# S3 Appendix: Homogeneity testing of within-stratum diversity components

If we scale our distance matrix **D** by its maximum entry $(d_{max}^2)$, then average the elements, both for the total collection and for any relevant subdivisions (e.g., within single species, within single populations, or even within single individuals), we can translate each of the resulting $Q$-values into a corresponding sum of squares. For the grand total collection of $(2N)$ alleles in the study, for the $(2N_{Sg})$ alleles within the $g^{th}$ species $(g = 1, \ldots, G)$, or for the $(2N_{Pk})$ alleles within the $k^{th}$ population $(k = 1, \ldots, K)$, the corresponding $Q$ - values are simple multiples of the corresponding sums of squares and variances:

$$Q = \left( {}_{sum}\mathbf{D}/4N^2 \right) = (SS \, / \, N) = \text{V} \cdot (2N - 1) \, / \, N \, , \qquad [S3.1a]$$

where SS and V are the grand total sum of squares and total variance, respectively. Similar extractions for the within-species, within-population, and within-individual $Q$-values yield

$$Q_{WSg} = \cdots = (SS_{WSg} \, / \, N_{Sg}) = \text{V}_{WSg} \cdot (2N_{Sg} - 1) \, / \, N_{Sg} \, , \qquad [S3.1b]$$

$$Q_{WPk} = \cdots = (SS_{WPk} \, / \, N_{Pk}) = \text{V}_{WPk} \cdot (2N_{Pk} - 1) \, / \, N_{Pk} \, , \qquad [S3.1c]$$

$$Q_{WIj} = \cdots = (SS_{WI\text{-}j} \, / \, N_{Ij}) = \text{V}_{WIj} \cdot (2N_{Ij} - 1) \, / \, N_{Ij} \, . \qquad [S3.1d]$$

Averages of the within-stratum components of [S3.1b - S3.1d] can be computed as needed.

We convert $Q$-values to their diversity analogues, and then use their maxima to scale each of those components from $(0 = \text{no diversity})$ to $(1 = \text{maximum diversity possible})$. Given the sampling realities of the study, those maxima take very simple forms,

$$Q^* = (2N - 1) \, / \, 2N \qquad\qquad Q^*_{WSg} = (2N_{Sg} - 1) \, / \, 2N_{Sg} \, ,$$

$$[S3.2]$$

$$Q^*_{WPk} = (2N_{Pk} - 1) \, / \, 2N_{Pk} \qquad Q^*_{WI-j} = (2N_{Ij} - 1) \, / \, 2N_{Ij} = (\tfrac{1}{2}) \, ,$$

recalling (for diploids) that $\left( N_{Ij} = 2 \right)$ alleles.

The net result of such translation is we translate all of our within-stratum diversity components into analogous within-stratum variance analogues,

$$\tilde{\gamma} = [\, Q / Q^* \,] = 2 \cdot V \qquad\qquad \tilde{\sigma}_{\mathrm{WS}g} = [\, Q_{\mathrm{WS}g} / Q^*_{\mathrm{WS}g} \,] = 2 \cdot V_{\mathrm{WS}g} \; ,$$

[S3.3]

$$\tilde{\alpha}_{\mathrm{WP}k} = [\, Q_{\mathrm{WP}k} / Q^*_{\mathrm{WP}k} \,] = 2 \cdot V_{\mathrm{WP}k} \qquad \tilde{\omega}_{\mathrm{WI}-j} = [\, Q_{\mathrm{WI}-j} / Q^*_{\mathrm{WI}-j} \,] = 2 \cdot V_{\mathrm{WI}\text{-}j} \; .$$

Given a collection of G species, we will typically compute a weighted average value for (pooled) within-species $\left(\tilde{\sigma}_{\mathrm{WS}g}\right)$ diversity, where the sampling weights ($f_{\mathrm{S}g}$-values) for the G species are defined as in Text Eq. [5]. Straightforward algebraic manipulation yields

$$\tilde{\sigma}_{\mathrm{WS}} = (f_{\mathrm{SA}} \cdot Q_{\mathrm{SA}} + \cdots + f_{\mathrm{SG}} \cdot Q_{\mathrm{SG}}) / (f_{\mathrm{SA}} \cdot Q^*_{\mathrm{SA}} + \cdots + f_{\mathrm{SG}} \cdot Q^*_{\mathrm{SG}}) = \cdots$$

[S3.4a]

$$= 2 \cdot (\eta_{\mathrm{SA}} \cdot V_{\mathrm{SA}} + \cdots + \eta_{\mathrm{SG}} \cdot V_{\mathrm{SG}}) = 2 \cdot V_{\mathrm{WS}} \; ,$$

where the sample-frame-dependent ($\eta$) - weights take the form

$$\eta_{\mathrm{S}g} = N_{\mathrm{S}g} \cdot (2N_{\mathrm{S}g} - 1) \; / \, [\, N_{\mathrm{SA}} \cdot (2\,N_{\mathrm{SA}} - 1) + \cdots + N_{\mathrm{SG}} \cdot (2N_{\mathrm{SG}} - 1) \,] \; , \qquad [S3.4b]$$

with the $\eta$-weights summing to unity. Thus, $(\tilde{\sigma}_{\mathrm{WS}} = 2 \cdot V_{\mathrm{WS}})$ is a weighted-average within-species scaled diversity component. Averaging of K within-population diversity components, using the sort of $f_{\mathrm{P}k}$-weights in Text Eq. [10], yields (weighted average) within-population translation,

$$\tilde{\alpha}_{\mathrm{WP}} = (f_{\mathrm{P1}} \cdot Q_{\mathrm{WP1}} + \cdots + f_{\mathrm{PK}} \cdot Q_{\mathrm{WPK}}) / (f_{\mathrm{P1}} \cdot Q^*_{\mathrm{WP1}} + \cdots + f_{\mathrm{PK}} \cdot Q^*_{\mathrm{WPK}}) = \cdots$$

[S3.5a]

$$= 2 \cdot (\eta_{\mathrm{P1}} \cdot V_{\mathrm{WP1}} + \cdots + \eta_{\mathrm{PK}} \cdot V_{\mathrm{WPK}}) = 2 \cdot V_{\mathrm{WP}} \; ,$$

where the sample-frame imposed population weights ($\eta$-values) now take the forms

$$\eta_{\mathrm{P}k} = N_{\mathrm{P}k} \cdot (2N_{\mathrm{P}k} - 1) \; / \, [\, N_{\mathrm{P1}} \cdot (2\,N_{\mathrm{P1}} - 1) + \cdots + N_{\mathrm{PK}} \cdot (2N_{\mathrm{PK}} - 1) \,] \; , \qquad [S3.5b]$$

with those ($\eta$-weights) again summing to unity, and yielding ($\tilde{\alpha}_{\mathrm{WP}} = 2 \cdot V_{\mathrm{WP}}$) as the weighted-average within-population scaled diversity.

Multi-stratum averages are convenient summaries, but there are situations where those averages conceal more than they reveal. We can test for heterogeneity of within-stratum diversity components with a non-parametric analogue of Bartlett's test of homogeneous within-stratum variances [1]. A test of homogeneous within-species diversity is tantamount to a test of homogeneous within-species variances, with a null hypothesis of the form:

$$(\sigma_{\widetilde{W}SA} = \cdots = \sigma_{\widetilde{W}SG}) \quad \Leftrightarrow \quad (V_{WSA} = \cdots = V_{WSG}) \ . \qquad [S3.6a]$$

The traditional Bartlett's test takes the form:

$$\text{Bartlett's Test} = [\ (2N_{SA} - 1) \cdot \ln (V_{WSA}) + \cdots + (2N_{SG} - 1) \cdot \ln (V_{WSG})\ ]$$
$$[S3.6b]$$
$$- (2N - G) \cdot \ln (V_{WS}) \ ,$$

which is typically compared with an asymptotic chi-square criterion [2]. The QDIVER routine (instead) permutes variance contributions of individual alleles (each measured from its own species mean) freely among species, computing a value of [S3.6b] for each randomization, and comparing the realized data outcome against the distribution of (say 999) permutational trials.

To deploy [S3.6b], we start by extracting the diversity contribution of each allele to the within-species (diversity, in variance form). Start with the $\mathbf{D}_{WSg}$-matrix, of dimension ($2N_{Sg}$ x $2N_{Sg}$), and scale it both by the largest element ($d^2_{max}$) and the maximum $(Q^*_{WSg})$-value, yielding

$$\mathbf{D}_{\widetilde{S}g} = 2N_{Sg} \cdot \mathbf{D}_{Sg} / (2\ N_{Sg} - 1) \cdot (d^2_{max}) \ , \qquad [S3.7a]$$

designed to scale $[\ 0 \le \sigma_{\widetilde{S}g} \le 1\ ]$ . We can extract the squared individual deviation of each allele for the ($g^{\text{th}}$)-species mean, by using a Gower transformation [3] of the distance matrix in [S3.7a] to its corresponding covariance matrix form. Conveniently, we only need the diagonal (variance) elements of that covariance matrix ($\mathbf{C}_g$), so for the $j^{\text{th}}$ allele, we compute:

$v_{jj} = (\frac{1}{2}) \cdot [\, j^{\text{th}} \text{ row aver. of } \mathbf{D}_{\tilde{S}g} + j^{\text{th}} \text{ column aver. of } \mathbf{D}_{\tilde{S}g} - \text{global aver. of } \mathbf{D}_{\tilde{S}g} \,]$ ,     [S3.7b]

and the summation of $(v_{jj})$, over all $(2N_{Sg})$ alleles within the $g^{\text{th}}$ species, is the variance within that species $(V_{WSg})$. We perform a separate transform for each of the G species under consideration, and from each of them extract the $(2N_{Sg})$ single-allele contributions to the within-species variation. We permute all $2N$ allelic $(v_{jj})$ values among species, in the numbers actually sampled, recompute the $V_{WSg}$ and the (pooled) $V_{WS}$ values for each of 999 trials, deploy the data as the $1,000^{\text{th}}$ randomization, and evaluate Eq. [S3.6b] for each trial.

We can test for homogeneous within-population diversities in exactly the same fashion. The null hypothesis is that the within-population diversities/variances are homogeneous

$$(\alpha_{\widetilde{WP1}} = \cdots = \alpha_{\widetilde{WPK}}) \quad \Leftrightarrow \quad (V_{WP1} = \cdots = V_{WPK}) \; .$$     [S3.8a]

To test this hypothesis, we again compute a Bartlett's test, now defined as

$$\text{Bartlett's Test} = [\, (2N_{P1} - 1) \cdot \ln (V_{WP1}) + \cdots + (2N_{PK} - 1) \cdot \ln (V_{WPK}) \,]$$
$$- (2N - K) \cdot \ln (V_{WP}) \; ,$$     [S3.8b]

where deviations of individual alleles are now measured from their population means. Again converting a distance matrix to its covariance matrix analogue for each population separately, and scaling, we obtain

$$\mathbf{D}_{\widetilde{P}k} = 2N_{Pk} \cdot \mathbf{D}_{Pk} / (2\,N_{Pk} - 1) \cdot (d_{max}^2) \; .$$     [S3.9a]

Thus, we scale $[\, 0 \leq \alpha_{\widetilde{P}k} \leq 1 \,]$, and then extract the squared individual deviation of each allele from its own population mean as

$v_{jj} = (\frac{1}{2}) \cdot [\, j^{\text{th}} \text{ row aver. of } \mathbf{D}_{\widetilde{P}k} + j^{\text{th}} \text{ column aver. of } \mathbf{D}_{\widetilde{P}k} - \text{global aver. of } \mathbf{D}_{\widetilde{P}k} \,]$ ,   [S3.9b]

and permute those $v_{jj}$-values freely among populations, computing individual within-population

variances, as well as the pooled within-population variance, using the observed values as the 1,000$^{th}$ randomization, and evaluating the null distribution of [S3.8b] for significance.

We treat scaled within-individual diversity in the same fashion, for situations where it would be useful to query whether within-individual diversity is homogeneous (or not) among individuals, by permuting $v_{jj}$-values among individuals. The larger point is that Homo- vs Heterogeneous (within-stratum diversity) is tantamount to Homo- vs Heterogeneous (within-stratum variance), so we have a convenient statistical comparison of homogeneous vs heterogeneous diversity hypotheses within different strata at any given level of the hierarchy.

## References

1. Bartlett MS, Properties of sufficiency and statistical tests. Proc Roy Stat Soc. 1937;160: 268-282.

2. Snedecor GW, Cochran WG, Statistical Methods, Eighth Edition, Iowa State Univ Press, Ames, Iowa (USA). 1989.

3. Gower JC, Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika. 1966;53: 315-328.