

Consider individual  $i$  and subpopulation  $k$  and let the estimated ancestry proportion of subpopulation  $k$  in individual  $i$  on chromosome  $c$  to be denoted as  $a_{ik}^c$ . We assume that for all  $c$  and  $c'$ ,  $w_{i,cc'} = w_i$ , that is, that the covariance between ancestry at any two pairs of chromosomes within an individual is the same, regardless of which pair of chromosomes we are considering. Denote

$$\text{cov}(a_{ik}^c, a_{ik}^{c'}) = w_{i,cc'} = w_i \quad (1)$$

$$\text{var}(a_{ik}^c) = \sigma_{ik}^2 \quad (2)$$

First, we find  $\text{cov}(\sum_{m \notin \{c\}} a_{ik}^m, \sum_{m \notin \{c'\}} a_{ik}^m)$ . In a small example where  $c = 2$  and  $c' = 3$  with  $m = 4$  we see that

$$\sum_{m \notin \{c\}} a_{ik}^m = a_1 + a_3 + a_4 \quad (3)$$

$$\sum_{m \notin \{c'\}} a_{ik}^m = a_1 + a_2 + a_3 \quad (4)$$

Then it follows that

$$\text{cov}(\sum_{m \notin \{c\}} a_{ik}^m, \sum_{m \notin \{c'\}} a_{ik}^m) = \text{cov}(a_1 + a_3 + a_4, a_1 + a_2 + a_3) \quad (5)$$

$$= \text{cov}(a_1, a_1) + \text{cov}(a_1, a_2) + \text{cov}(a_1, a_3) \quad (6)$$

$$+ \text{cov}(a_3, a_1) + \text{cov}(a_3, a_2) + \text{cov}(a_3, a_3) \quad (7)$$

$$+ \text{cov}(a_4, a_1) + \text{cov}(a_4, a_2) + \text{cov}(a_4, a_3) \quad (8)$$

$$= 2\sigma_{ik}^2 + 7w_i \quad (9)$$

$$= (m-2)\sigma_{ik}^2 + (m-2)^2w_i + (m-1)w_i \quad (10)$$

We next find  $\text{cov}(\sum_{m \notin \{c\}} a_{ik}^m, a_{ik}^{c'})$  under the same example where  $c = 2$ ,  $c' = 3$  and  $m = 4$

$$\text{cov}(\sum_{m \notin \{c\}} a_{ik}^m, a_{ik}^{c'}) = \text{cov}(a_1 + a_3 + a_4, a_3) \quad (11)$$

$$= \text{cov}(a_1, a_3) + \text{cov}(a_3, a_3) + \text{cov}(a_4, a_3) \quad (12)$$

$$= 2w_i + \sigma_{ik}^2 \quad (13)$$

$$= (m-2)w_i + \sigma_{ik}^2 \quad (14)$$

Now we have all the components to find  $cov(D_{ik}^c, D_{ik}^{c'})$ .

$$cov(D_{ik}^c, D_{ik}^{c'}) = cov(a_{ik}^{-c} - a_{ik}^c, a_{ik}^{-c'} - a_{ik}^{c'}) \quad (15)$$

$$= cov(a_{ik}^{-c}, a_{ik}^{-c'}) - cov(a_{ik}^{-c}, a_{ik}^{c'}) - cov(a_{ik}^c, a_{ik}^{-c'}) + cov(a_{ik}^c, a_{ik}^{c'}) \quad (16)$$

$$= cov(a_{ik}^{-c}, a_{ik}^{-c'}) - 2cov(a_{ik}^{-c}, a_{ik}^{c'}) + cov(a_{ik}^c, a_{ik}^{c'}) \quad (17)$$

$$= cov\left(\frac{1}{22} \sum_{m \notin \{c\}} a_{ik}^m, \frac{1}{22} \sum_{m \notin \{c'\}} a_{ik}^m\right) - 2cov\left(\frac{1}{22} \sum_{m \notin \{c\}} a_{ik}^m, a_{ik}^{c'}\right) + w_i \quad (18)$$

$$= \left(\frac{1}{22}\right)^2 cov\left(\sum_{m \notin \{c\}} a_{ik}^m, \sum_{m \notin \{c'\}} a_{ik}^m\right) - \frac{2}{22} cov\left(\sum_{m \notin \{c\}} a_{ik}^m, a_{ik}^{c'}\right) + w_i \quad (19)$$

$$= \left(\frac{1}{22}\right)^2 \{21(\sigma_{ik}^2 + 21w_i) + 22w_i\} - \frac{2}{22} \{21w_i + \sigma_{ik}^2\} + w_i \quad (20)$$

$$= \frac{1}{(m-1)^2} \{(m-2)\sigma_{ik}^2 + (m-2)^2w_i + (m-1)w_i\} \quad (21)$$

$$- \frac{2}{(m-1)} \{(m-2)w_i + \sigma_{ik}^2\} + w_i \quad (22)$$

where the first two terms are as found in Equations 10 and 14, respectively and  $m = 23$ . With some simplification and algebra, this equation becomes

$$cov(D_{ik}^c, D_{ik}^{c'}) = \frac{m}{(m-1)^2} (w_i - \sigma_{ik}^2) \quad (23)$$

Now, we are able to combine these results to calculate the entries in the variance-covariance matrix for our test statistics  $T_k^c$ .

$$cov(T_k^c, T_k^{c'}) = cov\left(\frac{\overline{D_k^c}}{\sigma_{ck}}, \frac{\overline{D_k^{c'}}}{\sigma_{c'k}}\right) \quad (24)$$

$$= \frac{1}{\sigma_{ck}\sigma_{c'k}} cov(\overline{D_k^c}, \overline{D_k^{c'}}) \quad (25)$$

$$= \frac{1}{n^2\sigma_{ck}\sigma_{c'k}} cov(D_{1k}^c + \dots + D_{nk}^c, D_{1k}^{c'} + \dots + D_{nk}^{c'}) \quad (26)$$

$$= \frac{1}{n^2\sigma_{ck}\sigma_{c'k}} \sum_{i=1}^n \frac{m}{(m-1)^2} (w_i - \sigma_i^2) \quad (27)$$

The diagonal entries correspond to  $var(T_k^c) = 1$ .