APPENDIX

Covariance matrix of residuals for observed allele frequencies of multiple linked markers

Derivation of the covariance between residuals for observed allele frequencies of marker j and l in the i^{th} family, with recombination rate r_{jl}, which is assumed known:

$$\text{cov}(e_{ij}^U, e_{il}^U) = \text{cov}(e_{ij}^L, e_{il}^L) = \text{cov}(se_{ij}^U + te_{ij}^U, se_{il}^U + te_{il}^U) = \text{cov}(se_{ij}^U, se_{il}^U),$$

$$\text{cov}(se_{ij}^U, se_{il}^U) = \text{cov}(f_{M_{ij}}^{UTrue}, f_{M_{il}}^{UTrue}) = \text{cov}(\frac{\#M_{ij}^U}{n}, \frac{\#M_{il}^U}{n}) = \frac{1}{n^2}\text{cov}(\#M_{ij}^U, \#M_{il}^U),$$

where $f_{M_{ij}}^{UTrue}$ and $f_{M_{il}}^{UTrue}$ are the true "M" allele frequencies from sire i for marker j and l in the upper pool, which are free of technical errors. The covariance between $f_{M_{ij}}^{UTrue}$ and $f_{M_{il}}^{UTrue}$ is equivalent to the covariance between the residuals because the expectations $(x_{ij}\beta_i)$ are constants. The $\#M_{ij}^U$ and $\#M_{il}^U$ are number of "M" alleles from sire i for marker j and l in the upper pool, and

$$\text{cov}(\#M_{ij}^U, \#M_{il}^U) = E(\#M_{ij}^U \cdot \#M_{il}^U) - E(\#M_{ij}^U)E(\#M_{il}^U).$$

Since covariance matrix is the same for residuals in the upper and lower pool, the following derivation will not specify upper or lower pools:

$$E(\#M_{ij}) = np_{M_{ij}} = n[(1 - r_j)p_{Q_i} + r_j(1 - p_{Q_i})],$$

$$E(\#M_{il}) = np_{M_{il}} = n[(1 - r_l)p_{Q_i} + r_l(1 - p_{Q_i})],$$

$$E(\#M_{ij} \cdot \#M_{il}) = E(\sum_{k_1=1}^{n} I_{M_{ijk_1}} \cdot \sum_{k_2=1}^{n} I_{M_{ilk_2}}) = E(\sum_{k_1=1}^{n}\sum_{k_2=1}^{n} I_{M_{ijk_1}} \cdot I_{M_{ilk_2}}),$$

where $I_{M_{ijk_1}}$ and $I_{M_{ilk_2}}$ are indicators for markers j and l of progeny k_1 and k_2. $I = 1$ if the progeny receives the "M" allele from sire i, $I = 0$ if the progeny gets the "m" allele.

$$E(\#M_{ij} \cdot \#M_{il}) = \sum_{k_1=1}^{n}\sum_{k_2=1}^{n} E(I_{M_{ijk_1}} \cdot I_{M_{ilk_2}}) = \sum_{k_1=1}^{n}\sum_{k_2=1}^{n} \Pr(I_{M_{ijk_1}} = 1 \text{ and } I_{M_{ilk_2}} = 1)$$

$$= \sum_{k_1=1}^{n}\sum_{k_2=k_1}^{n} \Pr(I_{M_{ijk_1}} = 1 \text{ and } I_{M_{ilk_2}} = 1) + \sum_{k_1=1}^{n}\sum_{k_2\neq k_1}^{n} \Pr(I_{M_{ijk_1}} = 1 \text{ and } I_{M_{ilk_2}} = 1).$$

1. Marker QTL order (M_j-Q-M_l):

If $k_1 = k_2$ (two individuals are the same):

$$\Pr(I_{M_{ijk_1}} = 1 \text{ and } I_{M_{ilk_2}} = 1) = (1 - r_j)(1 - r_l)p_{Q_i} + r_i r_j(1 - p_{Q_i}).$$

If $k_1 \neq k_2$ (individual k_1 and k_2 are different):

$$\Pr(I_{M_{ijk_1}} = 1 \text{ and } I_{M_{ilk_2}} = 1) = \Pr(I_{M_{ijk_1}} = 1\ ) \cdot \Pr(I_{M_{ijk_2}} = 1\ )$$
$$= [(1 - r_j)p_{Q_i} + r_j(1 - p_{Q_i})][(1 - r_l)p_{Q_i} + r_l(1 - p_{Q_i})],$$
$$E(\#M_{ij}, \#M_{il}) = n \cdot [(1 - r_j)(1 - r_l)p_{Q_i} + r_i r_j(1 - p_{Q_i})]$$
$$+ n(n - 1) \cdot [(1 - r_j)p_{Q_i} + r_j(1 - p_{Q_i})][(1 - r_l)p_{Q_i} + r_l(1 - p_{Q_i})],$$
$$\text{cov}(\#M_{ij}, \#M_{il}) = n \cdot (1 - r_j)(1 - r_l)p_{Q_i} + r_i r_j(1 - p_{Q_i})$$
$$+ n(n - 1) \cdot [(1 - r_j)p_{Q_i} + r_j(1 - p_{Q_i})][(1 - r_l)p_{Q_i} + r_l(1 - p_{Q_i})]$$
$$- n^2[(1 - r_j)p_{Q_i} + r_j(1 - p_{Q_i})][(1 - r_l)p_{Q_i} + r_l(1 - p_{Q_i})]$$
$$= n(1 - 2r_{jl})p_{Q_i}(1 - p_{Q_i}),$$

where $r_{jl} = (1-r_j)\cdot r_l + r_j\cdot(1-r_l)$, as implied by the Haldane mapping function.

Thus, $\text{cov}(se^U_{ij}, se^U_{il}) = \text{cov}(se^L_{ij}, se^L_{il}) = \frac{1}{n}(1 - 2r_{jl})p_{Q_i}(1 - p_{Q_i})$.

2. Marker QTL order (M_j -M_l-Q):

If $k_1 = k_2$ (two individuals are the same):

$$\Pr(I_{M_{ijk_1}} = 1 \text{ and } I_{M_{ilk_2}} = 1) = (1 - r_{jl})(1 - r_l)p_{Q_i} + (1 - r_{jl})r_l(1 - p_{Q_i}).$$

If $k_1 \neq k_2$ (individual k_1 and k_2 are different):

$$\Pr(I_{M_{ijk_1}} = 1 \text{ and } I_{M_{ilk_2}} = 1) = \Pr(I_{M_{ijk_1}} = 1\ ) \cdot \Pr(I_{M_{ilk_2}} = 1\ ),$$
$$\Pr(I_{M_{ilk_1}} = 1\ ) = [(1 - r_l)p_{Q_i} + r_l(1 - p_{Q_i})] = p_{M_{il}},$$
$$\Pr(I_{M_{ijk_2}} = 1\ ) = [(1 - r_{jl})p_{M_{il}} + r_{jl}(1 - p_{M_{il}})],$$
$$E(\#M_{ij} \cdot \#M_{il}) = n \cdot [(1 - r_j)(1 - r_{jl})p_{Q_i} + r_i r_j(1 - p_{Q_i})]$$
$$+ n(n - 1) \cdot p_{M_{il}}[(1 - r_{jl})p_{M_{il}} + r_{jl}(1 - p_{M_{il}})],$$
$$E(\#M_{il}) = np_{M_{il}} = n[(1 - r_l)p_{Q_i} + r_l(1 - p_{Q_i})],$$
$$E(\#M_{ij}) = np_{M_{ij}} = n[(1 - r_{jl})p_{M_{il}} + r_{jl}(1 - p_{M_{il}})],$$
$$\text{cov}(\#M_{ij} \cdot \#M_{il}) = n \cdot (1 - r_j)(1 - r_{jl})p_{Q_i} + r_i r_j(1 - p_{Q_i})$$
$$+ n(n - 1) \cdot p_{M_{il}}[(1 - r_{jl})p_{M_{il}} + r_{jl}(1 - p_{M_{il}})]$$
$$- n^2[(1 - r_l)p_{Q_i} + r_l(1 - p_{Q_i})][(1 - r_{jl})p_{M_{il}} + r_{jl}(1 - p_{M_{il}})].$$

Thus, $\text{cov}(se^U_{ij}, se^U_{il}) = \text{cov}(se^L_{ij}, se^L_{il}) = \frac{1}{n}(1 - 2r_{jl})p_{M_{il}}(1 - p_{M_{il}})$.