Lecture 4:
Test for Neutrality in Trait Evolution

Bruce Walsh
Second Bangalore School of
Population Genetics and Evolution
25 Jan- 6 Feb 2016

Outline

- Trait divergence under drift and mutation
- Lande's tests
  - drift only (constant variance test): short time scales
  - drift-mutation equilibrium (MDE test): long time scales
- Brownian and Ornstein-Uhlenbeck models
- F_{ST} vs Q_{ST} tests
- QTL-based tests: Orr's method
- Divergence in gene expression data: neutral or not?

Trait mean divergence

- Tests for whether an observed amount of divergence is consistent with drift appear in a number of venues
  - Divergence over time in a single population
    - Time series in the fossil record
  - Divergence between a set of populations
  - Divergence in the pattern of fixed QTL alleles
  - Divergence in levels of gene expression
- Drift should be tested, and rejected, before a selective explanation is offered

Drift in traits

- Genetic drift changes the frequency of alleles underlying traits
- This can, in turn, results in drift in the trait means over time
- Simple model: An additive locus with two alleles (a = effect, p(t) = frequency at generation t)
  - Mean = $2ap(t)$
  - Variance in the mean is
  - $E[(2ap(t))^2] - E[2ap(t)]^2 = 4a^2Var[p(t)]$
  - Var[p(t)] is the drift variance in allele frequnecy

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

Substituting into the previous expression and accounting for the initial sampling due to the founding pop ($1/N_{fo}$) gives

$$\sigma_B^2(t) = 4\sum_{i=1}^{n} a_i^2 p_i(0)[1-p_i(0)]\left\{\frac{1}{N_{fo}}+\left[1-\left(1-\frac{1}{2N_e}\right)^t\right]\right\}$$

$$= \left(\frac{1}{N_{fo}}+2f_t\right)\sigma_A^2(0) \qquad f_t = 1-\left(1-\frac{1}{2N_e}\right)^t$$

This accounts for the partitioning of the initial variance from drift. For large t (ignoring mutation), the between-group variance approaches twice the additive additive variance in the trait, V_B --> 2 V_A

The drift process also generates a correlation among the values of the mean *within* a given population over time (but no correlation *between* populations)

$$\sigma_B(t,t') = \left(\frac{1}{N_{fo}} + 2f_t\right)\sigma_A^2(0) \qquad \text{for } 0 < t < t'$$

The implication of this result is that the residuals in a regression of a mean over time are both *correlated* and *homoscedastic*, so that generalized least squares (GLS), not ordinary least squares (OLS), must be used in any regression-based analysis

Important to keep this in mind in the analysis of selection experiments! (Chapter 18)

Measures of divergence

- For two populations, can use absolute difference between means, *d*
- For *L* populations, use the between-group variance V_B, the variance in group means

$$V_B = \frac{1}{L-1}\sum_{i=1}^{L}(\overline{y}_i - \overline{y}_.)^2$$

- Slightly better to estimate using ANOVA

$$V_B(t) = \frac{MS_B - MS_W}{n_0}$$

Connection between d^2, V_B

- Consider V_B for $L = 2$

$$V_B = \left(\overline{y_1} - \frac{\overline{y_1} + \overline{y_2}}{2}\right)^2 + \left(\overline{y_2} - \frac{\overline{y_1} + \overline{y_2}}{2}\right)^2$$
$$= \frac{(\overline{y_1} - \overline{y_2})^2}{2} = \frac{d^2}{2}$$

- Hence, $2V_B = d^2$

Tests for excessive divergence

- Assuming t << N, and that mutation can be ignored, tests for excessive between-group variance relative to drift can be constructed.
- Use fact that estimates of variances are chi-squared distributed

$$\text{If } y_i \sim N(0, \mu), \quad \text{then for Var} = \frac{1}{n-1}\sum (y_i - \bar{y})^2,$$

$$(n-1)\text{Var} \sim \sigma^2 \chi^2_{n-1}$$

Lande's Constant Variance test

Expected: $V_B(t) = V_A(0)\, t/Ne$

$$F_{CV} = \frac{V_B(t)}{t \cdot V_A(0)/N_e}$$

$$V_B(t) = \frac{1}{L-1}\sum_{i=1}^{L}(\overline{y}_i - \overline{y}_.)^2 \sim \frac{\sigma_B^2(t)}{L-1}\chi^2_{L-1}$$

$$V_B(t) \sim \frac{t\sigma_A^2(0)/N_e}{L-1}\chi^2_{L-1}$$

Thus it follows that

$$\mathrm{F}_{CV} \sim \frac{\chi^2_{L-1}}{L-1} \sim F_{L-1,\infty}$$

More generally, if Var(A) is estimated, then the infinite degrees of freedom is replaced by the degrees of freedom in the design the estimated Var(A)

Example 12.2. Lande (1977) used Equation 12.9a to evaluate the results of a 12-year divergence experiment involving five populations of *Drosophila pseudoobscura* (Anderson 1973). Two of the populations had been maintained at 25°C, two at 27°C, and one at 16°C. They were then raised in two common environments (16 and 25°C) and measured for wing length. Estimates of the additive genetic variance for these two environments were 0.88 and 0.77, while the among-population variances were approximately 6.62 and 4.37 respectively. An approximate upper bound for the number of generations of divergence is $t = 150$, whereas the effective population size probably always exceeded $N_e = 1000$. The use of these extreme bounds gives conservative estimates of F_{CV}, making it more difficult to demonstrate diversifying selection on wing length. The resulting values (for the two environments) are

$$\mathrm{F}_{CV,1} = \frac{6.62}{150 \cdot 0.88/1000} = 50.15, \quad \text{and} \quad \mathrm{F}_{CV,2} = \frac{4.37}{150 \cdot 0.77/1000} = 37.84$$

Given that $\Pr(F_{4,\infty} > 4.6) = 0.001$, both values are highly significant. Thus, the hypothesis that the observed line divergence is solely attributable to random genetic drift can be rejected confidently. More likely, the different thermal conditions resulted in selection for different wing lengths.

$$\mathrm{F}_{CV} = \frac{V_B(t)}{t \cdot \mathrm{V}_A(0)/N_e}$$

Mutation and Drift

- When considering the variance introduced into a trait by new mutation, have to weight the mutation rate times the effect of a new mutation
- The result is the mutational variance V_m (also denoted by σ_m^2)
- This can be estimated by the rate of divergence between a set of inbred lines
- Typically, $V_m/V_E \sim 1/1000$

$$\sigma_B^2(t) = 2\sigma_m^2 t + 2\left(\sigma_A^2(0) - 2N_e\sigma_m^2\right)\left(1 - e^{-t/(2N_e)}\right)$$

At equilibrium,

Additive variance within a population is $2N_eV_m$

Divergence = $2V_mt$

These are the neutral trait analogs to the neutral allele values $4N_eu$ and $2ut$

Brownian Motion tests

- A standard model for divergence is Brownian motion,
  - $E[\Delta x] = 0$, $Var[\Delta x] = b$
  - Under this model, $x(t) \sim N(x_0, bt)$, where x_0 is the initial value
  - The variance becomes unbounded in t
- Lande (1976):
  - Assumed $b = V_A/N_e = h^2V_z/N_e$
- Turelli et al (1988)
  - V_A is function of N_e, with $V_A = 2N_eV_m$, giving
  - $b = V_A/N_e = 2N_eV_m/N_e = 2V_m$

Tests using *d* are thus based on the normal

- Let U ~ N(0,1) denote the unit normal.
- Pr(|U| $\leq$ 1.96) = 0.05

$$\Pr\left(|\mu(t)-\mu(0)|\leq d\right)=\Pr\left(\frac{|\mu(t)-\mu(0)|}{\sigma_t}\leq\frac{d}{\sigma_t}\right)=\Pr\left(|U|\leq\frac{d}{\sigma_t}\right)$$

Hence, tests of excessive divergence relative to drift (at p = 0.05) when

1.96 $\leq$ d/σ_t or d $\geq$ 1.96 σ_t

Tests of divergence

- Tests start with some assumed value of N_e or V_m
- Since divergence is caused by either small N_e or large V_m, directional selection on the mean suggested if too much divergence
  - Assumed N_e is too large or V_m too small
- Likewise, divergence is retarded with N_e large or V_m small. Hence, stabilizing selection indicated by too little divergence
  - Assumed N_e is too small or V_m too large

$$\Pr\left(|\mu(t)-\mu(0)| \le d\right) = \Pr\left(\frac{|\mu(t)-\mu(0)|}{\sigma_t} \le \frac{d}{\sigma_t}\right) = \Pr\left(|U| \le \frac{d}{\sigma_t}\right)$$

Where U ~ N(0,1) is a unit normal, where Pr(|U| $\leq$ 1.95) = 0.05

Lande's MDE test (mutation-drift equilibrium) finds the largest N_e that could account for the amount of observed divergence

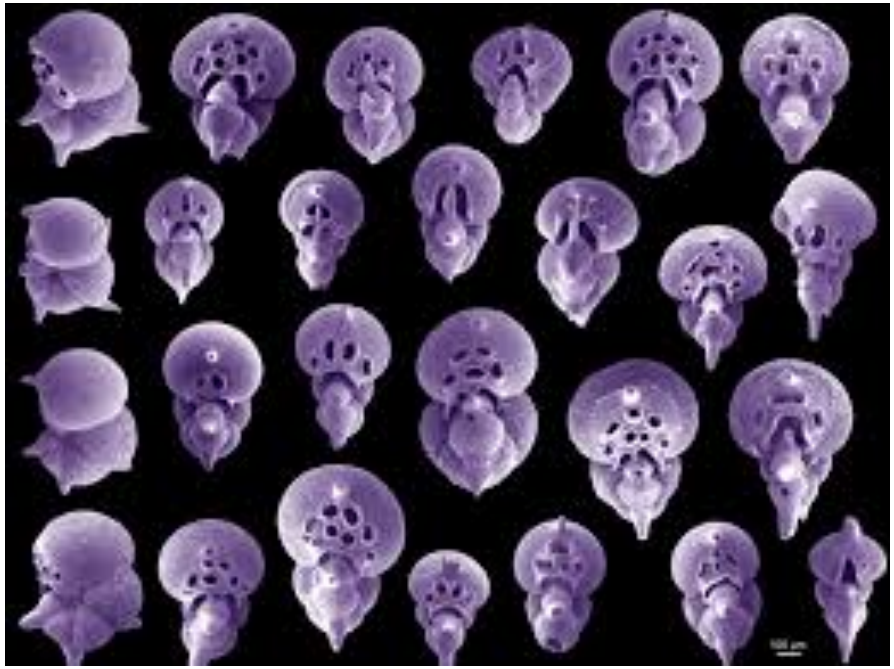
$$1.96 = \frac{d}{\sqrt{th^2\sigma_z^2/N_e}}, \quad \text{implying} \quad (1.96)^2 th^2\sigma_z^2 = N_e d^2$$

$$\widehat{N_e} = \frac{t \cdot h^2 \cdot 1.96^2}{d_*^2} = 3.84 \cdot \frac{t\,h^2}{d_*^2}$$

Example 12.3. Reyment (1982) observed a change of $1.49\sigma_z$ over roughly 5×10^5 generations in the size of a Cretaceous foraminifer. Taking a typical heritability value of 0.3, Equation 12.18b gives the largest population size consistent with this amount of divergence as

$$\widehat{N_e} = 3.84 \cdot \frac{t\,h^2}{d_*^2} = 3.84 \cdot \frac{5 \times 10^5 \cdot 0.3}{1.49^2} \simeq 260,000$$

However, paleontological data suggests that the effective population size was greater than 10^6, implying that drift was unlikely to account for such a rapid divergence. Assuming h^2 values of 0.5, 0.7, and 1 yields critical N_e values of 433,000; 607,000; and 867,000, so that only for assumed h^2 values close to one does the critical maximal size under drift approach the assumed size of $N_e > 10^6$.



Turelli et al. test

- V_A a function of N_e, so test using V_m
- Neutrality is not rejected (at a test of level α) if

$$\Pr\left[\left(\frac{L-1}{X_{1-\alpha/2,L-1}}\right)V_B(t) \leq 2t\sigma_m^2 \leq \left(\frac{L-1}{X_{\alpha/2,L-1}}\right)V_B(t)\right] = 1-\alpha$$

Not too large = directional

Not too small = stabilizing

Example 12.4. Let's return to Reyment's foraminifer data from Example 12.3. Using the original Lande model, we rejected the hypothesis that drift could have accounted for the divergence. Applying Equation 12.21a (and taking $h^2 \simeq 0.5$), the hypothesis of drift accounting for excessive divergence is not rejected when

$$h_m^2 > 0.10 \cdot 0.5 \cdot \frac{1.49^2}{5 \times 10^5} = 2.2 \times 10^{-7}$$

This is several orders of magnitude lower than typical values of the scaled mutation variance, and thus this pattern of divergence is not too excessive for drift.

Hence, N_e-based test rejects drift, when V_m-based fails to reject. One issue is that V_A is usually MUCH less than $2N_eV_m$, due to new mutations influencing the trait being slightly deleterious. Hence I prefer N_e-based tests.

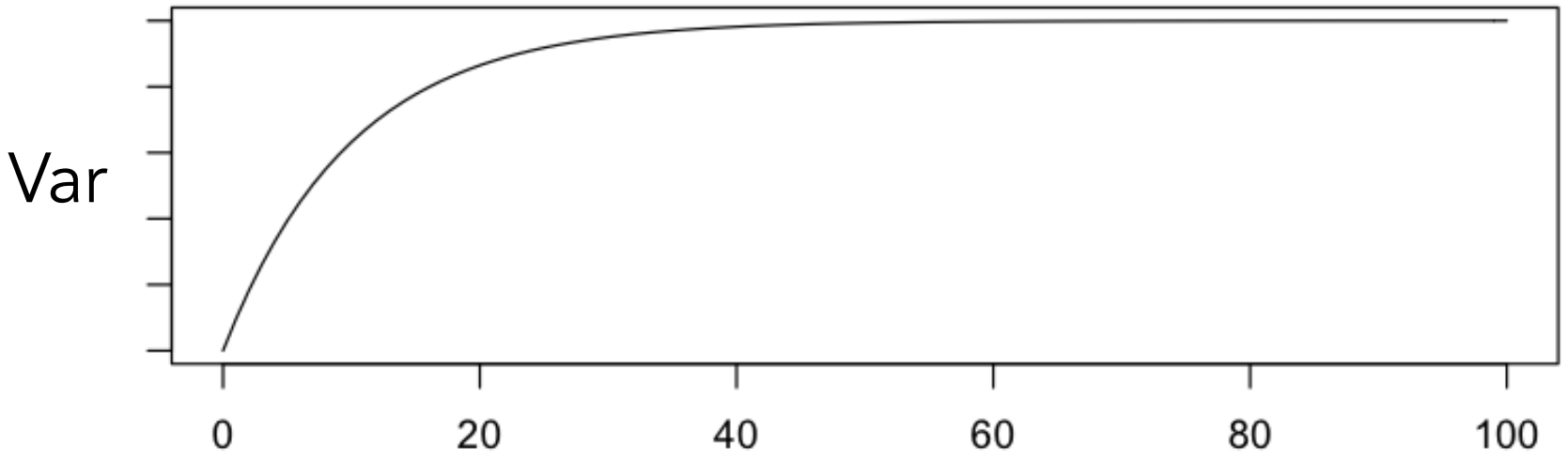
Ornstein-Uhlenbeck Models

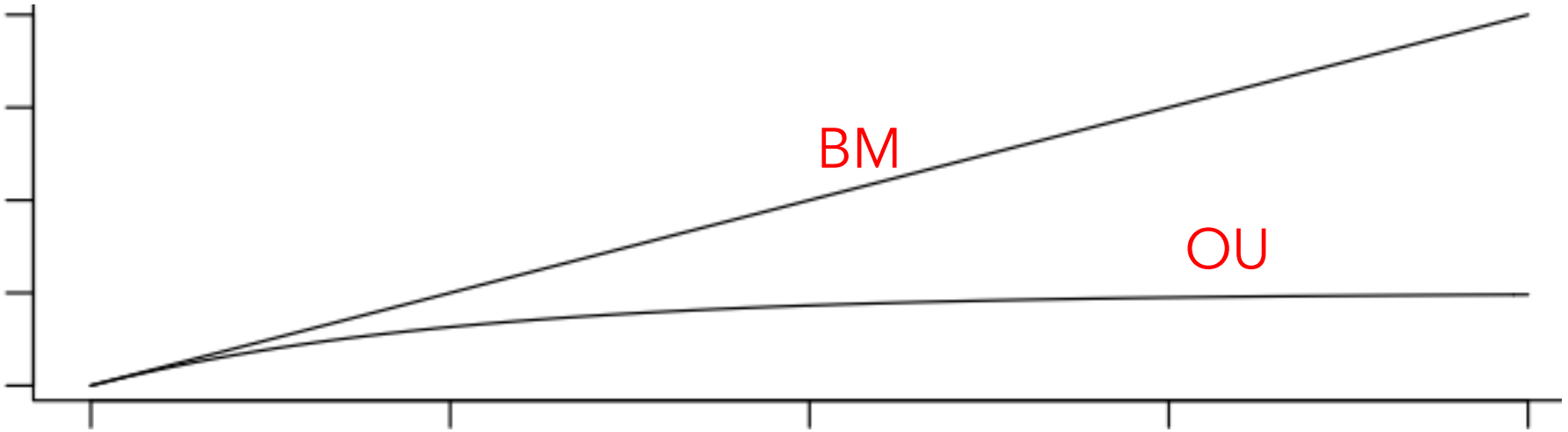
- Brownian motion with a restoring force towards a value θ (model for stabilizing selection)
  - $E[\Delta x] = -a(x-\theta)$, $Var[\Delta x] = b$
  - Under this model, x(t) is again normal, but

$$\mu_t = \theta + [x_o - \theta]\exp(-at)$$

$$\sigma_t^2 = \frac{b}{2a}[1 - \exp(-2at)]$$

Mean → θ, Variance → b/(2a)





Application to expression data

Under a Brownian-motion model, the expected divergence (measured by the between-group variance) scales linearly with divergence time t.

In contrast, under an Ornstein-Uhlenbeck process (drift countered by stabilizing selection), the divergence approaches an asymptotic value.

Bedford and Hartl (2009) used such a process to fit the pattern of expression divergence within a clade of seven species of *Drosophila*. In accordance with the OU model (and consistent with stabilizing selection), they found that the divergence variance does not linearly increase with time, but rather quickly approaches an asymptotic value.

F_{st} vs Q_{st}

Q_{ST}: Partitioning Additive Variance Over Subpopulations

Consider a quantitative trait with a purely additive-genetic basis, and denote genetic variance for the trait in the entire metapopulation under the assumption of panmixia by σ^2_G. From Table 11.3 (taking $f = Q_{ST}$), the within- and among-subpopulation 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

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

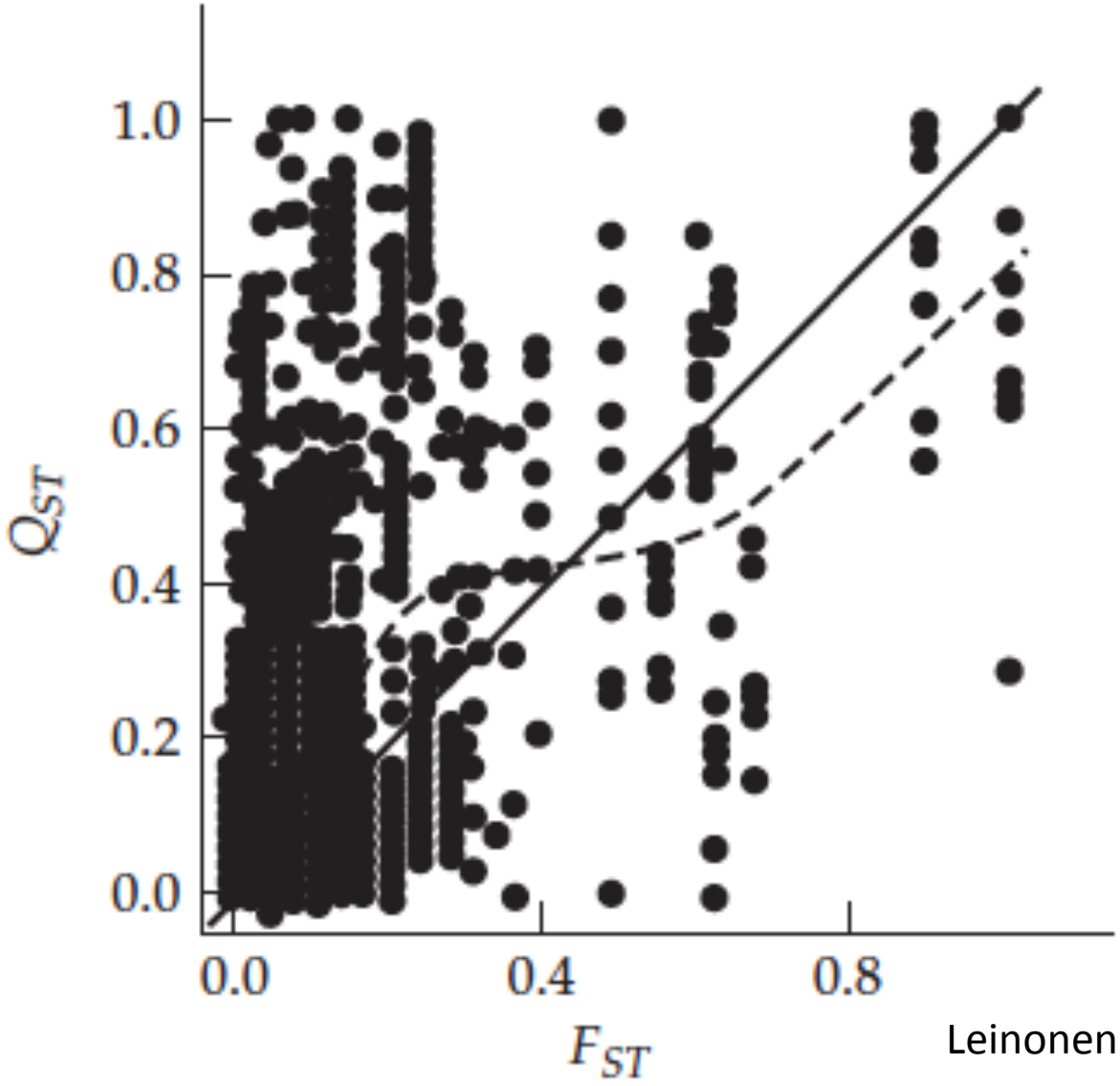
Table 12.2. Interpretation of Q_{ST} versus F_{ST} comparisons.

Observation	Interpretation
$Q_{ST} > F_{ST}$	**Divergent selection**: spatial variation in favored trait values
$Q_{ST} = F_{ST}$	Consistent with divergence expected under drift. Does not rule out selection, but does not support it either.
$Q_{ST} < F_{ST}$	**Convergent selection**: similar trait values favored over subpopulations.

Idea

- For additive traits, $V_A = 2a^2p(1-p)$
- F_{ST} also scales as p(1-p)
- Dominance causes Q_{ST} and F_{ST} to deviate under neutrality, with the directional of the inequality a function of the population structure.
- Requires estimates of genetic variances
  - Often done using phenotypic variances (P_{ST}), but this is VERY error-prone and should not be used

What do the data show?



Leinonen et al. (2008)

Lots! Of pitfalls with this approach

- Whitlock: 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 fact always chosen with previous knowledge of the likely results.

QTL-based tests: Orr's method

- When we have QTLs for a trait, we can also to for departures from randomness.
- Key idea: If divergence is due to only drift, the distribution of allelic effects is random (i.e, no directionality for selecting + or – QTL alleles)
- Two versions of Orr's test
  - QTL sign test for equal effects (QTLST-EE)
    - QTL alleles assumed to have equal (magnitude) effects
  - QTL sign test (QTLST)
    - QTL effects from some distribution
  - The ideal of either test is whether, conditional on a line being high, it has an excess number of + alleles

QTL sign test for equal effects (QTLST-EE)

The high lines must have at least [n/2] equal + alleles

$$[n/2] = \begin{cases} n/2+1 & \text{for } n \text{ even} \\ (n+1)/2 & \text{for } n \text{ odd} \end{cases}$$

$$\Pr(n_+ = k) = \binom{n}{k}(1/2)^k(1/2)^{n-k} = \binom{n}{k}(1/2)^n$$

The probability of seeing n_+ plus alleles, conditioned on at [n/2] plus alleles is

$$\Pr(n_+ \geq n_{+obs} \,|\, n_+ \geq [n/2]) = \sum_{i=n_{+obs}}^{n} \binom{n}{i} \Big/ \sum_{j \geq [n/2]}^{n} \binom{n}{j}$$

$$\Pr(n_+ \geq n_{+obs} \,|\, n_+ \geq [n/2]) = \sum_{i=n_{+obs}}^{n} \binom{n}{i} \Big/ \sum_{j \geq [n/2]}^{n} \binom{n}{j}$$

Example 12.6. True et al. (1997) found that none of the eight detected QTL for the posterior lobe in the male genitalia in a *Drosophila* cross were **antagonistic**, namely alleles in the opposite direction of the line value, such as low (minus) alleles in the high line or high (plus) alleles in the low line. Orr suggested that the equal effects model may not be unreasonable for this trait. Assuming this, we have $n = 8$, $[n/2] = 5$, and $n_{+obs} = 8$. Equation 12.31a gives

$$p = \sum_{i=8}^{8} \binom{8}{i} \Big/ \sum_{j \geq 5}^{8} \binom{8}{j} = \frac{\binom{8}{8}}{\binom{8}{5} + \binom{8}{6} + \binom{8}{7} + \binom{8}{8}} = \frac{1}{56 + 28 + 8 + 1} = 0.011$$

Anderson and Slatkin (2003) note that this test can be highly biased by trait ascertainment (where the investigator, often unconsciously, chooses traits showing excessive divergence, further biasing traits for an excess of plus alleles).

QTL sign test (QTLST)

Let R be the difference between lines, either the actual observed difference, or the difference based on summing the effects over all detected QTLs. With a distribution of QTL effects in hand, one can then conditional on the number of plus alleles given that the high line is R in excess of the low line.

$$p = \Pr(n_+ \geq n_{+obs} \,|\, G \geq R) = \sum_{i=n_{+obs}}^{n} \Pr(n_+ = i \,|\, G \geq R)$$

Two interesting applications of sign tests to particular problems were by Albertson et al. (2003) and Muir et al. (2014). Albertson examined traits in the massive species radiation occurring among cichlid fishes in the African rift lake. One striking feature of this radiation is extensive convergent evolution across lakes in feeding morphology, suggesting directional selection. The authors used QTLST (with effect sizes drawn from a gamma distribution) to examine this, crossing two wild species from Lake Malaw. Since most individual traits had less than six QTLs, they grouped traits, finding for jaw and teeth features that only 4 of the 46 QTLs where antagonistic, a highly significant p value, and supporting directional selection on these feeding traits.

Muir et al. (2014) used QTLST along with an estimated distribution of allelic effects in tomatoes (*Lycopersicon*) to examine leaf-related traits in wild species thought to be associated with adaptation to precipitation. They found no significant departure from neutrality for two leaf and two trichome (leaf hair) traits, but a significant departure from neutrality for two stomatal traits. They computed p values using both QTLST and QLTST-EE, and found (in agreement with Anderson and Slatkin 2003) that QTLST was more conservative, giving p about twice as large than those from QLTST-EE.

Divergence in expression

- Gene expression level is now widely used as a quantitative trait. A number of workers have tested for whether this pattern is neutral
  - Fitting Brownian vs. OU models to the data
  - Comparisons of within- vs. between-group variance
  - eQTL (expression QTLs)-based tests

Comparing within vs. between group variation

Expected values under neutrality

$$V_B \sim (2\tau\sigma_m^2) \cdot \chi^2_{L-1}/(L-1)$$ Between

$$Var \sim (2N_e\sigma_m^2) \cdot \chi^2_{k-1}/(k-1)$$ Within

Consider the test statistic

$$\mathrm{F}_{MDE'} = \frac{V_B}{Var} \cdot \left(\frac{2N_e\sigma_m^2}{2\tau\sigma_m^2}\right) = \frac{V_B}{Var} \cdot \left(\frac{N_e}{\tau}\right) \sim \frac{\chi^2_{L-1}/(L-1)}{\chi^2_{k-1}/(k-1)}$$

This follows an F distribution under netutrality

$$\mathrm{F}_{MDE'} \sim F_{L-1,k-1}$$

$$\mathrm{F}_{MDE'} = \frac{V_B}{Var} \cdot \left(\frac{2N_e\sigma_m^2}{2\tau\sigma_m^2} \right) = \frac{V_B}{Var} \cdot \left(\frac{N_e}{\tau} \right) \sim \frac{\chi^2_{L-1}/(L-1)}{\chi^2_{k-1}/(k-1)}$$

the critical ratio in Equation 12.33b is with respect to $2N_e/\tau$. Divergence was scored separately between *melanogaster* and each of the other two species ($L = 2$), with Var estimated from the four *melanogaster* lines ($k = 4$). The resulting upper and lower 2.5% values critical values follow first by noting that $\Pr(F_{1,3} \geq 17.4) = 0.025$ and $\Pr(F_{1,3} \leq 0.001) = 0.025$. The effective population size for *melanogaster* was estimated as $N_e = 3 \times 10^6$, while the total divergence times (in generations) were estimated as $\tau = 2t = 4.6 \times 10^7$ (*melanogaster* - *simulans*) and $\tau = 10.2 \times 10^7$ (*melanogaster* - *yakuba*). Hence the critical values for excessive divergence are

$$F_{c,mel-sim} = 17.4 \cdot \frac{4.6 \times 10^7}{6 \times 10^6} = 133.4, \quad F_{c,mel-yak} = 17.4 \cdot \frac{10.2 \times 10^7}{6 \times 10^6} = 195.8$$

Transcripts whose ratio of divergence to genetic variation exceeded these values are usually divergent. Using this criteria (as well as the lower threshold for too little divergence), of these remaining 527 genes, 464 where consistent with drift, while 63 were consistent with excessive divergence between at least one species pair.

Cis vs trans

- Cis vs trans: Historically
  - Cis refers to sites that must be on the same DNA molecule as the region it influences (DNA binding sites ,etc.)
  - Trans refers to diffusible factors (i.e., protein)
  - Cis-acting factors only operates on linked sites
  - Trans-acting factors can act throughout the genome
- eQTL mappers use
  - Cis: site close to the target gene whose expression is impacted
  - Trans: site is distance to the target of expression
  - *Local* and *distant* are better terms

Allele-specific expression (ASE)

- Occasionally, the gene products (alleles) can be disquieted, allow one to follow the specific expression of each.
- Use of hybrids when ASE is present can distinguish true cis from trans
- Take two parent lines that differ in expression for a particular locus
  - If this difference persists in the hybrid, at least part of the difference in expression is due to cis
  - If trans-acting factors determine the difference, then the expression levels in the F1 should reflect levels in the trans-supplying parent

eQTLs and sign tests

- Need at least six QTLs for Orr's sign test
- Typically, not that many eQTLs for a given transcript are detected. Hence, pool eQTLs over genes
- Bullard et al. (2010) used this approach in a cross of two closely related yeast species, *Saccharomyces cerevisiae* and *S. bayanus* . One key requirement is that each eQTL is independent, as a single eQTL that simultaneously influences k genes should be weighted as one change, not k changes in the same direction. The use of cis -regulatory alleles ensures independence over a set of loosely-linked genes. Bullard et al. accomplished this by only considering alleles showing ASE. An excessive number of up-regulated ASEs from one species indicates lineage-specific selection, and a number of pathways were detected showing this feature.

Example 12.8. A related approach to test for selection on specific transcripts using *cis* and *trans* data was suggested by Emerson et al. (2010), who combined their polymorphism data with divergence data from Tirosh et al. (2009) to examine the within- and between-species expression control in the yeasts *Saccharomyces cerevisiae* and *S. paradoxus*. They used a MacDonald-Kreitman approach (Chapter 10) by examining the fraction of *cis* and *trans* controlled transcripts measured within and between species. Their resulting contingency table

	Polymorphism	Divergence
Cis	396	1270
Trans	412	541

was highly significant, with *trans* polymorphisms being slightly more common than *cis*, but over twice as many *cis* sites fixed. Such a pattern could arise from either an excessive number of *cis* fixations between species, an excessive amount of *trans* polymorphism within a species, or a combination of both. Analogous to arguments with interpretation of MacDonald-Kreitman data discussed in Chapter 10, an excessive amount of polymorphism could arise from a high mutation rate for slightly deleterious alleles. This generates an excess amount of within-species polymorphism that does not transfer to between–species differences (as they are not fixed). Wittkopp et al. (2008), working with *Drosophila*, also noted an excess of *cis* sites being fixed over *trans* sites, and suggested that both fixation of some *cis* mutations by directional selection coupled with a larger number of slightly deleterious *trans* alleles likely underlies this pattern.