Supplementary Information for manuscript "Coalescent theory of migration network motifs"

NICOLAS ALCALA[*, 1]**, AMY GOLDBERG**[*, †]**, UMA RAMAKRISHNAN**[‡] **and NOAH A. ROSENBERG**[*]

[*]Department of Biology, Stanford University, Stanford, CA 94305-5020, USA, [†]Department of Evolutionary Anthropology, Duke University, Durham, NC 27708-9976, USA, [‡]National Centre for Biological Sciences, Tata institute of Fundamental Research, Bangalore, Karnataka 560065, India
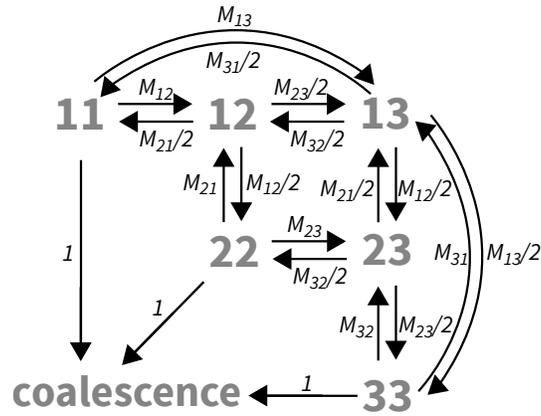
Figure S1 State diagram of the Markov chain representing the coalescent process of two lineages sampled in $K = 3$ subpopulations. States appear in gray and correspond to those presented in Figure 2; transition rates between states appear in black. M_{ij} corresponds to the scaled migration rate between subpopulations i and j. This diagram applies to all motifs with $K = 3$ subpopulations—motifs 4 to 7 in Figure 1. For example, motif 4 corresponds to the case where $M_{12} = M_{21} = M_{13} = M_{31} = M_{23} = M_{32} = 0$, and motif 5 corresponds to the case where $M_{12} = M_{21} = M$ and $M_{13} = M_{31} = M_{23} = M_{32} = 0$.
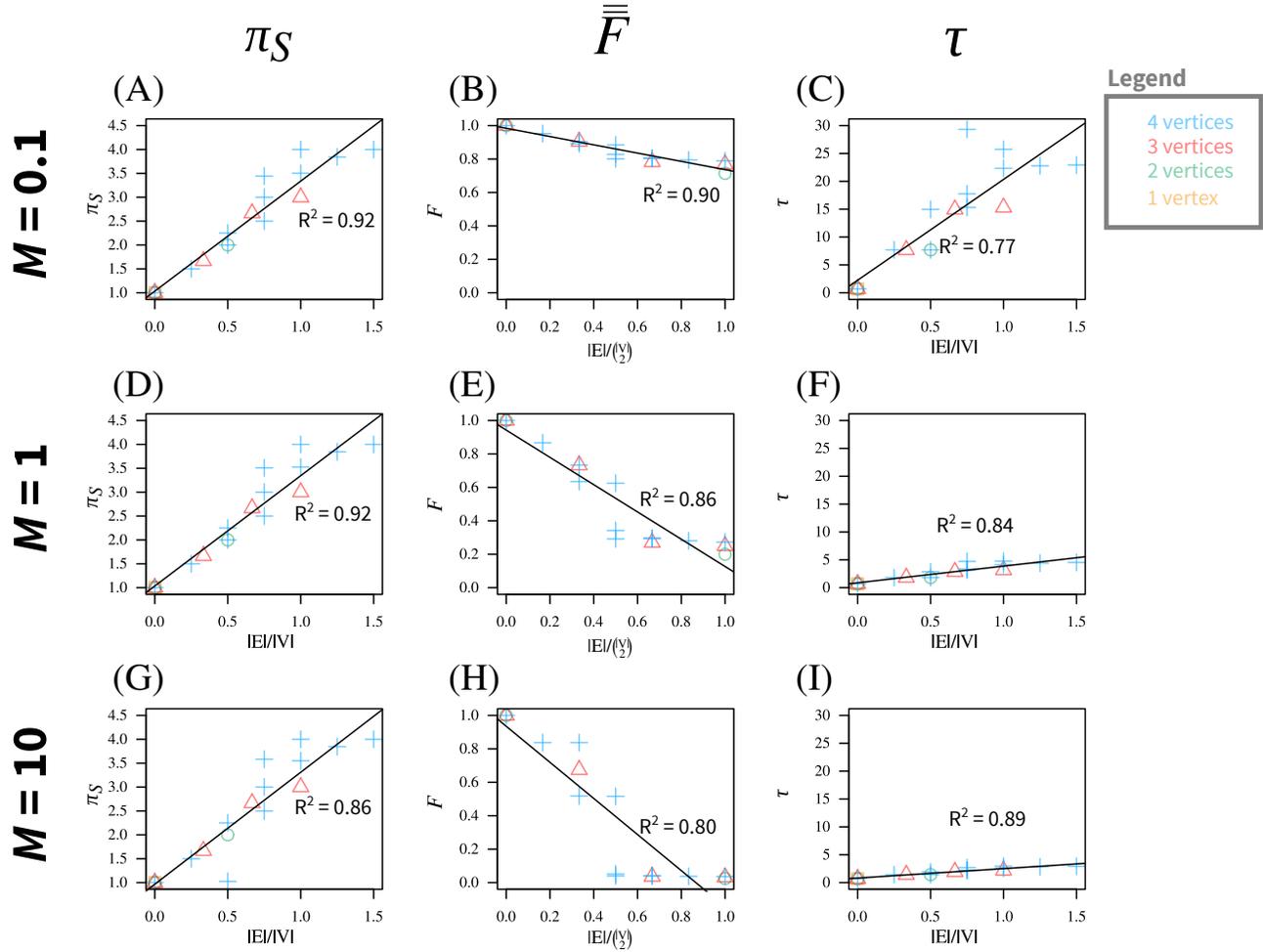
Figure S2 Relationships between π_S, $\bar{\bar{F}}$, τ, and network metrics, for different scaled migration rates M. π_S is computed from eq. 4, $\bar{\bar{F}}$ from eq. 6, and τ from eq. 7. The network metrics correspond to the metrics most highly correlated with the population-genetic statistics, as determined in Figure 3: the mean vertex degree $|E|/|V|$ for π_S, the motif density $|E|/\binom{|V|}{2}$ for $\bar{\bar{F}}$, and the mean vertex degree $|E|/|V|$ for τ. (A) $M = 0.1$, π_S. (B) $M = 0.1$, $\bar{\bar{F}}$. (C) $M = 0.1$, τ. (D) $M = 1$, π_S. (E) $M = 1$, $\bar{\bar{F}}$. (F) $M = 1$, τ. (G) $M = 10$, π_S. (H) $M = 10$, $\bar{\bar{F}}$. (I) $M = 10$, τ.
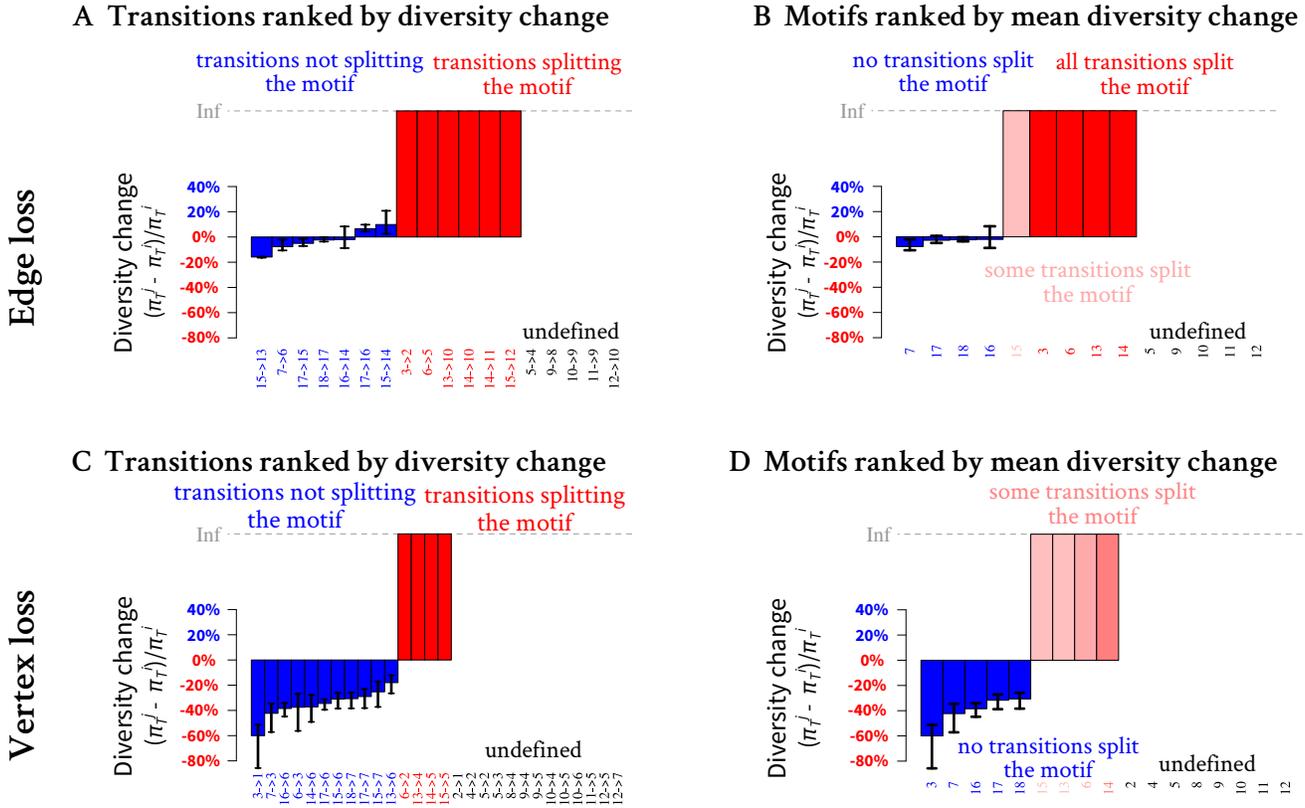
Figure S3 Change of total nucleotide diversity $\Delta\pi_{T,ij}=(\pi_T^j-\pi_T^i)/\pi_T^i$ following a transition from motif i to motif j, for all transitions involving the loss of a single edge or a single vertex. (A) Motif transitions involving an edge loss, ranked by $\Delta\pi_{T,ij}$ value. (B) Motifs ranked from largest to smallest diversity loss following edge loss. For each motif, the mean loss or gain is computed across all possible transitions to another motif. (C) Motif transitions involving a vertex loss, ranked by $\Delta\pi_{T,ij}$ value. (D) Motifs ranked from largest to smallest mean diversity loss following vertex loss. In all panels, $\Delta\pi_{T,ij}$ values assume $M = 1$, and black horizontal bars represent minimum and maximum values for M in $(0,\infty)$. Values of π_T^i and π_T^j are computed from Table S1; minima and maxima of $\Delta\pi_{T,ij}$ are obtained numerically. "Undefined" values correspond to the case where $\pi_T^j=\pi_T^i = \infty$.

Table S1 Exact expected total coalescence times, and weak and strong migration limits, for all motifs with 1 to 4 subpopulations. The total coalescence time of a motif is $\bar{t}_T = (1/K^2) \sum_{i=1}^{K} \sum_{j=1}^{K} \bar{t}_{ij}$, where coalescence times $\bar{t}_{ij}$ are computed from Tables 2, 3, and 4. When the motif has multiple disconnected components, the total coalescence times for all components are provided; the total coalescence time of the whole motif is then infinite. For example, motif 4 corresponds to 3 isolated subpopulations, leading to 3 components with coalescence time 1, so the value reported in the table is $\{1,1,1\}$. When $M \to 0$, between-subpopulation coalescence times tend to infinity; the weak-M column gives the Maclaurin series of $\bar{t}_T$ in M, which gives a sense of the rate of increase of coalescence time as M nears 0. The strong-M column corresponds to the asymptotic approximation of $\bar{t}_T$ when M tends to infinity. The a column corresponds to the value of the scaling factor a obtained when writing $\bar{t}_T$ in the form $aK[1 + f(M)]$, with K corresponding to the number of subpopulations in a component and $f(M)$ a rational function of M.

Motif	$\bar{t}_T$	$\bar{t}_T$ under weak M	$\bar{t}_T$ under strong M	a
	1	1	1	1
	$\{1,1\}$	$\{1,1\}$	$\{1,1\}$	$\{1,1\}$
	$2\left(1+\frac{1}{4M}\right)$	$\frac{1}{2M}$	2	1
	$\{1,1,1\}$	$\{1,1,1\}$	$\{1,1,1\}$	$\{1,1,1\}$
	$\{2\left(1+\frac{1}{4M}\right),1\}$	$\{\frac{1}{2M},1\}$	$\{2,1\}$	$\{1,1\}$
	$\frac{8}{3}\left(1+\frac{1}{2M}\right)$	$\frac{4}{3M}$	$\frac{8}{3}$	$\frac{8}{9}$
	$3\left(1+\frac{4}{9M}\right)$	$\frac{4}{3M}$	3	1
	$\{1,1,1,1\}$	$\{1,1,1,1\}$	$\{1,1,1,1\}$	$\{1,1,1,1\}$
	$\{2\left(1+\frac{1}{4M}\right),1,1\}$	$\{\frac{1}{2M},1,1\}$	$\{2,1,1\}$	$\{1,1,1\}$
	$\{\frac{8}{3}\left(1+\frac{1}{2M}\right),1\}$	$\{\frac{4}{3M},1\}$	$\{\frac{8}{3},1\}$	$\{\frac{8}{9},1\}$
	$\{2\left(1+\frac{1}{4M}\right),2\left(1+\frac{1}{4M}\right)\}$	$\{\frac{1}{2M},\frac{1}{2M}\}$	$\{2,2\}$	$\{1,1\}$
	$\{3\left(1+\frac{4}{9M}\right),1\}$	$\{\frac{4}{3M},1\}$	$\{3,1\}$	$\{1,1\}$
	$3\left(1+\frac{5}{8M}\right)$	$\frac{15}{8M}$	3	$\frac{3}{4}$
	$\frac{18}{5}\left(1+\frac{229M+260}{300M^2+336M}\right)$	$\frac{39}{14M}$	$\frac{18}{5}$	$\frac{9}{10}$
	$\frac{32}{9}\left(1+\frac{21680M^2+54791M+33660}{34560M^3+87040M^2+53248M}\right)$	$\frac{935}{416M}$	$\frac{32}{9}$	$\frac{8}{9}$
	$4\left(1+\frac{5}{8M}\right)$	$\frac{5}{2M}$	4	1
	$\frac{50}{13}\left(1+\frac{488M+507}{832M^2+864M}\right)$	$\frac{325}{144M}$	$\frac{50}{13}$	$\frac{25}{26}$
	$4\left(1+\frac{9}{16M}\right)$	$\frac{9}{4M}$	4	1
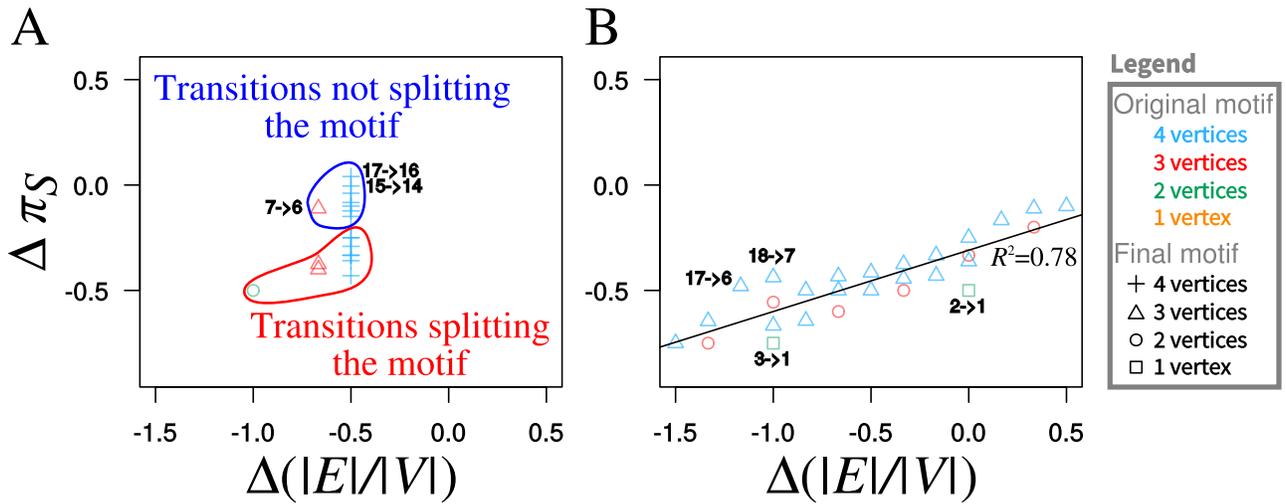
Figure S4 Relationships between the change of diversity π_S resulting from a motif transition, $\Delta\pi_S$, and the change in mean vertex degree, $\Delta(|E|/|V|)$ for $M = 1$. π_S is computed from eq. 4. (A) Edge loss. (B) Vertex loss. All transitions involving loss of a single edge or vertex are shown.

Supplementary File S1. Maxima spreadsheet to compute solutions to eqs. 9-13.

Table S2 Numerical approximation of expected within-subpopulation coalescence times for all motifs with 1 to 4 subpopulations. Values are computed from Tables 2, 3, and 4; when values depend on M, ranges of possible coalescence times are computed by minimizing and maximizing the coalescence times from Tables 2, 3, and 4 with respect to M, considering M in $(0, \infty)$. Values with no decimal places or a single decimal place are exact; when three decimal places are shown, the value is an approximation.

Motif	$\bar{t}_{11}$	$\bar{t}_{22}$	$\bar{t}_{33}$	$\bar{t}_{44}$
motif 1	1	-	-	-
motif 2	1	1	-	-
motif 3	2	2	-	-
motif 4	1	1	1	-
motif 5	1	2	2	-
motif 6	2.667	2.667	2.667	-
motif 7	3	3	3	-
motif 8	1	1	1	1
motif 9	1	2	1	2
motif 10	2.667	2.667	1	2.667
motif 11	2	2	2	2
motif 12	3	3	1	3
motif 13	3	3	3	3
motif 14	(3.142,3.6]	[3.6,3.715)	(3.142,3.6]	[3.6,3.715)
motif 15	(3.555,3.590)	(3.548,3.556)	(3.307,3.556)	(3.548,3.556)
motif 16	4	4	4	4
motif 17	(3.833,3.847)	(3.846,3.852)	(3.833,3.847)	(3.846,3.852)
motif 18	4	4	4	4

Table S3 Numerical approximation of the expected value of $\alpha = (1/F_{ST} - 1)/M$ for all motifs with 1 to 4 subpopulations, so $F_{ST} \approx 1/(1 + \alpha M)$. For a given motif, the range of possible α values gives a sense of how much the expression of F_{ST} can be approximated by a simple expression of the form $1/(1 + \alpha M)$. In addition, for each motif, α quantifies the influence of migration on F_{ST} values: when $1/\alpha \gg M$, $F_{ST} \approx 1$, whereas when $1/\alpha \ll M$, $F_{ST} \approx 0$. Values are computed from Tables 2, 3, and 5, using the formula $\alpha = (1/F_{ST} - 1)/M$; when α values depend on the value of M, ranges are computed by minimizing and maximizing α with respect to M, considering M in $(0, \infty)$. Values with no decimal places or a single decimal place are exact; when three decimal places are shown, the value is an approximation.

Motif	$\bar{F}_{12}$	$\bar{F}_{13}$	$\bar{F}_{14}$	$\bar{F}_{23}$	$\bar{F}_{24}$	$\bar{F}_{34}$
motif 2	0	-	-	-	-	-
motif 3	4	-	-	-	-	-
motif 4	0	0	-	0	-	-
motif 5	0	0	-	4	-	-
motif 6	3.2	2	-	0	-	-
motif 7	3	3	-	3	-	-
motif 8	0	0	0	0	0	0
motif 9	0	0	0	0	4	0
motif 10	3.2	0	3.2	0	2	0
motif 11	4	0	0	0	0	4
motif 12	3	0	3	0	3	0
motif 13	3	3	3	2	2	2
motif 14	(3.158,3.2]	(1.129,1.176)	1.5	1.5	(2.222,2.261)	(3.158,3.2]
motif 15	(2.608,2.614)	(2.962,2.989)	(2.608,2.614)	(1.747,1.752)	3	(1.747,1.752)
motif 16	2.667	2	2.667	2.667	2	2.667
motif 17	(2.711,2.713)	2	(2.711,2.713)	(2.711,2.713)	2.667	(2.711,2.713)
motif 18	2.667	2.667	2.667	2.667	2.667	2.667