CHAPTER 10

Beyond Individual Loci

Field cricket, Example 10.2

> Population geneticists recently have devoted much attention to the topic of gametic disequilibrium. The analysis of multiple-locus genotypic distributions can provide a sensitive measure of selection, genetic drift, and other factors that influence the genetic structure of populations.
>
> **(David W. Foltz et al. 1982, p. 80)**

> We thus pass from a point-theory to a strand-theory.
>
> **(R.A. Fisher 1965, p. 95)**

We have so far considered one locus at a time. Population genetic models become much more complicated when two or more loci are considered simultaneously (Slatkin 2008; Sved & Hill 2018). Many of our genetic concerns in conservation can be dealt with from the perspective of individual loci. Nevertheless, there are a variety of situations in which we must concern ourselves with the interactions between multiple loci, especially when using genomic data where many markers are no longer independent of each other. For example, genetic drift in small populations can generate nonrandom associations between genotypes at multiple loci. Therefore, the consideration of multilocus genotypes can provide powerful methods for detecting the effects of genetic drift in natural populations. It is more important than ever to understand the interpretation of multilocus genotypes to take advantage of the great power in genomic techniques that allow the description of genotypes at thousands of loci.

In addition, genotypes over many loci can be used to identify individuals genetically because the genotype of each individual (with the exception of identical twins or clones) is genetically unique if enough loci are considered. This genetic "fingerprinting" capability has many potential applications in understanding populations, estimating population size (Chapter 18), and applying genetics to problems in forensics (Chapter 22).

The nomenclature of multilocus genotypes is particularly messy and often inconsistent. It is difficult to find any two papers (even by the same author!) that use the same symbols and nomenclature for multilocus genotypes. Therefore, we have made a special effort to use the simplest possible nomenclature and symbols that are consistent as possible with previous usage in the literature.

The term **linkage disequilibrium** is commonly used to describe the nonrandom association between alleles at two loci (Box 10.1). However, this term is misleading because unlinked loci can be in so-called "linkage disequilibrium." Things are complicated enough without using misnomers that lead to additional confusion when considering multilocus models. The term **gametic disequilibrium** is a much more descriptive and appropriate term to use in this situation. We have chosen to use gametic disequilibrium in order to reduce confusion.

Conservation and the Genomics of Populations, Third Edition. Fred W. Allendorf, *et al.*, Oxford University Press.

Box 10.1 Linkage or gametic disequilibrium?

The terminology to describe the nonrandom association (i.e., correlation) between genotypes or alleles at multiple loci within a population has changed over the years. Lewontin & Kojima (1960, p. 459) introduced the term "linkage disequilibrium." They used "disequilibrium" because, if there is any recombination at all ($r > 0$), genotypes at two loci will eventually be randomly associated "in the absence of any evolutionary pressure such as selection." The term "linkage" apparently was chosen because the early treatments of this problem assumed that the loci were linked (e.g., Geiringer 1944).

As the theory of two-locus systems developed, it quickly became clear that the term linkage disequilibrium was inappropriate, and confusing because unlinked loci can be nonrandomly associated in many situations (e.g., small population size, population subdivision, hybridization, etc.). The term "gametic disequilibrium" began to be used in the literature within a few years (e.g., Fraser 1967). The theory of two loci is based upon analysis of gametic frequencies (Section 10.1); thus, this term is much more appropriate. Crow & Kimura (1970) used the phrase "gametic phase imbalance," but this phrase is awkward and unwieldy.

Gametic disequilibrium became common in the literature through the 1980s and peaked about 1990 (Hedrick 1987; Lewontin 1988). However, since then, linkage disequilibrium has become overwhelmingly more common. This has contributed to increasing confusion in the literature. The authors of this book have reviewed papers in which the authors test for "linkage disequilibrium," and incorrectly state that they tested for linkage. Some authors have distinguished between linkage and gametic disequilibrium with the view that nonrandom associations between syntenic loci is linkage disequilibrium, and it is gametic disequilibrium if the loci are on different chromosomes. However, in many cases, it is not known if the loci are syntenic or not.

The authors of this book have had some entertaining discussions on which term we should use. One of us, SNA, started a lively Twitter thread on this topic, which has led to some interesting suggested alternative terms (e.g., **interlocus allelic love**). In conclusion, we have decided to use the term gametic disequilibrium because we believe that reducing misunderstanding is more important than following convention.

We will first examine general models describing associations between loci and their evolutionary dynamics from generation to generation. We will then explore the various evolutionary forces that cause nonrandom associations between loci to come about in natural populations (genetic drift, natural selection, population subdivision, and hybridization). Finally, we will compare various methods for estimating associations between loci in natural populations.

10.1 Gametic disequilibrium

We now focus our interest on the behavior of two autosomal loci considered simultaneously under all of our Hardy–Weinberg (HW) equilibrium assumptions. We know that each locus individually will reach a neutral equilibrium in one generation under HW conditions. Is this true for two loci considered jointly? We will see shortly that the answer is no.

Loci on different chromosomes will be unlinked ($r = 0.5$) so that heterozygotes at both loci (*AaBb*) will produce all four gametes (*AB*, *Ab*, *aB*, and *ab*) in equal frequencies (Box 4.1; Chapter 4). Two loci that are close together on the same chromosome are generally linked so that the frequency of the parental gamete types (*AB* and *ab* in Figure 10.1) will be greater than the frequency of the nonparental gametes ($r < 0.5$). Some loci on the same chromosome can be far enough apart so that there is enough recombination to produce equal frequencies of all four gametes so they are unlinked ($r = 0.5$). Two loci that are on the same chromosome are **syntenic**, whether they are linked ($r < 0.5$) or unlinked ($r = 0.5$).

Allele frequencies are insufficient to describe genetic variation at multiple loci. Fortunately, however, we do not have to keep track of all possible genotypes. Rather, we can use the gamete frequencies to describe nonrandom associations between alleles at different loci. For example, in the case of

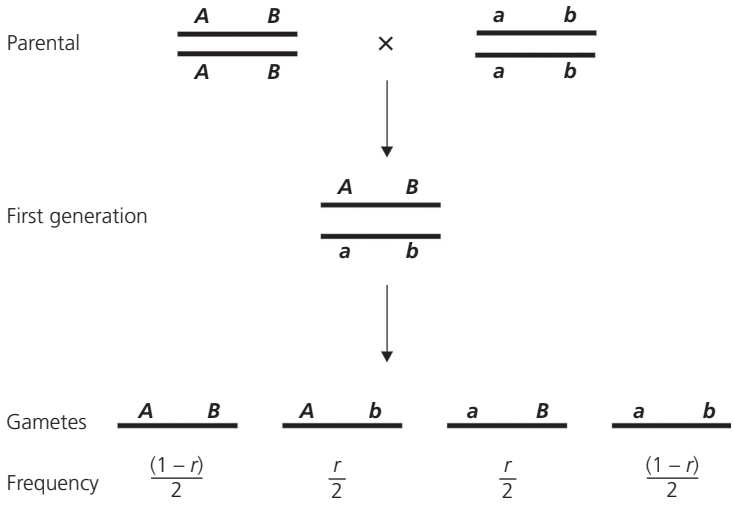

Figure 10.1 Outline of gamete formation in F_1 hybrids between two parents homozygous for different alleles at two loci. The gametes produced by the F_1 hybrids are affected by the rate of recombination (r). These four gametes will be equally frequent (25% each) for unlinked loci ($r = 0.5$). There will be an excess of parental gametes (*AB* and *ab* in this case) if the loci are linked ($r < 0.5$).

two loci that each has two alleles, there are just two allele frequencies, but there are nine different genotype frequencies. However, we can describe this system with just four gamete frequencies.

Let G_1, G_2, G_3, and G_4 be the frequencies of the four gametes *AB*, *Ab*, *aB*, and *ab*, respectively, as shown below. If the alleles at these loci are associated randomly then the expected frequency of any gamete type will be the product of the frequencies of its two alleles:

Gamete	Frequency	(10.1)
AB	$G_1 = (p_1)\,(p_2)$	
Ab	$G_2 = (p_1)\,(q_2)$	
aB	$G_3 = (q_1)\,(p_2)$	
ab	$G_4 = (q_1)\,(q_2)$	

where (p_1; q_1) and (p_2; q_2) are frequencies of the alleles (*A*; *a*) and (*B*; *b*), at locus *1* and *2*, respectively. The expected frequencies of two-locus genotypes in a random mating population can then be found as shown in Table 10.1.

D is used as a measure of the deviation from random association between alleles at the two loci (Lewontin & Kojima 1960). D is known as the coefficient of gametic disequilibrium and is defined as:

$$D = (G_1\,G_4) - (G_2\,G_3) \qquad (10.2)$$

Table 10.1 Genotypic array for two loci showing the expected genotypic frequencies in a random mating population.

	AA	***Aa***	***aa***
BB	${G_1}^2$	$2G_1G_3$	${G_3}^2$
Bb	$2G_1G_2$	$2G_1G_4 + 2G_2G_3$	$2G_3G_4$
bb	${G_2}^2$	$2G_2G_4$	${G_4}^2$

or:

$$D = G_1 - p_1\,p_2 \qquad (10.3)$$

If alleles are associated at random in the gametes (as in Equation 10.1), then the population is in gametic equilibrium and $D = 0$. If D is not equal to zero, the alleles at the two loci are not associated at random with respect to each other, and the population is said to be in gametic disequilibrium (Example 10.1). For example, if a population consists only of a 50:50 mixture of the gametes *AB* and *ab*, then:

$$G_1 = 0.5$$

$$G_2 = 0.0$$

$$G_3 = 0.0$$

$$G_4 = 0.5$$

and:

$$D = (0.5)\ (0.5) - (0.0)\ (0.0) = +0.25$$

Example 10.1 Genotypic frequencies with and without gametic disequilibrium

Let us consider two loci at which allele frequencies are $p_1 = 0.4$ ($q_1 = 1 - p_1 = 0.6$) and $p_2 = 0.7$ ($q_2 = 1 - p_2 = 0.3$) in two populations. The two loci are randomly associated in one population, but show maximum nonrandom association in the other. The gametic frequency values below show the case of random association of alleles at the two loci (gametic equilibrium, $D = 0$) and the case of maximum positive disequilibrium ($D = +0.12$; see Section 10.1.1 for an explanation of the maximum value of D).

Gamete	$D = 0$	D(max)
A B	$(p_1)(p_2) = 0.28$	0.40
A b	$(p_1)(q_2) = 0.12$	0.00
a B	$(q_1)(p_2) = 0.42$	0.30
a b	$(q_1)(q_2) = 0.18$	0.30

In a random mating population, the following genotypic frequencies will result in each case as shown below. The expected genotypic frequencies with $D = 0$ are shown without brackets, and the expected genotypic frequencies with maximum positive gametic disequilibrium are shown in square brackets:

	AA	*Aa*	*aa*	Total
BB	0.08 [0.16]	0.24 [0.24]	0.18 [0.09]	0.49 [0.49]
Bb	0.07 [0]	0.20 [0.24]	0.15 [0.18]	0.42 [0.42]
bb	0.01 [0]	0.04 [0]	0.03 [0.09]	0.09 [0.09]
Total	0.16 [0.16]	0.48 [0.48]	0.36 [0.36]	

Notice that each locus is in HW proportions in the populations either with or without gametic disequilibrium.

The amount of gametic disequilibrium (i.e., the value of D) will decay from generation to generation as a function of the rate of recombination (r, see Box 4.1) between the two loci.

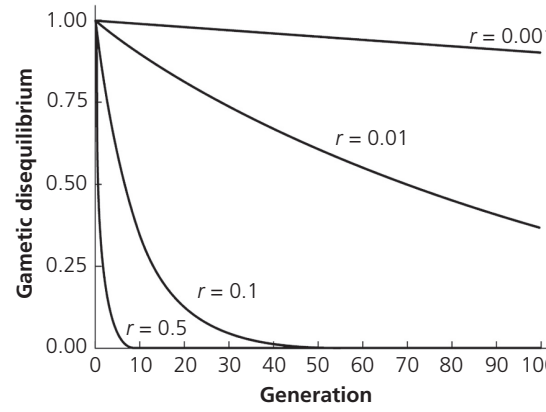

Figure 10.2 Expected decay of gametic disequilibrium (D_t/D_0) with time for various amounts of recombination (r) between the loci from Equation 10.3.

$$D' = D(1 - r) \tag{10.4}$$

So that after t generations:

$$D_t = D_0(1 - r)^t \tag{10.5}$$

If the two loci are not linked (i.e., $r = 0.5$), the value of D_t will be halved each generation until equilibrium is reached at $D = 0$. Linkage ($r < 0.5$) will delay the rate of decay of gametic disequilibrium. Nevertheless, D eventually will be equal to zero, as long as there is some recombination ($r > 0$) between the loci. However, if the two loci are tightly linked, it will take many generations for them to reach gametic equilibrium (Figure 10.2).

We therefore expect that nonrandom associations of genotypes between loci (i.e., gametic disequilibrium) would be much more frequent between tightly linked loci. For example, Zapata & Alvarez (1992) summarized observed estimates of gametic disequilibrium between five allozyme loci in several natural populations of *Drosophila melanogaster* on the second chromosome. The effective frequency of recombination is the mean of recombination rates in females and males. Only pairs of loci with less than 15% recombination showed consistent evidence of gametic disequilibrium. In contrast, recent studies of species of conservation interest have found much greater gametic disequilibrium, even between loci on different chromosomes (Section 10.8).

10.1.1 Other measures of gametic disequilibrium

D is a less than ideal measure of the relative amount of disequilibrium at different pairs of loci because the possible values of D are constrained by allele frequencies at both loci. The largest possible positive value of D is either p_1q_2 or p_2q_1, whichever value is smaller; and the largest negative value of D is the lesser value of p_1p_2 or q_1q_2. We can see that the largest positive value of D occurs when G_1 is maximum. p_1 is equal to G_1 plus G_2, and p_2 is equal to G_1 plus G_3. Therefore, the largest possible value of G_1 is the smaller of p_1 and p_2. We can see this in Example 10.1 in which the largest positive value of D occurs when G_1 is equal to p_1, which is less than p_2. Once the values of G_1, p_1, and p_2 are set, all of the other gamete frequencies must follow.

This allele frequency constraint of D reduces its value for comparing the amount of gametic equilibrium for the same loci in different populations or for different pairs of loci in the same population. For example, consider two pairs of loci in complete gametic disequilibrium. In case 1, both loci are at allele frequencies of 0.5, while in case 2, both loci are at allele frequencies of 0.9. The following gamete frequencies result:

	Frequencies	
Gamete	**Case 1**	**Case 2**
AB	0.5	0.9
Ab	0.0	0.0
aB	0.0	0.0
ab	0.5	0.1

The value of D in case 1 will be +0.25, while it will be +0.09 in case 2.

Several other measures of gametic disequilibrium have been proposed that are useful for various purposes (Hedrick 1987). A useful measure of gametic disequilibrium should have the same range regardless of allele frequencies. This will allow comparing the amount of disequilibrium among pairs of loci with different allele frequencies.

Lewontin (1964) suggested using the parameter D' to circumvent the problem of the range of values being dependent upon the allele frequencies:

$$D' = \frac{D}{D_{\max}} \tag{10.6}$$

Thus, D' ranges from zero to one for all allele frequencies. However, even D' is not independent of allele frequencies, and, therefore, is not an ideal measure of gametic disequilibrium (Lewontin 1988). Nevertheless, the D' coefficient is a useful tool for the estimation and comparison of the extent of overall disequilibrium among many pairs of multiallelic loci (Zapata 2000).

The correlation coefficient (R) between alleles at the two loci also has been used to measure gametic disequilibrium.

$$R = \frac{D}{(p_1q_1p_2q_2)^{1/2}} \tag{10.7}$$

R has a range of values between –1.0 and +1.0. However, this range is reduced somewhat if the two loci have different allele frequencies. Both D' and R will decay from generation to generation by a rate of $(1 - r)$, as does D, because they are both functions of D.

10.1.2 Associations between cytoplasmic and nuclear genes

Just as with multiple nuclear genes, nonrandom associations between nuclear loci and **mitochondrial DNA (mtDNA)** genotypes may occur in populations, as shown in the following table where *A* and *a* are alleles at a nuclear locus and *M* and *m* are haplotypes at a mtDNA locus.

Gamete	Frequency
AM	G_1
Am	G_2
aM	G_3
am	G_4

Again, D is a measure of the amount of gametic equilibrium and is defined as in Equation 10.2. D between nuclear and cytoplasmic genes will decay at a rate of one-half per generation, just as for two

Example 10.2 Cytonuclear disequilibrium in a hybrid zone of field crickets

Hybrid zones occur where two genetically distinct taxa are sympatric and hybridize to form at least partially fertile progeny (Section 13.2.3). Observations of the distribution of multilocus genotypes within hybrid zones and the patterns of introgression across hybrid zones can provide insight into the patterns of mating and the fitnesses of hybrids that may contribute to barriers to gene exchange between taxa.

Harrison & Bogdanowicz (1997) describe gametic disequilibrium in a hybrid zone between two species of field crickets, *Gryllus pennsylvanicus* and *G. firmus*. These two species hybridize in a zone that extends from New England to Virginia in the USA. Analyses of four anonymous nuclear loci, allozymes, mtDNA, and morphology at three sites in Connecticut indicate that nonrandom associations between nuclear markers, between nuclear and mtDNA (Figure 10.3), and between genotypes and morphology persist primarily because of more frequent matings between parental types. That is, the crickets at these three sites in this hybrid zone appear to be primarily parental with a few F_1 individuals and even fewer later generation hybrids.

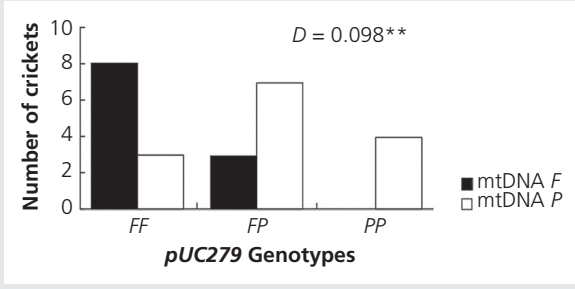

Figure 10.3 Gametic disequilibrium between mtDNA and a nuclear locus (*pUC279*) in a hybrid zone between two species of field crickets, *Gryllus pennsylvanicus* (P) and *Gryllus firmus* (F). The mtDNA from *G. firmus* (*F*) is significantly more frequent for homozygotes (*FF*) for the *G. firmus* nuclear allele. ** $P < 0.01$. Redrawn from Harrison & Bogdanowicz (1997).

These two species of field crickets are genetically similar. There are no fixed diagnostic differences at allozyme loci, and more than 50 anonymous nuclear loci had to be screened to find four that were diagnostic. These two taxa meet the criteria for species according to some **species concepts** but not others. Regardless, the long-term persistence of parental types throughout an extensive hybrid zone indicates that these species are clearly distinct biological units.

unlinked nuclear genes. That is,

$$D' = D(0.5) \tag{10.8}$$

and, therefore,

$$D_t = D(0.5)^t \tag{10.9}$$

For an empirical example of nonrandom association between nuclear and mtDNA loci, see Example 10.2.

10.2 Small population size

Nonrandom associations between loci will be generated by sampling effects in small populations. We can see this readily in the extreme case of a bottleneck of a single individual capable of reproducing by selfing because a maximum of only two gamete types can occur within a single individual. Conceptually, we can imagine the four gamete frequencies to be analogous to four alleles at a single locus. Changes in gamete frequencies from generation to generation caused by drift will often result in nonrandom associations between alleles at different loci. The expected value of D due to drift is zero. Nevertheless, drift-generated gametic disequilibria may be great and are equally likely to be positive or negative in sign. For example, genome-wide investigations in humans have found that large blocks

of gametic disequilibrium occur throughout the genome in human populations. These blocks of disequilibrium are thought to have arisen during an extreme population bottleneck that occurred some 25,000–50,000 years ago (Reich et al. 2001).

Gametic disequilibrium produced by a single generation of drift may take many generations to decay. Therefore, we would expect substantially more drift-generated gametic disequilibrium between closely linked loci. In fact, the expected amount of disequilibrium for closely linked loci is:

$$E(R^2) \approx \frac{1}{1+4Nr} \tag{10.10}$$

where R^2 is the square of the correlation coefficient (R) between alleles at the two loci (Equation 10.7) (Hill & Robertson 1968; Ohta & Kimura 1969). For unlinked loci, the following value of R^2 is expected (Weir & Hill 1980):

$$E(R^2) \approx \frac{1}{3N} \tag{10.11}$$

Guest Box 10 discusses the use of genotype frequencies at many loci to estimate effective population size in natural populations using Equation 10.11.

10.3 Natural selection

Let us examine the effects of natural selection with constant fitnesses at two loci each with two alleles. We will designate the fitness of a genotype to be w_{ij}, where i and j are the two gametes that join to form a particular genotype. There are two genotypes that are heterozygous at both loci (*AB/ab* and *Ab/aB*); we will assume that both double heterozygotes have the same fitness (i.e., $w_{23} = w_{14}$).

	AA	*Aa*	*aa*
BB	w_{11}	w_{13}	w_{33}
Bb	w_{12}	$w_{23} = w_{14}$	w_{34}
bb	w_{22}	w_{24}	w_{44}

The frequency of the *AB* gamete after one generation of selection will be:

$$G_{1'} = \frac{G_1(G_1w_{11} + G_2w_{12} + G_3w_{13} + G_4w_{14}) - rw_{14}D}{\bar{w}} \tag{10.12}$$

where $\bar{w}$ is the average fitness of the population. We can simplify this expression by defining $\bar{w}_i$ to be the average fitness of the ith gamete.

$$\bar{w}_i = \sum_{j=1}^{4} G_j w_{ij} \tag{10.13}$$

and then:

$$\bar{w} = \sum_{i=1}^{4} G_i \bar{w}_i \tag{10.14}$$

and:

$$G_{1'} = \frac{G_1\bar{w}_1 - rw_{14}D}{\bar{w}} \tag{10.15}$$

We can derive similar recursion equations for the other gamete frequencies.

$$G_{2'} = \frac{G_2\bar{w}_2 + rw_{14}D}{\bar{w}}$$

$$G_{3'} = \frac{G_{3'}\bar{w}_3 + rw_{14}D}{\bar{w}} \tag{10.16}$$

$$G_{4'} = \frac{G_4\bar{w}_4 - rw_{14}D}{\bar{w}}$$

There are no general solutions for selection at two loci. That is, there is no simple formula for the equilibria and their stability. However, a number of specific models of selection have been analyzed. The simplest of these is the additive model where the fitness effects of the two loci are summed to yield the two-locus fitnesses. Another simple case is the multiplicative model where the two-locus fitnesses are determined by the product of the individual locus fitnesses. In both of these cases, heterozygous advantage at each locus is necessary and sufficient to ensure stable polymorphisms at both loci.

In some cases, the multilocus fitness cannot be predicted by either the additive or multiplicative combination of fitnesses at individual loci (Phillips 2008). Such interaction between loci is referred to as **epistasis** (i.e., the interaction of different loci such that the multiple locus phenotype is different than that predicted by simply combining the effects of each individual locus). The study of epistasis, or interactions between genes, is fundamentally important to understanding the structure and function of genetic pathways and the evolutionary dynamics of complex genetic systems.

A detailed examination of the effects of natural selection at two loci, including epistasis, is beyond the scope of our consideration. Interested readers are directed to appropriate population genetics sources (e.g., Hartl & Clark 1997; Phillips 2008; Hedrick 2011). We will consider two situations of selection at multiple loci that are particularly relevant for conservation.

10.3.1 Genetic hitchhiking

Natural selection at one locus can affect closely linked loci in many ways. Let us first consider the case where directional selection occurs at one locus (B) and the second locus is selectively neutral (A). The following fitness set results:

	AA	Aa	aa
BB	w_{11}	w_{11}	w_{11}
Bb	w_{12}	w_{12}	w_{12}
bb	w_{22}	w_{22}	w_{22}

where $w_{11} < w_{12} < w_{22}$.

Imagine that the favored b allele is a new mutation at the B locus. In this case, the selective advantage of the b allele may carry along either the A or a allele, depending upon which allele is initially associated with the b mutation. This is known as **genetic hitchhiking** and will result in a so-called **selective sweep**. The magnitude of this effect depends on the **selection differential**, the amount of recombination (r), and the initial gametic array (Figure 10.4). A selective sweep will reduce the amount of variation at loci that are tightly linked to the locus under selection (Figure 10.5).

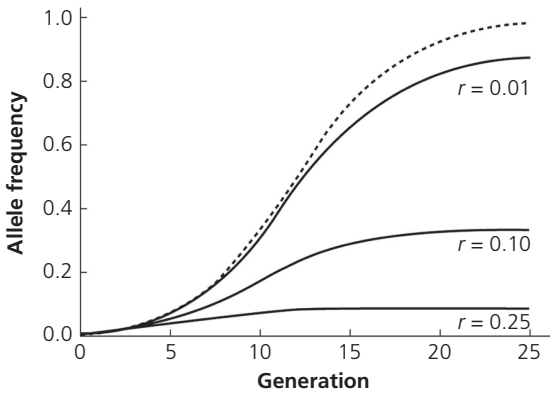

Figure 10.4 Effect of hitchhiking on a neutral locus that is initially in complete gametic disequilibrium with a linked locus that is undergoing directional selection ($w_{11} = 1.0$; $w_{12} = 0.75$; $w_{22} = 0.5$). r is the recombination rate between the two loci. The dashed line shows the expected change at the selected locus.

Figure 10.5 A selective sweep for a mutation (red) that quickly goes to fixation by natural selection. Neutral variants (blue) closely linked to the site under selection will also go to fixation through hitchhiking. This results in reduced genetic variation in the region around the selected site. As genetic distance increases, recombination will act to reduce this effect. From Sætre & Ravinet (2019).

For example, Kardos et al. (2015b) detected a selective sweep in bighorn sheep at *RXFP2*, a gene that strongly affects horn size in domestic sheep. The massive horns carried by bighorn rams appear to have evolved in part via strong positive selection at *RXFP2* in the past 2 million years since their divergence from domestic sheep.

10.3.2 Associative overdominance

Selection at one locus also can affect closely linked neutral loci when the genotypes at the selected locus are at an equilibrium allele frequency. Consider the case of heterozygous advantage where, using the previous fitness array, $w_{11} = w_{22} = 1.0$ and $w_{12} = (1 + s)$. The effective fitnesses at the A locus are affected by selection at the B locus (s) and D; the marginal fitnesses are the average fitness at the A locus considering the two-locus genotypes. These would be the estimated fitnesses at the A locus if only that locus were observed. If D is zero then all the genotypes at the A locus will have the same fitness. However, if there is gametic disequilibrium (i.e., D is not equal to zero), then heterozygotes at

the *A* locus will experience a selective advantage because of selection at the *B* locus.

This effect has been called **associative overdominance** (Ohta 1971) or **pseudo-overdominance** (Carr & Dudash 2003). This pattern of selection has also been called marginal overdominance (Hastings 1981). However, **marginal overdominance** has more generally been used for the situation where genotypes experience multiple environments and different alleles are favored in different environments (Wallace 1968). This situation can lead to an overall greater fitness of heterozygotes even though they do not have a greater fitness in any single environment.

Heterozygous advantage is not necessary for linked loci to experience associative overdominance. Heterozygous individuals at a selectively neutral locus will have higher average fitnesses than homozygotes if the locus is in gametic disequilibrium with a locus having deleterious recessive alleles (Ohta 1971).

We can see this with the genotypic arrays in Example 10.1. Let us assume that the *b* allele is a recessive lethal (i.e., fitness of the *bb* genotype is zero). In the case of gametic equilibrium ($D = 0$), exactly q_2^2 ($0.3 \times 0.3 = 0.09$) of genotypes at the *A* locus have a fitness of zero. Thus, the mean or marginal fitness at the *A* locus is $1 - 0.09 = 0.91$. However, in the case of maximum positive disequilibrium, only the *aa* genotypes have reduced fitness because the *AA* and *Aa* genotypes do not occur in association with the *bb* genotype. Thus, the fitness of *AA*, *Aa*, and *aa* are 1, 1, and 0.75. There are many more *aa* than *AA* homozygotes in the population; therefore, *Aa* heterozygotes have greater fitness than the mean of the homozygotes.

10.3.3 Genetic draft

We saw in Section 10.3.1 that directional selection at one locus can reduce the amount of genetic variation at closely linked loci following a selective sweep. This is a special case of a more general effect in which selection at one locus will reduce the effective population size of linked loci. This has been termed the **Hill–Robertson effect** (Hey 2000) because it was first discussed in a paper that considered the effect of linkage between two loci under selection (Hill & Robertson 1966). Observations with *Drosophila* have found that regions of the genome with less recombination tend to be less genetically variable as would be expected with the Hill–Robertson effect (Charlesworth 1996).

This effect has potential importance for conservation genetics. For example, we would expect a strong Hill–Robertson effect for mtDNA where there is no recombination. A selective sweep of a mutant with some fitness advantage could quickly fix a single haplotype and therefore greatly reduce genetic variation. Therefore, low variation at mtDNA may not be a good indicator of the effective population size experienced by the nuclear genome.

Gillespie (2001) has presented an interesting consideration of the effects of hitchhiking on regions near a selected locus. He has termed this effect **genetic draft** and has suggested that the stochastic effects of genetic draft may be more important than genetic drift in large populations. In general, it would reduce the central role thought to be played by effective population size in determining the amount of genetic variation in large populations. The potential effects of genetic draft seem to not be important for the effective population sizes usually of concern in conservation genetics.

10.4 Population subdivision

Population subdivision will generate nonrandom associations (gametic disequilibrium) between alleles at multiple loci if the allele frequencies differ among subpopulations at both loci. This is an extension to two loci of the Wahlund principle, the excess of homozygotes caused by population subdivision at a single locus, to two loci (Section 9.1) (Sinnock 1975). In general, for *k* equal-sized subpopulations:

$$D = \bar{D} + cov(p_1, p_2) \qquad (10.17)$$

where $\bar{D}$ is the average *D* value within the *k* subpopulations (Nei & Li 1973; Prout 1973).

This effect is important when two or more distinct subpopulations are collected in a single sample. For example, many populations of fish living in lakes consist of several genetically distinct subpopulations that reproduce in different tributary streams. Thus, a single random sample taken of the fish living in the lake will comprise several separate

demes. Makela & Richardson (1977) have described the detection of multiple genetic subpopulations by an examination of gametic disequilibrium among many pairs of loci.

Cockerham & Weir (1977) introduced a composite measure of gametic disequilibrium that partitions gametic disequilibrium into two components: the usual measure of gametic disequilibrium, D, plus an added component that is due to the nonrandom union of gametes caused by population subdivision (D_B).

$$D_C = D + D_B \tag{10.18}$$

In a random mating population, D and D_C will have the same value. We will see in the next section that the composite measure is of special value when estimating gametic disequilibrium from population samples. Campton (1987) has provided a helpful discussion of the derivation and use of the composite gametic disequilibrium measure.

10.5 Hybridization

Hybridization between populations, subspecies, or species will result in gametic disequilibrium. Figure 10.1 can be viewed as the resulting genotypes and gametes in the first two generations of hybridization. The F_1 hybrid will be heterozygous for all loci at which the two taxa differ. The gametes produced by the F_1 hybrid will depend on the linkage relationship of the two loci. If the two loci are unlinked, then all four gametes will be produced in equal frequencies because of recombination.

Table 10.2 shows the genotypes produced by hybridization between two taxa that are fixed for different alleles at two unlinked loci. This assumes that the two taxa are equally frequent and mate at random. We can see here that gametic disequilibrium (D) will be reduced by exactly one-half each generation. For unlinked loci, recombination will eliminate the association between loci in heterozygotes. However, only one-half of the population in a random mating population will be heterozygotes in the first generation. Recombination in the two homozygous genotypes will not have any effect. Therefore, gametic disequilibrium (D) will be reduced by exactly one-half each generation. A similar effect will occur in later generations even though more genotypes will be present. That is, recombination will only affect the frequency of gametes produced in individuals that are heterozygous at both loci (*AaBb*).

Gametic disequilibrium will decay at a rate slower than one-half per generation if the loci are linked. Tight linkage will greatly delay the rate of decay of D. For example, it will take an expected 69 generations for D to be reduced by one-half if there is 1% recombination between loci (Equation 10.5).

Gametic disequilibrium also will decay at a slower rate if the population does not mate at random and there is positive assortative mating of the parent types. This will reduce the frequency of double heterozygotes in which recombination can act to

Table 10.2 Expected genotype frequencies and coefficient of gametic disequilibrium (D) in a random mating hybrid swarm.

	Genotype frequencies				
Genotypes	**Parental**	**First generation**	**Second generation**	**Third generation**	**Equilibrium**
AABB	0.500	0.250	0.141	0.098	0.063
AABb			0.094	0.118	0.125
AAbb			0.016	0.035	0.063
AaBB			0.094	0.118	0.125
AaBb		0.500	0.312	0.267	0.250
Aabb			0.094	0.118	0.125
aaBB			0.016	0.035	0.063
aaBb			0.094	0.118	0.125
aabb	0.500	0.250	0.141	0.098	0.063
D	—	+0.250	+0.125	+0.063	0.000

reduce gametic disequilibrium. We can see this using Equation 10.18. In this case, the D component of the composite measure (D_C) will decline at the expected rate, but D_B will persist depending upon the amount of assortative mating. Random mating in a hybrid population can be detected by testing for HW proportions at individual loci.

These two alternative explanations of persisting gametic disequilibrium in a hybrid can be distinguished. Assortative mating will affect all pairs of loci (including cytoplasmic and nuclear associations) while the effect of linkage will differ between pairs depending upon their rate of recombination. Example 10.2 describes the multilocus genotypes in a natural hybrid zone between two species of crickets. In this case, most genotypes are similar to the parental taxa and gametic disequilibrium persists over all loci because of assortative mating. Forbes & Allendorf (1991) have described a hybrid swarm in which mating is at random (all loci are in HW proportions), but gametic disequilibrium persists at linked loci (Example 10.3).

We will examine hybridization and its genotypic effects again in Chapter 13 when we consider the effects of hybridization on conservation.

10.6 Estimation of gametic disequilibrium

There is no simple way to estimate gametic equilibrium values from population data (Kalinowski & Hedrick 2001; Barton 2011; Hui & Burt 2020). As described in the next section, even the simplest case of two alleles at a pair of loci is complicated. Estimation becomes more difficult for loci that have more than two alleles. There are a total of $n(n-1)/2$ pairwise combinations of loci if we examine n loci. So with 10 loci, each with just 2 alleles, there are a total of 45 combinations of two-locus gametic equilibrium values to estimate.

10.6.1 Two loci with two alleles each

Let us consider the simplest case of two alleles at a pair of loci (see genotypic array in Table 10.1). The gamete types (e.g., *AB* or *Ab*) cannot be observed directly but must be inferred from the diploid genotypes. For example, *AABB* individuals can only result from the union of two *AB* gametes, and *AABb* individuals can only result from the union of an *AB* gamete and an *Ab* gamete. Similar inferences of gametic types can be made for all individuals that are homozygous at one or both loci. In contrast, gamete frequencies cannot be inferred from double heterozygotes (*AaBb*) because they may result from either union of *AA* and *bb* gametes or *Ab* and *aB* gametes. Consequently, gametic disequilibrium cannot be calculated directly from diploids.

Several methods are available to estimate gametic disequilibrium values in natural populations when the two gametic types of double heterozygotes cannot be distinguished. The simplest way is to ignore them, and simply estimate D from the remaining eight genotypic classes. The problem with this method is that double heterozygous individuals may represent a large proportion of the sample (Example 10.3), and their exclusion from the estimate will result in a substantial loss of information.

The best alternative is the **expectation maximization** (EM) algorithm, which provides a maximum likelihood estimate of gamete frequencies assuming random mating (Hill 1974). We previously used the EM approach in the case of a null allele where not all genotypes could be distinguished at a single locus (Section 5.4.2). This approach uses an iteration procedure along with the maximum likelihood estimate of the gamete frequencies:

$$\hat{G}_1 = [\frac{1}{2N}][2N_{11} + N_{12} + N_{21} + \frac{N_{22}\hat{G}_1(1-\hat{p}_1-\hat{p}_2+\hat{G}_1)}{\hat{G}_1(1-\hat{p}_1-\hat{p}_2-\hat{G}_1)+(\hat{p}_1-\hat{G}_1)(\hat{p}_2-\hat{G}_1)}] \quad (10.19)$$

where N is the sample size, N_{11} is the number of *AABB* genotypes observed, N_{12} is the number of *AABb* genotypes observed, and N_{21} is the number of *AaBB* genotypes observed (Hedrick 2011, p. 585). This expression is not as opaque as it first appears. The first three sums in the right-hand parentheses are the observed numbers of the G_1 gametes in genotypes that are homozygous for at least one locus. The fourth value is the expected number of copies of the G_1 gamete in the double heterozygotes.

Example 10.3 Gametic disequilibrium in a hybrid swarm

Forbes & Allendorf (1991) studied gametic disequilibrium in a **hybrid swarm** of cutthroat trout (Figure 10.6). They observed the following genotypic distribution between two closely linked diagnostic allozyme loci. At both loci, the upper-case allele (*A* and *B*) designates the allele fixed in the Yellowstone cutthroat trout and the lower-case allele (*a* and *b*) is fixed in westslope cutthroat trout. The expected genotypes with **gametic equilibrium** ($D = 0$) are presented in parentheses. There is a large excess of both parental gamete types (*AABB*) and (*aabb*). The allele frequencies at the two loci are $p_1 = 0.589$ and $p_2 = 0.518$:

	LDH-A2			
ME-4	***AA***	***Aa***	***aa***	**Total**
BB	7	0	0	7
	(2.6)	(3.6)	(1.3)	
Bb	3	12	0	15
	(4.8)	(6.8)	(4.9)	
bb	0	1	5	6
	(1.2)	(2.2)	(1.0)	
Total	10	13	5	

Figure 10.6 Westslope cutthroat trout in Lake Rogers, Montana. Photo courtesy of John Ashley.

The estimated value of D in this case is 0.213 using the **expectation maximization** (EM) method described in Section 10.6.1, and $D' = 1.000$. The estimated gamete frequencies are presented below:

Gamete	$D = 0$	$D = 0.213$
A B	$(p_1)(p_2) = 0.305$	0.518
A b	$(p_1)(q_2) = 0.284$	0.071
a B	$(q_1)(p_2) = 0.213$	0.000
a b	$(q_1)(q_2) = 0.198$	0.411

Thus, we see that even though mating is at random in this hybrid swarm, gametic disequilibrium persists for many generations when loci are linked.

Bilton et al. (2018) have presented a method for estimating pairwise gametic disequilibrium in random mating populations with genomic data using the method of Hill (1974). Their method takes into account errors resulting from undercalled heterozygotes because of allelic dropout, as well as sequencing errors.

We need to make an initial estimate of gamete frequencies and then iterate using this expression. Our initial estimate can either be the estimate of gamete frequencies with $D = 0$, or we can use the procedure described in the previous paragraph to initially estimate D from the remaining eight genotypic classes. The other three gamete frequencies can be solved directly once we estimate G_1 and the single-locus allele frequencies. Iteration can sometimes converge on different gamete values depending upon the initial gamete frequencies (Excoffier & Slatkin 1995). Kalinowski & Hedrick (2001) present a detailed consideration of the implications of this problem when analyzing datasets with multiple loci.

It is crucial to remember that the EM algorithm assumes random mating and HW proportions. The greater the deviation from expected HW proportions, the greater the probability that this iteration will not converge on the maximum likelihood estimate. Stephens et al. (2001) have provided an algorithm to estimate gamete frequencies that assumes that the gametes in the double heterozygotes are likely to be similar to the other gametes in the samples. This method is likely to be less sensitive to nonrandom mating in the population being sampled.

10.6.2 More than two alleles per locus

The numbers of possible multilocus genotypes expand rapidly when we consider more than two alleles per locus. For example, there are six genotypes and three gametes types at a single locus with three alleles. Therefore, there are $6 \times 6 = 36$ diploid genotypes and $3 \times 3 = 9$ possible combinations of gametes at two loci each with three alleles. D values for each pair of alleles at two loci can be estimated and tested statistically (Kalinowksi & Hedrick 2001). The EM iteration procedure is more likely to converge to a value other than the maximum likelihood solution as the number of alleles per locus increases. Therefore, it is important to initiate the iteration from many different starting points with highly polymorphic samples.

10.7 Strand theory: Junctions and chromosome segments

The ability to sequence large sections of chromosomes provides the opportunity to interpret multiple locus genetic data using entirely new conceptual approaches (Thompson 2018). It is now possible to use sequence data to identify chromosomal segments originating from different ancestral chromosomes. Junctions are the points at which the chromosome of origin changes because of historical recombination events during meiosis. For example, Figure 10.7 shows the inheritance and transmission of a single ancestral chromosome over two generations. One junction in each meiotic event has resulted in a chromosome with segments originating from three different ancestral chromosomes.

Chromosomal strand theory was developed by R.A. Fisher (1949) in the context of understanding

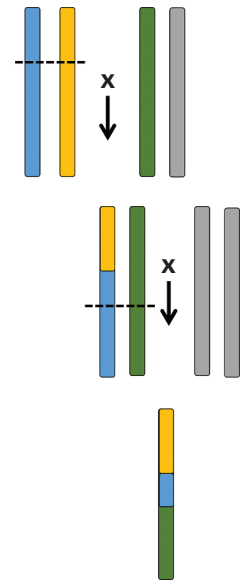

Figure 10.7 The inheritance and transmission of a chromosome over two generations. One junction (dashed lines) in each meiotic event has resulted in a final chromosome with segments originating from three different ancestral chromosomes. The gray shaded chromosomes do not contribute to the resulting chromosome.

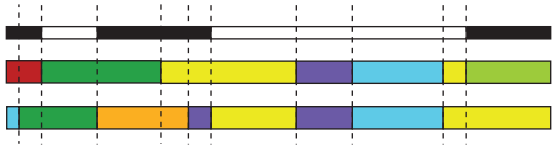

Figure 10.8 Two homologous chromosomes in a single individual sampled from a population a few generations after founding. Different colors represent different ancestral chromosomes. The dotted vertical lines indicate the location of junctions. The two white sections of the bar above the chromosomes indicate regions originating from the same ancestral chromosomes. These are regions of the genome where this individual is IBD (see Section 17.2) because of inbreeding. Redrawn from Chapman & Thompson (2002).

the effects of close inbreeding in lines of laboratory mice (Thompson 2018). For example, Figure 10.8 shows two hypothetical chromosomes sampled from a random mating population a few generations after founding (Chapman & Thompson 2002). It is astounding that this theory was envisioned and developed over 50 years before techniques were available to provide empirical data that could be used to apply this theory to real populations.

The number and distributions of lengths of these segments provide the opportunity to estimate a variety of fundamental population genetic parameters: inbreeding coefficients, migration rates, and effective population size. We will use this conceptual approach in understanding gene flow and hybridization (Chapter 13) and inbreeding (Chapter 17).

10.7.1 Microhaplotypes

The conceptual framework of strands rather than points can also be applied to the interpretation of multiple **single nucleotide polymorphisms (SNPs)**. Most SNPs are bi-allelic and thus individually have somewhat limited power in comparison with highly polymorphic microsatellite loci that often have many alleles. However, multiple SNPs that occur within the same small region can be genotyped jointly from high-throughput short-read DNA sequences to derive multi-allelic microhaplotype markers (Kidd et al. 2014). Microhaplotypes have been defined as a locus with two or more SNPs that occur within a short segment of DNA (e.g., 200 base pairs (bp)) that can be covered by a single sequence read and collectively define a multiallelic locus (Kidd & Speed 2015).

Figure 10.9 shows a hypothetical example of the interpretation of three bi-allelic SNPs within a 40-bp region. There are eight possible microhaplotypes in a region containing three bi-allelic SNPs ($2 \times 2 \times 2 = 8$). However, only four of these possible eight microhaplotypes are present in this sample of four individuals. If these SNPs were in gametic equilibrium, then all eight microhaplotypes would be equally frequent. However, strong gametic disequilibrium is often present between closely linked SNPs.

Microhaplotypes are especially useful in situations where multiple allele loci are more powerful. For example, Box 12.1 considers the use of loss of alleles to detect population bottlenecks. Baetscher et al. (2018) have shown that the use of microhaplotypes in the kelp rockfish, a nearshore marine fish, provides large increases in power to identify kin relationships from the same amount of DNA sequence data. Kidd et al. (2015) demonstrate that microhaplotypes are extremely valuable for a number of forensic applications.

10.8 Multiple loci and conservation

Understanding multiple locus genotypes is especially important in conservation because small population size will generate nonrandom relationships between loci. Substantial gametic disequilibrium has been found even between unlinked pairs of loci on different chromosomes in many species (Example 10.4; Bensch 2006; Slate & Pemberton 2007). This is perhaps not unexpected. We saw in Section 10.2 that small population size in itself can produce substantial amounts of gametic disequilibrium. In addition, the rate of hybridization between subpopulations in many species has also increased because of human activities (Slate & Pemberton 2007; see Chapter 13). Thus, many of these populations that are of conservation interest might have substantial gametic disequilibrium because of hybridization, population subdivision, or small population size.

The interpretation of multilocus genotypes is becoming increasingly important for conservation because of the ability to screen many loci. The more loci examined, the more pairs of loci we are likely to

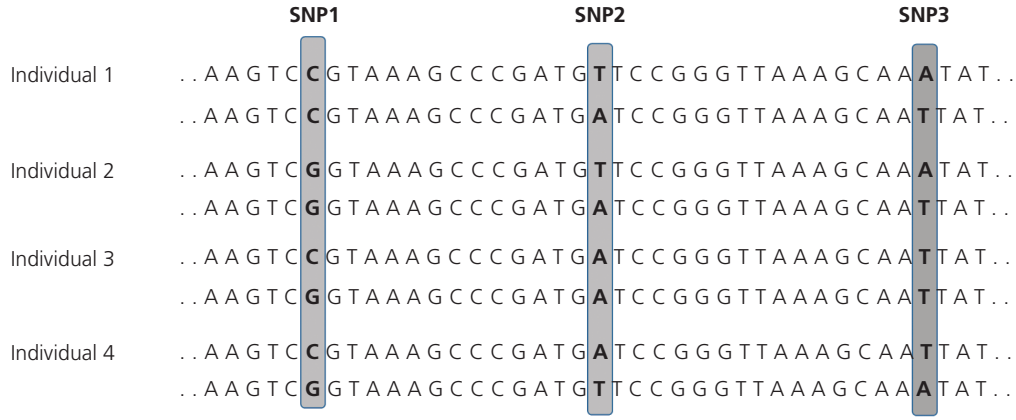

Figure 10.9 Hypothetical example of the interpretation of multiple closely linked SNPs as microhaplotypes. Above is shown the sequences of 40 bp in four individuals with three bi-allelic SNPs. The genotypes of the four individuals are CTA/CAT, GTA/GAT, CAT/GAT, and CAT/GTA at the individuals SNPs. Below are shown all eight possible microhaplotypes for the three SNPs. Only four of these possible haplotypes are actually present. The frequencies of these four microhaplotypes in this sample of four individuals are in the far right column.

Example 10.4 Extensive gametic disequilibrium at microsatellite loci in the Siberian jay

Li & Merilä (2010) estimated gametic disequilibrium between 103 microsatellite loci in a semi-isolated population of Siberian jay from western Finland. This subpopulation has been the subject of a long-term field study for over 35 years.

A linkage map for this population was constructed from pedigrees through direct field observations in combination with verification of parentage using microsatellite genotypes (Jaari et al. 2009). Recombination rates were estimated by the examination of 311 progeny fathered by 85 males and mothered by 95 females. A total of 107 microsatellite loci were assigned to one Z chromosome-specific and nine autosomal linkage groups. Ten loci could not be assigned to any linkage group. Six of the loci were found to be sex linked; three of these were in a pseudoautosomal region found on both the Z and W chromosomes, and three were Z chromosome-specific. As has been found in many species (Otto & Payseur 2019), there was less recombination in males than in females. On average, there was 28% greater recombination in females than in males. Figure 10.10 shows the comparative linkage map for one of the autosomal linkage groups.

A total of 97 autosomal and the 6 sex-linked loci were genotyped to estimate gametic disequilibrium in the wild population (Li & Merilä 2010). As expected, the amount of gametic disequilibrium between pairs of linked loci declined as the rate of recombination increased (Figure 10.11). Unlike the data from *Drosophila* described in Section 10.1, substantial gametic disequilibrium was found between pairs of loci separated by much more than 15% recombination. Significant ($P < 0.05$) gametic disequilibrium was even found in 83% of unlinked marker pairs on different chromosomes. As expected, the amount of gametic disequilibrium between pairs of loci on different chromosomes and unlinked pairs of loci on the same chromosome was quite similar: $D' = 0.356$ versus $D' = 0.354$.

Continued

Example 10.4 *Continued*

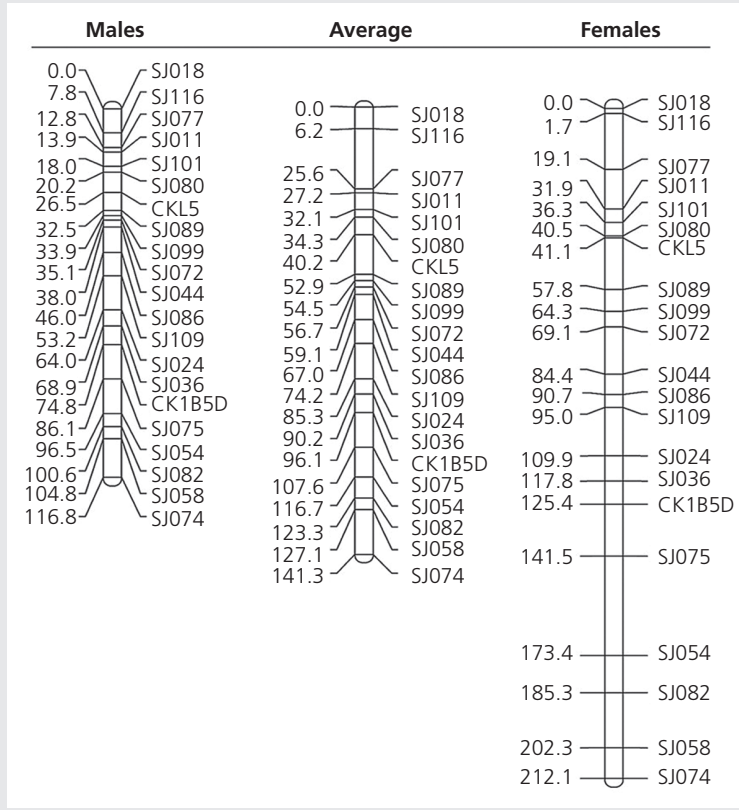

Figure 10.10 Linkage group 2 of the Siberian jay for males, females, and the average recombination rates of males and females. The names of the loci are on the right, and the total map distances on the left in centimorgans (cM). There is greater recombination for this linkage group in females than in males, as indicated by the greater distances between loci in the female. From Jaari et al. (2009).

The overall amount of gametic disequilibrium in the population is surprisingly high. This gametic disequilibrium probably results at least partially from the small effective population size of this population (N_e = 170, Fabritius 2010). In addition, pedigree analysis over many generations revealed five different extended family groups in this population. Such subdivision is expected to increase gametic disequilibrium, as we saw in Section 10.4.

The substantial gametic disequilibrium in this population has important implications, regardless of its cause. These observations also emphasize again how misleading it is to use the term "linkage disequilibrium" to refer to nonrandom associations between loci, as we discussed in Box 10.1. In this case, most pairs of loci found to be in "linkage" disequilibrium are actually not linked.

Continued

Example 10.4 *Continued*

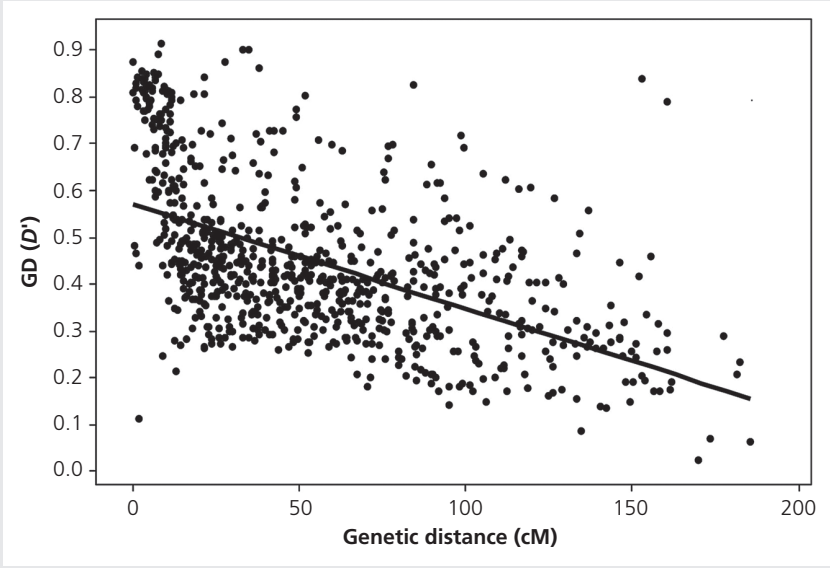

Figure 10.11 Gametic disequilibrium, as measured by D' (Equation 10.6), for syntenic pairs of loci separated by different amounts of recombination in the Siberian jay. Data from Li & Merilä (2010).

sample that are on the same chromosome and are in gametic disequilibrium.

In addition, advances in data analysis have revealed the presence in many species of chromosomal inversions that suppress recombination (Section 3.1.6). Such inversions have been found to be associated with local adaptation and life history variation in many species (Wellenreuther & Bernatchez 2018; Box 15.2). For example, Petrou et al. (2021) genotyped 6,718 SNP loci in over 1,000 Pacific herring from spawning aggregations along the Pacific Coast of North America. Overall genome-wide differentiation was low (F_{ST} = 0.014), but 116 outlier loci were detected (mean F_{ST} = 0.11) that were strongly correlated with time of spawning. Plots of gametic disequilibrium versus physical distance indicated the presence of major inversions on four (7, 8, 12, and 15) of the 26 pairs of chromosomes (Figure 10.12). Many of the outlier loci responsible for local adaptation were found within these inversions.

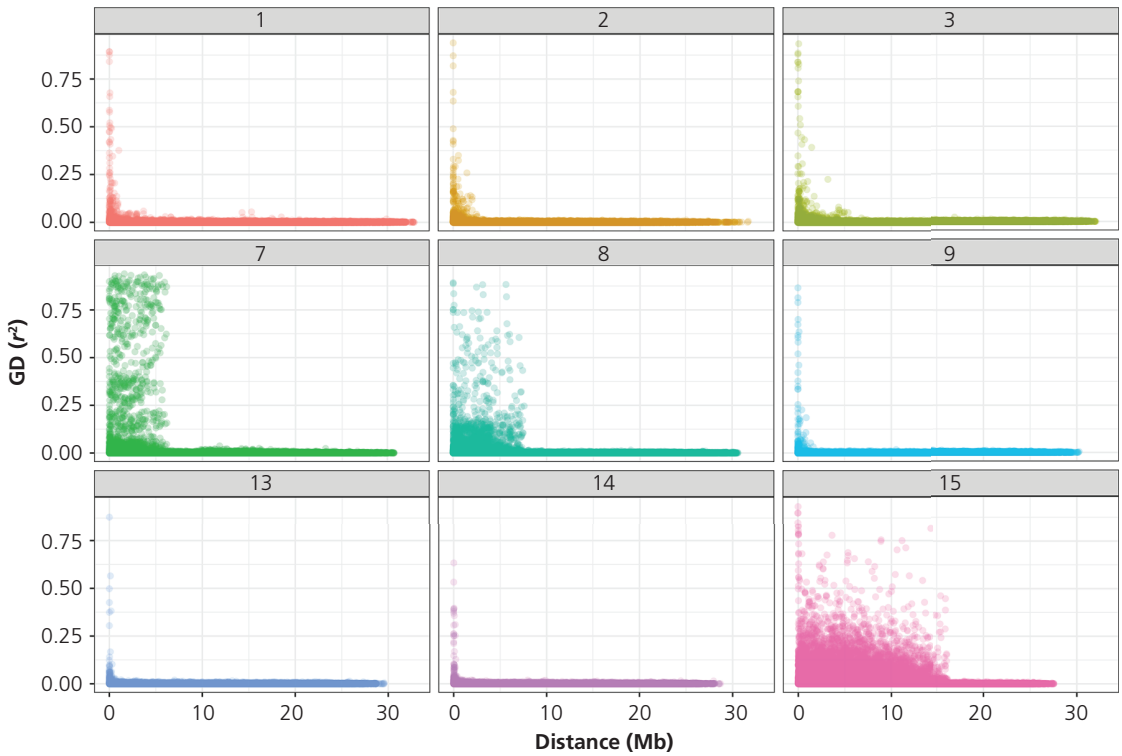

Figure 10.12 GD (r^2) versus the physical distance (Mb) between pairs of SNP loci in Pacific herring (Petrou et al. 2021). Nine of the 26 pairs of chromosomes in this species are shown. The average genome-wide recombination rate is 2.54 cM/Mb in the closely related Atlantic herring (Peterrsson et al. 2019). Thus, 10 Mb corresponds approximately to 25 cM. High gametic disequilibrium over long physical distances is present on chromosomes 7, 8, and 15 because of chromosomal inversions. F_{ST} outlier loci associated with spawning time among 23 spawning aggregations tended to map within these inversions. Figure courtesy of Eleni Petrou.

Guest Box 10 Estimation of effective population size using gametic disequilibrium with genomic data

Robin S. Waples

For many decades, use of genetic data to estimate N_e focused on the temporal method, which uses changes in allele frequency in temporally spaced samples (Waples 1989). Because gametic disequilibrium arises by drift in all finite populations at a rate inversely proportional to N_e (Equation 10.11), genetic estimates of effective size also can be obtained from individual samples. Initially, the gametic disequilibrium method was thought to have little practical relevance because of the high variance associated with gametic disequilibrium at individual pairs of loci (Hill 1981). However, the bias-corrected *LDNE* method developed by Waples and Do (2008) showed that robust estimates can be obtained by combining data for multiple pairs of loci, and within a few years Palstra and Fraser (2012) reported that most genetic estimates of N_e used single-sample methods, either *LDNE* or the sibship method of Wang (2009).

Extensive computer simulations have demonstrated the following regarding the gametic disequilibrium method:

- Precise estimates can be obtained with even modest amounts of data (10–20 microsatellite loci or ~100 SNPs) for relatively small populations ($N_e \leq$ a few hundred; Waples and Do 2010).
- Estimating N_e is challenging in large populations because the genetic signal (proportional to $1/N_e$) is weak (Waples 2016; Marandel et al. 2019), but this can be overcome to some extent using large numbers of individuals and incorporating life history information, which can place an upper bound on $\hat{N}_e$ (Waples et al. 2018).
- Missing data reduces precision but does not cause bias, provided that missingness is independent of genotype (Peel et al. 2013).
- Rare alleles can upwardly bias $\hat{N}_e$ but this can be controlled by setting an allele frequency cutoff (Waples & Do 2010).
- Effects of age structure can be accounted for based on the species' life history traits (Waples et al. 2014).
- In metapopulations (see Chapter 18), the gametic disequilibrium method estimates local N_e unless migration rate is relatively high (Waples & England 2011; Gilbert & Whitlock 2015). For continuously distributed populations, samples taken from within the breeding window estimate Wright's neighborhood size (Neel et al. 2013).
- The gametic disequilibrium method can rapidly detect small N_e associated with a bottleneck (England et al. 2010).

Genomics-scale datasets create new opportunities, as well as future challenges for the gametic disequilibrium method. For example, linkage downwardly biases estimates of contemporary N_e, unless it is accounted for based on genome size (Waples et al. 2020). On the other hand, as detailed linkage information becomes more readily available for nonmodel species, it is increasingly feasible to estimate N_e back in time, based on analogs to Equation 10.11 that include a term for recombination rate (Hollenbeck et al. 2016; Lehnert et al. 2019; Santiago et al. 2020). For large numbers of SNPs (10^3–10^6), the theoretical precision of the gametic disequilibrium method becomes arbitrarily high; however, these pairwise comparisons are not independent, and exactly how much this reduces precision is difficult to quantify (Waples et al. 2016). Preliminary results suggest that in most cases use of more than a few thousand SNPs does not help much to reduce variance of $\hat{N}_e$.