Modeling X-Linked Ancestral Origins in Multiparental Populations

Chaozhi Zheng[1]

Biometris, Wageningen University and Research Centre, The Netherlands

ABSTRACT The models for the mosaic structure of an individual's genome from multiparental populations have been developed primarily for autosomes, whereas X chromosomes receive very little attention. In this paper, we extend our previous approach to model ancestral origin processes along two X chromosomes in a mapping population, which is necessary for developing hidden Markov models in the reconstruction of ancestry blocks for X-linked quantitative trait locus mapping. The model accounts for the joint recombination pattern, the asymmetry between maternally and paternally derived X chromosomes, and the finiteness of population size. The model can be applied to various mapping populations such as the advanced intercross lines (AIL), the Collaborative Cross (CC), the heterogeneous stock (HS), the Diversity Outcross (DO), and the *Drosophila* synthetic population resource (DSPR). We further derive the map expansion, density (per Morgan) of recombination breakpoints, in advanced intercross populations with L inbred founders under the limit of an infinitely large population size. The analytic results show that for X chromosomes the genetic map expands linearly at a rate (per generation) of two-thirds times $1 - 10/(9L)$ for the AIL, and at a rate of two-thirds times $1 - 1/L$ for the DO and the HS, whereas for autosomes the map expands at a rate of $1 - 1/L$ for the AIL, the DO, and the HS.

There have been recently designed quantitative trait locus (QTL) mapping populations with either multiple parents to increase the genetic diversity of the founder population, or many intercross generations to improve the mapping resolution by accumulating historical recombination events. Some examples include the Collaborative Cross (CC) (Churchill *et al.* 2004), the advanced intercross lines (AIL) (Darvasi and Soller 1995), the heterogeneous stock (HS) (Mott *et al.* 2000), the diversity outcross (DO) (Svenson *et al.* 2012), and the *Drosophila* synthetic population resource (DSPR) (King *et al.* 2012). The CC can be regarded as a set of eight-way recombinant inbred lines (RIL) by sibling mating, where eight founders of each line are permuted.

The genomes of individuals in QTL mapping populations are random mosaics of the founders' genomes. The QTL mapping generally necessitates the reconstruction of these genome blocks along two homologous chromosomes of a sampled individual from available genotype data. Such reconstruction is often performed under a hidden Markov model (HMM) with the latent state being the pair of ancestral origins at a locus, where the transition probability of ancestral origins between two loci, or the two-locus diplotype (two-haplotype) probabilities are required.

Modeling ancestral origins along a pair of autosomal chromosomes has been well developed recently. Broman (2012a) extended the approach of Haldane and Waddington (1931) from the two-way to four- and eight-way RIL by sibling mating and provided recipes for calculating autosomal two-locus diplotype probabilities numerically. Johannes and Colome-Tatche (2011) derived autosomal two-locus diplotype probabilities for the two-way RIL by selfing. Zheng *et al.* (2014) described a general modeling framework for ancestral origins that can be applied to autosomes in various mapping populations such as the RIL by selfing or sibling mating and the AIL.

A special treatment is required for modeling ancestral origins along a pair of X chromosomes. Haldane and Waddington (1931) derived the recurrence relations of the X-linked two-locus diplotype probabilities for the two-way RIL by sibling mating and the bi-parental repeated parent-offspring mating, and their closed form solutions for the final

homozygous lines. Broman (2005) extended the solutions to the two- and three-locus haplotype probabilities for the two, four, or eight-way RIL by sibling mating. Broman (2012b) derived the X-linked two-locus haplotype probabilities in advanced intercross populations including the AIL, the HS, and the DO, assuming an infinitely large population size.

In this paper, we extend our previous work (Zheng *et al.* 2014) to model the ancestral origins along a pair of X chromosomes in a finite mapping population. This extension also builds on the theory of junctions in inbreeding (Fisher 1949, 1954). A junction is defined as a boundary point of genome blocks on chromosomes where two distinct ancestral origins meet, and the boundary points that occur at the same location along multiple chromosomes are counted as a single junction. The map expansion is the expected junction density (per Morgan) on a maternally or paternally derived X chromosome, denoted by R^m or R^p, respectively. We denote by ρ^{mp} the overall junction density along the XX chromosomes of a female, and it can be used as a measure of X-linked QTL mapping resolution (Darvasi and Soller 1995; Weller and Soller 2004).

The key feature of this extension is to account for the asymmetry between maternally and paternally derived X chromosomes because the latter did not experience any crossover events with Y chromosomes. We first present a model framework for X-linked ancestral origins, where the recurrent relations are derived for various junction densities including the map expansions R^m and R^p. Then, we derive the closed form solutions for these expected densities in mapping populations including the RIL by sibling mating, the AIL, the HS, the DO, and the DSPR; they are evaluated by forward simulation studies. Lastly, we discuss the model assumptions and the implications of the analytic results on haplotype reconstructions and breeding designs.

A MODEL FOR X-LINKED ANCESTRAL ORIGINS

Assumptions and notation

Consider a dioecious population with two separate sexes: homogametic females with sex chromosomes XX and heterogametic males with sex chromosomes XY. There are no recombination events between X and Y, and thus we ignore the pseudoautosomal regions on the XY chromosomes. As in most mammals and some insects (*Drosophila*), some flowering species, such as white campion (*Silene latifolia*), papaya (*Carica papaya*), and asparagus (*Asparagus officinalis*), have the XY sex determination system (Ming and Moore 2007). The dioecious population was founded in generation 0, and it has nonoverlapping generations. There are no natural or artificial selections since the founder population. The mating schemes of producing the next generation are random, and they may vary from one generation to the next. The assignments of offspring genders are assumed to be independent of mating schemes.

The ancestral origins along two homologous autosomes have been modeled as a continuous time Markov chain (CTMC) (Zheng *et al.* 2014). We extend the approach to account for the asymmetry of XX chromosomes, using superscript m (p) for maternally (paternally) derived genes or chromosomes. See Supporting Information, Table S1 for a list of symbols used in this paper. Let $\mathbf{O}(x) = (O^m(x), O^p(x))$ be the ordered pair of the ancestral origins at location x along the two X chromosomes of a randomly sampled female. The ancestral origin process $\mathbf{O}(x)$ is assumed to follow a CTMC, where x is the time parameter in unit of Morgan. We assign a unique ancestral origin to the X chromosomes of each inbred founder, or to each X chromosome of each outbred founder. Multiple genes, within or between loci, are identical by descent (IBD) if they have the same ancestral origins. Let

L be the number of possible ancestral origins that $O^m(x)$ or $O^p(x)$ may take. L may be less than the number of inbred founders if some male founders did not produce daughters to pass down their X chromosomes. For example, $L = 3$ for the four-way RIL by sibling mating since one of the founder mating pairs produces only one son (Figure 1A).

The L possible ancestral origins are assumed to be exchangeable, so that we focus on the changes of ancestral origins. See Figure 1B and the relevant part of Discussion on the exchangeability assumption. The initial distribution of $\mathbf{O}(0)$ at the leftmost locus $x = 0$ is specified by $\alpha^{mp}(11)$, a probability that the two ancestral origins are the same (IBD) at a locus. Let $\alpha^{mp}(12) = 1 - \alpha^{mp}(11)$ be the non-IBD probability. Given either IBD or non-IBD at the locus, the ancestral origin pair $\mathbf{O}(0)$ takes each of the possible combinations with equal probability.

The transition rate matrix of the CTMC can be constructed from the expected densities $J^{mp}(abcd)$ of all the junction types $(abcd)$ along the two X chromosomes of a female. The junction type $(abcd)$ denotes the four-gene IBD configuration $(abcd)$ on both sides of a junction, where ab (cd) is on the left-hand (right-hand) side, haplotype ac (bd) is on the first (second) chromosome, and the same integers denote IBD. Figure 1C illustrates the seven types of junctions: (1112), (1121), (1122), (1211), (1213), (1222), and (1232) for $L \geq 3$, where the two types (1213) and (1232) do not exist for $L = 2$. We do not define junction types for the eight two-locus configurations (1111), (1123), (1212), (1221), (1223), (1231), (1233), and (1234), because there are either zero or no less than two junctions between the two loci. Figure 1D shows the transition rate matrix of the CTMC in the four-way RIL by sibling mating. Figure 1E shows the relationships between the expected densities $J^{mp}(abcd)$ and the transition rates, and they are derived based on the interpretation that $J^{mp}(abcd)\Delta d$ is the two-locus diplotype probability, in the limit that the genetic distance Δd (in Morgan) between two loci goes to zero.

The map expansions R^m and R^p and the overall expected junction density ρ^{mp} are given by

$$R^m = J^{mp}(1121) + J^{mp}(1122) + J^{mp}(1222) + J^{mp}(1232), \qquad (1)$$

$$R^p = J^{mp}(1112) + J^{mp}(1122) + J^{mp}(1211) + J^{mp}(1213), \qquad (2)$$

$$\rho^{mp} = R^m + R^p - J^{mp}(1122), \qquad (3)$$

similar to those for autosomes (Zheng *et al.* 2014) except that $R^m \neq R^p$ for X chromosomes. We have $J^{mp}(1112) = J^{mp}(1211)$ and $J^{mp}(1121) = J^{mp}(1222)$, since the junction densities do not depend on the direction of chromosomes. In contrast to the single-locus two-gene non-IBD probability $\alpha^{mp}(12)$, the ordering of the superscripts in $J^{mp}(abcd)$ generally does matter, that is, $J^{mp}(abcd) \neq J^{pm}(abcd)$ except for the junction type (1122). In addition, we have $J^{mp}(1213) = J^{pm}(1232)$ (see Figure 1C). Thus, the CTMC of X-linked ancestral origins can be described by one non-IBD probability $\alpha^{mp}(12)$ and the five expected junction densities R^m, R^p, $J^{mp}(1122)$, $J^{mp}(1232)$, and $J^{pm}(1232)$, under the exchangeability assumption of the L possible ancestral origins.

Single-locus non-IBD probabilities

The calculation of the expected junction densities necessitates the introduction of the probabilities for the two- and three-gene IBD configurations at a single locus. All the following derivations of the recurrence relations for these probabilities are based on the Mendelian inheritance of X-linked genes: a paternally derived gene must be a copy of the maternally derived gene in a male of the previous
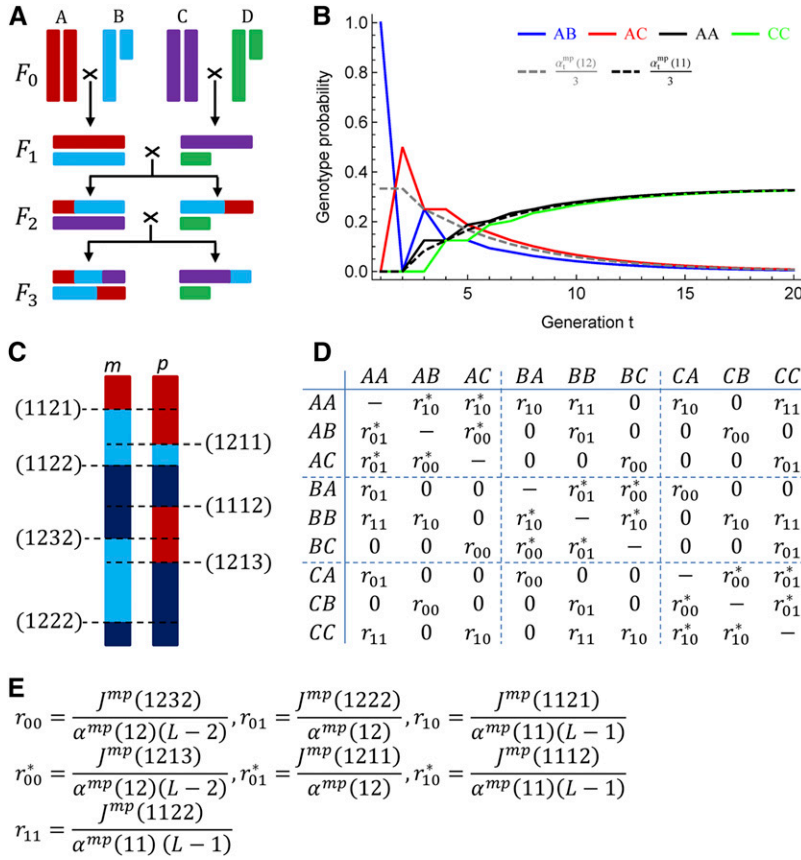
Figure 1 The continuous time Markov chain (CTMC) of X-linked ancestry blocks in the four-way recombinant inbred lines (RILs) by sibling mating. (A) One realization of ancestry blocks in the four-way RIL with generation up to F_3. The sex chromosomes of the four inbred founders are represented by different colors and labeled as A, B, C, and D. The short bars denote Y chromosomes. The ancestral origin D is impossible in the X chromosomes of generation $t \geq 1$. (B) Evaluation of the exchangeability assumption by one-locus genotype probabilities. The gray dashed line refers to the average genotype probability for one particular non-IBD genotype AB, AC, or BC; the black dashed line is for one particular IBD genotype AA, BB, or CC. Note that the ancestral origins A and B are exchangeable, but the ancestral origin C is not exchangeable with either A or B. (C) Schematics of the seven junction types along the maternally (left) and paternally (right) derived X chromosomes. (D) The rate matrix of the CTMC for the four RILs in (A). The diagonal elements are given so that row sums are zero. The rate matrix is determined by the seven basic rates, each corresponding to one of the seven junction types. The subscripts of the basic rates denote the IBD (1) or non-IBD (0) states on the left- and right-hand sides of the junctions, and the rates with superscript * refer to the transitions on the paternally derived chromosome. (E) The general relationships between the basic rates and the expected densities for the seven types of junctions, with $L = 3$ for the four-way RIL in (A).

generation, and a maternally derived gene has equal probability of being a copy of either the maternally derived gene or the paternally derived gene in a female of the previous generation.

In a dioecious mapping population, the single-locus two-gene probabilities of IBD configuration (ab) depend on whether or not the two homologous genes are in a single individual. Thus, we denote by $\beta^{mm}(ab)$, $\beta^{mp}(ab)$, and $\beta^{pp}(ab)$ the two-gene probability of IBD configuration (ab), given that the two homologous genes are in two distinct individuals in generation t and have parental origins mm, mp, and pp, respectively (Figure 2A); it holds that $\beta^{pm}(ab) = \beta^{mp}(ab)$.

The recurrence relations of the two-gene non-IBD probabilities are derived by tracing the parental origins of two homologous genes from generation $t \geq 1$ into the previous generation, and they are given by

$$\beta_t^{mm}(12) = s_t^m \frac{1}{2}\alpha_{t-1}^{mp}(12)$$
$$+ \left(1 - s_t^m\right)\left[\frac{1}{4}\beta_{t-1}^{mm}(12) + \frac{1}{2}\beta_{t-1}^{mp} + \frac{1}{4}\beta_{t-1}^{pp}\right] \quad (4a)$$

$$\beta_t^{mp}(12) = \frac{1}{2}\beta_{t-1}^{mm}(12) + \frac{1}{2}\beta_{t-1}^{mp}(12) \quad (4b)$$

$$\beta_t^{pp}(12) = \left(1 - s_t^p\right)\beta_{t-1}^{mm}(12) \quad (4c)$$

$$\alpha_t^{mp}(12) = \beta_t^{mp}(12) \quad (4d)$$

where equation (4d) holds immediately after one generation of random mating, although it may not hold in the founder population at $t = 0$. In equation (4a), the first term on the right-hand side refers to the scenario that the two genes with parent origins mm in gen-

eration t come from a single female of the previous generation with the probability s_t^m, and with probability $1/2$ that they come from different genes of the female. In equation (4b), the two genes with parental origins mp cannot merge because they must come from one male and one female of the previous generation. In equation (4c), the two genes with parental origins pp in generation t come from a single male of the previous generation with the probability s_t^p; if so, they must merge because there is only one X chromosome in a male.

We introduce the single-locus three-gene probabilities of IBD configuration (abc). Let $\beta^{mmm}(abc)$, $\beta^{mmp}(abc)$, $\beta^{mpp}(abc)$, and $\beta^{ppp}(abc)$ be the probabilities of IBD configuration (abc), given that the three homologous genes are in three distinct individuals in generation t and have parental origins mmm, mmp, mpp, and ppp, respectively (Figure 2B). Similarly, we define $\alpha^{mmp}(abc)$ and $\alpha^{mpp}(abc)$ for three homologous genes in two distinct individuals. The ordering of the superscripts does not matter for these three-gene probabilities, for example, $\beta^{mmp}(abc) = \beta^{mpm}(abc) = \beta^{pmm}(abc)$.

The recurrence relations of the three-gene non-IBD probabilities are derived by tracing the parental origins of three homologous genes from generation $t \geq 1$ into the previous generation, and they are given by

$$\beta_t^{mmm}(123) = 3q_t^m \frac{1}{2}\left[\frac{1}{2}\alpha_{t-1}^{mmp}(123) + \frac{1}{2}\alpha_{t-1}^{mpp}(123)\right] + \left(1 - s_t^m - 2q_t^m\right)$$
$$\times \left[\frac{1}{8}\beta_{t-1}^{mmm}(123) + \frac{3}{8}\beta_{t-1}^{mmp}(123) + \frac{3}{8}\beta_{t-1}^{mpp}(123)\right.$$
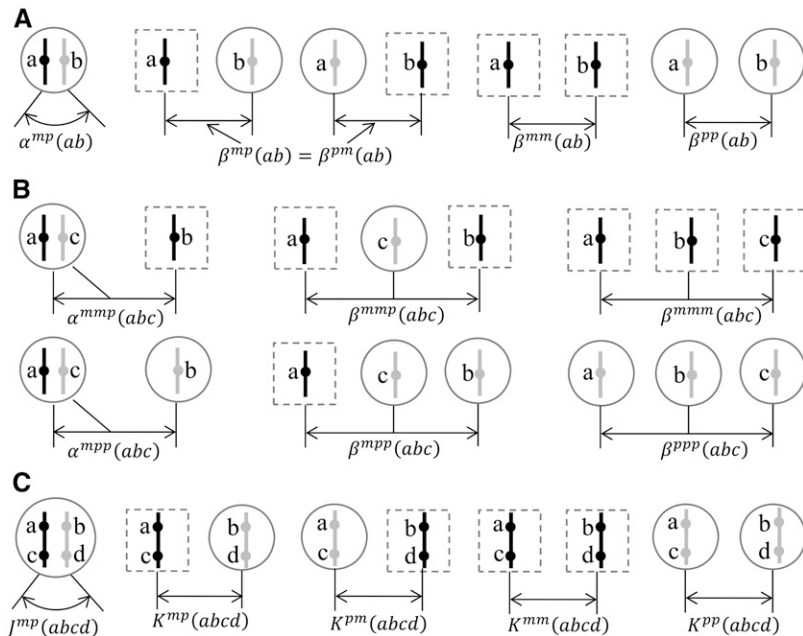$$\left. + \frac{1}{8}\beta_{t-1}^{ppp}(123)\right] \quad (5a)$$

Figure 2 Schematics of (A) the probabilities of the two-gene IBD configurations, (B) the probabilities of the three-gene IBD configurations, and (C) the expected junction densities. Circles denote females, and dashed rectangles for males or females. Black vertical lines denote the maternally derived X chromosomes, and gray vertical lines for the paternally derived. Dots denote genes on chromosomes.

$$\beta_t^{mmp}(123) = s_t^m \frac{1}{2}\alpha_{t-1}^{mmp}(123) + (1 - s_t^m)$$

$$\times \left[\frac{1}{4}\beta_{t-1}^{mmm}(123) + \frac{1}{2}\beta_{t-1}^{mmp}(123) + \frac{1}{4}\beta_{t-1}^{mpp}(123)\right]$$

$$(5b)$$

$$\beta_t^{mpp}(123) = (1 - s_t^p)\left[\frac{1}{2}\beta_{t-1}^{mmm}(123) + \frac{1}{2}\beta_{t-1}^{mmp}(123)\right] \qquad (5c)$$

$$\beta_t^{ppp}(123) = (1 - s_t^p - 2q_t^p)\beta_{t-1}^{mmm}(123) \qquad (5d)$$

$$\alpha_t^{mmp}(123) = \beta_t^{mmp}(123) \qquad (5e)$$

$$\alpha_t^{mpp}(123) = \beta_t^{mpp}(123) \qquad (5f)$$

where q_t^m is the coalescence probability of three maternally derived genes in generation t that a particular pair of genes come from a single female of the previous generation and the third comes from another female of the previous generation, and similarly q_t^p for three paternally derived genes. The equations (5e, 5f) hold immediately after one generation of random mating, although they may not hold in the founder population at $t = 0$.

The derivations of the recurrence equations (5a–5d) for the three-gene non-IBD probabilities are similar to equations (4a–4c) for the two-gene non-IBD probabilities. In equation (5a), the pre-factor 3 denotes that each of the three possible pairs of genes may come from a single female of the previous generation; the term $(1 - s_t^m - 2q_t^m)$ is the probability that the three maternally derived genes in generation t come from three distinct females of the previous generation, and it is obtained by the probability $1 - s_t^m$ that one pair of genes come from two distinct females minus the probability $2q_t^m$ that the third gene and either gene of the pair come from a single female of the previous generation. Similarly, the term $(1 - s_t^p - 2q_t^p)$ in equation

(5d) is the probability that the three paternally derived genes in generation t come from three distinct males of the previous generation.

Expected junction densities

We derive the recurrence relations for R^m, R^p, $J^{mp}(1122)$, $J^{mp}(1232)$, and $J^{pm}(1232)$. The recurrence relation for R^m follows from the theory of junctions (Fisher 1954): a new junction is formed whenever a recombination event occurs between two X chromosomes that are non-IBD at the location of a crossover. The recurrence relations for the map expansions R^m and R^p are given by

$$R_t^m = \frac{1}{2}R_{t-1}^m + \frac{1}{2}R_{t-1}^p + \alpha_{t-1}^{mp}(12), \qquad (6a)$$

$$R_t^p = R_{t-1}^m, \qquad (6b)$$

where equation (6b) follows directly from no recombination events occurring between the XY chromosomes in a male of the previous generation.

To measure differential map expansions between maternally and paternally derived chromosomes, we define $R_t^X = (2R_t^m + R_t^p)/3$ and $R_t^- = (R_t^m - R_t^p)/2$, and their recurrence relations are given by

$$R_t^X = R_{t-1}^X + \frac{2}{3}\alpha_{t-1}^{mp}(12), \qquad (7a)$$

$$R_t^- = -\frac{1}{2}R_{t-1}^- + \frac{1}{2}\alpha_{t-1}^{mp}(12), \qquad (7b)$$

according to the recurrence equations (6a, 6b). If there are equal numbers of males and females in the population, a randomly chosen X chromosome is maternally derived with probability 2/3, and it is paternally derived with probability 1/3. Thus R_t^X can be interpreted as the map expansion on a randomly chosen X chromosome.

For comparisons, we denote by R_t^A the map expansion on a random chosen autosome, and and its recurrence relation is given by (MacLeod *et al.* 2005; Zheng *et al.* 2014)

$$R_t^A = R_{t-1}^A + \alpha_{t-1}^{AA}(12), \tag{8}$$

where $\alpha_t^{AA}(12)$ refers to the non-IBD probability between two homologous autosomal genes in an individual. The equations (7a, 8) show that the map expansion R_t^X for an X chromosome is two-thirds R_t^A for an autosome if the non-IBD probability $\alpha_t^{AA}(12)$ for autosomes is the same as $\alpha_t^{mp}(12)$ for XX chromosomes, and the sex ratio is 1.

In addition to $J_t^{mp}(abcd)$ and $J_t^{pm}(abcd)$, we define $K_t^{mm}(abcd)$, $K_t^{mp}(abcd)$, $K_t^{pm}(abcd)$, and $K_t^{pp}(abcd)$ for haplotypes ac and bd that are in two distinct individuals and have parental origins mm, mp, pm, and pp, respectively (Figure 2C). The contributions to the junctions in the current generation come from either the existing junctions at the previous generation, or a new junction via a crossover event. In the following, we focus on the formation of a new junction, because the contributions of the existing junctions in the previous generation are similar to those for the two-gene non-IBD probabilities in the recurrence equations (4a–4c).

The schematics of the recurrence relations for junction types (1232) and (1122) are shown in Figure S1. The ancestry transitions of type (1122) occur on both haplotypes ac and bd at exactly the same location, and thus a new junction of type (1122) can be formed only by duplicating a chromosome segment. It holds that $J_t^{mp}(1122) = J_t^{pm}(1122)$ and $K_t^{mp}(1122) = K_t^{pm}(1122)$ because of the symmetry of type (1122). We have

$$K_t^{mm}(1122) = s_t^m \left[\frac{1}{2} J_{t-1}^{mp}(1122) + \frac{1}{4} R_{t-1}^m + \frac{1}{4} R_{t-1}^p \right]$$
$$+ \left(1 - s_t^m \right) \left[\frac{1}{4} K_{t-1}^{mm}(1122) + \frac{1}{2} K_{t-1}^{mp}(1122) \right. \tag{9a}$$
$$\left. + \frac{1}{4} K_{t-1}^{pp}(1122) \right],$$

$$K_t^{mp}(1122) = \frac{1}{2} K_{t-1}^{mm}(1122) + \frac{1}{2} K_{t-1}^{mp}(1122), \tag{9b}$$

$$K_t^{pp}(1122) = s_t^p R_{t-1}^m + \left(1 - s_t^p \right) K_{t-1}^{mm}(1122), \tag{9c}$$

$$J_t^{mp}(1122) = K_t^{mp}(1122), \tag{9d}$$

for $t \geq 1$, where equation (9d) may not hold in the founder population at $t = 0$, the first term on the right-hand side of equation (9a) refers to the scenario that both haplotypes ac and bd come from a single female of the previous generation, and the first term on the right-hand side of equation (9c) refers to the scenario that both haplotypes are the duplicated copies of the maternally derived X chromosome in a male of the previous generation (Figure S1A). According to equations (6a, 6b) and equations (9a–9d), the overall expected density ρ^{mp} in equation (3) does not depend on the three-gene non-IBD probabilities.

The ancestry transition of type (1232) occurs on haplotype ac. A new junction of type (1232) is formed whenever the two parental chromosomes of haplotype ac and the parental chromosome of haplotype bd are distinct and have the IBD configuration (123) at the location of the crossover. We have

$$K_t^{mm}(1232) = s_t^m \left[\frac{1}{4} J_{t-1}^{mp}(1232) + \frac{1}{4} J_{t-1}^{pm}(1232) \right]$$
$$+ \left(1 - s_t^m \right) \left\{ \frac{1}{4} \left[K_{t-1}^{mm}(1232) + \alpha_{t-1}^{mmp}(123) \right] \right.$$
$$\left. + \frac{1}{4} \left[K_{t-1}^{mp}(1232) + \alpha_{t-1}^{mpp}(123) \right] \right\}$$
$$+ \left(1 - s_t^m \right) \left\{ \frac{1}{4} \left[K_{t-1}^{pm}(1232) + \alpha_{t-1}^{mmp}(123) \right] \right. \tag{10a}$$
$$\left. + \frac{1}{4} \left[K_{t-1}^{pp}(1232) + \alpha_{t-1}^{mpp}(123) \right] \right\},$$

$$K_t^{mp}(1232) = \frac{1}{2} \left[K_{t-1}^{mm}(1232) + \alpha_{t-1}^{mmp}(123) \right]$$
$$+ \frac{1}{2} \left[K_{t-1}^{pm}(1232) + \alpha_{t-1}^{mmp}(123) \right], \tag{10b}$$

$$K_t^{pm}(1232) = \frac{1}{2} K_{t-1}^{mm}(1232) + \frac{1}{2} K_{t-1}^{mp}(1232), \tag{10c}$$

$$K_t^{pp}(1232) = \left(1 - s_t^p \right) K_{t-1}^{mm}(1232), \tag{10d}$$

$$J_t^{mp}(1232) = K_t^{mp}(1232), \tag{10e}$$

$$J_t^{pm}(1232) = K_t^{pm}(1232), \tag{10f}$$

for $t \geq 1$, where equations (10e, 10f) may not hold in the founder population at $t = 0$. A new junction is formed at the rate $\alpha_{t-1}^{mmp}(123)$ ($\alpha_{t-1}^{mpp}(123)$), given that the parental chromosome of haplotype bd is maternally (paternally) derived. The density $K_t^{pm}(1232)$ in equation (10c) has no contributions of a new junction because there are no crossover events occurring between the XY chromosomes in the father of haplotype ac (Figure S1B). We denote by $K_t^{mp+}(1232) = [K_t^{mp}(1232) + K_t^{pm}(1232)]/2$, and $K_t^{mp-}(1232) = [K_t^{mp}(1232) - K_t^{pm}(1232)]/2$, and their recurrence relations are given in Appendix A. Both $K_t^{mp-}(1232)$ and R_t^- measure the asymmetry between maternally and paternally derived X chromosomes.

Model evaluation by simulations

To evaluate the theoretical predications of non-IBD probabilities and expected junction densities, we perform simulation studies with the same model assumptions: random mating with discrete generations, no natural selections, and no genetic interferences, except that the ancestral origins along chromosomes do not follow Marker assumptions. Instead, the genome ancestral origins are simulated forwardly by first generating a pedigree according to a given breeding design, and then dropping on the pedigree the distinct founder genome labels (ancestral origins) that are assigned to the whole X chromosomes of each complete inbred founder or to each X chromosome of each outbred founder. The X chromosomes of each descendant gamete are specified as a list of the labeled segments determined by chromosomal crossovers.

For a mapping population with the particular breeding design, the realized junction densities and IBD probabilities are saved for all individuals in each generation in each simulation replicate, and they are averaged over in total 2×10^4 replicates. Various mating schemes are used in simulating breeding pedigrees. We denote by RM1 the random mating where each sampling of two randomly chosen individuals with opposite genders produces *one* offspring, and RM2 the random mating where each sampling of two randomly chosen

individuals with opposite genders produces *two* offspring. We combine these mating schemes with -NE if each parent contributes a Poisson distributed number of gametes to the next generation, and -E if each parent contributes exactly two gametes. Thus, we have four random mating schemes, RM1-NE, RM1-E, RM2-NE, and RM2-E. The sibling mating belongs to RM2-E with population size 2, and the exclusively pairing in 2^n-way ($n \geq 1$) crosses can be regarded as a special case of random mating without inbreeding. The genders are assigned randomly, independent of mating schemes.

APPLICATION TO QTL MAPPING POPULATIONS

Multistage populations

For mapping populations with stage-wise constant mating schemes, we derive analytic expressions of the non-IBD probabilities and the expected junction densities for constructing CTMC of X-linked ancestral origins, according to the recurrence relations. The closed form solutions are obtained by linking results of each subsequent stage via the initial conditions. The general results for a population with constant random mating are derived in Appendix A, where three scenarios are considered: finite population of size ≥ 6, sibling-mating population of size 2, and large population of size $\gg 6$. Table S2 gives the coalescence probabilities of X chromosomes for various mating schemes, similar to Table 1 of Zheng *et al.* (2014) for autosomes. Table S3 summarizes the results for X chromosomes in a sibling-mating population, and Table S4 for autosomes; they are necessary for dioecious breeding populations with a stage of inbreeding by

sibling mating such as the CC and the DSPR. We use the superscripts of A denoting the quantities for autosomes.

We derive the analytic expressions of $\alpha_t^{mp}(12)$, R_t^X, R_t^m, $J_t^{mp}(1122)$, $J_t^{mp+}(1232)$, and $J_t^{mp-}(1232)$ in the mapping populations of the RIL, the AIL, and the DO, and they are given in Table 1, Table 2, and Table 3, respectively. These results are necessary for constructing the CTMC of ancestral origins along the XX chromosomes of a female; only the expression of R_t^m is needed for the maternal derived X chromosome of a male. For comparisons, the autosomal results for $\alpha_t^{AA}(12)$, R_t^A, $J_t^{AA}(1122)$, and $J_t^{AA}(1232)$ are included. The results for the AIL, the DO, and the DSPR are derived under the assumption of a large population size in the intercross stage. We evaluate this assumption in the DSPR, because the evaluation results hold similarly for the AIL and the DO. In addition, the map expansions R_t^X and R_t^A are given explicitly under the assumption of an infinitely large intercross population size, which may be used as a simple measure of QTL mapping resolution.

Many breeding populations can be divided into three stages: mixing, intercross, and inbreeding, such as the RIL by sibling mating, the CC, and the DSPR. There is no inbreeding stage for the AIL, the HS, and the DO. We denote by U the number of intercross generations, V the number of inbreeding generations, and N the intercross population size. Let $\mathcal{M}_F$ and $\mathcal{M}_I$ denote the random mating schemes for mixing and intercross stages, respectively. We choose the mixing stage to consist of one generation of random mating, so that the non-IBD probabilities and the expected junction densities in the F_1 population do not depend on whether genes or haplotypes are in distinct individuals.

■ **Table 1 Results for X chromosomes in the 2^n-way RIL by sibling mating in the last generation $g = U + V + 1$, where $U = 0$ for $n = 1$ and $U = n - 2$ for $n \geq 2$**

Quantity	Theoretical Prediction
(A) 2 ways sibling	
$\alpha_g^{mp}(12)$	$\dfrac{5+\sqrt{5}}{10}(\lambda_1)^V + conjugate$
R_g^X	$\dfrac{8}{3} - \dfrac{20+8\sqrt{5}}{15}(\lambda_1)^V + conjugate$
R_g^m	$\dfrac{8}{3} - \dfrac{2}{3}\left(-\dfrac{1}{2}\right)^V - \dfrac{5+3\sqrt{5}}{5}(\lambda_1)^V + conjugate$
$J_g^{mp}(1122)$	$\dfrac{8}{3} + \dfrac{1}{3}\left(-\dfrac{1}{2}\right)^V - \left(\dfrac{3+\sqrt{5}}{2} + \dfrac{5+\sqrt{5}}{10}V\right)(\lambda_1)^V + conjugate$
$J_g^{mp+}(1232)$	0
$J_g^{mp-}(1232)$	0
(B) $2^n (n \geq 2)$ ways sibling	
$\alpha_g^{mp}(12)$	$\dfrac{5+3\sqrt{5}}{10}(\lambda_1)^V + conjugate$
R_g^X	$\dfrac{2}{3}(U+6) - \dfrac{30+14\sqrt{5}}{15}(\lambda_1)^V + conjugate$
R_g^m	$\dfrac{2}{3}(U+6) + \dfrac{1}{3}\left(R_{U+1}^m - R_{U+1}^p\right)\left(-\dfrac{1}{2}\right)^V - \dfrac{10+4\sqrt{5}}{5}(\lambda_1)^V + conjugate$
$J_g^{mp}(1122)$	$\dfrac{2}{3}(U+6) - \dfrac{1}{6}\left(R_{U+1}^m - R_{U+1}^p\right)\left(-\dfrac{1}{2}\right)^V$
	$\qquad - \left[\dfrac{10+4\sqrt{5}}{5} + \dfrac{1+\sqrt{5}}{4}R_{U+1}^m + \dfrac{5+\sqrt{5}}{20}R_{U+1}^p + \dfrac{5+3\sqrt{5}}{10}V\right](\lambda_1)^V + conjugate$
$J_g^{mp+}(1232)$	$-\left(\dfrac{1}{2}\right)^V + \left[\dfrac{5+3\sqrt{5}}{10} + \dfrac{1+\sqrt{5}}{4}R_{U+1}^m + \dfrac{5+\sqrt{5}}{20}R_{U+1}^p\right](\lambda_1)^V + conjugate$
$J_g^{mp-}(1232)$	$\left(\dfrac{1}{2}\right)^{V+1} + \left(R_{U+1}^p - R_{U+1}^m + 1\right)\left(-\dfrac{1}{2}\right)^{V+1}$

The eigenvalues $\lambda_1 = (1+\sqrt{5})/4$ and $\lambda_2 = (1-\sqrt{5})/4$. The map expansions R_{U+1}^p and R_{U+1}^m are given by equations (11a, 11b). The conjugate is given by replacing $\sqrt{5}$ with $-\sqrt{5}$ from the terms involving λ_1. For example, the conjugate term for R_g^X in (A) is given by $-(20-8\sqrt{5})(\lambda_2)^V/15$. RIL, recombinant inbred line.

■ **Table 2 Results for the AIL in the last generation $g = U + 1$**

Quantity	Theoretical Prediction
(A) X chromosomes	
$\alpha_g^{mp}(12)$	$\left(1 - \dfrac{10}{9L}\right)(\lambda_1)^U + \dfrac{2}{9L}\left(-\dfrac{1}{2}\right)^U + \dfrac{8}{9L}\left(\dfrac{1}{4}\right)^U$
R_g^X	$\dfrac{8}{9L} + \dfrac{2}{3}\left(1 - \dfrac{10}{9L}\right)\dfrac{1 - (\lambda_1)^U}{1 - \lambda_1} - \dfrac{8}{81L}\left(-\dfrac{1}{2}\right)^U - \dfrac{64}{81L}\left(\dfrac{1}{4}\right)^U$
R_g^m	$\dfrac{2}{9} + \dfrac{52}{81L} + \dfrac{2}{3}\left(1 - \dfrac{10}{9L}\right)\dfrac{1 - (\lambda_1)^U}{1 - \lambda_1} - \left(\dfrac{2}{9} + \dfrac{20}{81L} + \dfrac{4}{27L}U\right)\left(-\dfrac{1}{2}\right)^U - \dfrac{32}{81L}\left(\dfrac{1}{4}\right)^U$
$J_g^{mp}(1122)$	$\dfrac{2}{9} + \dfrac{52}{81L} + \dfrac{2}{3}\left(1 - \dfrac{10}{9L}\right)\dfrac{1 - (\lambda_1)^U}{1 - \lambda_1} - \left[\dfrac{2}{9} + \dfrac{52}{81L} + \left(\dfrac{2}{3} + \dfrac{16}{27}s\right)\left(1 - \dfrac{10}{9L}\right)U\right](\lambda_1)^U$
$J_g^{mp+}(1232)$	$\left(1 - \dfrac{2}{L}\right)\left(1 - \dfrac{4}{3L}\right)\dfrac{(\lambda_1)^U - (\lambda_4)^U}{s} + \left(1 - \dfrac{2}{L}\right)\left(-\dfrac{1}{9} + \dfrac{76}{81L}\right)(\lambda_1)^U + \left(1 - \dfrac{2}{L}\right)\dfrac{16}{81L}(\lambda_4)$
	$+ \left(1 - \dfrac{2}{L}\right)\left(\dfrac{1}{9} - \dfrac{4}{81L} + \dfrac{8}{27L}U\right)\left(-\dfrac{1}{2}\right)^U + \left(1 - \dfrac{2}{L}\right)\left(-\dfrac{88}{81L} + \dfrac{16}{27L}U\right)\left(\dfrac{1}{4}\right)^U$
$J_g^{mp-}(1232)$	$\left(1 - \dfrac{2}{L}\right)\left(\dfrac{1}{3} - \dfrac{4}{9L}\right)(\lambda_4)^U - \left(1 - \dfrac{2}{L}\right)\left(\dfrac{1}{3} + \dfrac{4}{9L}\right)\left(-\dfrac{1}{2}\right)^U + \left(1 - \dfrac{2}{L}\right)\dfrac{8}{9L}\left(\dfrac{1}{4}\right)^U$
(B) Autosomes	
$\alpha_g^{AA}(12)$	$\left(1 - \dfrac{1}{L}\right)(\lambda_1^A)^{U-1}$
R_g^A	$1 + \left(1 - \dfrac{1}{L}\right)\dfrac{1 - (\lambda_1^A)^{U-1}}{1 - \lambda_1^A}$
$J_g^{AA}(1122)$	$[1 - (\lambda_1^A)^{U-1}] + \left(1 - \dfrac{1}{L}\right)\left[\dfrac{1 - (\lambda_1^A)^{U-1}}{1 - \lambda_1^A} - (U - 1)(\lambda_1^A)^{U-2}\right]$
$J_g^{AA}(1232)$	$\left(1 - \dfrac{2}{L}\right)(\lambda_1^A)^{U-1} + \left(1 - \dfrac{1}{L}\right)\left(1 - \dfrac{2}{L}\right)\dfrac{(\lambda_1^A)^{U-1} - (\lambda_4^A)^{U-1}}{s^A}$

The eigenvalues $\lambda_1 = 1 - s/3$ and $\lambda_4 = 1 - s$ for X chromosomes, and for autosomes $\lambda_1^A = 1 - s^A/2$ and $\lambda_4^A = 1 - 3s^A/2$. AIL, advanced inter-cross lines.

The general derivation procedure is as follows. First, we derive the initial conditions in the F_1 population for the intercross stage, according to the genetic compositions of the founder population F_0. Second, we substitute the obtained initial conditions into the theorems of Appendix A3 under the assumption of a large intercross population size. Alternatively, the theorems of Appendix A1 may be used for a finite intercross population. Lastly, if there is a stage of inbreeding by sibling mating, we substitute analytic expressions in the F_{U+1} population into the theorems of Appendix A2 to obtain the results in the last generation $g = U + V + 1$.

RIL

The 2^n-way RIL by sibling mating can be regarded as a three-stage mapping population without the intercross stage for $n \leq 2$. All the founders are fully inbred, and the intercross mating scheme is exclusively pairing so that inbreeding is completely avoided. Thus $R_1^m = R_1^p = 0$, and the non-IBD probability $\alpha_1^{mp}(12) = 1$ during the intercross stage $1 \leq t \leq U + 1$, where $U = 0$ for $n = 1$ and $U = n - 2$ for $n \geq 2$. According to the recurrence equations (6a, 6b), it holds

$$R_{U+1}^m = \dfrac{2}{9}\left[1 + 3U - \left(-\dfrac{1}{2}\right)^U\right], \qquad (11a)$$

$$R_{U+1}^p = R_U^m, \qquad (11b)$$

and $R_{U+1}^X = 2U/3$. Furthermore, it is straightforward to obtain $\beta_{U+1}^{mp}(12) = 1$, $\beta_{U+1}^{mm}(12) = \alpha_{U+1}^{mmp}(123) = \delta_{n \geq 2}$, $K_{U+1}^{mp}(1122) = K_{U+1}^{mp}(1122) = 0$, $K_{U+1}^{mm}(1232) = K_{U+1}^{mp}(1232) = R_{U+1}^m$, and

$K_{U+1}^{pm}(1232) = R_{U+1}^p$, where the indicator $\delta_{n \geq 2} = 1$ if $n \geq 2$ and 0 otherwise, since the two maternally derived genes at $t = 1$ must come from the inbred female founder for the two-way RIL.

Substituting the initial conditions in the F_{U+1} population into Table S3, we obtain the results for the RIL in the last generation $t = U + V + 1$ shown in Table 1. The non-IBD probabilities $\alpha_t^{mp}(12)$ for X chromosomes are the same as those for autosomes (Table 2 of Zheng *et al.* 2014). Thus, we show analytically that the map expansion R^X for the X chromosome is two-thirds that of the autosome for the 2^n-way ($n \geq 1$) RIL, according to equations (7a, 8). Broman (2012a) has verified this two-thirds rule via Maxima for the 2^n-way RIL up to $n = 98$.

Figure 3 shows that these theoretical predictions fit very well with the forward simulation results for the two- and eight-way RIL by sibling mating. The differential densities R_t^- and $J_t^{mp-}(1232)$ decay very fast with generation t and show some oscillations in the beginning generations. The overall expected junction density ρ_t^{mp} reaches the maximum in the same generation for autosomes.

AIL

We consider a multiparental AIL population that is founded by $L/2$ inbred females and $L/2$ inbred males. A unique ancestral origin is assigned to each inbred founder's genomes so that the two-gene non-IBD probabilities $\alpha_0^{mp}(12) = 0$ and $\beta_0^{mm}(12) = \beta_0^{mp}(12) = \beta_0^{pp}(12) = 1$, and similarly for the three-gene non-IBD probabilities $\alpha_0^{mmp}(123) = \alpha_0^{mpp}(123) = 0$, $\beta_0^{mmm}(123) = \beta_0^{mmp}(123) = \beta_0^{mpp}(123) = \beta_0^{ppp}(123) = 1$ if they exist.

Quantity	Theoretical Prediction
(A) X chromosomes	
$\alpha_g^{mp}(12)$	$\left(1 - \frac{1}{L}\right)(\lambda_1)^U$
R_g^X	$R_0^X + \frac{2}{3}\alpha_0^{mp}(12) + \frac{2}{3}\left(1 - \frac{1}{L}\right)\frac{1 - (\lambda_1)^U}{1 - \lambda_1}$
R_g^m	$R_0^X + \frac{2}{3}\alpha_0^{mp}(12) + \frac{2}{9}\left(1 - \frac{1}{L}\right) + \frac{2}{3}\left(1 - \frac{1}{L}\right)\frac{1 - (\lambda_1)^U}{1 - \lambda_1}$
	$\quad - \left[\frac{2}{9}\left(1 - \frac{1}{L}\right) + \frac{1}{6}(R_0^m - R_0^P) - \frac{1}{3}\alpha_0^{mp}(12)\right]\left(-\frac{1}{2}\right)^U$
$J_g^{mp}(1122)$	$R_0^X + \frac{2}{3}\alpha_0^{mp}(12) + \frac{2}{9}\left(1 - \frac{1}{L}\right) + \frac{2}{3}\left(1 - \frac{1}{L}\right)\frac{1 - (\lambda_1)^U}{1 - \lambda_1}$
	$\quad - \left[R_0^X + \frac{2}{3}\alpha_0^{mp}(12) + \frac{2}{9}\left(1 - \frac{1}{L}\right) + \left(\frac{2}{3} + \frac{16}{27}s\right)\left(1 - \frac{1}{L}\right)U\right](\lambda_1)^U$
$J_g^{mp+}(1232)$	$\left(1 - \frac{1}{L}\right)\left(1 - \frac{2}{L}\right)\frac{(\lambda_1)^U - (\lambda_4)^U}{s} + \left(1 - \frac{2}{L}\right)\left[-\frac{1}{9}\left(1 - \frac{1}{L}\right) + R_0^X + \frac{2}{3}\alpha_0^{mp}(12)\right](\lambda_1)^U$
	$\quad + \left(1 - \frac{2}{L}\right)\left[\frac{1}{9}\left(1 - \frac{1}{L}\right) + \frac{1}{12}(R_0^m - R_0^P) - \frac{1}{6}\alpha_0^{mp}(12)\right]\left(-\frac{1}{2}\right)^U$
$J_g^{mp-}(1232)$	$\frac{1}{3}\left(1 - \frac{1}{L}\right)\left(1 - \frac{2}{L}\right)(\lambda_4)^U + \left(1 - \frac{2}{L}\right)\left[-\frac{1}{3}\left(1 - \frac{1}{L}\right) - \frac{1}{4}(R_0^m - R_0^P) + \frac{1}{2}\alpha_0^{mp}(12)\right]\left(-\frac{1}{2}\right)^U$
(B) Autosomes	
$\alpha_g^{AA}(12)$	$\left(1 - \frac{1}{L}\right)(\lambda_1^A)^U$
R_g^A	$R_0^A + \alpha_0^{AA}(12) + \left(1 - \frac{1}{L}\right)\frac{1 - (\lambda_1^A)^U}{1 - \lambda_1^A}$
$J_g^{AA}(1122)$	$[R_0^A + \alpha_0^{AA}(12)][1 - (\lambda_1^A)^U] + \left(1 - \frac{1}{L}\right)\left[\frac{1 - (\lambda_1^A)^U}{1 - \lambda_1^A} - U(\lambda_1^A)^{U-1}\right]$
$J_g^{AA}(1232)$	$[R_0^A + \alpha_0^{AA}(12)]\left(1 - \frac{2}{L}\right)(\lambda_1^A)^U + \left(1 - \frac{1}{L}\right)\left(1 - \frac{2}{L}\right)\frac{(\lambda_1^A)^U - (\lambda_4^A)^U}{s^A}$

The eigenvalues $\lambda_1 = 1 - s/3$ and $\lambda_4 = 1 - s$ for X chromosomes, and for autosomes $\lambda_1^A = 1 - s^A/2$ and $\lambda_4^A = 1 - 3s^A/2$. DO, diversity outcross.

The F_1 population of size N is produced by mating scheme $\mathcal{M}_F =$ RM1-NE or RM2-NE. According to Table S2, the coalescence probabilities $s_1^m = s_1^P = 2/L$ and $q_1^m = q_1^P = (2/L)(1 - 2/L)$ for mating scheme RM1-NE, and they hold approximately for RM2-NE with large population size $N \gg 6$. Thus, the two-gene non-IBD probabilities at $t = 1$ are given by $\beta_1^{mm}(12) = \beta_1^{PP}(12) = 1 - 2/L$ and $\alpha_1^{mp}(12) = \beta_1^{mp}(12) = 1$ according to the recurrence equations (4a–4d), and the three-gene non-IBD probabilites at $t = 1$ are given by $\beta_1^{mmm}(123) = \beta_1^{PPP}(123) = (1 - 2/L)(1 - 4/L)$ and $\alpha_1^{mmp}(123) = \alpha_1^{mpp}(123) = \beta_1^{mmp}(123) = \beta_1^{mpp}(123) = 1 - 2/L$ according to the recurrence equations (5a–5f). In addition, no junctions can be formed from inbred founders so that it holds that $R_1^m = R_1^P = 0$, $K_1^{mm}(1122) = K_1^{mp}(1122) = K_1^{pp}(1122) = 0$, and $K_1^{mm}(1232) = K_1^{mp}(1232) = K_1^{pm}(1232) = K_1^{pp}(1232) = 0$.

The F_1 population is maintained for U generations with constant size N and sex ratio 1. Assuming that the intercross population size is large ($N \gg 6$), all the two- and three-gene coalescence probabilities at $t \geq 2$ are approximately equal and are denoted by s, and they are determined by the intercross mating scheme $\mathcal{M}_I$ according to Table S2. Substituting the initial conditions in the F_1 population into the theorems of Appendix A3, we obtain in Table 2 the results for X chromosomes in the AIL in the last generation $t = U + 1$. Table 2 also shows the results for autosomes, which are derived according to Zheng *et al.* (2014).

As shown in Table 2, the non-IBD probabilities $\alpha_t^{mp}(12)$ for X-chromosomes are unequal to those for autosomes, and thus the map expansions generally do not satisfy the two-thirds rule. According to the map expansions R_t^X and R_t^A in Table 2, we derive their approximations under the limit of an infinitely large population size ($N \to \infty$) so that the coalescence probability goes to zero ($s \to 0$),

$$R_{U+1}^X \approx \frac{8}{9L} + \frac{2}{3}\left(1 - \frac{10}{9L}\right)U, \quad (12a)$$

$$R_{U+1}^A \approx 1 + \left(1 - \frac{1}{L}\right)(U - 1), \quad (12b)$$

where the last two terms for R_t^X in Table 2 are small and thus ignored. The equations (12a, 12b) show that the two-thirds rule is approximately valid for a large number L of founder lines. The map expansion of equation (12b) for $L = 2$ is consistent with the previous results (Darvasi and Soller 1995; Liu *et al.* 1996; Winkler *et al.* 2003; Broman 2012b).

The left panels of Figure 4 show for the AIL that the theoretical predictions fit very well with the forward simulation results, where $\mathcal{M}_F = $ RM1-NE, $\mathcal{M}_I = $ RM1-E, $L = 8$, and $N = 100$. Within $U = 20$ intercross generations, the non-IBD probability $\alpha_t^{mp}(12)$ decreases slowly with generation t, the differential map expansion
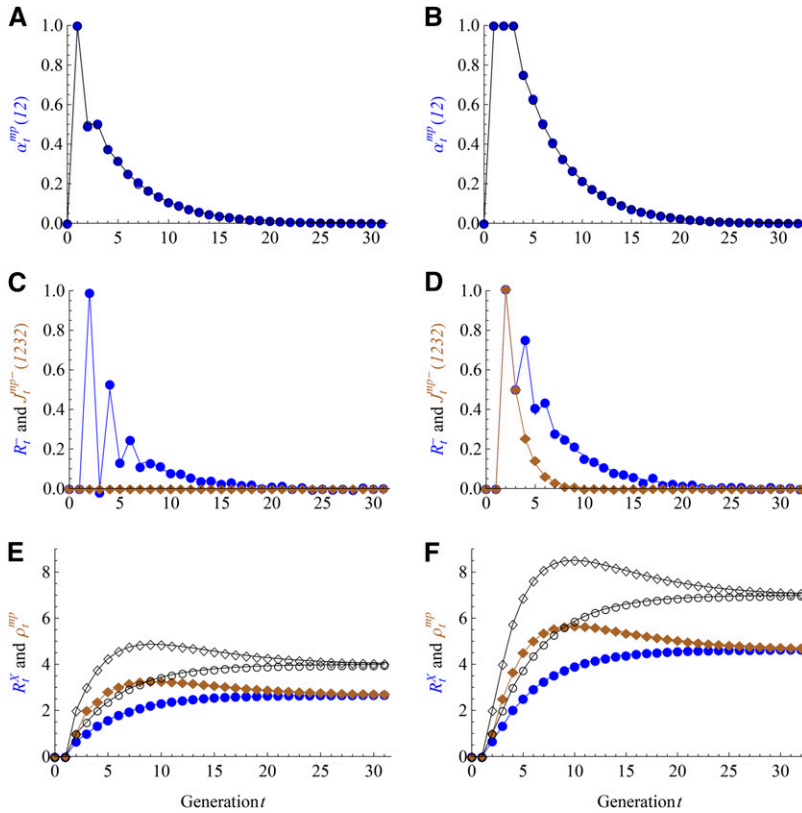
Figure 3 Results of the 2^n-way recombinant inbred lines (RILs) with by sibling mating for $n = 1$ (left panels) and $n = 3$ (right panels). The filled symbols refer to the results for X chromosomes, the empty symbols for autosomes, and lines for the theoretical predictions in Table 1. The non-IBD probabilities $\alpha_t^{mp}(12)$ for X chromosomes and autosomes are overlapped with each other. The brown filled diamonds refer to $J_t^{mp-}(1232)$ in (C) and (D) and ρ_t^{mp} in (E) and (F).

R_t^- remains almost constant after a few generations of oscillations, and the map expansions in equations (12a, 12b), shown as thick red lines in Figure 4, are very good approximations.

HS and DO

The HS and the DO differ from the AIL only in the genetic compositions of the founder population. The N progenitors of the DO at $t = 0$ were sampled independently from pre-CC lines at a variety of different generations. Each pre-CC line is produced by the RIL by sibling mating from $L = 8$ randomly permuted founder strains. Let q_k denote the proportion of the pre-CC progenitors that were in generation k. Thus, for a random progenitor, it holds $\alpha_0^{mp}(12) = \sum_k q_k \alpha_k^{mp}(12|\text{pre-CC})$ and $R_0^O = \sum_k q_k R_k^O(\text{pre-CC})$, where $\alpha_k^{mp}(12|\text{pre-CC})$ and $R_k^O(\text{pre-CC})$ for $O = m, p$ can be obtained from Table 1. Because the founder stains are exchangeable, we obtain $\beta_0^{mm}(12) = \beta_0^{mp}(12) = \beta_0^{pp}(12) = 1 - 1/L$, $\alpha_0^{mmp}(123) = \alpha_0^{mpp}(123) = \alpha_0^{mp}(12)(1 - 2/L)$, and $\beta_0^{mmm}(123) = \beta_0^{mmp}(123) = \beta_0^{mpp}(123) = \beta_0^{ppp}(123) = (1 - 1/L)(1 - 2/L)$, and because recombination crossovers are independent among different pre-CC lines, the between-individual expected junction densities at $t = 0$ are given by $K_0^{mm}(1122) = K_0^{mp}(1122) = K_0^{pp}(1122) = 0$, $K_0^{mm}(1232) = K_0^{mp}(1232) = R_0^m(1 - 2/L)$, and $K_0^{pp}(1232) = K_0^{pm}(1232) = R_0^p(1 - 2/L)$, where $1 - 2/L$ refers to the probability that the third ancestral origin on haplotype bd is different from the two ancestral origins on haplotype ac where the ancestry transition occurs. The within-individual expected junction densities at $t = 0$ are not required in the following derivations.

The F_1 population of size N is produced by random mating with equal sex ratio. Assuming that the population size $N \gg 6$, the coalescence probabilities at $t = 1$ are approximated to be zero. According to the recurrence equations for the two- and three-gene non-IBD probabilities, the between-individual probabilities did not change and

the within-individual non-IBD probabilities at $t = 1$ equal to the corresponding between-individual probabilities. In addition, we have $R_1^m = R_0^m/2 + R_0^p/2 + \alpha_0^{mp}(12)$, $R_1^p = R_0^m$, $K_1^{mm}(1122) = K_1^{mp}(1122) = K_1^{pp}(1122) = 0$, $K_1^{mm}(1232) = K_1^{mp}(1232) = [R_0^m/2 + R_0^p/2 + \alpha_0^{mp}(12)](1 - 2/L)$, $K_1^{pp}(1232) = K_1^{pm}(1232) = R_0^m(1 - 2/L)$, according to the recurrence equations for the expected junction densities.

Similar to the intercross stage of the AIL, we obtain in Table 3 the results for X chromosomes in the DO in the last generation $t = U + 1$ by substituting the initial conditions in the F_1 population into the theorems of Appendix A3. Table 3 also shows the results for autosomes, which are derived according to Zheng *et al.* (2014). Under the limit of an infinitely large population size ($N \rightarrow \infty$), we obtain from Table 3

$$R_{U+1}^X \approx R_0^X + \frac{2}{3}\alpha_0^{mp}(12) + \frac{2}{3}\left(1 - \frac{1}{L}\right)U, \qquad (13a)$$

$$R_{U+1}^A \approx R_0^A + \alpha_0^{AA}(12) + \left(1 - \frac{1}{L}\right)U, \qquad (13b)$$

showing that the two-thirds rule is valid under such an approximation since $\alpha_0^{mp}(12) = \alpha_0^{AA}(12)$ and $R_0^X = 2R_0^A/3$ for progenitors drawn from the RIL (Table 1). The map expansion in equation (13b) for $L = 8$ is the same as the one obtained by Broman (2012b).

The right panels of Figure 4 show for the HS that the theoretical predictions fit very well with the forward simulation results, where $\mathcal{M}_F = \mathcal{M}_I = \text{RM1-E}$, the $N = 100$ individuals in the F_0 population were sampled independently from CC funnels at the same generation $t = 3$. The results are similar to those for the AIL with the same L shown in the left panels of Figure 4. For X chromosomes, the non-IBD probabilities in the DO are larger than those in the AIL, and thus in the DO the map expands at a higher rate than that for the AIL, see equations (12a, 13a).
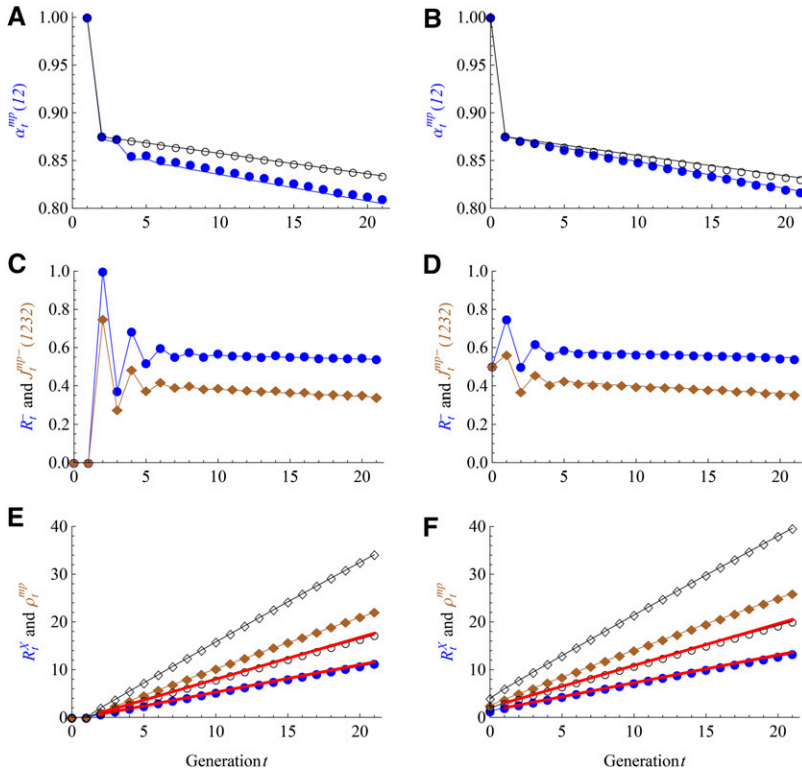
Figure 4 Results of the AIL (left panels) and the HS (right panels) with $L = 8$ and $N = 100$. The random mating schemes $\mathcal{M}_F$ = RM1-NE for the AIL and RM1-E for the HS, and $\mathcal{M}_I$ = RM1-E for both populations. The symbols and lines are the same as those in Figure 3. The theoretical predictions refer to Table 2 for the AIL and Table 3 for the DO. The additional red lines denote the map expansions under the large size assumption, given by equations (12a, 12b) for the AIL and equations (13a, 13b) for the HS.

DSPR

The DSPR RILs were derived from two synthetic populations, each created independently by adding the multiparental AIL with an inbreeding stage by sibling mating (King *et al.* 2012). For example, we derive the analytic expressions of the map expansions in one synthetic population with L founder strains. We assume that $\beta_{U+1}^{mm}(12) = \beta_{U+1}^{mp}(12)$, which holds in a non-inbreeding population and approximately in a large population (*e.g.*, $N \geq 100$) with a large number of intercross generations (*e.g.*, $U \geq 6$). According to the map expansions in Table S3, we have

$$R_{U+V+1}^X \approx R_{U+1}^X$$
$$+ \alpha_{U+1}^{mp}(12)\left[4 - \frac{30 + 14\sqrt{5}}{15}(\lambda_1)^V - \frac{30 - 14\sqrt{5}}{15}(\lambda_2)^V\right],$$

(14a)

$$R_{U+V+1}^A \approx R_{U+1}^A$$
$$+ \alpha_{U+1}^{AA}(12)\left[6 - \frac{15 + 7\sqrt{5}}{5}(\lambda_1)^V - \frac{15 - 7\sqrt{5}}{5}(\lambda_2)^V\right],$$

(14b)

where $\lambda_1 = \left(1 + \sqrt{5}\right)/4$ and $\lambda_2 = \left(1 - \sqrt{5}\right)/4$, and R_{U+1}^X and $\alpha_{U+1}^{mp}(12)$ are given in Table 1, Table 2, or Table 3 if the F_{U+1} population is the last generation of the RIL, the AIL, or the DO, respectively.

We evaluate the large size assumption for various random mating schemes by simulation studies of the DSPR. Figure 5 shows the fitting of the theoretical predictions with the forward simulation results for the intercross size $N = 20$, 50, and 100, where the mating schemes $\mathcal{M}_F$ = RM1-NE and $\mathcal{M}_I$ = RM1-E (RM1-NE) for the left (right) panels. The theoretical predictions are obtained by combining the results for the AIL (Table 2) with those for the sibling-mating pop-

ulation (Table S3), assuming the large size ($N \gg 6$). The relative worse fitting for the differential densities R_t^- and $J_t^{mp-}(1232)$ is probably attributable to the limited number (2×10^4) of simulation replicates. The theoretical fitting becomes improved with increased size N, and it is very good for $N = 100$ within the range of $U = 20$ intercross generations. The fitting for RM1-E is better than RM1-NE because in the former case the two-gene coalescence probabilities are always equal to the three-gene probabilities (Table S2), independent of the size N. Figure S2 shows similar results for the random mating scheme RM2, except that the expected junction densities are slightly smaller. Figure S3 and Figure S4 show that the large size assumption is less sensitive for autosomes, and the fittings are very good even for $N = 20$.

DISCUSSION

We have extended our previous framework of modeling ancestral origin processes from autosomes to X chromosomes, and thus the same assumptions such as exchangeability of ancestral origins, Markov properties and random mating also apply (Zheng *et al.* 2014). The deviations from Markov properties result in larger variances in the IBD-tract length and the junction densities, which have been shown to be acceptable (Chapman and Thompson 2003; Martin and Hospital 2011). The random mating indicates that our approach does not apply to breeding populations with marker-assisted selections.

In contrast to the previous approaches (Haldane and Waddington 1931; Broman 2012a), the exchangeability assumption of ancestral origins greatly reduces model complexity, because the number of possible junction types does not depend on the number of founders for $L \geq 3$ whereas the number of diplotype states increases very fast with L. The assumption affects the rate matrix of the Markov model, but not the expected junction densities where only changes of ancestral origins matter. The exchangeability is a good approximation for the AIL- or the multiparent advanced generation inter-cross (*i.e.*, MAGIC)-type
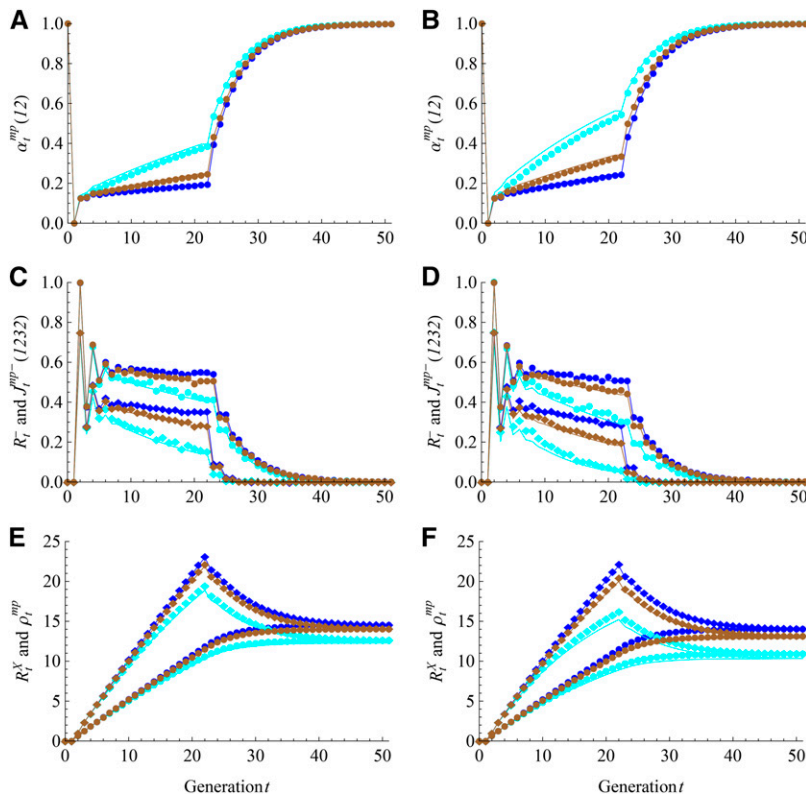
Figure 5 Results of the DSPR for X chromosomes with $L = 8$ and $N = 20$ (cyan), 50 (brown), and 100 (blue). The random mating schemes M_F = RM1-NE for all panels and M_I = RM1-E (RM1-NE) for the left (right) panels. The lines denote the theoretical predictions under the large size assumption. The filled circles refer to R_t^- in panels C and D, and R_t^X in panels E and F; the filled diamonds refer to $J_t^{mp-}(1232)$ in (C) and (D) and ρ_t^{mp} in (E) and (F).

populations with random mating, but it does not hold for the multiway RIL by sibling mating.

However, the exchangeability assumption is not critical for the application of our results to haplotype reconstructions from genotype data. The genomes of the individuals collected in the last generation have been well mixed by random chromosomal segregations over many generations. This is demonstrated in Figure 1A for the four-way RIL by sibling mating, where a female A and a male B was crossed, and a female C and a male D was crossed, and then a daughter from A × B and a son from C × D was crossed. The X chromosome of the founder D is lost in F_1. The genotype probabilities for AB and AC are different and given in the Table 2 of Broman (2012a), although the sum of the genotype probabilities for AB, AC, and BC is equal to $\alpha_t^{mp}(12)$ in Table 1. Figure 1B shows that the genotype probability for AB or AC becomes close to the average probability $\alpha_t^{mp}(12)/3$ as generation t increases. Furthermore, in the beginning generations when the asymmetry among ancestral origins is large, there are fewer number of recombination breakpoints, and thus more marker data per genome block are available to estimate ancestral origins. As a result, *a priori* equal weights of ancestral origins have little effects.

An HMM is under development for reconstructing ancestral origins for both autosomes and X chromosomes from marker data, using the present model and the previous one (Zheng *et al.* 2014) as the prior distribution. The previously implemented HMM methods, such as GAIN (Liu *et al.* 2010) and HAPPY (Mott *et al.* 2000), were developed for autosomes, and they do not account for the asymmetry between maternally and paternally derived X chromosomes.

The closed form expressions for non-IBD probabilities and various expected junction densities have been derived for stage-wise mapping populations. They provide the complete information for constructing the CTMC along two X chromosomes but also the guides for designing a new population in terms of X-linked QTL mapping resolutions. For advanced intercross populations such as the AIL, the HS, and the DO under the assumption of a large intercross size, the map expands linearly at a rate proportional to the inverse of the number L of inbred founders, which is robust to intercross mating schemes. For the RIL and the inbreeding stage of the DSPR, the map expansion slows down with increasing level of inbreeding. The overall junction density ρ^{mp} for the DSPR decreases after one generation of the inbreeding stage by sibling mating, whereas for the RIL it reaches the maximum in the middle of inbreeding by sibling mating. These conclusions can also be applied to autosomes. Thus the most effective way of improving mapping resolutions is to increase the number U of intercross generations in a large population ($N \geq 5U$, empirically).

LITERATURE CITED

Broman, K., 2005 The genomes of recombinant inbred lines. Genetics 169: 1133–1146.

Broman, K. W., 2012a Genotype probabilities at intermediate generations in the construction of recombinant inbred lines. Genetics 190: 403–412.

Broman, K. W., 2012b Haplotype probabilities in advanced intercross populations. G3 (Bethesda) 2: 199–202.

Chapman, N., and E. Thompson, 2003 A model for the length of tracts of identity by descent in finite random mating populations. Theor. Popul. Biol. 64: 141–150.

Churchill, G., D. Airey, H. Allayee, J. Angel, A. Attie *et al.*, 2004 The collaborative cross, a community resource for the genetic analysis of complex traits. Nat. Genet. 36: 1133–1137.

Darvasi, A., and M. Soller, 1995 Advanced intercross lines, an experimental population for fine genetic-mapping. Genetics 141: 1199–1207.

Fisher, R., 1949 *The Theory of Inbreeding*. Oliver and Boyd, London.

Fisher, R., 1954 A fuller theory of junctions in inbreeding. Heredity 8: 187–197.

Haldane, J., and C. Waddington, 1931 Inbreeding and linkage. Genetics 16: 357–374.

Johannes, F., and M. Colome-Tatche, 2011 Quantitative epigenetics through epigenomic perturbation of isogenic lines. Genetics 188: 215–227.

King, E. G., S. J. Macdonald, and A. D. Long, 2012 Properties and power of the *Drosophila* Synthetic Population Resource for the routine dissection of complex traits. Genetics 191: 935–949.

Liu, E. Y., Q. Zhang, L. McMillan, F. Pardo-Manuel de Villena, and W. Wang, 2010 Efficient genome ancestry inference in complex pedigrees with inbreeding. Bioinformatics 26: i199–i207.

Liu, S., S. Kowalski, T. Lan, I. Feldmann, and A. Paterson, 1996 Genome-wide high-resolution mapping by recurrent intermating using *Arabidopsis thaliana* as a model. Genetics 142: 247–258.

MacLeod, A. K., C. S. Haley, and J. A. Woolliams, 2005 Marker densities and the mapping of ancestral junctions. Genet. Res. 85: 69–79.

Martin, O. C., and F. Hospital, 2011 Distribution of parental genome blocks in recombinant inbred lines. Genetics 189: 645–654.

Ming, R., and P. H. Moore, 2007 Genomics of sex chromosomes. Curr. Opin. Plant Biol. 10: 123–130.

Mott, R., C. Talbot, M. Turri, A. Collins, and J. Flint, 2000 A method for fine mapping quantitative trait loci in outbred animal stocks. Proc. Natl. Acad. Sci. USA 97: 12649–12654.

Svenson, K. L., D. M. Gatti, W. Valdar, C. E. Welsh, R. Y. Cheng *et al.*, 2012 High-resolution genetic mapping using the mouse diversity outbred population. Genetics 190: 437–447.

Weller, J., and M. Soller, 2004 An analytical formula to estimate confidence interval of qtl location with a saturated genetic map as a function of experimental design. Theor. Appl. Genet. 109: 1224–1229.

Winkler, C., N. Jensen, M. Cooper, D. Podlich, and O. Smith, 2003 On the determination of recombination rates in intermated recombinant inbred populations. Genetics 164: 741–745.

Zheng, C. Z., M. P. Boer, and F. A. van Eeuwijk, 2014 A general modeling framework for genome ancestral origins in multiparental populations. Genetics 198: 87–101.

Communicating editor: E. Huang

APPENDIX A

RESULTS FOR CONSTANT RANDOM MATING POPULATIONS

We introduce some matrix-vector notations to facilitate the derivations. Denote by $A \odot B$ the element-by-element multiplication of the two matrices A and B, and by $A \oslash B$ the element-by-element division of the two matrices. Denote by $\boldsymbol{x}_{i-j}^t = (x_i^t, x_{i+1}^t, \ldots, x_j^t)$ the element-wise power where the subscripts of the natural numbers $i \leq j$, and by default $\boldsymbol{x}_{i-j} = (x_i, x_{i+1}, \ldots, x_j)$. Let $\mathbf{1}$ be a vector with appropriate length and all the elements being 1. Let $\Lambda(\boldsymbol{x})$ be the diagonal matrix with the diagonal elements being the vector $\boldsymbol{x}$. Denote by $[\boldsymbol{x}, \ldots, \boldsymbol{y}]$ the matrix with row vectors $\boldsymbol{x}, \ldots, \boldsymbol{y}$ of equal length. Denote by superscript T the transpose of a vector or matrix.

The closed form expressions for the two- and three-gene non-IBD probabilities and the expected junction densities are derived for populations with constant size and random mating schemes. The coalescence probabilities are thus constant, and set $s_t^m = s^m$, $s_t^p = s^p$, $q_t^m = q^m$, and $q_t^p = q^p$. We first consider a finite population, number of males $N_m \geq 1$ and number of females $N_f \geq 3$, so that all the two- and three-gene non-IBD probabilities exist. Then consider an example of small population size, a sibling-mating population with one male and one female ($N_f = N_m = 1$), where the non-IBD probabilities $\beta^{pp}(12)$, $\beta^{mmm}(123)$, $\beta^{mmp}(123)$, $\beta^{mpp}(123)$, and $\beta^{ppp}(123)$ do not exist. Lastly, we consider a large population under the limit that the size $N_f = N_m \gg 3$.

A1 Finite population

Definition A1.1. The finite population refers to a population of constant number $N_m \geq 1$ of males and number $N_f \geq 3$ of females, maintained by random mating, and the initial population satisfies $\alpha_0^{mp}(12) = \beta_0^{mp}(12)$, $\alpha_0^{mmp}(123) = \beta_0^{mmp}(123)$, and $\alpha_0^{mpp}(123) = \beta_0^{mpp}(123)$.

Proposition A1.2. *Denote by* $\boldsymbol{\beta}_t(12) = (\beta_t^{mm}(12), \beta_t^{mp}(12), \beta_t^{pp}(12))^T$ *the two-gene non-IBD probability in a finite population. According to equations (4a–4c), it holds*

$$\boldsymbol{\beta}_t(12) = \boldsymbol{T}_{12} \boldsymbol{\beta}_{t-1}(12) \tag{A1.1}$$

where

$$\boldsymbol{T}_{12} = \begin{pmatrix} \frac{1}{4}(1 - s^m) & \frac{1}{2} & \frac{1}{4}(1 - s^m) \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 1 - s^p & 0 & 0 \end{pmatrix}. \tag{A1.2}$$

Premise A1.3. *The eigenvalues of* $\boldsymbol{T}_{12}$ *in a finite population, denoted by* λ_k, $k \in \{1, 2, 3\}$ *in the decreasing order of their absolute values, are distinct with multiplicities 1, and none of them is 0, 1, or* $-1/2$.

Theorem A1.4. *The two-gene non-IBD probability* $\boldsymbol{\beta}_t(12)$ *in a finite population is given by*

$$\boldsymbol{\beta}_t(12) = \boldsymbol{P} \Lambda(\boldsymbol{\lambda}_{1-3}^t)(\boldsymbol{C}_{1-3})^T, \tag{A1.3}$$

where

$$\boldsymbol{P} = \left[2\boldsymbol{\lambda}_{1-3} - 1, \mathbf{1}, (1 - s^p)(2\boldsymbol{\lambda}_{1-3} - 1) \oslash \boldsymbol{\lambda}_{1-3} \right], \tag{A1.4}$$

$$(\boldsymbol{C}_{1-3})^T = \boldsymbol{P}^{-1} \boldsymbol{\beta}_0(12). \tag{A1.5}$$

Proof. It holds

$$\beta_t^{mm}(12) = \boldsymbol{B}_{1-3}(\boldsymbol{\lambda}_{1-3}^t)^T, \tag{A1.6a}$$

$$\beta_t^{mp}(12) = \boldsymbol{C}_{1-3}(\boldsymbol{\lambda}_{1-3}^t)^T, \tag{A1.6b}$$

$$\beta_t^{pp}(12) = \boldsymbol{D}_{1-3}(\boldsymbol{\lambda}_{1-3}^t)^T, \tag{A1.6c}$$

where the constant coefficients $\boldsymbol{B}_{1-3}$, $\boldsymbol{C}_{1-3}$, and $\boldsymbol{D}_{1-3}$ are to be solved. Substituting $\beta_t^{mm}(12)$ and $\beta_t^{mp}(12)$ of equations (A1.6a, A1.6b) into the recurrence equation (4b), we obtain

$$B_k = (2\lambda_k - 1)C_k, k \in \{1, 2, 3\}. \tag{A1.7}$$

Substituting $\beta_t^{mm}(12)$ and $\beta_t^{pp}(12)$ of equations (A1.6a, A1.6c) into the recurrence equation (4c), we obtain

$$D_k = (1 - s^p)\frac{B_k}{\lambda_k}, k \in \{1, 2, 3\}, \tag{A1.8}$$

Substituting B_k and D_k of equations (A1.7, A1.8) into equation (A1.6a–A1.6c), we obtain

$$\boldsymbol{\beta}_t(12) = \boldsymbol{P}\Big[\boldsymbol{C}_{1-3}\big(\boldsymbol{\lambda}_{1-3}^t\big)^T\Big], \tag{A1.9}$$

where $\boldsymbol{P}$ is given by in equation (A1.4), and the constant coefficient $\boldsymbol{C}_{1-3}$ is determined by the initial condition $\boldsymbol{\beta}_0(12)$ and it is given by equation (A1.5).

Proposition A1.5. *Denote by* $\boldsymbol{\beta}_t(123) = (\beta_t^{mmm}(123), \beta_t^{mmp}(123), \beta_t^{mpp}(123), \beta_t^{ppp}(123))^T$ *the three-gene non-IBD probability in a finite population. According to equations (5a–5d), it holds*

$$\boldsymbol{\beta}_t(123) = \boldsymbol{T}_{123}\boldsymbol{\beta}_{t-1}(123) \tag{A1.10}$$

where

$$\boldsymbol{T}_{123} = \begin{pmatrix} \frac{1}{8}(1 - s^m - 2q^m) & \frac{3}{8}(1 - s^m) & \frac{3}{8}(1 - s^m) & \frac{1}{8}(1 - s^m - 2q^m) \\ \frac{1}{4}(1 - s^m) & \frac{1}{2} & \frac{1}{4}(1 - s^m) & 0 \\ \frac{1}{2}(1 - s^p) & \frac{1}{2}(1 - s^p) & 0 & 0 \\ 1 - s^p - 2q^p & 0 & 0 & 0 \end{pmatrix}. \tag{A1.11}$$

Premise A1.6. *The eigenvalues of* $\boldsymbol{T}_{123}$ *in a finite population, denoted by* $\lambda_k, k \in \{4, 5, 6, 7\}$ *in the decreasing order of their absolute values, are distinct with multiplicities 1, and none of them is* $0, 1, -1/2,$ *or* $\lambda_k, k \in \{1, 2, 3\}$.

Theorem A1.7. *The three-gene non-IBD probability* $\boldsymbol{\beta}_t(123)$ *in a finite population is given by*

$$\boldsymbol{\beta}_t(123) = \boldsymbol{Q}\boldsymbol{\Lambda}\big(\boldsymbol{\lambda}_{4-7}^t\big)\big(\boldsymbol{C}_{4-7}\big)^T, \tag{A1.12}$$

where

$$\boldsymbol{Q} = \Big[\boldsymbol{a}_{4-7}, \boldsymbol{1}, (1 - s^p)(\boldsymbol{a}_{4-7} + 1)\oslash(2\boldsymbol{\lambda}_{4-7}), (1 - s^p - 2q^p)\boldsymbol{a}_{4-7}\oslash\boldsymbol{\lambda}_{4-7}\Big], \tag{A1.13}$$

$$\big(\boldsymbol{C}_{4-7}\big)^T = \boldsymbol{Q}^{-1}\boldsymbol{\beta}_0(123). \tag{A1.14}$$

and for $k \in \{4, 5, 6, 7\}$

$$a_k = \frac{8(\lambda_k)^2 - 4\lambda_k - (1 - s^m)(1 - s^p)}{(1 - s^m)(1 - s^p + 2\lambda_k)}. \tag{A1.15}$$

Proof. It holds

$$\beta_t^{mmm}(123) = \boldsymbol{A}_{4-7}\big(\boldsymbol{\lambda}_{4-7}^t\big)^T, \tag{A1.16a}$$

$$\beta_t^{mmp}(123) = \boldsymbol{C}_{4-7}\big(\boldsymbol{\lambda}_{4-7}^t\big)^T, \tag{A1.16b}$$

$$\beta_t^{mpp}(123) = \boldsymbol{B}_{4-7}\big(\boldsymbol{\lambda}_{4-7}^t\big)^T, \tag{A1.16c}$$

$$\beta_t^{ppp}(123) = \boldsymbol{D}_{4-7}\big(\boldsymbol{\lambda}_{4-7}^t\big)^T, \tag{A1.16d}$$

where the constant coefficients $\boldsymbol{A}_{4-7}, \boldsymbol{B}_{4-7}, \boldsymbol{C}_{4-7},$ and $\boldsymbol{D}_{4-7}$ are to be solved. Substituting $\beta_t^{mmm}(123)$ and $\beta_t^{ppp}(123)$ of equations (A1.16a, A1.16d) into the recurrence equation (5d), we obtain

$$D_k = (1 - s^p - 2q^p)\frac{A_k}{\lambda_k}, k \in \{4, 5, 6, 7\}. \tag{A1.17}$$

Substituting $\beta_t^{mmm}(123), \beta_t^{mmp}(123),$ and $\beta_t^{mpp}(123)$ of equations (A1.16a–A1.16c) into the recurrence equation (5c), we obtain

$$B_k = \frac{(1 - s^p)(A_k + C_k)}{2\lambda_k}, k \in \{4, 5, 6, 7\}. \tag{A1.18}$$

Substituting $\beta_t^{mmm}(123)$, $\beta_t^{mmp}(123)$, and $\beta_t^{mpp}(123)$ of equations (A1.16a–A1.16c) into the recurrence equation (5b), and substituting B_k of equation (A1.18), we obtain

$$A_k = a_k C_k, k \in \{4, 5, 6, 7\}, \tag{A1.19}$$

where a_k is given by equation (A1.15). Substituting A_k, B_k, and D_k of equations (A1.17–A1.19) into equations (A1.16a–A1.16d), we obtain

$$\boldsymbol{\beta}_t(123) = \boldsymbol{Q}\left[\boldsymbol{C}_{4-7}\left(\boldsymbol{\lambda}_{4-7}^t\right)^T\right], \tag{A1.20}$$

where $\boldsymbol{Q}$ is given by equation (A1.13), and the constant coefficient $\boldsymbol{C}_{4-7}$ is determined by the initial condition $\boldsymbol{\beta}_0(123)$ and it is given by equation (A1.14).

Theorem A1.8. *The map expansions in a finite population are given by*

$$R_t^X = C_8 - \frac{2}{3}\left[\boldsymbol{C}_{1-3} \oslash (1 - \boldsymbol{\lambda}_{1-3})\right]\left(\boldsymbol{\lambda}_{1-3}^t\right)^T, \tag{A1.21}$$

$$R_t^m = C_8 + C_9\left(-\frac{1}{2}\right)^t + \boldsymbol{C}_{10-12}\left(\boldsymbol{\lambda}_{1-3}^t\right)^T, \tag{A1.22}$$

$$R_t^p = C_8 - 2C_9\left(-\frac{1}{2}\right)^t + \boldsymbol{C}_{10-12}\left(\boldsymbol{\lambda}_{1-3}^{t-1}\right)^T, \tag{A1.23}$$

where $\boldsymbol{C}_{1-3}$ is given by equation (A1.5), and

$$C_8 = R_0^X + \frac{2}{3}\left[\boldsymbol{C}_{1-3} \oslash (1 - \boldsymbol{\lambda}_{1-3})\right], \tag{A1.24}$$

$$C_9 = R_0^m - C_8 - \boldsymbol{C}_{10-12}\mathbf{1}^T, \tag{A1.25}$$

$$\boldsymbol{C}_{10-12} = -2(\boldsymbol{C}_{1-3} \odot \boldsymbol{\lambda}_{1-3}) \oslash \left[(1 - \boldsymbol{\lambda}_{1-3}) \odot (2\boldsymbol{\lambda}_{1-3} + 1)\right]. \tag{A1.26}$$

Proof. According to equation (7a), R_t^X can be obtained from the accumulative summation of the non-IBD probability β_t^{mp} of equation (A1.6b), and we have

$$R_t^X = R_0^X + \frac{2}{3}\sum_{k=1}^{3} C_k \frac{1 - (\lambda_k)^t}{1 - \lambda_k}, \tag{A1.27}$$

which is equivalent to equation (A1.21) with the stationary map expansion C_8 being given by equation (A1.24). The eigenvalues for the transition matrix of the linear recurrence equations (4a–4c) and equations (6a, 6b) are 1, $-1/2$, $\boldsymbol{\lambda}_{1-3}$, and thus the map expansions R^m and R^p can be expressed in the forms of equations (A1.22, A1.23), where the constant coefficients $\boldsymbol{C}_{10-12}$ are determined by calculating $R_t^X = 2R_t^m/3 + R_t^p/3$ from equations (A1.22, A1.23) and comparing the result with equation (A1.21), and C_9 is determined by the initial condition R_0^m. $\square$

Theorem A1.9. *Denote by $\boldsymbol{K}_t(1122) = (K_t^{mm}(1122), K_t^{mp}(1122), K_t^{pp}(1122))^T$ the expected density of junction type (1122) in a finite population, and it holds*

$$\boldsymbol{K}_t(1122) = C_8 + (1, -1/2, -2)^T C_9\left(-\frac{1}{2}\right)^t + \boldsymbol{W}\boldsymbol{\Lambda}\left(\boldsymbol{\lambda}_{1-3}^t\right)(\boldsymbol{C}_{18-20})^T$$

$$+ \boldsymbol{P}\boldsymbol{\Lambda}\left(\boldsymbol{\lambda}_{1-3}^t\right)\left[(\boldsymbol{C}_{15-17})^T + (\boldsymbol{C}_{18-20})^T t\right] \tag{A1.28}$$

where $\boldsymbol{C}_{8-12}$ are given by equations (A1.24–A1.26), $\boldsymbol{P}$ is given by equation (A1.4),

$$\boldsymbol{W} = \left[2\boldsymbol{\lambda}_{1-3}, \boldsymbol{0}, \left(s^p\boldsymbol{\Phi}_{1-3}^{-1} + (1 - s^p)\right) \oslash \boldsymbol{\lambda}_{1-3}\right], \tag{A1.29}$$

$$(\boldsymbol{C}_{15-17})^T = \boldsymbol{P}^{-1}\left[\boldsymbol{K}_0(1122) - C_8 - (1, -1/2, -2)^T\right.$$

$$\left.\times C_9 - \boldsymbol{W}(\boldsymbol{C}_{18-20})^T\right], \tag{A1.30}$$

$$\boldsymbol{C}_{18-20} = \boldsymbol{\Phi}_{1-3} \odot \boldsymbol{C}_{10-12}, \tag{A1.31}$$

and for $k \in \{1, 2, 3\}$

$$\Phi_k = \frac{s^m(\lambda_k + 1) + (1 - s^m)s^p}{(6 - 2s^m)\lambda_k^2 + (6 - 2s^m - 4s^p + 4s^m s^p)\lambda_k - 3(1 - s^m)(1 - s^p)}. \tag{A1.32}$$

Proof. The eigenvalues for the transition matrix of the linear recurrence equations (4a–4c), equations (6a, 6b), and equations (9a–9c) are 1, $-1/2$, and duplicated $\boldsymbol{\lambda}_{1-3}$. It holds

$$K_t^{mm}(1122) = B_{13} + B_{14}\left(-\frac{1}{2}\right)^t + \boldsymbol{B}_{15-17}\left(\boldsymbol{\lambda}_{1-3}^t\right)^T + \boldsymbol{B}_{18-20}t\left(\boldsymbol{\lambda}_{1-3}^t\right)^T, \tag{A1.33a}$$

$$K_t^{mp}(1122) = C_{13} + C_{14}\left(-\frac{1}{2}\right)^t + \boldsymbol{C}_{15-17}\left(\boldsymbol{\lambda}_{1-3}^t\right)^T + \boldsymbol{C}_{18-20}t\left(\boldsymbol{\lambda}_{1-3}^t\right)^T, \tag{A1.33b}$$

$$K_t^{pp}(1122) = D_{13} + D_{14}\left(-\frac{1}{2}\right)^t + \boldsymbol{D}_{15-17}\left(\boldsymbol{\lambda}_{1-3}^t\right)^T + \boldsymbol{D}_{18-20}t\left(\boldsymbol{\lambda}_{1-3}^t\right)^T, \tag{A1.33c}$$

where the constant coefficients $\boldsymbol{B}_{13-20}$, $\boldsymbol{C}_{13-20}$, and $\boldsymbol{D}_{13-20}$ are to be solved. Substituting $K_t^{mm}(1122)$ and $K_t^{mp}(1122)$ of equations (A1.33a, A1.33b) into the recurrence equation (9b), we obtain

$$\begin{aligned}
B_{13} &= C_{13}, \\
B_{14} &= -2C_{14}, \\
B_{14+k} &= (2\lambda_k - 1)C_{14+k} + 2\lambda_k C_{17+k}, k \in \{1, 2, 3\}, \\
B_{17+k} &= (2\lambda_k - 1)C_{17+k}, k \in \{1, 2, 3\}.
\end{aligned} \tag{A1.34}$$

Substituting $K_t^{mm}(1122)$ and $K_t^{pp}(1122)$ of equations (A1.33a, A1.33c) and R_t^m of equation (A1.22) into the recurrence equation (9c), and substituting $\boldsymbol{B}_{13-20}$ of equation (A1.34), we obtain

$$\begin{aligned}
D_{13} &= C_{13} + s^p(C_8 - C_{13}), \\
D_{14} &= -2s^p C_9 + 4(1 - s^p)C_{14}, \\
D_{14+k} &= (1 - s^p)\frac{2\lambda_k - 1}{\lambda_k}C_{14+k} + s^p\frac{C_{9+k}}{\lambda_k} + (1 - s^p)\frac{C_{17+k}}{\lambda_k}, k \in \{1, 2, 3\}, \\
D_{17+k} &= (1 - s^p)\frac{2\lambda_k - 1}{\lambda_k}C_{17+k}, k \in \{1, 2, 3\},
\end{aligned} \tag{A1.35}$$

Substituting $K_t^{mp}(1122)$, $K_t^{mm}(1122)$ and $K_t^{pp}(1122)$ of equations (A1.33a–A1.33c) and R_t^m and R_t^p of equations (A1.22, A1.23) into the recurrence equation (9a), and substituting $\boldsymbol{B}_{13-20}$ and $\boldsymbol{D}_{13-20}$ of equations (A1.34, A1.35), we obtain $\boldsymbol{C}_{18-20}$ in equation (A1.31) and

$$C_{13} = C_8, \tag{A1.36a}$$

$$C_{14} = -\frac{1}{2}C_9. \tag{A1.36b}$$

The constant coefficient $\boldsymbol{C}_{15-17}$ is determined by the initial condition $\boldsymbol{K}_0(1122)$.

    Definition A1.10. Denote by

$$K_t^{mp+}(1232) = \frac{K_t^{mp}(1232) + K_t^{pm}(1232)}{2}, \tag{A1.37}$$

$$K_t^{mp-}(1232) = \frac{K_t^{mp}(1232) - K_t^{pm}(1232)}{2}, \tag{A1.38}$$

and thus

$$K_t^{mp}(1232) = K_t^{mp+}(1232) + K_t^{mp-}(1232), \tag{A1.39}$$

$$K_t^{pm}(1232) = K_t^{mp+}(1232) - K_t^{mp-}(1232). \tag{A1.40}$$

Proposition A1.11. *According to Definition A1.10, the recurrence equations (10a–10d) in a finite population are transformed into*

$$K_t^{mm}(1232) = (1-s^m)\frac{1}{4}K_{t-1}^{mm}(1232) + \frac{1}{2}K_{t-1}^{mp+}(1232) + (1-s^m)\frac{1}{4}K_{t-1}^{pp}(1232)$$
$$+ (1-s^m)\frac{1}{2}\left[\beta_{t-1}^{mmp}(123) + \beta_{t-1}^{mpp}(123)\right], \tag{A1.41a}$$

$$K_t^{mp+}(1232) = \frac{1}{2}K_{t-1}^{mm}(1232) + \frac{1}{2}K_{t-1}^{mp+}(1232) + \frac{1}{2}\beta_{t-1}^{mmp}(123), \tag{A1.41b}$$

$$K_t^{pp}(1232) = (1-s^p)K_{t-1}^{mm}(1232), \tag{A1.41c}$$

$$K_t^{mp-}(1232) = -\frac{1}{2}K_{t-1}^{mp-}(1232) + \frac{1}{2}\beta_{t-1}^{mmp}(123). \tag{A1.41d}$$

Theorem A1.12. *Denote by* $\mathbf{K}_t(1232) = (K_t^{mm}(1232), K_t^{mp+}(1232), K_t^{pp}(1232))^T$ *the expected density of junction type (1232) in a finite population, and it holds*

$$\mathbf{K}_t(1232) = \mathbf{P}\boldsymbol{\Lambda}(\boldsymbol{\lambda}_{1-3}^t)(\mathbf{C}_{21-23})^T + \mathbf{W}\boldsymbol{\Lambda}(\boldsymbol{\lambda}_{4-7}^t)(\mathbf{C}_{24-27})^T, \tag{A1.42}$$

where $\mathbf{P}$ *is given by equation (A1.4),*

$$\mathbf{W} = \left[-\boldsymbol{\Phi}_{4-7}^{-1} + (2\boldsymbol{\lambda}_{4-7} - 1), \mathbf{1}, (1-s^p)\left(-\boldsymbol{\Phi}_{4-7}^{-1} + (2\boldsymbol{\lambda}_{4-7} - 1)\right) \oslash \boldsymbol{\lambda}_{4-7}\right], \tag{A1.43}$$

$$(\mathbf{C}_{21-23})^T = \mathbf{P}^{-1}\left[\mathbf{K}_0(1232) - \mathbf{W}(\mathbf{C}_{24-27})^T\right], \tag{A1.44}$$

$$\mathbf{C}_{24-27} = \boldsymbol{\Phi}_{4-7} \odot \mathbf{C}_{4-7}, \tag{A1.45}$$

and for $k \in \{4,5,6,7\}$

$$\Phi_k = \frac{\frac{1}{2}(\lambda_k)^2 + \frac{1}{8}[(1-s^p)(a_k+1)/\lambda_k + 1](1-s^m)\lambda_k - \frac{1}{8}(1-s^m)(1-s^p)}{f_{12}(\lambda_k)}, \tag{A1.46}$$

where $f_{12}(x)$ *is the characteristic polynomial of the transition matrix* $\mathbf{T}_{12}$ *of equation (A1.2).*

Proof. The eigenvalues for the transition matrix of the linear recurrence equations (5a–5d) and equations (A1.41a–A1.41c) are λ_k, $1 \leq k \leq 7$. It holds

$$K_t^{mm}(1232) = \mathbf{B}_{21-27}(\boldsymbol{\lambda}_{1-7}^t)^T, \tag{A1.47a}$$

$$K_t^{mp+}(1232) = \mathbf{C}_{21-27}(\boldsymbol{\lambda}_{1-7}^t)^T, \tag{A1.47b}$$

$$K_t^{pp}(1232) = \mathbf{D}_{21-27}(\boldsymbol{\lambda}_{1-7}^t)^T, \tag{A1.47c}$$

where the constant coefficients $\mathbf{B}_{21-27}$, $\mathbf{C}_{21-27}$, and $\mathbf{D}_{21-27}$ are to be solved. Substituting $K_t^{mm}(1232)$ and $K_t^{pp}(1232)$ of equations (A1.47a, A1.47c) into the recurrence equation (A1.41c), we obtain

$$D_{20+k} = (1-s^p)\frac{B_{20+k}}{\lambda_k}, 1 \leq k \leq 7, \tag{A1.48}$$

Substituting $K_t^{mp+}(1232)$ and $K_t^{mm}(1232)$ of equations (A1.47a, A1.47b) and β_t^{mmp} of equation (A1.16b) into the recurrence equation (A1.41b), we obtain

$$B_{20+k} = \begin{cases} (2\lambda_k - 1)C_{20+k} & k \in \{1,2,3\}, \\ -C_k + (2\lambda_k - 1)C_{20+k} & k \in \{4,5,6,7\}. \end{cases} \tag{A1.49}$$

The theorem follows by substituting $K_t^{mm}(1232)$, $K_t^{mp+}(1232)$ and $K_t^{pp}(1232)$ of equations (A1.47a–A1.47c) and β_t^{mmp} and β_t^{mpp} of equations (A1.16b, A1.16c) into the recurrence equation (A1.41a), and substituting $\mathbf{B}_{21-27}$ and $\mathbf{D}_{21-27}$ of equations (A1.48, A1.49), where the constant coefficient $\mathbf{C}_{21-23}$ in equation (A1.44) is determined by the initial condition $\mathbf{K}_0(1232)$. $\square$

Theorem A1.13. *The expected density* $K_t^{mp-}(1232)$ *is given by*

$$K_t^{mp-}(1232) = A_{28}\left(-\frac{1}{2}\right)^t + \boldsymbol{A}_{24-27}\left(\boldsymbol{\lambda}_{4-7}^t\right)^T, \tag{A1.50}$$

where

$$\boldsymbol{A}_{24-27} = \boldsymbol{C}_{4-7} \oslash (2\boldsymbol{\lambda}_{4-7}+1), \tag{A1.51}$$

$$A_{28} = K_0^{mp-}(1232) - \boldsymbol{A}_{24-27}\mathbf{1}^T. \tag{A1.52}$$

Proof. The eigenvalues of the transition matrix of the linear recurrence equations (5a–5d) and equation (A1.41d): $-1/2$, λ_k, $k \in \{4,5,6,7\}$. Thus equation (A1.50) holds, where $\boldsymbol{A}_{24-27}$ is determined by putting equations (A1.50, A1.16b) into equation (A1.41d), and the constant coefficient A_{28} is determined by the initial condition $K_0^{mp-}(1232)$. □

Corollary A1.14. *The overall expected junction density ρ_t^{mp} of a female in a finite population is given by*

$$\rho_t^{mp} = C_8 - \frac{1}{2}C_9\left(-\frac{1}{2}\right)^t + \left[\boldsymbol{C}_{10-12}\odot\left(1+\boldsymbol{\lambda}_{1-3}^{-1}\right) - \boldsymbol{C}_{15-17} - \boldsymbol{C}_{18-20}t\right]\left(\boldsymbol{\lambda}_{1-3}^t\right)^T, \tag{A1.53}$$

where $\boldsymbol{C}_{8-12}$ are given by equations (A1.24–A1.26), and $\boldsymbol{C}_{15-20}$ by equations (A1.30, A1.31).

Proof. The corollary follows by substituting R_t^m and R_t^p of equations (A1.22, A1.23) and $K_t^{mp}(1122)$ of equation (A1.33b) into equation (3).

A2 Sibling-mating population

Definition A2.1. The sibling-mating population refers to a population of constant size 2, one male and one female, maintained by sibling mating, and the initial population satisfies $\alpha_0^{mp}(12) = \beta_0^{mp}(12)$. In such a population, the coalescence probability $s^m = 1$, and the coalescence probabilities s^p, q^m, q^p are set to zero.

Definition A2.2. In a sibling-mating population, the non-IBD probabilities $\beta_t^{pp}(12)$, $\beta_t^{mmm}(123)$, $\beta_t^{mmp}(123)$, $\beta_t^{mpp}(123)$, and $\beta_t^{ppp}(123)$ are set to zero since they do not exist. Similarly the expected junction densities $K^{pp}(1122)$ and $K^{pp}(1232)$ are set to zero.

Proposition A2.3. Denote by $\boldsymbol{\beta}_t(12) = (\beta_t^{mm}(12), \beta_t^{mp}(12))$ the two-gene non-IBD probability in a sibling-mating population. According to equations (4a, 4b), it holds

$$T_{12} = \begin{pmatrix} 0 & \dfrac{1}{2} \\ \dfrac{1}{2} & \dfrac{1}{2} \end{pmatrix}, \tag{A2.1}$$

and its eigenvalues are given by

$$\lambda_1 = \frac{1+\sqrt{5}}{4}, \lambda_2 = \frac{1-\sqrt{5}}{4}. \tag{A2.2}$$

Definition A2.4. In a sibling-mating population, the conjugate is obtained by replacing $\sqrt{5}$ with $-\sqrt{5}$ from the terms involving λ_1.

Theorem A2.5. *The two-gene non-IBD probability $\boldsymbol{\beta}_t(12)$ in a sibling-mating population is given by*

$$\beta_t^{mm}(12) = \left[\frac{5-\sqrt{5}}{10}\beta_0^{mm}(12) + \frac{\sqrt{5}}{5}\beta_0^{mp}(12)\right](\lambda_1)^t + conjugate \tag{A2.3}$$

$$\beta_t^{mp}(12) = \left[\frac{\sqrt{5}}{5}\beta_0^{mm}(12) + \frac{5+\sqrt{5}}{10}\beta_0^{mp}(12)\right](\lambda_1)^t + conjugate \tag{A2.4}$$

Proof. Similar to Theorem A1.4, it holds

$$\boldsymbol{\beta}_t(12) = P\Lambda\left(\boldsymbol{\lambda}_{1-2}^t\right)(\boldsymbol{C}_{1-2})^T, \tag{A2.5}$$

where

$$P = [2\boldsymbol{\lambda}_{1-2}-1, \mathbf{1}], \tag{A2.6}$$

$$(\boldsymbol{C}_{1-2})^T = P^{-1}\boldsymbol{\beta}_0(12). \tag{A2.7}$$

The theorem follows by substituting P an C_{1-2} into equation (A2.5). □

Theorem A2.6. *The three-gene non-IBD probability in a sibling-mating population is given by*

$$\alpha_t^{mmp}(123) = \alpha_0^{mmp}(123)\left(\frac{1}{2}\right)^t. \tag{A2.8}$$

Proof. According to the recurrence equations (5b, 5e),

$$\alpha_t^{mmp}(123) = \frac{1}{2}\alpha_{t-1}^{mmp}(123). \tag{A2.9}$$

Thus it holds

$$\alpha_t^{mmp}(123) = C_4(\lambda_4)^t, \tag{A2.10}$$

where

$$\lambda_4 = 1/2, \tag{A2.11}$$

$$C_4 = \alpha_0^{mmp}(123). \tag{A2.12}$$

Theorem A2.7. *The map expansions in a sibling-mating population are given by*

$$R_t^X = C_8 - \left[\frac{10 + 6\sqrt{5}}{15}\beta_0^{mm}(12) + \frac{20 + 8\sqrt{5}}{15}\beta_0^{mp}(12)\right](\lambda_1)^t + conjugate, \tag{A2.13}$$

$$R_t^m = C_8 + C_9\left(-\frac{1}{2}\right)^t - \left[\frac{5 + \sqrt{5}}{5}\beta_0^{mm}(12) + \frac{5 + 3\sqrt{5}}{5}\beta_0^{mp}(12)\right](\lambda_1)^t + conjugate \tag{A2.14}$$

$$R_t^p = C_8 - 2C_9\left(-\frac{1}{2}\right)^t - \left[\frac{4\sqrt{5}}{5}\beta_0^{mm}(12) + \frac{10 + 2\sqrt{5}}{5}\beta_0^{mp}(12)\right](\lambda_1)^t + conjugate \tag{A2.15}$$

where C_8 and C_9 are given by

$$C_8 = \frac{1}{3}\left(2R_0^m + R_0^p\right) + \frac{4}{3}\left(\beta_0^{mm}(12) + 2\beta_0^{mp}(12)\right), \tag{A2.16}$$

$$C_9 = \frac{1}{3}\left(R_0^m - R_0^p\right) + \frac{2}{3}\left(\beta_0^{mm}(12) - \beta_0^{mp}(12)\right), \tag{A2.17}$$

Proof. Similar to Theorem A1.8, the map expansions in a sibling-mating population are given by

$$R_t^X = C_8 - \frac{2}{3}\left[\boldsymbol{C}_{1-2} \oslash (1 - \boldsymbol{\lambda}_{1-2})\right]\left(\boldsymbol{\lambda}_{1-2}^t\right)^T, \tag{A2.18}$$

$$R_t^m = C_8 + C_9\left(-\frac{1}{2}\right)^t + \boldsymbol{C}_{10-11}\left(\boldsymbol{\lambda}_{1-2}^t\right)^T, \tag{A2.19}$$

$$R_t^p = C_8 - 2C_9\left(-\frac{1}{2}\right)^t + \boldsymbol{C}_{10-11}\left(\boldsymbol{\lambda}_{1-2}^{t-1}\right)^T, \tag{A2.20}$$

where

$$C_8 = R_0^X + \frac{2}{3}\left[\boldsymbol{C}_{1-2} \oslash (1 - \boldsymbol{\lambda}_{1-2})\right], \tag{A2.21}$$

$$C_9 = R_0^m - C_8 - \boldsymbol{C}_{10-11}\mathbf{1}^T, \tag{A2.22}$$

$$\boldsymbol{C}_{10-11} = -2(\boldsymbol{C}_{1-2} \odot \boldsymbol{\lambda}_{1-2}) \oslash \left[(1 - \boldsymbol{\lambda}_{1-2}) \odot (2\boldsymbol{\lambda}_{1-2} + 1)\right]. \tag{A2.23}$$

The theorem follows by substituting $\boldsymbol{\lambda}_{1-2}$ of equation (A2.2) and $\boldsymbol{C}_{1-2}$ of equation (A2.7) into the aforementioned equations.

Theorem A2.8. *The expected density $\boldsymbol{K}_t(1122) = (K_t^{mm}(1122), K_t^{mp}(1122))^T$ in a sibling-mating population is given by*

$$K_t^{mm}(1122) = C_8 + C_9\left(-\frac{1}{2}\right)^t$$

$$+ \left[-\frac{5+\sqrt{5}}{5}\beta_0^{mm}(12) - \frac{5+4\sqrt{5}}{5}\beta_0^{mp}(12)\right](\lambda_1)^t$$

$$+ \left[\frac{5-\sqrt{5}}{10}K_0^{mm}(1122) + \frac{\sqrt{5}}{5}K_0^{mp}(1122) - \frac{1}{2}R_0^m - \frac{\sqrt{5}}{10}R_0^p\right](\lambda_1)^t \tag{A2.24}$$

$$+ \left[-\frac{5-\sqrt{5}}{10}\beta_0^{mm}(12) - \frac{\sqrt{5}}{5}\beta_0^{mp}(12)\right]t(\lambda_1)^t + conjugate$$

$$K_t^{mp}(1122) = C_8 - \frac{1}{2}C_9\left(-\frac{1}{2}\right)^t$$

$$+ \left[-\frac{5+3\sqrt{5}}{10}\beta_0^{mm}(12) - \frac{3+\sqrt{5}}{2}\beta_0^{mp}(12)\right](\lambda_1)^t$$

$$+ \left[\frac{\sqrt{5}}{5}K_0^{mm}(1122) + \frac{5+\sqrt{5}}{10}K_0^{mp}(1122) - \frac{1+\sqrt{5}}{4}R_0^m - \frac{5+\sqrt{5}}{20}R_0^p\right](\lambda_1)^t \tag{A2.25}$$

$$+ \left[-\frac{\sqrt{5}}{5}\beta_0^{mm}(12) - \frac{5+\sqrt{5}}{10}\beta_0^{mp}(12)\right]t(\lambda_1)^t + conjugate$$

where C_8 and C_9 are given by equations (A2.16, A2.17), respectively.

Proof. Similar to Theorem A1.9, the expected density $K_t(1122)$ in a sibling-mating population is given by

$$K_t(1122) = C_8 + (1, -1/2)^T C_9\left(-\frac{1}{2}\right)^t + W\Lambda\left(\lambda_{1-2}^t\right)(C_{18-19})^T$$

$$+ P\Lambda\left(\lambda_{1-2}^t\right)\left[(C_{15-16})^T + (C_{18-19})^T t\right] \tag{A2.26}$$

where C_8 and C_9 are given by equations (A2.16, A2.17), respectively, and P is given by equation (A2.6), and

$$W = [2\lambda_{1-2}, 0], \tag{A2.27}$$

$$(C_{15-16})^T = P^{-1}\left[K_0(1122) - C_8 - (1, -1/2)^T C_9 - W(C_{18-19})^T\right], \tag{A2.28}$$

$$C_{18-19} = \Phi_{1-2} \odot C_{10-11}, \tag{A2.29}$$

and for $k \in \{1, 2\}$

$$\Phi_k = \frac{1}{4\lambda_k}. \tag{A2.30}$$

The theorem follows by substituting λ_{1-2} of equation (A2.2) and C_{10-11} of equation (A2.23).

Theorem A2.9. *The expected densities of type (1232) in a sibling-mating population are given by*

$$K_t^{mm}(1232) = \left[\frac{5-\sqrt{5}}{10}K_0^{mm}(1232) + \frac{\sqrt{5}}{5}K_0^{mp+}(1232) + \frac{5+\sqrt{5}}{10}\alpha_0^{mmp}(123)\right](\lambda_1)^t$$

$$- \alpha_0^{mmp}(123)\left(\frac{1}{2}\right)^t + conjutate \tag{A2.31}$$

$$K_t^{mp+}(1232) = \left[\frac{\sqrt{5}}{5}K_0^{mm}(1232) + \frac{5+\sqrt{5}}{10}K_0^{mp+}(1232) + \frac{5+3\sqrt{5}}{10}\alpha_0^{mmp}(123)\right](\lambda_1)^t$$

$$- \alpha_0^{mmp}(123)\left(\frac{1}{2}\right)^t + conjutate, \tag{A2.32}$$

$$K_t^{mp-}(1232) = \left[K_0^{mp-}(1232) - \frac{1}{2}\alpha_0^{mmp}(123)\right]\left(-\frac{1}{2}\right)^t + \frac{1}{2}\alpha_0^{mmp}(123)\left(\frac{1}{2}\right)^t \tag{A2.33}$$

Proof. Denote by $\boldsymbol{K}_t(1232) = (K_t^{mm}(1232), K_t^{mp+}(1232))^T$. Similar to Theorem A1.12, the expected density $\boldsymbol{K}_t(1232)$ in a sibling-mating population is given by

$$\boldsymbol{K}_t(1232) = \boldsymbol{P}\boldsymbol{\Lambda}(\boldsymbol{\lambda}_{1-2}^t)(\boldsymbol{C}_{21-22})^T + \boldsymbol{W}\boldsymbol{\Lambda}(\boldsymbol{\lambda}_{4-4}^t)(\boldsymbol{C}_{24-24})^T, \tag{A2.34}$$

where $\boldsymbol{P}$ is given by equation (A2.6), and

$$\boldsymbol{W} = \left[-\boldsymbol{\Phi}_{4-4}^{-1} + (2\boldsymbol{\lambda}_{4-4} - 1), \mathbf{1}\right], \tag{A2.35}$$

$$(\boldsymbol{C}_{21-22})^T = \boldsymbol{P}^{-1}\left[\boldsymbol{K}_0(1232) - \boldsymbol{W}(\boldsymbol{C}_{24-24})^T\right], \tag{A2.36}$$

$$\boldsymbol{C}_{24-24} = \boldsymbol{\Phi}_{4-4} \odot \boldsymbol{C}_{4-4}, \tag{A2.37}$$

and for $k = 4$

$$\Phi_k = \frac{\frac{1}{2}\lambda_k}{f_{12}(\lambda_k)}, \tag{A2.38}$$

where $f_{12}(x)$ is the characteristic polynomial of the transition matrix $\boldsymbol{T}_{12}$ of equation (A2.1). The equations (A2.31, A2.32) follow by substituting $\boldsymbol{\lambda}_{1-2}$ of equation (A2.2), λ_4 of equation (A2.11), and C_4 of equation (A2.12).

Similar to Theorem A1.13, the expected density $K_t^{mp-}(1232)$ in a sibling-mating population is given by

$$K_t^{mp-}(1232) = A_{28}\left(-\frac{1}{2}\right)^t + \boldsymbol{A}_{24-24}(\boldsymbol{\lambda}_{4-4}^t)^T \tag{A2.39}$$

where

$$\boldsymbol{A}_{24-24} = \boldsymbol{C}_{4-4} \oslash (2\boldsymbol{\lambda}_{4-4} + 1), \tag{A2.40}$$

$$A_{28} = K_0^{mp-}(1232) - \boldsymbol{A}_{24-24}. \tag{A2.41}$$

Corollary A2.10. *The overall expected junction density ρ_t^{mp} in a sibling-mating population is given by*

$$\begin{aligned}
\rho_t^{mp} = C_8 &- \frac{1}{2}C_9\left(-\frac{1}{2}\right)^t \\
&+ \left[-\frac{5 + 7\sqrt{5}}{10}\beta_0^{mm}(12) - \frac{3 + \sqrt{5}}{2}\beta_0^{mp}(12)\right](\lambda_1)^t \\
&+ \left[-\frac{\sqrt{5}}{5}K_0^{mm}(1122) - \frac{5 + \sqrt{5}}{10}K_0^{mp}(1122)\right. \\
&+ \left.\left[\frac{\sqrt{5}}{5}\beta_0^{mm}(12) + \frac{5 + \sqrt{5}}{10}\beta_0^{mp}(12)\right]t(\lambda_1)^t + conjugate,
\end{aligned} \tag{A2.42}$$

where C_8 and C_9 are given by equations (A2.16, A2.17), respectively.

Proof. The corollary follows by substituting R_t^m and R_t^p of equations (A2.14, A2.15) and $K_t^{mp}(1122)$ of equation (A2.25) into equation (3).

A3 Large population

Definition A3.1. The large population refers to a population with large and equal number $N_m = N_f \gg 3$ of males and females, maintained by random mating, and the initial population satisfies $\alpha_0^{mp}(12) = \beta_0^{mp}(12)$, $\alpha_0^{mmp}(123) = \beta_0^{mmp}(123)$, and $\alpha_0^{mpp}(123) = \beta_0^{mpp}(123)$. In such a large population, the coalescence probabilities are set to $s^m = s^q = q^m = q^p = s$.

Proposition A3.2. *Denote by $\boldsymbol{\beta}_t(12) = (\beta_t^{mm}(12), \beta_t^{mp}(12), \beta_t^{pp}(12))$ the two-gene non-IBD probability in a large population, and the eigenvalues of the transition matrix $\boldsymbol{T}_{12}$ in equation (A1.2) are given by*

$$\lambda_1 = 1 - \frac{s}{3}, \ \lambda_2 = -\frac{1}{2}, \lambda_3 = \frac{1}{4}. \tag{A3.1}$$

where λ_1 is approximated to the first order of s to keep it smaller than 1, and λ_2 and λ_3 are approximated to zero order of s.

Theorem A3.3. *The two-gene non-IBD probabilities in a large population are given by*

$$\boldsymbol{\beta}_t(12) = \boldsymbol{P}\boldsymbol{\Lambda}\big(\boldsymbol{\lambda}_{1-3}^t\big)\big(\boldsymbol{C}_{1-3}\big)^T, \tag{A3.2}$$

where

$$\boldsymbol{P} = \begin{bmatrix} 1 & -2 & -1/2 \\ 1 & 1 & 1 \\ 1 & 4 & -2 \end{bmatrix}, \tag{A3.3}$$

$$\big(\boldsymbol{C}_{1-3}\big)^T = \frac{1}{9}\begin{bmatrix} 4 & 4 & 1 \\ -2 & 1 & 1 \\ -2 & 4 & -2 \end{bmatrix}\boldsymbol{\beta}_0(12). \tag{A3.4}$$

$$\tag{A3.5}$$

Proof. The theorem follows directly from Theorem A1.4, where

$$\boldsymbol{P} = \big[2\boldsymbol{\lambda}_{1-3} - \boldsymbol{1}, \boldsymbol{1}, (1-s)(2\boldsymbol{\lambda}_{1-3} - 1) \oslash \boldsymbol{\lambda}_{1-3}\big], \tag{A3.6}$$

$$\big(\boldsymbol{C}_{1-3}\big)^T = \boldsymbol{P}^{-1}\boldsymbol{\beta}_0(12). \tag{A3.7}$$

The matrix $\boldsymbol{P}$ is approximated to the zero order of s. $\square$

Proposition A3.4. *Denote by* $\boldsymbol{\beta}_t(123) = (\beta_t^{mmm}(123), \beta_t^{mmp}(123), \beta_t^{mpp}(123), \beta_t^{ppp}(123))^T$ *the three-gene non-IBD probability in a large population, and the eigenvalues of the transition matrix* $\boldsymbol{T}_{123}$ *of equation (A1.11) are given by*

$$\lambda_4 = 1 - s, \lambda_5 = -\frac{1}{2}, \lambda_6 = \frac{1}{4}, \lambda_7 = -\frac{1}{8}, \tag{A3.8}$$

where λ_4 *is approximated to the first order of s to keep it smaller than 1, and* $\boldsymbol{\lambda}_{5-7}$ *is approximated to zero order of s.*

Theorem A3.5. *The three-gene non-IBD probability* $\boldsymbol{\beta}_t(123)$ *in a large population is given by*

$$\boldsymbol{\beta}_t(123) = \boldsymbol{Q}\boldsymbol{\Lambda}\big(\boldsymbol{\lambda}_{4-7}^t\big)\big(\boldsymbol{C}_{4-7}\big)^T, \tag{A3.9}$$

where

$$\boldsymbol{Q} = \begin{bmatrix} 1 & -\dfrac{3}{s} & -1 & -\dfrac{1}{2} \\ 1 & 1 & 1 & 1 \\ 1 & \dfrac{3}{s} & 0 & -2 \\ 1 & \dfrac{6}{s} & -4 & 4 \end{bmatrix}, \tag{A3.10}$$

$$\big(\boldsymbol{C}_{4-7}\big)^T = \boldsymbol{Q}^{-1}\boldsymbol{\beta}_0(123). \tag{A3.11}$$

Proof. The theorem follows directly from Theorem A1.7, where

$$a_k = \frac{8(\lambda_k)^2 - 4\lambda_k - (1-s)^2}{(1-s)(1-s+2\lambda_k)}, k \in \{4,5,6,7\},$$

$$\boldsymbol{Q} = \big[\boldsymbol{a}_{4-7}, \boldsymbol{1}, (1-s)(\boldsymbol{a}_{4-7} + 1) \oslash (2\boldsymbol{\lambda}_{4-7}), (1-3s)\boldsymbol{a}_{4-7}\oslash\boldsymbol{\lambda}_{4-7}\big],$$

where the elements of $\boldsymbol{Q}$ are approximated to be the leading order of s and up to the zero order of s. $\square$

Theorem A3.6. *The map expansions in a large population are given by*

$$R_t^X = R_0^X + \frac{2}{3}C_1\frac{\big[1 - (\lambda_1)^t\big]}{1 - \lambda_1} + \frac{4}{9}C_2\big[1 - (\lambda_2)^t\big] + \frac{8}{9}C_3\big[1 - (\lambda_3)^t\big], \tag{A3.12}$$

$$R_t^m = C_8 + C_{10}\frac{1 - (\lambda_1)^t}{1 - \lambda_1} + (C_9 + C_{11}t)(\lambda_2)^t + C_{12}(\lambda_3)^t, \tag{A3.13}$$

$$R_t^p = C_8 + C_{10}\frac{1 - (\lambda_1)^{t-1}}{1 - \lambda_1} + [C_9 + C_{11}(t-1)](\lambda_2)^{t-1} + C_{12}(\lambda_3)^{t-1}, \tag{A3.14}$$

where C_{1-3} is given by equation (A3.4), and

$$C_8 = R_0^X + \frac{2}{3(2\lambda_1 + 1)}C_1 + \frac{4}{9}C_2 + \frac{8}{9}C_3, \tag{A3.15}$$

$$C_9 = R_0^m - C_8 - C_{12}, \tag{A3.16}$$

$$C_{10} = \frac{2\lambda_1}{2\lambda_1 + 1}C_1, \quad C_{11} = -\frac{2}{3}C_2, \quad C_{12} = -\frac{4}{9}C_3. \tag{A3.17}$$

Remark. As $s \to 0$ given t, $[1 - (\lambda_1)^t]/(1 - \lambda_1) \to t$, that is, the map expansions are dominant by the linear term proportional to t; as $t \to \infty$ given s, the map expansions asymptotically go to the constant $C_8 + C_{10}/(1 - \lambda_1)$.

Proof. The proof is similar to Theorem A1.8, except that the eigenvalue λ_2 or $-1/2$ has multiplicity 2. □

Theorem A3.7. *The expected density* $\boldsymbol{K}_t(1122) = (K_t^{mm}(1122), K_t^{mp}(1122), K_t^{pp}(1122))$, *is given by*

$$\boldsymbol{K}_t(1122) = C_8 + C_{10}\frac{1 - \lambda_1^t}{1 - \lambda_1} - \left[1 + \left(\frac{2}{9}, \frac{8}{9}, -\frac{4}{9}\right)s\right]C_{10}t\lambda_1^t + \boldsymbol{P}\boldsymbol{\Lambda}(\boldsymbol{\lambda}_{1-3}^t)(\boldsymbol{C}_{15-17})^T, \tag{A3.18}$$

where C_8, C_{10}, *and* $\boldsymbol{P}$ *are given by equation (A3.15), equation (A3.17), and equation (A3.3), respectively, and*

$$\boldsymbol{C}_{15-17} = \boldsymbol{P}^{-1}[\boldsymbol{K}_0(1122) - C_8]. \tag{A3.19}$$

Proof. The eigenvalues for the transition matrix of the linear recurrence equations (4a–4c), equations (6a, 6b), and equations (9a–9c) are 1, triplicated $\lambda_2 = -1/2$, and duplicated λ_1 and λ_3, under the limit $s \to 0$. It holds

$$K_t^{mm}(1122) = B_{13} + \left(B_{15} - \frac{C_{10}}{1 - \lambda_1}\right)(\lambda_1)^t + \boldsymbol{B}_{16-17}(\boldsymbol{\lambda}_{2-3}^t)^T + \boldsymbol{B}_{18-20}t(\boldsymbol{\lambda}_{1-3}^t)^T + B_{14}t^2(\lambda_2)^t, \tag{A3.20a}$$

$$K_t^{mp}(1122) = C_{13} + \left(C_{15} - \frac{C_{10}}{1 - \lambda_1}\right)(\lambda_1)^t + \boldsymbol{C}_{16-17}(\boldsymbol{\lambda}_{2-3}^t)^T + \boldsymbol{C}_{18-20}t(\boldsymbol{\lambda}_{1-3}^t)^T + C_{14}t^2(\lambda_2)^t, \tag{A3.20b}$$

$$K_t^{pp}(1122) = D_{13} + \left(D_{15} - \frac{C_{10}}{1 - \lambda_1}\right)(\lambda_1)^t + \boldsymbol{D}_{16-17}(\boldsymbol{\lambda}_{2-3}^t)^T + \boldsymbol{D}_{18-20}t(\boldsymbol{\lambda}_{1-3}^t)^T + D_{14}t^2(\lambda_2)^t, \tag{A3.20c}$$

where C_{10} is given by equation (A3.17), and the constant coefficients $\boldsymbol{B}_{13-20}$, $\boldsymbol{C}_{13-20}$, and $\boldsymbol{D}_{13-20}$ are to be solved. Substituting $K_t^{mm}(1122)$ and $K_t^{mp}(1122)$ of equations (A3.20a, A3.20b) into the recurrence equation (9b), we obtain

$$\begin{aligned}
B_{13} &= C_{13}, \\
B_{14} &= (2\lambda_2 - 1)C_{14}, \\
\boldsymbol{B}_{15-17} &= (2\boldsymbol{\lambda}_{1-3} - 1)\odot\boldsymbol{C}_{15-17} + 2\boldsymbol{\lambda}_{1-3}\odot\boldsymbol{C}_{18-20} + (2C_{10}, 2\lambda_2 C_{14}, 0), \\
\boldsymbol{B}_{18-20} &= (2\boldsymbol{\lambda}_{1-3} - 1)\odot\boldsymbol{C}_{18-20} + (0, 4\lambda_2 C_{14}, 0).
\end{aligned} \tag{A3.21}$$

Substituting $K_t^{mm}(1122)$ and $K_t^{pp}(1122)$ of equations (A3.20a, A3.20c) and R_t^m of equation (A3.13) into the recurrence equation (9c), and substituting $\boldsymbol{B}_{13-20}$ of equation (A3.21), we obtain

$$D_{13} = C_{13} + s\left[\left(C_8 + \frac{C_{10}}{1 - \lambda_1} - C_{13}\right)\right],$$

$$D_{14} = \frac{(1 - s)(2\lambda_2 - 1)}{\lambda_2} C_{14},$$

$$\boldsymbol{D}_{15-17} = [(1 - s)(2\boldsymbol{\lambda}_{1-3} - 1) \oslash \boldsymbol{\lambda}_{1-3}] \odot \boldsymbol{C}_{15-17} + (1 - s)\boldsymbol{C}_{18-20} \oslash \boldsymbol{\lambda}_{1-3} \tag{A3.22}$$

$$+ \left(\frac{1 - 2s}{\lambda_1} C_{10}, \frac{s}{\lambda_2} C_9 - \frac{s}{\lambda_2} C_{11} - \frac{1 - s}{\lambda_2} C_{14}, \frac{s}{\lambda_3} C_{12}\right),$$

$$\boldsymbol{D}_{18-20} = [(1 - s)(2\boldsymbol{\lambda}_{1-3} - 1) \oslash \boldsymbol{\lambda}_{1-3}] \odot \boldsymbol{C}_{18-20} + \left(0, \frac{s}{\lambda_2} C_{11} + 2\frac{1 - s}{\lambda_2} C_{14}, 0\right).$$

Substituting $K_t^{mp}(1122)$, $K_t^{mm}(1122)$ and $K_t^{pp}(1122)$ of equations (A3.20a–A3.20c) and R_t^m of equation (A3.13) into the recurrence equation (9a), substituting $B_{13} - B_{20}$ of equation (A3.21), and substituting the exact $D_{13} - D_{20}$ of equation (A3.22), we obtain

$$C_{13} = C_8 + \frac{C_{10}}{1 - \lambda_1}, C_{18} = -\left(1 + \frac{8}{9}s\right)C_{10}, C_{14} = C_{19} = C_{20} = 0, \tag{A3.23}$$

where the constant coefficients C_{14}, C_{19}, and C_{20} are approximated by keeping up to the zero order of s, while C_{18} is approximated by keeping up to the first order of s. The theorem follows by substituting $\boldsymbol{B}_{13-20}$, $\boldsymbol{C}_{13-14}$, $\boldsymbol{C}_{18-20}$, and $\boldsymbol{D}_{13-20}$ of equations (A3.21–A3.23) into equations (A3.20a–A3.20c), where the constant coefficient $\boldsymbol{C}_{15-17}$ of equation (A3.19) is determined by the initial conditions $\boldsymbol{K}_0(1122)$. $\square$

Theorem A3.8. *Denote by* $\boldsymbol{K}_t(1232) = (K_t^{mm}(1232), K_t^{mp+}(1232), K_t^{pp}(1232))^T$ *the expected density of junction type* (1232) *in a large population, and it holds*

$$\boldsymbol{K}_t(1232) = C_4 \frac{\lambda_1^t - \lambda_4^t}{s} + \boldsymbol{P}\boldsymbol{\Lambda}\left(\boldsymbol{\lambda}_{1-3}^t\right)(\boldsymbol{C}_{21-23})^T + \boldsymbol{W}_t, \tag{A3.24}$$

where $\boldsymbol{P}$ *is given by equation (A3.3),* $\boldsymbol{C}_{4-7}$ *is given by equation (A3.11), and*

$$(\boldsymbol{C}_{21-23})^T = \boldsymbol{P}^{-1}[\boldsymbol{K}_0(1232) - \boldsymbol{W}_0], \tag{A3.25}$$

$$\boldsymbol{W}_t = \begin{bmatrix} \frac{1}{3} & -\frac{2}{3s} & -\frac{7}{9} & -\frac{4}{9} \\ 0 & 0 & 0 & -\frac{4}{9} \\ -\frac{1}{3} & -\frac{4}{3s} & -\frac{20}{9} & \frac{32}{9} \end{bmatrix} \left[\left(\lambda_1^t, \lambda_2^t, \lambda_3^t, \lambda_7^t\right) \odot \boldsymbol{C}_{4-7}\right]^T$$

$$+ \begin{bmatrix} -\frac{4}{3s} & -\frac{2}{9} \\ \frac{2}{3s} & \frac{4}{9} \\ \frac{8}{3s} & -\frac{8}{9} \end{bmatrix} \left[\left(t\lambda_2^t, t\lambda_3^t\right) \odot \boldsymbol{C}_{5-6}\right]^T. \tag{A3.26}$$

Remark. In equation (A3.24), $(\lambda_1^t - \lambda_4^t)/s \to 2t/3$ as $s \to 0$; the term $\boldsymbol{W}_t$ of equation (A3.26) is independent of s since $1/s$ in the matrices is canceled with s of C_5 (see equation (A3.11)).

Proof. In a large population, the seven eigenvalues for the transition matrix of the linear recurrence equations (5a–5d) and equations (A1.41a–A1.41c) are $\lambda_1, \lambda_4, \lambda_2 = \lambda_5 = -1/2, \lambda_3 = \lambda_6 = 1/4$, and λ_7. It holds

$$K_t^{mm}(1232) = \boldsymbol{B}_{21-23}\left(\boldsymbol{\lambda}_{1-3}^t\right)^T + B_{24}\left[(\lambda_4)^t - (\lambda_1)^t\right] + \boldsymbol{B}_{25-26}t\left(\boldsymbol{\lambda}_{2-3}^t\right)^T + B_{27}(\lambda_7)^t, \tag{A3.27a}$$

$$K_t^{mp+}(1232) = \boldsymbol{C}_{21-23}\left(\boldsymbol{\lambda}_{1-3}^t\right)^T + C_{24}\left[(\lambda_4)^t - (\lambda_1)^t\right] + \boldsymbol{C}_{25-26}t\left(\boldsymbol{\lambda}_{2-3}^t\right)^T + C_{27}(\lambda_7)^t, \tag{A3.27b}$$

$$K_t^{pp}(1232) = \boldsymbol{D}_{21-23}\left(\boldsymbol{\lambda}_{1-3}^t\right)^T + D_{24}\left[(\lambda_4)^t - (\lambda_1)^t\right] + \boldsymbol{D}_{25-26}t\left(\boldsymbol{\lambda}_{2-3}^t\right)^T + D_{27}(\lambda_7)^t, \tag{A3.27c}$$

where the constant coefficients $\boldsymbol{B}_{21-27}$, $\boldsymbol{C}_{21-27}$, and $\boldsymbol{D}_{21-27}$ are to be solved. Substituting $K_t^{mp+}(1232)$ and $K_t^{mm}(1232)$ of equations (A3.27a, A3.27b) and β_t^{mmp} of equation (A1.16b) into the recurrence equation (A1.41b), we obtain

$$B_{21} = (2\lambda_1 - 1)C_{21} + 2(\lambda_4 - \lambda_1)C_{24} - C_4,$$
$$\boldsymbol{B}_{22-23} = [2\boldsymbol{\lambda}_{2-3} - 1]\odot\boldsymbol{C}_{22-23} + 2\boldsymbol{\lambda}_{2-3}\odot\boldsymbol{C}_{25-26} - \boldsymbol{C}_{5-6}, \tag{A3.28}$$
$$\boldsymbol{B}_{24-27} = [2\boldsymbol{\lambda}_{4-7} - 1]\odot\boldsymbol{C}_{24-27} + (-C_4, 0, 0, -C_7).$$

Substituting $K_t^{mm}(1232)$ and $K_t^{pp}(1232)$ of equations (A3.27a, A3.27c) into the recurrence equation (A1.41c), we obtain

$$D_{21} = \frac{(1-s)(2\lambda_1 - 1)}{\lambda_1}C_{21} - \frac{(1-s)(\lambda_1 - \lambda_4)}{\lambda_1\lambda_4}C_{24} - \frac{1-s}{\lambda_4}C_4,$$

$$\boldsymbol{D}_{22-23} = [(1-s)(2\boldsymbol{\lambda}_{2-3} - 1) \oslash \boldsymbol{\lambda}_{2-3}] \odot \boldsymbol{C}_{22-23} + (1-s)(\boldsymbol{C}_{25-26} - \boldsymbol{C}_{5-6}) \oslash \boldsymbol{\lambda}_{2-3}, \tag{A3.29}$$

$$\boldsymbol{D}_{24-27} = [(1-s)(2\boldsymbol{\lambda}_{4-7} - 1)\oslash \boldsymbol{\lambda}_{4-7}] \odot \boldsymbol{C}_{24-27} + \left(-\frac{1-s}{\lambda_4}C_4, 0, 0, -\frac{1-s}{\lambda_7}C_7\right),$$

Substituting $K_t^{mm}(1232)$, $K_t^{mp+}(1232)$ and $K_t^{pp}(1232)$ of equations (A3.27a–A3.27c) and β_t^{mmp} and β_t^{mpp} of equations (A1.16b–A1.16c) into the recurrence equation (A1.41a), and substituting $\boldsymbol{B}_{21-27}$ and $\boldsymbol{D}_{21-27}$ of equations (A3.28, A3.29) we obtain

$$C_{24} = -\frac{B_4 + 2C_4}{3s}, C_{25} = \frac{1}{9}(2B_5 + C_5), C_{26} = -\frac{4}{9}(B_6 - C_6), C_{27} = -\frac{4}{81}(4B_7 + 17C_7). \tag{A3.30}$$

According to equation (A3.10), we have

$$\boldsymbol{B}_{4-7} = \left(1, \frac{3}{s}, 0, -2\right) \odot \boldsymbol{C}_{4-7}, \tag{A3.31}$$

and thus

$$\boldsymbol{C}_{24-27} = \left(-\frac{1}{s}, \frac{2}{3s}, \frac{4}{9}, -\frac{4}{9}\right) \odot \boldsymbol{C}_{4-7}, \tag{A3.32}$$

where $\boldsymbol{C}_{4-7}$ is given by equation (A3.11). The constant coefficient $\boldsymbol{C}_{21-23}$ is determined by the initial condition $\boldsymbol{K}_0(1232)$, and it is given by equation (A3.25). The theorem follows by writing equations (A3.27a–A3.27c) in the matrix form while keeping only the leading order of s.

Theorem A3.9. *The expected density $K_t^{mp-}(1232)$ in a large population is given by*

$$K_t^{mp-}(1232) = \frac{1}{3}C_4(\lambda_4)^t + A_{28}(\lambda_5)^t + \frac{2}{3}C_6(\lambda_6)^t + \frac{4}{3}C_7(\lambda_7)^t, \tag{A3.33}$$

where C_4, C_6, and C_7 are given by equation (A3.11), and

$$A_{28} = K_0^{mp-}(1232) - \frac{1}{3}C_4 - \frac{2}{3}C_6 - \frac{4}{3}C_7. \tag{A3.34}$$

Proof. In a large population, the eigenvalues of the transition matrix of the linear recurrence equations (5a–5d) and equation (A1.41d) are λ_4, duplicated $\lambda_5 = -1/2$, λ_6 and λ_7. It holds

$$K_t^{mp-}(1232) = A_{24}(\lambda_4)^t + (A_{28} + A_{25}t)(\lambda_5)^t + A_{26}(\lambda_6)^t + A_{27}(\lambda_7)^t, \tag{A3.35}$$

where $\boldsymbol{A}_{24-28}$ is to be solved. Substituting $K_t^{mp-}(1232)$ of equation (A3.35) and β_t^{mmp} of equation (A3.9) into the recurrence equation (A1.41d), and we obtain

$$A_{24} = \frac{1}{3}C_4, A_{25} = -C_5, A_{26} = \frac{2}{3}C_6, A_{27} = \frac{4}{3}C_7, \tag{A3.36}$$

and A_{28} is determined by the initial condition $K_0^{mp-}(1232)$ and it is given in equation (A3.34). The theorem follows by substituting equation (A3.36) into equation (A3.35), and approximating A_{25} to zero since the leading term of C_5 is the first order of s, see equation (A3.11).

Corollary A3.10. *The overall expected unction density ρ_t^{mp} in a large population is given by*

$$\rho_t^{mp} = C_8 + C_{10}\frac{1 - \lambda_1^{t-1}}{1 - \lambda_1} + \left[-C_{15} + \left(1 + \frac{8}{9}s\right)C_{10}t\right]\lambda_1^t$$
$$+ \left(-C_{16} - C_9 - \frac{4}{3}C_2 + \frac{2}{3}C_2t\right)\lambda_2^t + \left(-C_{17} - \frac{20}{9}C_3\right)\lambda_3^t, \tag{A3.37}$$

where $\boldsymbol{C}_{1-3}$, $\boldsymbol{C}_{8-10}$, and $\boldsymbol{C}_{15-17}$ are given by equation (A3.4), equations (A3.15–A3.17), and equation (A3.19), respectively.

Proof. The corollary follows by substituting R_t^m and R_t^p of equations (A3.13, A3.14) and $K_t^{mp}(1122)$ of equation (A3.20b) into equation (3).