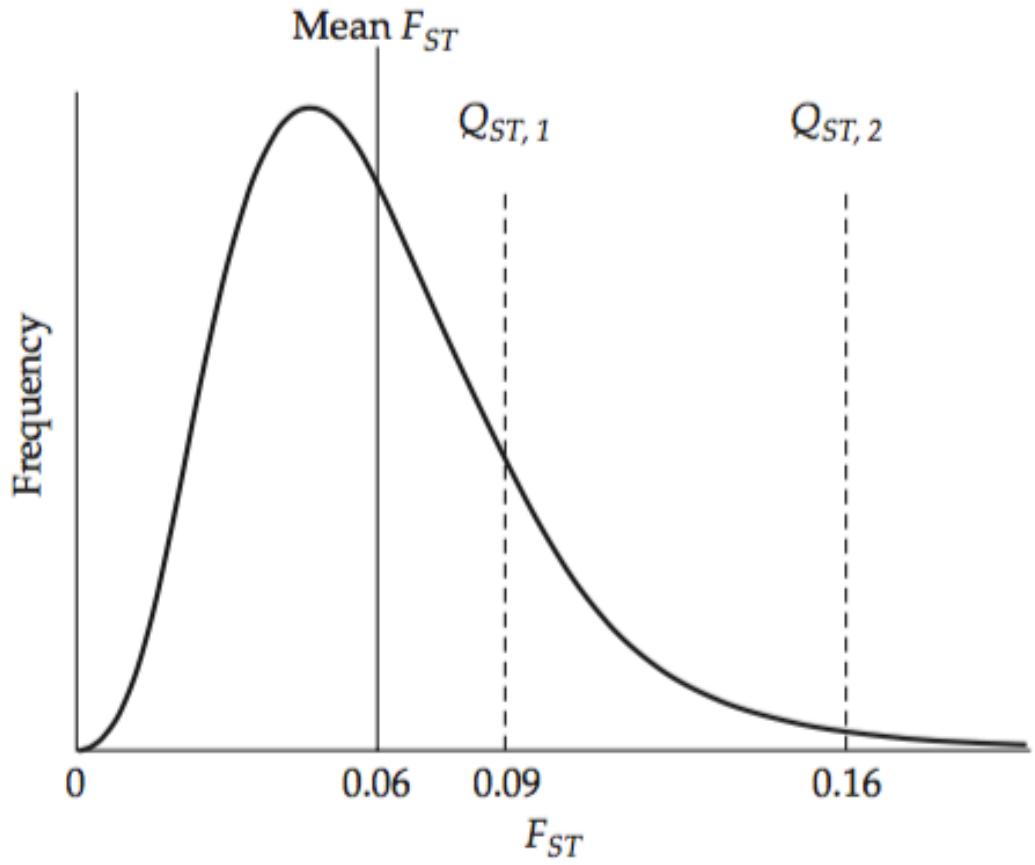


Figure 12.8 When $\overline{F}_{ST}$ is small, the Q_{ST} distribution for a neutral, completely additive trait should approximately follow the Lewontin-Krakauer distribution (Equations 9.10a and 12.28a). In this example, two traits, one with $Q_{ST} = 0.09$, and a second with $Q_{ST} = 0.16$ are both larger than $\overline{F}_{ST} = 0.06$, but only trait 2 is significant. (After Whitlock 2008.)

The requirement that $\overline{\overline{F}}_{ST}$ is small arises (in part) from χ^2 being defined over $(0, \infty)$, while Q_{ST} is restricted to $(0, 1)$. Hence, the approximation given by Equation 12.28a assumes that there is essentially no probability in the upper tail of a χ^2 above a critical value,

$$\Pr\left(\frac{\overline{\overline{F}}_{ST}}{n_d - 1}\chi^2_{n_d-1} > 1\right) = \Pr\left(\chi^2_{n_d-1} > \frac{n_d - 1}{\overline{\overline{F}}_{ST}}\right) \simeq 0 \qquad (12.28b)$$

To achieve this condition, Whitlock (2008) recommended an upper limit of $\overline{\overline{F}}_{ST} < 0.1$. For example, with $n_d = 2$, 5, and 10, the probabilities in Equation 12.28b (with $\overline{\overline{F}}_{ST} = 0.1$) become 0.002, $4{\cdot}10^{-8}$, and $2{\cdot}10^{-15}$, respectively.

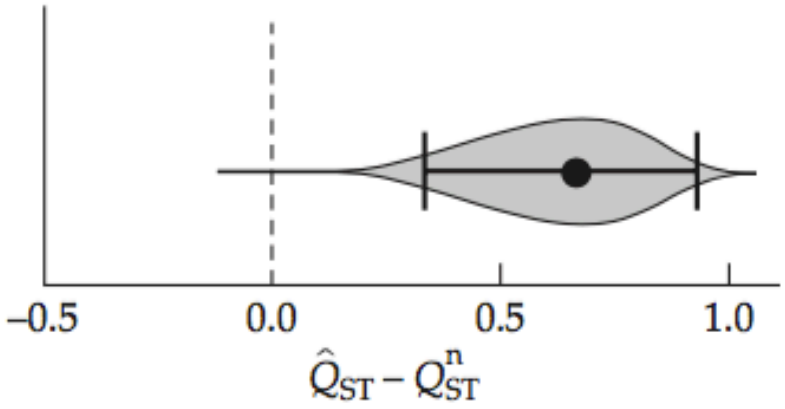


Figure 12.9 A violin plot for the distribution of the difference $(\widehat{Q}_{ST} - Q^n_{ST})$ for body length in the sea-run brown trout (*Salmo trutta*), using the resampling scheme suggested by Whitlock and Guillaume (2009), and detailed in Example 12.9. The width of the "violin" indicates the probability mass in that interval, the dot denotes the highest posterior probability, and the error bars the 95% credibility interval. Here this interval is completely above zero, demonstrating that $\widehat{Q}_{ST}$ is significantly in excess of its predicted neutral value given $\overline{F}_{ST}$. (After Rogell et al. 2012.)

Power

Insight into power is obtained by asking, under the null, how often the ratio Q_{ST}/F_{ST} exceeds some value, δ. Rearranging Equation 12.28a yields

$$\Pr\left(\frac{Q_{ST}}{\overline{F}_{ST}} > \delta\right) = \Pr\left(\frac{(n_d-1)Q_{ST}}{\overline{F}_{ST}} > \delta(n_d-1)\right) = \Pr\left(\chi^2_{n_d-1} > \delta(n_d-1)\right) \quad (12.28c)$$

Consider $n_d = 2$, as occurs when comparing two populations. How much larger must the true value of Q_{ST} be than the true value F_{ST} for this difference to be significant at the $\alpha = 0.05$ level? Because tests involving Q_{ST} are two-sided (either too large or too small being of interest), and $\Pr(\chi^2_1 > 5.02) = 0.025$, Equation 12.28c gives the critical value as $\delta = 5.02$. Hence, Q_{ST} must be in excess of 5 times $\overline{F}_{ST}$ to be significant at the 5% level. For $n = 10$, $\Pr(\chi^2_9 > 19.03) = 0.025$, or $\delta = 19.03/3 = 2.1$, and hence only a two-fold difference is required for significance. The same logic can be used to obtain the critical value when $Q_{ST} < F_{ST}$. For example, because $\Pr(\chi^2_9 < 2.7) = 0.025$, a value of Q_{ST} less than one third of $\overline{F}_{ST}$ (2.7/9 = 0.3) is significant at the 5% level when $n_d = 10$.

P_{ST}

P_{ST}: Approximating Q_{ST} with Phenotypic Data

Because of the requirement for assays in a common-garden arena, true joint studies of Q_{ST} and F_{ST} are not common. Pujol et al. (2008) noted that roughly half of the wild population studies they reviewed were not based on estimated additive variances. Instead, a phenotypic-based proxy for Q_{ST} was used, where within- and/or among-population *phenotypic* variances replace the more challenging estimates of additive variation. The former can easily be obtained via a standard ANOVA (e.g., Holand et al. 2011), while the latter require a series of parent-offspring or sib rearings in a common environment. A modification of this purely phenotypic approach is to use

$$\widehat{\sigma}^2_{GB} = c\,\widehat{\sigma}^2_{PB}, \qquad \widehat{\sigma}^2_{GW} = h^2\,\widehat{\sigma}^2_{PW} \tag{12.27a}$$

where c reflects the fact that only part of an observed phenotypic difference in means may be genetic (Merilä 1997; Leinonen et al. 2006; Sæther et al. 2007; Brommer 2011). Substitution into Equation 12.26 yields the P_{ST} statistic of Leinonen et al. (2006),

$$\widehat{P}_{ST,L} = \frac{\widehat{\sigma}^2_{PB}}{2(h^2/c)\,\widehat{\sigma}^2_{PW} + \widehat{\sigma}^2_{PB}} \tag{12.27b}$$

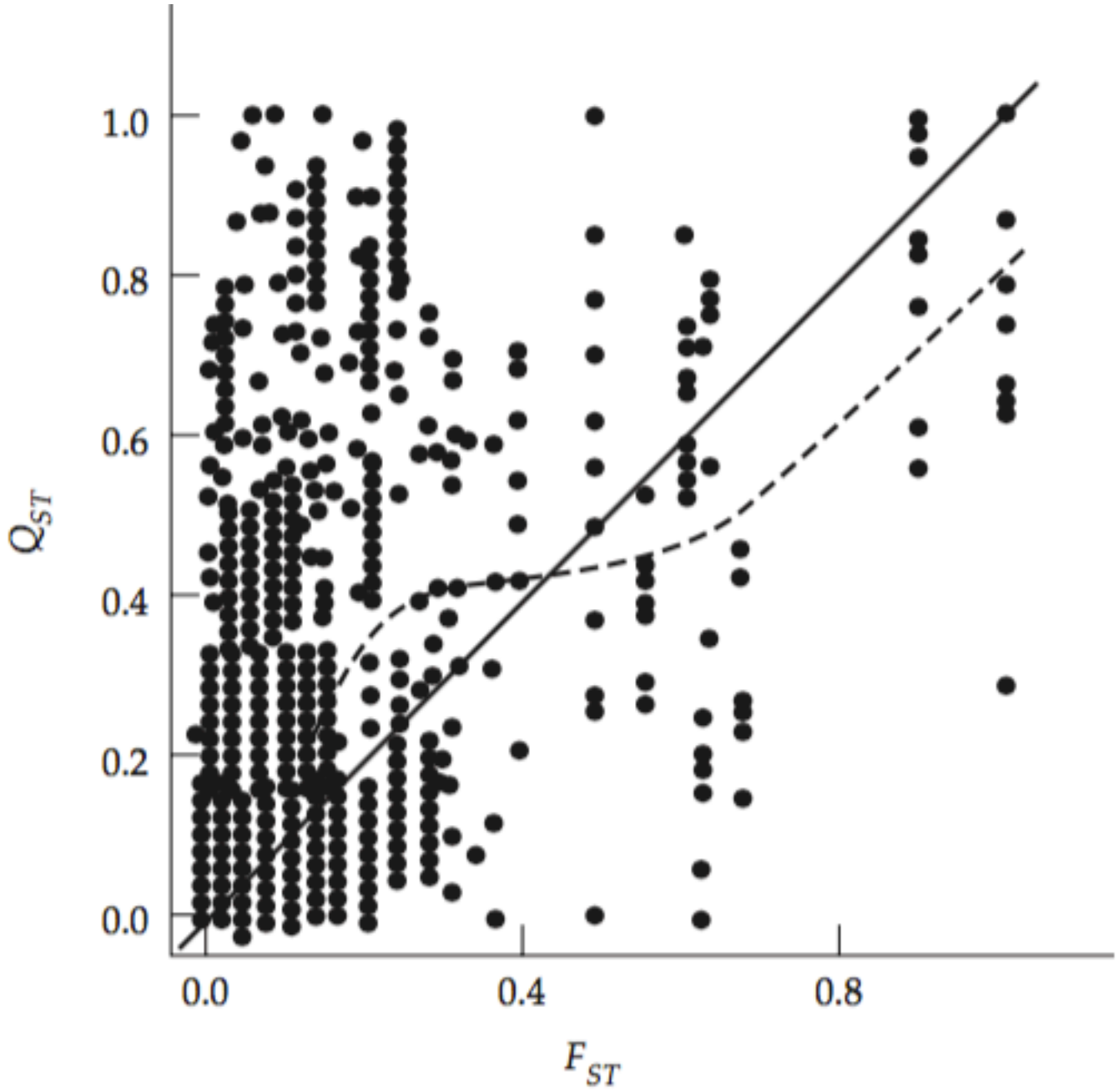
When $c = h^2$, this reduces to Equation 12.26a, with phenotypic variances replacing their genetic counterparts. Holand et al. (2011) suggested doing a sensitivity analysis by varying

Empirical Data

Results from the large number of Q_{ST} vs. F_{ST} comparisons from natural populations were summarized by Merilä and Crnokrak (2001), McKay and Latta (2002), and Leinonen et al. (2008, 2013). Values of Q_{ST} and F_{ST} are positively correlated, with $r = 0.24$ (Leinonen et al. 2013). Thus, there is a modest tendency for the structure of quantitative-trait variation to parallel the population structure for neutral alleles. The striking finding is that $Q_{ST} > F_{ST}$ for $\sim 70\%$ of all traits, which, taken at face value, suggested that diversifying selection was very widespread (Figure 12.10). Conversely, values of $Q_{ST} < F_{ST}$ are rare, despite the bias in this direction for neutral traits under a variety of conditions (discussed above), suggesting that persistent stabilizing or uniform selection is far less common.

One potential explanation for this trend of $Q_{ST} > F_{ST}$ was offered by Miller et al. (2008). They found that the variance of Q_{ST} is significantly larger than that for F_{ST} and noted a strong positive correlation in the data between Q_{ST} and the difference $(Q_{ST} - F_{ST})$. Hence, populations with larger Q_{ST} values tend to also have greater departures from F_{ST}. In particular, they noted that if more variable traits are overrepresented in the sampling process, this generates outliers of Q_{ST}, given the latter's larger variance, which in turn generates excessive $(Q_{ST} - F_{ST})$ values, even under neutrality. Whitlock (2008) further stressed this concern:

> It will always be possible to choose a set of traits that have higher than average Q_{ST} values. Traits chosen in this way cannot reliably be used to infer the extent of spatially heterogeneous selection. Examination of the traits chosen for many Q_{ST} studies makes one wonder whether traits are in fact always chosen with previous knowledge of the likely results.


1.0
0.8
0.6
0.4
0.2
0.0
Q_{ST}
0.0
0.4
0.8
F_{ST}

A second source of bias in comparisons of Q_{ST} and F_{ST} was noted by Edelaar and Björklund (2011) and Edelaar et al. (2011). Markers with high mutation rates underestimate F_{ST}, and the most widely used markers in early Q_{ST}/F_{ST} studies, microsatellites, have high mutation rates. As shown in Figure 12.11, there is a strong positive relationship between the polymorphism level of a marker (with highly polymorphic markers having higher mutation rates) and the excess values of Q_{ST} over F_{ST}. Note that most of this trend is driven by studies employing microsatellites, with allozyme studies showing an excess of Q_{ST} largely independent of their polymorphism level.

Thus, the striking trend of $Q_{ST} > F_{ST}$ is certainly inflated by ascertainment bias, and somewhat inflated by the use of highly polymorphic markers (which is a more recent trend), making it difficult to make any general statement about how commonly diversifying selection structures quantitative traits in subdivided populations. As noted by Whitlock (2008), "While useful, Q_{ST} is a crude measure of the genetic differentiation of a trait caused by local adaptation."

Closing Comments: Q_{ST}, F_{STQ}, and Linkage Disequilibrium

Tests comparing F_{ST} values at candidate loci against the distribution of F_{ST} values at putatively neutral markers were discussed at length in Chapter 9. Comparisons of Q_{ST} to F_{ST} are a step removed, in that, ideally, we would like to contrast the F_{STQ} value (the average F_{ST} value for loci underlying our focal trait) against the genome-wide F_{ST} neutral standard. Given the near impossibility of locating all such causative loci, we have instead been using Q_{ST}, as with an additive trait, this should track the F_{STQ} values at the underlying causative loci. However, as is detailed in Chapters 16 and 24, allele-frequency changes are not the only route through which genetic variances (and hence the components of Q_{ST}) can change. Selection-generated gametic-phase disequilibrium (LD)—*even among unlinked loci*—can have a dramatic effect, even in situations where little allele-frequency change occurs. This impact of LD on Q_{ST} was stressed first by Latta (1998, 2005), and later by Le Corre and Kremer (2003, 2012; Kremer and Le Corre 2012). Because Q_{ST} is based on variance components, it can be influenced by linkage disequilibrium, which generates covariances between alleles at different loci, either inflating or deflating the resulting variances. When this happens, the values of Q_{ST} and F_{STQ} can become decoupled, and (as we will see) Q_{ST} can have more power to detect selection than F_{STQ} (even presuming we could locate all the underlying loci).

Thus, while a significant departure of Q_{ST} from the background value of F_{ST} is usually taken as indicating a shift in the F_{STQ} values at the underlying trait loci, this is only strictly correct when linkage disequilibrium is absent. Even in cases where selection induces little allele-frequency change (and hence little shift in F_{STQ} relative to the background F_{ST}), selection-induced disequilibrium (i.e., shifts in gamete, as opposed to allele, frequencies) can still generate a significant Q_{ST} signal. In particular, under the infinitesimal model, there is essentially no shift in the allele frequencies at underlying loci ($F_{STQ} \simeq F_{ST}$), but there can be a substantial change in the genetic variances due to selection-induced LD (Chapters 16 and 24), and hence a perturbation of Q_{ST} away from F_{STQ}. In such a setting, a direct comparison of F_{STQ} to the genome-wide F_{ST} standard would not reveal any evidence of selection, *but* a comparison of Q_{ST} (with its LD-shifted variance components) against F_{ST} might. Hence, under polygenic sweep conditions (Chapter 8), an appropriately performed Q_{ST} test might detect selection signatures missed by allele-frequency based tests.

To expand on this point, we need to consider how the within- and among-population LD (Ohta 1982) impact Q_{ST}. Letting the subscript x denote either within- or among-population values ($x = w$ and $x = a$, respectively), we can express the variances comprising Q_{ST} as

$$\sigma_x^2 = \sigma_{x,0}^2 + d_x = (1 + \phi_x)\sigma_{x,0}^2, \quad \text{where} \quad \phi_x = \frac{d_x}{\sigma_{x,0}^2} \tag{12.30a}$$

where $\sigma_{x,0}^2$ is the linkage equilibrium value, d_x is the disequilibrium contribution generated by covariance among alleles at different loci (Equations 16.1 and 16.2), and ϕ_x is the ratio of the disequilibrium contribution to the linkage-equilibrium (i.e., genic) variance (note that ϕ_x is negative when d_x is negative). As discussed in Chapter 16, stabilizing or directional selection within a population generates negative d, so we often expect negative within-population LD (negative values of d_w and ϕ_w).

Turning to the among-population LD, Latta (1998) noted that if each population is under stabilizing selection for a different optimum value (θ), then for an additive trait where the

population means have reached their optimal values,

$$d_a = \sigma_\theta^2 - 2F_{STQ}\,\sigma_A^2 \tag{12.30b}$$

where σ_θ^2 is the variance in the optimum value over populations, and σ_A^2 is the expected additive genetic variation if the populations were to be randomly mated to form a single, panmictic, population (in linkage equilibrium). With nearly uniform selection (the variance in θ values over demes is small) and reduced migration (so that F_{STQ} is large), Equation 12.30b gives a negative covariance ($d_a, \phi_a < 0$) between trait-increasing alleles at different loci across demes, reducing the among-group variance σ_{GB}^2 below its linkage-equilibrium value. Conversely, if diversifying selection is strong (σ_θ^2 is large) and gene flow is high (F_{STQ} is small), a positive covariance is expected ($d_a, \phi_a > 0$), and σ_{GB}^2 is inflated relative to its value in the absence of LD. Thus, Q_{ST} often magnifies the effect of selection over what is expected from changes in F_{STQ} alone, with significant changes in Q_{ST} (relative to F_{ST}) possible even when little differentiation has occurred at the underlying QTLs ($F_{STQ} \simeq F_{ST}$).

For a completely additive trait, Le Corre and Kremer (2003) quantified the influence of LD on Q_{ST} by noting that the relationship between Q_{ST} (based on variance components) and F_{STQ} (based on the underlying loci) is given by

$$Q_{ST} = \frac{(1+\phi_a)F_{STQ}}{(\phi_a - \phi_w)F_{STQ} + 1 + \phi_w} \tag{12.30c}$$

where ϕ_x is given by Equation 12.30a. Note that Q_{ST} equals F_{STQ} only when the among- and within-population LD values are equal ($\phi_a = \phi_w$). Using Equation 12.30c, Kremer and Le Corre (2012) showed that $Q_{ST} > F_{STQ}$ when $\phi_a > \phi_w$. Given that stabilizing selection within populations generates negative values of ϕ_w, while diversifying selection (variation in the optimum over populations) generates positive values of ϕ_a (Equation 12.30b), this combination amplifies the signal in Q_{ST} over that generated from F_{STQ}. As $Q_{ST} > F_{ST}$ is the signal for divergent selection (Table 12.2), while our last result implies that $Q_{ST} > F_{STQ} > F_{ST}$, the impact of LD is to *magnify* the impact of divergent selection over that expected from allele-frequency change alone (F_{STQ}). Again, the salient point is that even if the difference between F_{STQ} and F_{ST} is small, the difference between Q_{ST} and F_{ST} can still be large.

Hence, while Q_{ST}-based tests are fraught with complications, if properly performed (which is no small feat), they may actually be more powerful than a scan for F_{ST} outliers at known candidate genes for the trait of interest (Chapter 9). While F_{ST}-based scans are trait independent, knowledge of the potential target trait or traits allows Q_{ST}, and thus further information from LD, to be exploited. We return to this point below when considering certain trait-augmented marker-based tests.

Table 12.2 Interpretation of Q_{ST} versus F_{ST} comparisons.

Observation	Interpretation
$Q_{ST} > F_{ST}$	**Divergent selection**: spatial variation in trait values in excess of neutral expectation.
$Q_{ST} = F_{ST}$	Consistent with divergence expected under drift. Does not rule out selection, but does not support it either.
$Q_{ST} < F_{ST}$	**Convergent selection**: spatial variation in trait values less than neutral expectation. Similar trait values are favored over populations.

Example 12.7. Using candidate genes in the photoperiod pathway (detected in *Arabidiopsis thaliana*), Ma et al. (2010) explored whether variation in these loci is involved in growth cessation in populations of European aspen (*Populus tremula*) across a latitudinal gradient in Sweden. Their population sample consisted of 10 trees from each of 12 sites (spanning roughly ten degrees of latitude), scoring 113 SNPs from 23 photoperiod genes and 93 SNPs from 21 random control genes. Six of the photoperiod SNPs showed significant associations with growth cessation (with no evidence of epistatic interactions between these detected loci). While four of the photoperiod SNPs showed a significant correlation with latitude, the F_{ST} values for the photoperiod and control groups of SNPs were not significantly different (0.018 vs. 0.016, corresponding to F_{STQ} and F_{ST}, respectively), although photoperiod SNPs showed a significantly greater variance in F_{ST} values relative to the control SNPs. None of the individual SNPs showed significant F_{ST} departures from the control loci, so that even when using candidate genes in a known pathway that is likely to be under selection, no signature of selection was observed in individual F_{STQ} values. However, as a group, the photoperiod SNPs showed a significant excess of pairs of alleles from different loci correlated with each other (i.e., showed LD), while no such pattern was seen with the control SNPs. Further, the highest five of the allelic pairs correlated between loci also involved either one (or both) alleles (SNPs) that showed significant clines with latitude. Thus, while selection in this study did not seem to generate a significant departure between F_{STQ} and F_{ST}, it did generate among-population covariances.

Summary

- Q_{ST} is trait-specific
- Often approximated by using P_{ST} (very problematic)
- Ascertainment bias issues, more variable traits usually chosen (i.e., a nonrandom set)
- Q_{ST} does not necessarily track F_{STQ}
  - This is a good thing, as small allele-frequency shifts result in little change in F_{STQ}
  - However, LD generated by selection can significantly change Q_{ST} even in the face of NO F_{STQ} change