Modeling the Distribution of Crossovers and Interference with Mice DNA Data

by

© **Shuhana Shahnaz**

A practicum submitted to the School of Gradate Studies in partial fulfillment of the requirements for the degree of Master of Applied Statistics.

Department of Mathematics and Statistics

Memorial University

April 2016

St. John's, Newfoundland and Labrador, Canada

Abstract

Chiasma and crossover are two related biological processes of great importance in the understanding genetic variation. The study of these processes is straightforward in organisms where all products of meiosis are recovered and can be observed. This is not the case in mammals. Our understanding of these processes depends on our ability to model them. In this study I describe the biological processes that underline chiasma and crossover as well as the two main inference problems associated with these processes: i) in mammals we only recover one of the four products of meiosis and, ii) in general, we do not observe where the crossovers actually happen, but we find an interval containing type-2 censored information. NPML estimate was proposed and used in this work and used to compare chromosome length and chromosome expansion through the crosses.

Acknowledgement

I would like to express my gratitude to the many people who helped me to finish this thesis. First, I would like to thank Professor JC Loredo-Osti for providing me the opportunity of taking part in Master of Applied Statistics program. I am very grateful for his help, guidance, kindness and financial support throughout my entire program of study. I would also like to acknowledge my thesis examiners Dr. Taraneh Abarin and Dr. Yildiz Yilmaz. I am gratefully indebted to them for their valuable comments on this thesis. I would also like to thank Dr. Elena de La Casa-Esperón for providing me with the data and suggesting the problem. My greatest gratitude goes to Nada Rose who patiently revised and corrected my thesis for English. I would also like to thank Dr. Chris Radford, Dr. Zhaozhi Fan, Leonce Morrissey, Jennifer Dawe, Kimberly Lucas, Lawrence Greening, Ros English and my classmates at Department of Mathematics and Statistics for their continuous support. My deepest gratitude goes to my parents for their love and unconditional support throughout my life and my studies. I must express my very profound gratitude to my husband for providing me with support and encouragement throughout my years of study and through the process of researching and writing this thesis. Finally, a special thanks goes to my daughter who sacrificed a lot of her childhood time giving me space to study. This accomplishment would not have been possible without their help. Thank you.

Statement of contribution

This thesis is written by Shuhana Shahnaz, under the supervision of Dr. JC Loredo Osti. Memorial University as a whole has a big contribution behind the thesis writing process. The problem stated on this thesis was suggested by Dr. Elena de La Casa-Esperón and she also provided the data.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Mouse genetics research probably started in the 1860s. Mouse genetics is a valuable tool to understand fundamental biological processes like recombination, crossing over and interference in mammals. Mice are ideal for modelling complex human diseases as well as for drug efficacy testing. A lot of study has been done on the genetics of humans, mice and other organisms. Few of the related studies are described in brief in the following paragraphs.

In humans, strong positive crossover interference has been found. A study on 10 control men using methods which allow for direct identification of frequency and location of crossovers in specific chromosomes of pachytene cells were used [23]. Fitting the frequency distribution of intercrossover distances using the gamma model revealed that interference level varies significantly across the whole genome in the human male. Significant interindividual, inter- and intra-chromosomal variation in interference levels was observed in human males, with inter-arm interference levels providing the major contribution to interchromosomal crossover interference. Interference was observed to act across the centromere and smaller chromosomes exhibited stronger interference [23]. Other studies revealed that there is a possibility of two

crossover pathways in humans [13]. A study on both sexes, where the genotype data were taken from more than $8,000$ polymorphisms in 8 CEPH (Center for the Study of Human Polymorphisms) families, found evidence that both sexes have the same level of interference and there is evidence of individual variation in interference among the female. This study also showed that the gamma model fit far better than the four versions of the count-location model concluding that the gamma model is better for modeling interference. Comparing the five models, analysis of intercrossover distances showed better results than analysis of crossover counts [2].

Several works have been done in mouse genetics. Analysis of the distribution of immunofluorescent foci of mice, the foci that is visualized when protein complexes are involved in crossover, found strong interference among MLH1 foci in pachytene, among MSH4 foci in late zygotene and among replication protein A (RPA) foci. MSH4 foci and RPA foci both mark interhomolog recombinational interactions, most of which do not yield crossovers in the mouse. Interference in mice is not specific to crossovers and crossover interference occurs in two successive steps [4].

To understand the crossover process and factors that determine the location and relative activity of hotspots locations, 5,472 crossover events along mouse Chromosome 1 arising in 6,028 meioses of male and female reciprocal F1 hybrids of C57BL/6J and CAST/EiJ mice were observed [27]. It was found that a small number of the most active hotspots were responsible for the majority of crossovers. Overall female mice have a higher number of crossovers than the males. Regional crossover rate, hotspot position and activities are different in males and females. Crossover also depends on regional positioning relative to the chromosomal ends and local gene content, parental imprinting, and hotspot position [27]. From similar dataset it was found that crossover is regulated on at least three levels, chromosome-wide, regional, and at individual hotspots, and the regulation levels depend on sex and genetic background

and not on gene content [1]. Shorter genomic interference distance in females than in males is the reason for difference in crossover rates in both sexes in mice, but the fundamental processes that regulate positioning of multiple crossovers along the chromosomes appear to be the same in the meiosis of both sexes [28]. Comparison of crossover in meiocytes from XY sex-reversed and XO females with that from XX female and XY male mice reveals that rate and pattern of crossover in XY and XO oocytes were virtually identical to those in normal XX females, which indicates that sex, not genotype, is the primary determinant of meiotic recombination patterns [25].

Of the three components of interference, distribution of the number and location of chiasmata are well modeled but still a lot of work needs to be done for chromatid interference. Several observations suggest that interference depends on genetic distance rather than physical distance. Foss, Lange, Stahl and Steinberg (1993) suggested a model in which interference is related directly to genetic distance. It is well known that chiasma event (C) can be resolved either with crossing over (C_x) or without crossing over(C_o). This model predicted that tetrads with close double crossovers should be enriched for intervening C_o events compared to general population or tetrads with no crossovers. Predictions from this model have been compared with data from Drosophila, Neurospora and yeast. The model accurately predicts data from Drosophila and Neurospora but central prediction of the model is not fulfilled in yeast [7, 8]. The chi-square model (or gamma model) for the occurrence of crossovers, was first proposed in the 1940's as a description of interference that was mathematically understandable, but there was no biological basis. Afterwards, this model was justified from a biological perspective in different studies. Zhao, Speed and McPeek (1995) derived the probability for single spore or tetrad joint recombination pattern using the chi-square model under the assumption of no chromatid interference. Maximum likelihood is then used to estimate the chi-square parameter and genetic distances.

The model applied to the data of different organisms and gave some insight into the underlying crossover process [42].

Chromatid interference and chiasma interference are generally assumed to be absent in analyzing genetic data although crossover interference has been observed in almost all organisms studied. Zhao and Speed (1998) suggested Markov models for both chromatid interference and chiasma interference to capture the main features of the genetic data [41]. The model presented by Teucher, Brockmann, Rudolph, Swalve and Guiard (2000) assumes that chromatid interference acts only in the neighborhood of a chiasma. They applied the model to three sets of data and found their model obtained a better fit compared to the model that assumes no chromatid interference. When they allowed heterogeneity of chromatid interference in the model, a further improvement in fit was achieved [39].

In this study, five different crosses of mice are considered in an effort to understand crossover process and interference, as well as view interferences from different angles using various statistical tools. NPML estimate was proposed and used in this work and used to compare chromosome length and chromosome expansion through the crosses.

Some biology background that is important to understand this study is provided in the following chapter. This chapter will give brief description of meiosis, double strand break process, recombination, crossover rate and interference. In the next chapter methodology of modeling interference is depicted. This chapter includes various statistical and mechanistic models that have been suggested over the year as well as detail description of some statistical tools to identify the presence of interference. Chapter 4 describes the data sets used in this study and the way the data sets were edited. Next chapter provides the methods that was applied to analysis the data. Results of data analysis and comments on interference in the data are also made in

this chapter. The last chapter gives an over all idea about the study along with limitations.

Chapter 2

Biology background

2.1 Introduction

Genes are located at specific sites, named loci, along a chromosome. Alleles are defined as variants of a gene. In diploid organisms, except for the sex chromosomes, every locus has two alleles and this pair constitute the genotype at that autosomal locus. If the two alleles are identical, then we say it is a homozygous genotype; otherwise, it is a heterozygous genotype. Genotypes may not be observable, what is observable is phenotype. In the gene mapping studies, a number of marker loci on the same chromosome are typed and used as the basis for inference about a putative gene. When these loci are simultaneously followed in a pedigree, the phenomenon of recombination can often be observed [21]. In fact, recombination is the phenomenon on which gene mapping is based. Amongst sexual organisms recombination takes place during the formation of gametes at meiosis. In this chapter, I will review some of the basic concepts and processes related to recombination.

2.2 Meiosis

Meiosis is a process necessary for sexual reproduction and variability. It refers to a special type of two stage cell division, grossly divided in two stages: meiosis I and meiosis II. In a cell, the chromosomes come in pairs. For each pair set, one chromosome is of maternal origin and the other is of paternal origin. The chromosomes in each pair are called homologous chromosomes. When a cell has two complete sets of each chromosome it is called diploid. Before meiosis begins, the cell goes through interphase where the DNA of each chromosome is replicated resulting in each chromosome having two sister chromatids. These sister chromatids are identical copies of a single chromosome joined by a common centromere. The maternally and paternally derived sister pairs then align to form a bundle of four chromatids. There is evidence of a direct association between DNA replication and double strand break (DSB) formation with the key aspect of such an association being faithfulness of replication since an anomalous gamete may result in a defective organism. So, to prevent anomalies there are several replication checking mechanisms that monitor the process from the beginning of replication until resolution of DSB and their various intermediates [24, 38].

Meiosis I starts with prophase I, which is the longest meiotic phase and is characterized by the emergence of DSB and the undergoing of synapsis by chromosomes. During synapsis, homologous chromosomes from the male and female pair up to form the so-called bivalents. The bivalents exchange chromosomal segments between non-sister chromatids reciprocally (COs) or non-reciprocally (non-COs). In metaphase I, after crossing-over, the chromosomes line up in the middle of the cell. At this time, they can organize themselves in different ways. Crossovers, along with their orientation, give a huge amount of variation to DNA in meiosis. When the chromosomes are

done organizing themselves, the centrosomes move to either side of the cell and the spindle fibres are attached to the centromere of each of the homologous chromosome pair. During anaphase I, these pairs are pulled apart so that the homologous chromosomes are moved to either side. Then, in telophase I, new nuclei are formed in each side and the rest of the cell is divided during cytokinesis. This concludes meiosis I [43, 30].

Meiosis II starts with prophase II where there is no exchange of chromosomal segments but the nuclear envelope breaks down. In metaphase II the chromosomes line up again in the middle of the cells and spindles will attach to each of the centromeres. During anaphase II, centromeres are cleared and the sister chromatids pulled apart. Finally, during telophase II, cytokinesis takes place providing four new daughter cells. In a nut shell, meiosis consists of two consecutive cell divisions after chromosome replication in a cell containing sets of two chromosomes that results in the formation of four cells, each having only one set of single chromosomes, i.e. a set of four haploid cells. These cells are the gametes. Normally, only one of these gametes will go to the next generation; one from each parent. So, as a consequence of meiosis, the offspring are different from their parents as well as different from each other.

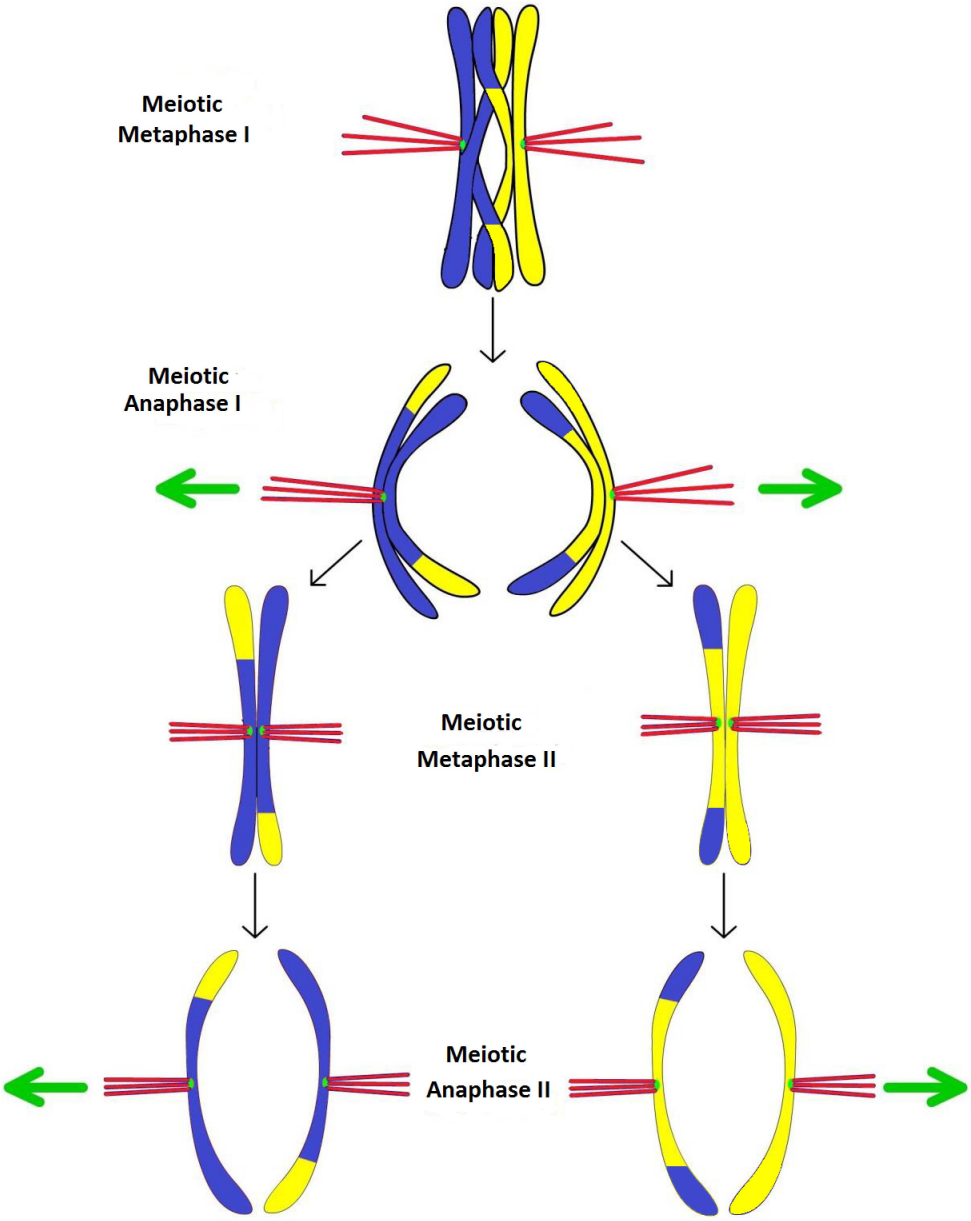


Figure 2.1: Scheme of Meiosis

2.3 Double Strand Break (DSB) process

DSB is a phenomenon that happens in early meiosis. A description of the molecular model for DSB and gap repair will give a clear idea about how the double strand gap is repaired and triggered the evolution of either crossover or non crossover event (see Figure 2.2). Gap repair using homologous chromosomes ensures that those genetic materials are restored. The process begins when a DSB appears in one homologue (a). Then, at the breaking point, an exonuclease (a type of enzyme which cleaves nucleotides) removes few nucleotides from 5' to 3' end (b).

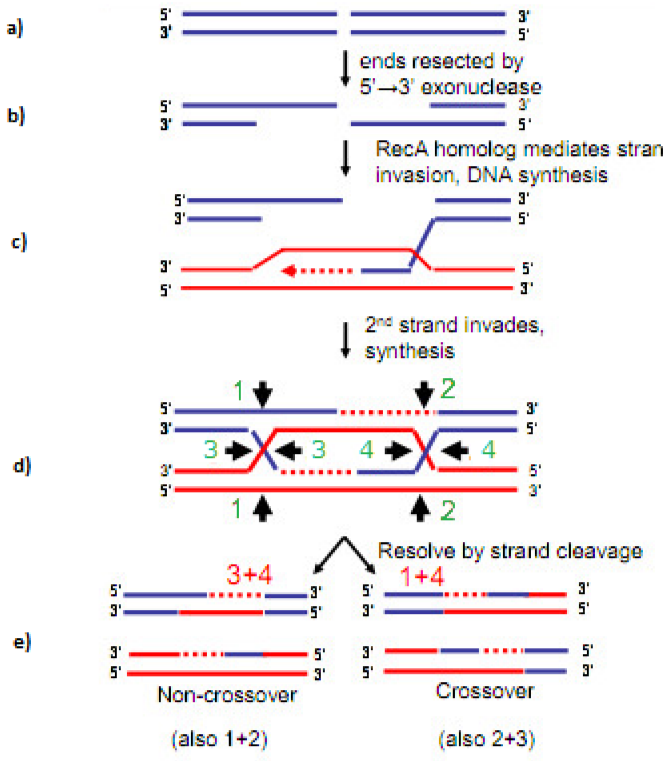


Figure 2.2: Double strand break repair crossover model

This chewing off stops when it reaches a particular hotspot sequence where a 3'-OH terminus will be produced. Then the 3' ends of both of the strands will come and check whether there is a complimentary base to pair with. If one of them finds it,

then the 3' terminus invades the complementary sequence in the homologue forming a displacement loop (D-loop), (c). This process is also known as single cell invasion. Then the D-loop will be enlarged by repair synthesis from the invading 3' end, until the other 3' end can anneal to complementary single stranded sequences. A second round of repair synthesis from the second 3' terminus completes the process of gap repair. The process of producing branches is denoted by dots in the figure. Thus gap repair by two rounds of single strand synthesis coupled with ligation of the break ends leads to connection of the homologies by two four-way DNA junctions; such an intermediate DNA structure is known as the double Holliday junction (d). Cleavage of the double Holliday junction leads to chiasma formation that can be resolved either by disengagement of the two pairs of strands, i.e. by cleavage at the two junctions resulting in a non crossover, or via its endonucleolytic cleavage mediated by resolvases resulting in a crossover event [17].

2.4 Recombination as a biological process

Recombination is a process in which chromosomes are split and randomly re-assorted during Anaphase I of meiosis. Genetic recombination makes changes in the genetic make up of the individuals. It begins with the appearance of DSBs which are mediated by Spo11 or some of its analogs. During synapsis, DSBs occur and these breaks are repaired by using homologous sequences as a template. As mentioned earlier, a DSB will result in either a reciprocal exchange, i.e. a crossover event, CO, or a gene conversion event, non-CO. If crossovers along each chromosome were absent, each gamete would be just one of the 2 parental homologes for each chromosome and with n pairs of homologes, there would be 2^n possible gametes. Furthermore,

in such a case, parental chromosomes would remain intact through generations with little or no variation. Crossover thus enhances the possibilities of genetic diversity by shuffling the origin (paternal or maternal) of blocks of DNA on the gamete. Most certainly this shuffling happens under selection constraints and it is well regulated: too little shuffling may reduce the adaptability to changing environments and too much shuffling may break the favorable associations of alleles. When considering that the sizes of genomes of various organisms can differ by factors of 100 or more, while their recombination rates vary typically only by factors no longer than 3, we might see the importance of a regulated recombination rate. The role that COs play for faithful homologous replication is very important since they ensure proper orientation and spindle force. Once the bipolar forces are balanced, successful segregation happens when homologous chromosomes separate the COs. For proper segregation the number and distribution of COs is important. When gametes are formed, it is crucial that enough COs are placed throughout the genome [32].

It is clear that DSBs are precursors of COs (reciprocal exchanges between non-sister chromatids) and non-CO (non-reciprocal exchanges), but there are transitory molecules/stages, like displacement loops and double Holliday junctions, in the DSB model of recombination that seem to be associated only to CO formation. It seems that the decision of forming a CO or a non-CO is made after DSB formation but before Holliday junctions and single-end invasion formation [32].

For non-CO, the synthesis dependent strand annealing diversion has been suggested. In general, DNA repair relies in using a non-sister chromatid as a template. Whenever this is not possible, sister chromatids may be used to carry on the repair, avoiding CO formation which may lead to error-prone separation [32].

The pattern of crossover positions along the chromosome gives clues about how

crossovers were formed. There are at least two distinct crossover formation pathways to achieve synapsis and repair the double-strand-breaks. Each one is mediated and monitored by a specific group of protein. These two pathways are i) interfering crossover that depends on the Msh4-Msh5 complex, where crossovers are subject to special regulation i.e., it controls the position of COs, and ii) non-interfering crossover, mediated by the Mus81-Eme1 complex, where crossovers are not subject to such regulation. This means that at least a small portion of crossovers may form without interference. Figure 2.3 shows how Mus81 protein acts earlier in the crossover reaction to cleave the D-loop. A species might have only the interfering or strictly the non-interfering crossover pathway or a combination of both, depending on how they recombine. When both interfering and non-interfering are present, it has been hypothesized that some of the crossovers (putatively, those in the pairing pathway and occurring in early meiosis) occur independently, i.e., in absence of interference, whereas the remaining ones (putatively, those crossovers in the disjunction pathway and occurring in late meiosis) are subject to interference [13]. The majority of crossover events are thought to be generated by an interference-sensitive pathway. Interference strength and the proportion of the non-interfering events may vary for different gene families. Some studies showed that the proportion of non-interference crossovers varies across species usually from 5 to 30% [9]. Both of these pathways as well as the non-CO pathway begin with DSBs.

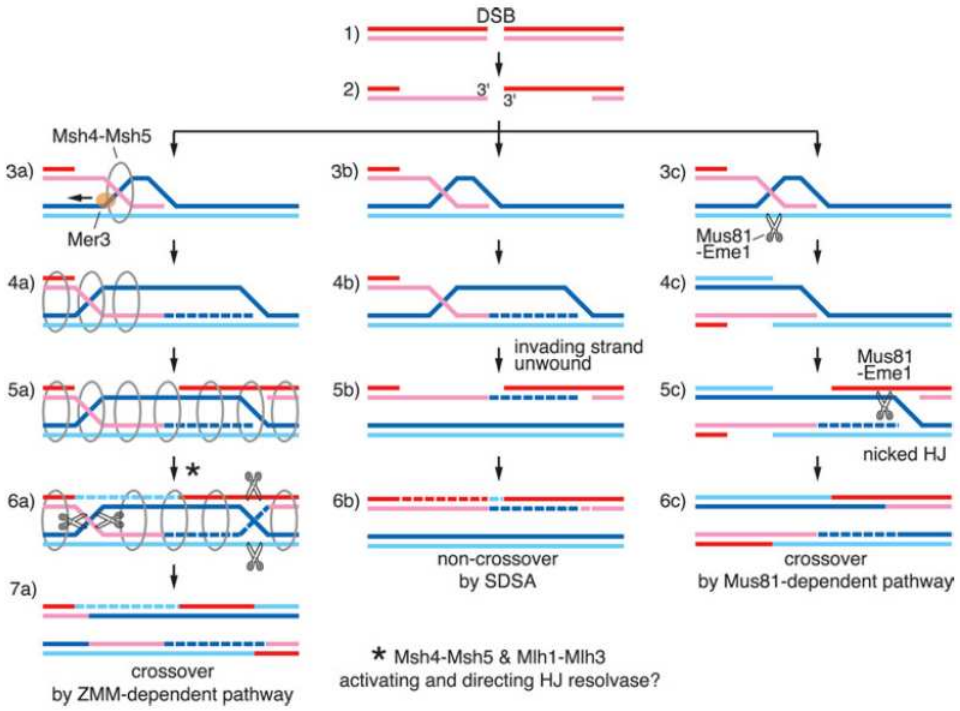


Figure 2.3: Interfering and non-interfering crossover pathways (picture is taken from a PhD thesis by Roy 2014 [32])

2.5 Crossover rates variation

Crossover rate varies among species and even within species. This suggests that sex and crossover interference play an important role in crossover rate. Mammals show a wide range of crossover rate and among mammalian species, mice possess an unusually low crossover rate [6]. Crossover rates follow similar patterns on most mouse chromosomes except for chromosome 19, which is the shortest one. Depending on the genetic background and sex, crossover rate and placement varies through the entire chromosome, at both the regional and local levels. The megabase scale control on crossover is considered as a regional level control. The regional distribution is determined by the positioning of the crossover with respect to centromere-telomere axis and interference. Positioning of individual hotspots of crossover is considered as a local level of control. In many species most crossover events occur in very limited intervals along the genome known as crossover hotspots. Hotspots do not affect the global crossover pattern. Most of the hotspots in mammals are 1 to 2 kb long and are unevenly distributed along the chromosome. Both sexes tend to have similar sets of hotspots, however, the level of activity of hotspots differs among regions for different sexes [27, 16, 1].

For some species crossover does not occur at all in one of the sexes. Crossover rates are significantly higher in females than in males in human, dog, pig, zebrafish, and most mouse strains. On the other hand, the sheep and wallaby male crossover rate is higher than that of the female. These differences are more pronounced in certain chromosomal regions. Mammalian meiosis differs with sex in various ways which could result in differences in crossover. Although meiosis takes much longer to complete and each stage of meiosis lasts longer in mammalian females than in males, temporal differences in initiation and progression of meiosis are likely to have

little impact on differences in crossover rate. On the other hand, it is almost certain that differences in pairing and synapsis of homologs, as well as differences in intensity of crossover interference in sexes, causes differences in crossover rate in males and females.

Despite the fact that some regions of a chromosome tend to recombine much more than others depending on sex, if there is no interference, the relationship between genetic and physical distance remains approximately monotonic. Genetic distance can be used to compare the genetic similarity between different species and divergence between different sub-species. Genetic distance between two loci is defined as the average number of crossovers within that interval per meiosis. Its unit is Morgan, although the centimorgan is more commonly used. In the absence of interference, one $Morgan$ indicates that the segment has on average one crossover per gamete. Physical distance refers to the count of consecutive base pairs between two positions in a DNA sequence (megabase). Its unit is base pairs (bp). A base pair is two chemical bases bonded to each other forming a rung of the DNA ladder. A $Morgan$ may contain from several hundreds of bp to a few mega base pairs (mbp), depending upon the region of the chromosome. The physical distance covered by a centimorgan also varies with species. Nonetheless, there is an almost constant relationship between physical and genetic distance despite the proved presence of crossover hotspots. [25, 1].

2.6 Crossover interference

Crossovers along chromosomes are not events that happen completely at random. In fact, they may tend to be evenly spaced. This departure from "absolute" randomness is called interference. There are two types of interference: crossover, or position

interference, and chromatid interference (CI). Chiasma interference and crossover interference refer to the same underlying process. Both of them refer to the suppression of the exchange of genetic material between paternal and maternal chromosomes of a given individual. This kind of interference is also commonly called genetic interference, but to avoid confusion, we will not use this term. Crossover interference is a process that occurs in many species, including mice. It affects the distribution of the number and location of crossovers. In general, it has been observed across species that formation of a crossover prohibits, or interferes, with the formation of a crossover in its neighborhood in both directions along a chromosome. From this phenomenon the process takes its name: crossover interference. Because of this interference, two crossovers tend not to be close to each other resulting in widely spaced and, to some extent, infrequent crossovers along chromosomes. Crossover interference can exercise its effect across whole chromosomes or be localized to some particular regions, like one of the chromosome arms. In different species, crossover interference acts over widely varying DNA blocks. In mice and humans, it is on the range of tens of *mbp*. Within a specific chromosome region, interference varies depending on the overall size and structure of the chromosome. The reason behind female mice chromosomes accommodating more multiple crossovers than male mice, is the shorter crossover interference distance in females compared to males when measured in *mbp*. For the same reason, triple crossovers are somehow common in female meiosis but very rare in males. Higher inter crossover distance indicates stronger interference. In the absence of interference, crossovers would have been distributed at random, i.e. their distribution would be uniform along the chromosome with respect to each other. However, because of crossover interference, there are constraints on the resolution of crossovers and even on the total number of crossovers. The basic constraint is that each pair of neighboring crossovers must have a certain number of non-crossover

events between them. During meiosis, each chromosome receives at least one crossover (the obligate crossover), but their density is limited by crossover interference which results in a strong tendency for shorter chromosomes to have more crossovers per unit of length than larger chromosomes do. Too many crossovers can be harmful from an evolutionary view point. Hence, it can be said that interference works as a mechanism to balance between too many or too few crossovers so that the mean number of crossovers per chromosome remains modest.

Chromatid interference refers to the suppression of exchanges by affecting the chances of a chromatid for participation in multiple exchanges or the chromatid sampling probabilities. It is a process where two nonsister chromatids forming a crossover, affect which two chromatids will be involved in the other crossovers in the same meiosis. [32, 25, 1, 7, 27].

Chapter 3

Methodology

To understand the models of crossover interference properly we need to understand recombination fraction, genetic map distance and genetic map function.

3.1 Recombination Fraction:

The recombination fraction between two loci on the same chromosome is the ratio of the number of recombinant gametes to the total number of gametes produced. Two loci on a gamete are said to be recombinant whenever there is an odd number of crossovers between them. Let us assume that the two loci are at positions a and b and the number of chiasmata occurring on the interval $[a, b]$ of the four strand chromatid bundle is denoted by $N_{[a,b]}$. In the absence of interference, Mather's formula defines the recombination fraction θ as,

$$\theta = \frac{1}{2}Pr(N_{[a,b]} > 0) = \frac{1}{2}[1 - Pr(N_{[a,b]} = 0)] \tag{3.1}$$

From this equation it is clear that θ lies between 0 and 1/2 and it is a non-decreasing function of $|b-a|$ [21].

3.2 Genetic map distance:

The genetic map distance or genetic distance, d, that separates two loci at positions a and b is defined as the expected number of crossovers on $[a, b]$ per gamete or equivalently as half the number of chiasmata on $[a, b]$ in the four strand bundle. Notice that the genetic distance is not a function of how physically far apart two loci are, but of how many crossovers we may expect to occur amongst them. The unit of this distance is Morgan. For a short interval, say, for (0,0.05), $\theta \approx d$, since in such a case, $E(N_{[a,b]}) \approx Pr(N_{[a,b]} > 0)$ [21].

3.3 Genetic map function:

Genetic map function, M, is a monotonic function that relates recombination fraction, θ, and genetic map distance, d, between pairs of loci along a chromosome by $\theta = M(d)$ [21]. From this relation, the genetic distance can be obtained by $d = M^{-1}(\theta)$. The simplest case in the map function could be $\theta = d$, with inverse $d = \theta$. This is quite satisfactory for small θ and d, say, in the interval $(0, 0.05)$.

The recombination fraction and genetic distance of an interval differ only when there is a nonzero chance of multiple crossovers occurring in the interval. With the increase in the interval size, the chance of crossover also increases. If we denote the distribution of the number of crossovers, N, in a particular interval by $q_0, q_1, q_2, ...$, where, $q_n = Pr(N = n)$, i.e., the expected proportion for a gamete of having n exchange points (crossovers) in the interval, then recombination fraction can be written as the probability of an odd number of crossovers in the interval,

$$\begin{aligned} \theta &= \Pr(N = 2n+1, \qquad \text{for } n \in \mathbb{N}) \\ &= q_1 + q_3 + q_5 + ... \end{aligned} \tag{3.2}$$

and the genetic distance, by definition becomes,

$$\begin{aligned} d &= E(Y) \\ &= q_1 + 2q_2 + 3q_3 + ... \end{aligned} \tag{3.3}$$

Suppose that in the interval the number of chiasmata on the four strand bundle follows Poisson distribution with mean d, i.e., under no interference,

$$q_n = \frac{e^{-d}d^n}{n!} \tag{3.4}$$

then, the recombination fraction can be expressed in terms of genetic distance as,

$$\theta = e^{-d} + \frac{e^{-d}d^3}{3!} + ... = \frac{1}{2}(1 - e^{-2d}) \tag{3.5}$$

and its inverse as,

$$d = -\frac{1}{2}\ln(1 - 2\theta) \tag{3.6}$$

This map function is known as Haldane's map function. Though this function entails no chiasma interference, it is widely used for its computational simplicity.

Absence of both chiasma interference and chromatid interference are necessary and sufficient conditions for the chiasma process to follow a Poisson process, which imply Haldane's formula is a simple consequence of assuming that chiasma process is a Poisson process [33, 22].

3.4 Modeling crossover interference

The analysis of crossover interference is complicated when there are two formation mechanisms involved; interfering and non-interfering. A number of statistical analyses has been done to find out the relationship between linkage map distance and crossover interference in creatures with both types of crossovers [34, 37, 7, 36, 35]. These analyses were based on the assumption that noninterfering crossovers are distributed along the chromosomes independently of the interfering crossovers and each other, while interfering crossovers are distributed with respect to each other according to an Erlang or Gamma distribution. The purpose of these analyses was to estimate the numbers of crossovers of both kinds and the values of indices that reflect the degree of modality in the frequency distribution of intercrossover distances for the interfering crossovers [34].

The traditional indicators of interference give an indication of the strength of interference without consideration of the relative frequencies of crossovers occurring from two pathways [34]. Because of this, traditional indicators may not capture the salient features of the process and can potentially provide misleading results. The work presented here considers the two crossover formation pathways in greater detail and tries to model it. Modeling crossover interference can provide insights that cannot be reached by looking only at inter crossover distributions. Various models have been suggested over the years, some statistical and others mechanistic.

Statistical models are

- Standard non-interference,
- Count location model of chiasma,
- Renewal process model of chiasma,

- One pathway and
- Two pathway

Mechanistic models are

- The polymerization model of chiasma interference and
- Stress model of chiasma interference

Each of these models has their own merits and caveats. However, the lack of a clear likelihood formulation for mechanistic models leads us to prefer the statistical models [32]. Statistical models are mainly based on the distribution of genetic distances between successive crossovers [9].

Statistical models considering chiasma interference and no chromatid interference:

Count location model [21]: This model considers the total number of chiasma located independently along the bundle according to some common continuous distribution.

Let the total number of chiasmata along the four strand bundle be N and its distribution $q_n = Pr(N = n)$. Say, the individual chiasmata are located independently following a distribution F and let λ be the expected number of chiasmata along the four strand bundle. Then the map length, d, of an interval $[a, b]$ can be shown to be,

$$d = \frac{1}{2}\lambda[F(b) - F(a)] \Rightarrow F(b) - F(a) \quad = \quad 2d\lambda^{-1}$$

If there are $k+1$ loci along a chromosome, then there would be k adjacent intervals which are, $I_1, I_2, ..., I_k$. Let us consider a subset, $S \subset \{1, 2..., k\}$ of intervals where the gamete is recombinant. Then the generating function of N can be written as,

$$Q(s) = \sum_{n=0}^{\infty} q_n s^n$$

Then, the Mather's formula for recombination fraction can be expressed as,

$$\begin{aligned}
\theta = \frac{1}{2} Pr(N_{[a,b]} > 0) &= \frac{1}{2}[1 - Pr(N_{[a,b]} = 0)] \\
&= \frac{1}{2} \sum_{n=0}^{\infty} Pr(N = n)[1 - Pr(N_{[a,b]} = 0 | N = n)] \\
&= \frac{1}{2} \sum_{n=0}^{\infty} q_n \{1 - [1 - F(b) + F(a)]^n\} \\
&= \frac{1}{2} - \frac{1}{2} Q(1 - 2\lambda^{-1} d) \qquad (3.7)
\end{aligned}$$

Renewal process model [42, 41]: This model considers that the chiasmata arise as a stationary renewal process.

If chiasma events are randomly distributed along the four strand bundle, then every chiasma event either resolves in a crossover C_x or not C_o. When a chiasma event resolves as a C_x, the next m chiasma events must resolve as C_o events, and after m C_o's the next chiasma event must resolve as a C_x, i.e., the chiasma events resolve in a sequence $...C_x(C_o)^m C_x(C_o)^m....$ Thus the model is represented in the form $C_x(C_o)^m$. Let $\mathscr{A}_1, \mathscr{A}_2, ..., \mathscr{A}_{(n+1)}$ be a set of markers along a chromosome. Then, there are $n+1$ parameters that need to be specified, namely, m and the genetic distances between each consecutive pair of markers, $d_l, d_2, ..., d_n$.

Because the chiasma events are assumed to be randomly distributed along the four-strand bundle and if the number of chiasma events, s in an interval of genetic

length d follows the Poisson distribution with parameter λ, the chance of getting s number of chiasma events is, $\frac{e^{-\lambda}\lambda^s}{s!}$, $for \quad s = 0, 1, ...$ If $p = m + 1$, then only $\frac{1}{p}$ of these chiasma events resolve as a crossover event. Under the assumption of no chromatid interference, each strand has probability $\frac{1}{2}$ of being involved in a crossover. So, for s chiasma events, on average each strand will have $\frac{s}{2p}$ crossovers. Since the genetic distance is defined to be the expected number of crossovers on a single strand, under no interference, the genetic distance d and the Poisson parameter λ are related by $d = \frac{\lambda}{2p}$ i.e., $\lambda = 2pd$.

Suppose markers $\mathscr{A}_1, \mathscr{A}_2, ..., \mathscr{A}_n$, are laid out from left to right, and the chiasma events also occur from left to right. To keep things simple only two markers, $\mathscr{A}_1$ and $\mathscr{A}_2$, are considered at first. As the process is stationary, the first chiasma event to the right of $\mathscr{A}_1$ has an equal chance of resolving as any of the $m+1$ elements of $C_x(C_o)^m$. Say, k_1 C_x's are between $\mathscr{A}_1$ and $\mathscr{A}_2$. Depending on the number of C_o's before the first C_x to the right of $\mathscr{A}_1$ and the number of C_o's between $\mathscr{A}_2$ and the nearest C_x to the left of it, k C_x's between the two markers can occur in p^2 possible ways. The number of C_o can vary anywhere from 0 to $p-1$. Therefore, the chance of k_1 number of C_x's between $\mathscr{A}_1$ and $\mathscr{A}_2$ with Poisson parameter λ_1 can be computed as

$$\frac{e^{-\lambda_1}}{p} \sum_{i=1}^{p} \sum_{j=0}^{p-1} \frac{\lambda_1^{pk_1-p+i+j}}{(pk_1 - p + i + j)!}. \tag{3.8}$$

The sum can be written in a matrix product form as

$$\frac{1}{p} \mathbb{1} \boldsymbol{D}_{k_1}(\lambda_1) \mathbb{1}', \ where \ \mathbb{1} = (1, 1, 1...., 1). \tag{3.9}$$

Each element in the first column of the above mentioned $\boldsymbol{D}_{k_1}(\lambda_1)$ matrix corresponds to the last chiasma event between $\mathscr{A}_1$ and $\mathscr{A}_2$ being a C_x; the second column corresponds to the last chiasma event being the first C_o after the k_1th C_x, the jth column to the jth C_o after the k_1th C_x. Therefore, the sum of the jth column multiplied by $\frac{1}{p}$ is the probability that there are k_1 crossovers between $\mathscr{A}_1$ and $\mathscr{A}_2$, and the last chiasma event is the $(j-1)$th C_o after the k_1th C_x. These probabilities can be defined as

$$(p_{k_1}^1 p_{k_1}^2 ... p_{k_1}^p) = \frac{1}{p} \mathbb{1} \boldsymbol{D}_{k_1}(\lambda_1). \tag{3.10}$$

Now consider three markers, $\mathscr{A}_1, \mathscr{A}_2$ and , $\mathscr{A}_3$ and let the first chiasma event to the right of $\mathscr{A}_2$ be the lth C_o after a C_x, then the probability of k_2 crossovers between $\mathscr{A}_2$ and $\mathscr{A}_3$ with Poisson parameter λ_2 is

$$e^{-\lambda_2} \sum_{i=1}^{p} \frac{\lambda_2^{pk_2+l-i}}{(pk_2+l-i)!}. \tag{3.11}$$

Since the probability of l C_o's between $\mathscr{A}_2$ and the first C_x and the probability of the last chiasma event between $\mathscr{A}_1$ and $\mathscr{A}_2$ is the $p-l-1$th chiasma event after a C_x is the same and the probability is $p_{k_1}^{p-l-1}$, then the chance of k_1 crossovers between $\mathscr{A}_1, \mathscr{A}_2$ and k_2 crossovers between $\mathscr{A}_2$ and $\mathscr{A}_3$ can be written as

$$e^{-\lambda_2} \sum_{i=1}^{p} p_{k_1}^{p+l-1} \sum_{j=0}^{p-1} \frac{\lambda_2^{pk_2-p+i+j}}{(pk_2-p+i+j)!}. \tag{3.12}$$

In matrix form it can be written as

$$(p_{k_1}^1 p_{k_1}^2 ... p_{k_1}^p)\mathbb{1}\boldsymbol{D}_{k_2}(\lambda_2)\mathbb{1}'. \tag{3.13}$$

From equation (3.3) and (3.6),

$$\frac{1}{p}\mathbb{1}\boldsymbol{D}_{k_1}(\lambda_1)\boldsymbol{D}_{k_2}(\lambda_2)\mathbb{1}'. \tag{3.14}$$

Generalizing the probability for n intervals can be written as

$$\frac{1}{p}\mathbb{1}\boldsymbol{D}_{k_1}(\lambda_1)\boldsymbol{D}_{k_2}(\lambda_2)...\boldsymbol{D}_{k_n}(\lambda_n)\mathbb{1}'. \tag{3.15}$$

This model is also called χ^2 model since the probability distribution of this model is a scaled version of χ^2 distribution.

Statistical models considering chromatid interference:

One pathway: Considers renewal process based on the sampling probability for every chromatid.

Two pathway [13, 14]: Considers a mixture of renewal processes but also takes into account the two pathway of chiasma formation.

Let $d_0, d_1, ..., d_n$ be the intercrossover distances along the four strand bundle of an infinitely long chromosome, where $d_0 + d_1 + ... + d_n = L$ is the length of the chromosome. Since χ^2 distribution is a special case of gamma distribution, then the map length, d, can be written as gamma distributed with rate λ and shape parameter p (where $p = m + 1$ as denoted in the earlier model. m is known as the interference parameter in this model) as

$$f(d|\lambda, p) = \frac{\lambda^p}{\Gamma(p)} d^{p-1} e^{-\lambda d}, \quad d > 0. \tag{3.16}$$

Suppose that in disjunction pathway the probability of occurring a crossover is q. Since the number of crossingover on a tetrad is twice that on a single product, then the rate becomes $2qp$ and the density for intercrossover distance is, $f(d|2qp, p)$.

The distribution of the length to the first crossover depends on the stationarity and not on the start of the chromosome. The probability density of the length to one of the ends, d can be expressed as

$$g(d|q, m) = 2q(1 - F(d|2qp, p)), \quad d > 0. \tag{3.17}$$

where F is the cumulative distribution function of d.

The distribution of the length from the last crossover to the end of the tetrad can be calculated as a right censored distribution as follows

$$1 - F(d|2qp, p), \quad d > 0. \tag{3.18}$$

We do not observe tetrad data, rather we observe single meiotic products. Assuming no chromatid interference, each chromatid has probability $\frac{1}{2}$ of getting each crossover. Then the density for the distances between observed crossovers or intercrossover distance is

$$f^*(d|q,m) = \sum_{k=1}^{\infty} \frac{f(d|2qp, kp)}{2^k}, \quad d > 0. \tag{3.19}$$

The starting has probability $\frac{1}{2}$ of getting the first crossover and then comes the additional intercrossover distances. Using the properties of convolutions the density for the distance from the start of the chromosome to the first crossover is

$$\begin{aligned} g^*(d|q,m) &= \frac{1}{2}g(d|q,m) + \sum_{k=1}^{\infty} \frac{1}{2^{k+1}} \int_d^{\infty} g(x|q,m) f(d-x|2qp, kp) dx \\ &= q(1 - F^*(d|q,m)), \quad d > 0. \end{aligned} \tag{3.20}$$

The distribution of the distance from the last crossover to the end of the chromosome is obtained as a right-censored observation of the intercrossover distances and can be written as

$$1 - F^*(d|q,m), \quad d > 0. \tag{3.21}$$

The probability of having no crossovers, or not having a first crossover, is a right-censored observation from the density of the distribution for the distance to the first crossover and is as follows:

$$1 - G^*(d|q,m), \tag{3.22}$$

where G^* is the cumulative distribution function of g^*.

If we consider all 2^n possible ways to assign the crossovers to the pairing and

disjunction types, then we have to consider two sets of intercrossover distances. Say, $x_0, x_1, ..., x_k$ indicates intercrossover distances for disjunction type of crossovers and $y_0, y_1, ..., y_j$ for pairing type of crossovers, where $j + k = n + 1$ and $x_0 + x_1 + ... + x_k = y_0 + y_1 + ... + y_j = L$.

The probability of the intercrossover distances for disjunction type of crossovers, where $q = 1 - \nu$ i.e., the probability that a randomly chosen crossover being interference is

$$Pr(x_0, x_1, ..., x_k|(1-\nu), m) = \begin{cases} 1 - G^*(d|(1-\nu), m), & \text{if } k = 0 \\ g^*(x_0|(1-\nu), m)[1 - F^*(x_1|(1-\nu), m)], & \text{if } k = 1 \\ g^*(x_0|(1-\nu), m)[\prod_{i=1}^{j-1} f^*(x_i|(1-\nu), m)][1 - F^*(x_k|(1-\nu), m)], & \\ \text{otherwise.} & \end{cases}$$

The probability of the intercrossover distances for pairing type of crossovers where $q = \nu$ i.e., the probability that a randomly chosen crossover being interference free is

$$Pr(y_0, y_1, ..., y_j|\nu, 0) = \begin{cases} 1 - G^*(d|\nu, 0), & \text{if } j = 0 \\ g^*(y_0|\nu, 0)[1 - F^*(y_1|\nu, 0)], & \text{if } j = 1 \\ g^*(y_0|\nu, 0)[\prod_{i=1}^{j-1} f^*(y_i|\nu, 0)][1 - F^*(y_j|\nu, 0)], & \text{otherwise.} \end{cases}$$

Then considering the both pathways and all the 2^n possible divisions of the n crossovers, the probability of the observed pattern would be

$$Pr(d_0, d_1, ..., d_n|\nu, m) = \sum_{(x_0, x_1, ..., x_k), (y_0, y_1, ..., y_j)} Pr(x_0, x_1, ..., x_k|\nu, m) Pr(y_0, y_1, ..., y_j|\nu, 0)$$

Likelihood function can be written as the product of the probabilities of the individual meiotic patterns as,

$$\mathscr{L}(\nu, m|data) = \prod_i Pr[d_0(i), d_1(i), ..., d_n(i)|\nu, m] \tag{3.23}$$

This likelihood function can be maximized to estimate the parameters m and ν.

If the genetic length is known, then both the shape and scale parameters of the distributions do not vary freely. As the genetic length is known then the likelihood function (3.17) becomes a function of only one variable, which is ν and can be optimized more easily.

Standard non-interference models assumes that there is no interference, which is unrealistic. Still this model is widely used as this defect is partially compensated for by its computational simplicity. Count location and renewal process models takes chiasma interference into account but assumes that there is no chromatid interference. These models are better than standard non-interference model as chiasma interference has been observed in almost all organisms. Renewal model was built from biological perspective [21]. Zhao, Speed and McPeek [1995] compared the likelihood ratio statistics of count location model and renewal process model for various organisms and showed that renewal process model performed better than the count location model. But renewal process model has some limitations e.g., it assumes that the parameter m is integer, this model considers crossovers but does not include gene conversions [42]. One pathway and two pathway models consider renewal process as well as both chiasma and chromatid interference. Since interfering and non interfering pathways co-exist in most organisms and two pathway model incorporates both pathways, it is

better to use two pathway model when a data has both pathways of chiasma formation [32]. The distribution of the distance between consecutive crossover that will be discussed in Section 3.6, is an easier way to estimate average genetic distance but does not consider interference. On the other hand, the non-parametric maximum likelihood estimator found through quadratic programming which will be discussed in Section 3.7.3, gives self-consistent estimate considering both chiasma and chromatid interference and it also satisfies certain important constrains. This approach is even better than Turnbull's self-consistent algorithm (Section 3.7.2) as it satisfies all KKT conditions as mentioned in Section 3.7.3.

Mechanistic models:

The polymerization model of chiasma interference [18]:

This model postulates that chiasmata points move as polymerizing signals along the chromosome. It proposes that early structures implied in crossover formation (like double strand breaks) randomly attach amongst and along the synaptonemal complexes of meiotic nuclei. Once attached, each structure has an equal chance per unit of time of initiating bidirectional polymerization reaction. These structures initiating a polymerization reaction will eventually promote reciprocal exchange and finally mature into crossovers. Furthermore, the growing polymers extend from their initiation site blocking the binding of early structures to the synaptonemal complex. As these polymers grow, bound structures that have not yet initiated such a reaction continue to have the opportunity to do so until they are ejected by the advance of a polymer initiated at a nodule located elsewhere on the same chromosome. The ejected structures move into the surrounding medium, where they are either degraded, reattach to an available site on the same synaptonemal complex, or reattach to a site on a different complex. A chromosome that received only a single structure would retain the structure since it could not be ejected, and would be guaranteed to have a

single late nodule and thus a single crossover. The number of chromosome arms with zero crossovers is initially determined by a Poisson distribution based on the average number of early structures. In this model, there are more early structures than late nodules, thus there are fewer chromosome arms with zero crossovers than there would be if the number expected was based solely on the average number of late nodules. The number of chromosome arms with zero crossovers may be further reduced by the relocation of ejected early structures onto chromosome arms that were initially void of structures. These features of our model ensure that virtually all bivalent arms will eventually obtain at least one late nodule, provided a moderate excess of early structures is synthesized. This model was partly inspired by the proposal of Rasmussen and Holm (1978) for a redistribution from random crossover nodules associated with the synaptonemal complex at zygotene to the nonrandom nodules observed.

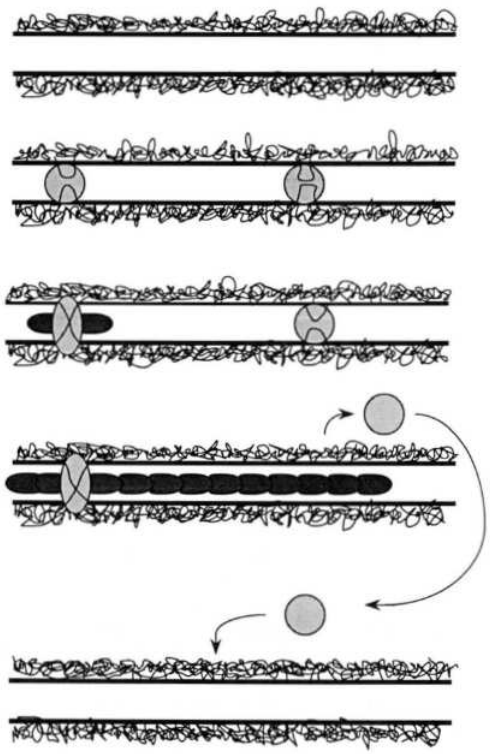


Figure 3.1: A speculative drawing of the model. Early structures (circular) bind randomly to the synaptonemal complex. Some initiate polymerization reactions thus becoming late nodules (oval). The growing polymers eject early structures. The ejected early structures are either degraded or bind to the synaptonemal complex that is free of polymer.

Stress model of chiasma interference [19]:

Based on the fact that in physical systems any local increase or decrease in mechanical stress at one position automatically tends to redistribute outward from that point, this model proposes that chiasma/crossover happens as a result of the stress generated by the normal expansion/contraction of the chromosome. i.e., the model postulates that chiasma/crossover forms to release the stress locally. Stress relief in the neighborhood of a chiasma point dissipates along the chromosome in both directions down the synaptonemal complex axis. The authors use a beam/film system to simulate the distribution of "crack" locations that can be used to model experimentally the observed distribution of meiotic chiasmata/crossovers.

The Beam/film system can easily be explained by a real life model. Consider an elastic metal beam which is coated on one side by a thin film of ceramic (Figure- 3.2 A (1)). The film is not smooth along the edges. If this object is heated then, because of the difference in thermal expansion of the two materials, the metal beam will try to expand more than the ceramic film. This will force the ceramic film to stretch beyond its natural thermal expansion and cause a tensile stress within the film. The film then resists this stretching tendency. As a result, a balancing compressive stress within the metal plate is generated. But if the film is too thin compared to the beam, then the film cannot resist the stretch anymore resulting in a crack nucleation at the edge flaws. Once triggered, the crack extends to the entire width and height of the film (Figure-3.2 A (2)) which relieves the stress locally on either side of the affected site (Figure-3.2 B (2)). As the beam is very big compared to the film, it absorbs some of the relived stress and spreads some of the stress through its elasticity in such a way that the stress decreases gradually with distance from the site forming stress relief domains (Figure-3.2 B (3)). Cracks that occur due to additional stress tend to occur outside the stress relief domains of prior crack(s). Thus occurrence of a crack at one

position interferes with the occurrence of a crack nearby and the end result is that the cracks are found almost evenly distributed along the beam.

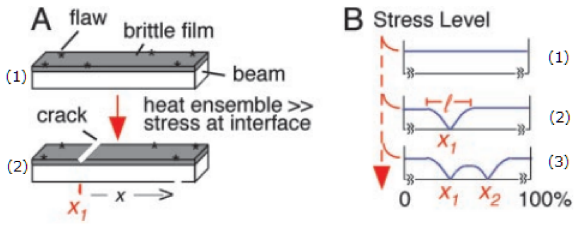


Figure 3.2: The beam/film system.

Crossovers arise by stochastic selection of a few sites from a much larger array of undifferentiated precursor interaction sites. These precursor interactions can be taken as flaws and designation of certain interactions to be crossovers can be compared as a stress-promoted process corresponding to crack formation. Thus, when the key parameters are set at optimal values in this model, the number and distribution of cracks matches closely with those of crossovers. The three general properties that can be explained by the crack formation process in the beam/film system are: 1) Ensuring an obligatory event: adequate stress causes the first obligatory crack to happen at a unique site, just like every pair of homolog always receives at least one crossover. 2) Self-limiting spatial domains: unlike the stress relief domain that is discussed in the previous paragraph, chromosomes have different domains such as centromeres, imprinted regions, heterochromatic domains, etc. These domains are also self-limiting and thus can arise without unique genetic designation of boundaries similar to the domains described in the beam/film system. 3) Interference and even spacing: because of the interference, when there are two or more crossovers, they are not randomly distributed along the chromosome; rather they are evenly distributed just as in the real life example, when more and more cracks occur, they tend to find

the gaps between previous cracks showing a tendency to be evenly placed along the beam.

Chromosome function is governed by internal mechanical forces. The following points will describe the potential origin and effects of physical forces within chromosomes.

Origin of DNA/chromatin expansional force: any change in the chromatin fiber can expand the chromatin. When such expansion happens inside a cell where it is constrained by either external features, such as other chromatin, or by an internal network of intersegment interlinks, the chromatin will push on the constraining components and those constraints will also push back. Such pushing causes the expansion of chromatin and may give it a longer outline length.

Predicted effects: the stress along the length of chromatin tends to destabilize the DNA duplex and alter its natural qualities. Because of this stress the chromatin fiber could have a bulking shape (Figure-3.3 A) or smooth bending shape (Figure-3.3 B). This alteration could promote basic processes such as DNA replication, crossover, etc. and organizational processes such as wrapping of DNA around the nucleosome core, programmed chromatin loops, etc.

If the expansion happens while two chromatin are touching to each other, then a chromatin pressure will be created. It could raise two very opposite tendencies. Two chromatins could either tend to push one another apart, creating a force for separation, or they could tend to push into one another, creating a force for conjunction (Figure-3.3 C). According to the state of chromatin fiber either or both of the effects can happen. The separation force is responsible for separation of sister chromatids or homolog, separation of unrelated chromosomes into distinct territories, etc. Conjunction force, however, is responsible for intermingling of chromosome territories, linages between different chromatins or chromatin regions, etc.

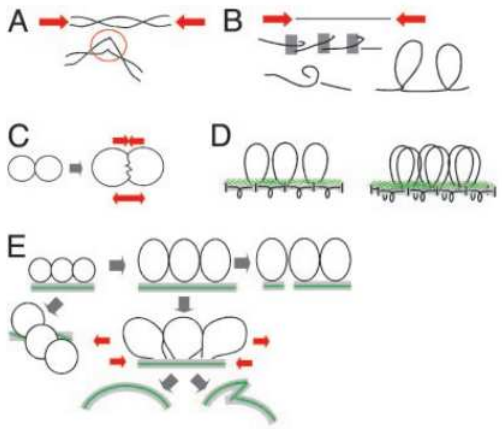


Figure 3.3: Effects of mechanical forces within chromosomes

After prophase, the chromatin is organized into a linear array of loops and its base is formed by proteins giving it a form of geometric and structural axis (Figure-3.3 D). If expansion happens in these adjacent loops, they cannot move apart as they are attached to the underlying axis. In this case, the axis will feel the stretch (Figure-3.3 E), then it will attempt to extend and ultimately twist or break. This will give the chromatin loops more space to expand causing the axis to bend and writhe. This is the reason for the axial coiling of late-stage chromosomes.

The mitotic and meiotic programs are results of sequential global chromatin expansion and contraction. From the appearance of discrete compact individualized chromosomes, it is seen that the global chromatin expansion and contraction cycle happens as chromatin fiber is well folded at mid-G_1, unfolds during S phase and G_2 and becomes compact at prophase. But videomicrographs of living Haemanthus chromosomes show that chromosomes increase in volume during prometaphase, decrease during metaphase, increase at preanaphase and again decrease during anaphase. At telophase, chromosomes expand before completely disappearing but fold into compact

fibers at early G_1. Thus mitotic consists of four sequential cycles of expansion and contraction.

The meiotic program also starts with chromatin expansion/contraction during $G_1 - S$/prophase. After that meiotic prophase is longer than that of mitotic. Measurements of chromatin volume and electron microscope thin sections show that chromatin increases at leptotene, decreases at zygotene, increases at early-mid pachytene, decreases at late pachytene, increases at the diffuse stage and again decreases at diplotene. The latter three cycles of meitotic and meiotic are analogous. Meiosis is believed to be initiated by programmatic triplication of the last three cycles of the mitotic program.

A correlation has also been found between the change in histone $H3$ phosphorylation and global chromatin expansion and contraction in mitotic and meiotic.

In addition to the above mentioned effects, chromatin expansion could drive and direct topoisomerase to alter DNA supercoiling, catenation or decatenation. Mechanical forces within chromatin could alter the structural properties in associated proteins or RNAs. Microtubule-mediated bipolar spindle forces could create tension on centromere regions of a chromosome, which governs the beginning of anaphase.

Polymerization model can produce the regulation of recombination observed in both genetic and cytological experiments. This model is simple and biologically reasonable [18]. It explains interference and obligatory crossover rule [32]. The advantage of stress model is it ensures crossover interference and homoeostasis. But this model has a very complex implementation as certain heterogeneities must be introduced in the mechanical properties [32].

3.5 Poisson Process:

A stochastic process $\{X(t), t \in T\}$ is a collection of random variables. When t is interpreted as time, $X(t)$ is called the state of the process at time t.

A stochastic process $\{N(t), t \geq 0\}$ is said to be a counting process when N(t) satisfies:

(i) $N(t)$ is the total number of 'events' that have occurred up to time t.

(ii) $N(t)$ is integer valued.

(iii) $N(t) \geq 0$.

(iv) If $s < t$, then $N(s) \leq N(t)$.

(v) For $s < t$, $N(t) - N(s)$ equals the number of events that have occurred in the interval $(s, t]$.

The counting process $\{N(t), t \geq 0\}$ is said to be a Poisson process with rate λ ($\lambda > 0$), if:

(i) $N(0) \geq 0$.

(ii) The process has stationary and independent increments, which means the numbers of events that occur in disjoint time intervals are independent and the distribution of the number of events that occur in any interval of time depends only on the length of the time interval.

(iii) $P[N(h) = 1] = \lambda h + o(h)$.

(iv) $P[N(h) \geq 2] = o(h)$.

The definition of Poisson process with rate λ implies that the number of events in any interval of length t is Poisson distributed with mean λt. That is, for all $s, t \geq 0$,

$$P\{N(t+s) - N(s) = n\} = e^{-\lambda t}\frac{(\lambda t)^n}{n!}, \quad n = 0, 1, ...$$

Another important feature to be considered here is the distribution of interarrival

time. Let T_n denote the time between $(n-1)$st and the nth event. The sequence $\{T_n, n \geq 1\}$ is called the interarrival times. When given $T_n = t$, number of events is Poisson distributed with mean λt in an interval of t, then $T_n, n = 1, 2, ...$ are independent indentically distributed exponential random variables with mean $\frac{1}{\lambda}$ and the waiting time until the nth event is, $S_n = \sum_{i=1}^{n} T_i, \;\; n \geq 1$, S_n has a gamma distribution with parameters n and λ [31].

Assuming that the number of chiasmata on an interval has a Poisson distribution implies that the distance between two successive chiasmata follows exponential distribution and the total interval, or distance, follows gamma distribution. Also, in a very short interval the chance of chiasmata being more than one is extremely rare. This holds if there is no interference. Under no interference, both chiasma and crossover processes are Poisson processes.

Now we will show that the crossover process is also a Poisson process. Say, under no interference chiasma process X is a Poisson process with mean λ. This also implies that the chiasma happening across disjoint intervals are mutually independent Poisson random variables. For a chromosomal interval t of length the distribution of $X(t)$ is Poisson with mean λt. Since under no chromatid interference each chiasma results in a crossover independently in a given gamete with probability 0.5, then crossover process $Y(t)$ consists of a Bernoulli thinning of $X(t)$ with $p = 0.5$ which means that crossover process $Y(t)$ is a thinned version of the chiasma process. In a Poisson process we can label each point with a mark and use these marks to distinguish between different events that have happened at a certain point. This is called marked Poisson process. Each chiasma point can be marked according to whether the chromatids participate in exchange or not so that when a gamete is sampled (observed), we identify as a crossover only the marked points where the chromatid participation is an exchange and consequently the crossover process is a thinned version of the

chiasmata process. Under no chiasma and chromatid interference the crossover process is a thinned Poisson process. We prove this with the help of a probability generating function.

Let, $X(t) \sim P(\lambda t)$ and each of these events be of type 1 with probability ω. $Z(t_1), Z(t_2), Z(t_3), ...$ is a sequence of independent Bernoulli variables with $P(Z_i = 1) = \omega$ and let,

$$Y(t) = \sum_{i=1}^{X(t)} Z(t_i)$$

.

Then,

$$Y(t) = \sum_{i=1}^{\infty} Z(t_i)\, I_{[i \leq X(t)]}$$

By independence of $Z(t_i)$ and $X(t)$ for every i,

$$E(Y(t)) = \sum_{i=1}^{\infty} E(Z(t_i))E(I_{[i \leq X(t)]}) = \omega \sum_{i=1}^{\infty} E(I_{[i \leq X(t)]}) = \omega E(t) = \omega \lambda t$$

By using the probability generating function,

$$E[s^{Y(t)}] = \sum_{n=0}^{\infty} E[s^{Y(t)} \mid X(t) = n]\ \frac{e^{-\lambda t}(\lambda t)^n}{n!}, \quad s \epsilon [0, 1]$$

where,

$$E[s^{Y(t)} \mid X(t) = n] = E[s^{Z(t_1)}]^n = (1 + \omega(s-1))^n$$

Then,

$$E[s^{Y(t)}] = \sum_{n=0}^{\infty} (1+\omega(s-1))^n\ \frac{e^{-\lambda t}(\lambda t)^n}{n!} = \exp\left(\lambda t(1 + \omega(s-1))\right) \exp\left(-\lambda t\right) = \exp\left(\omega \lambda t(s-1)\right)$$

So $Y(t)$ follows poisson distribution with mean $\lambda \omega t$.

Since the crossover process is a thinned version of the chiasma process, the intensity of the chiasma process is twice that of the crossover process.

Map function plays an important role in genetics for two reasons: i) genetic map distances are additive by definition, whereas recombination fractions are not; ii) recombination fractions are much easier to estimate from data as it refers only to features of chromosomes at the end points of intervals. On the other hand, to estimate a map distance, the information about crossovers in the interval between two loci is required and surely, this information cannot be observed [33]. Nonetheless, information about crossover is contained in recombination and through statistical techniques such information can be reached.

We can think of the distance between two crossover events along the chromosome as a waiting distance for an event to happen. In this way we have a stochastic process which is of similar structure to those time-to-event processes whose analysis techniques are used in survival analysis. If we could observe the distance between two successive crossovers, we would be able to estimate the distribution of the genetic distance between any pair of crossover sites. Unfortunately, we do not observe the crossover sites themselves, but their location between flanking markers. This means that our data set is really a set of double censored observations. We obtained our double censored data in the following way. Suppose, G is the distance between consecutive crossovers. For inter-crossover distance the data is type-2 double censored. To see this, we are considering two consecutive crossover events: the first one happening at some point in between the loci positions m_1 and m_2 and the next one happening in between m_3 and m_4, with $m_1 < m_2 < m_3 < m_4$. Thus , the inter crossover distance is in the interval, $[m_3 - m_2, m_4 - m_1)$, i.e, in general, $d_i \, \epsilon \, [L_i, U_i)$. After getting the data in a double censored data format, we would like to estimate the distribution of the distance.

3.6 Distribution of the distance between consecutive crossover in the absence of interference

Suppose we have a random sample of individuals for which a double crossingover on a particular chromosome has been observed. Let n is the number of consecutive crossovers, i.e., the sample size. When there is no interference of any kind, the distance between consecutive crossovers is exponential, so estimation of its distribution is reduced to the estimation of the parameter of the exponential distribution with double censored data. As before, suppose that the length between a double crossover on a chromosome is d. Then, as per the discussion of the previous sections, it can be said that d follows exponential distribution with parameter λ. We do not observe d but we know that when a double recombination happens, $d_i \epsilon [L_i, U_i)$. To estimate the parameter by the maximum likelihood method, we proceed as follows. Say, $d_i \, \epsilon \, [L_i, U_i), i = 1, 2, ...n$.

Then the probability that d_i lies between L_i and U_i is,

$$\begin{aligned} \pi_i(\lambda) &= \Pr\left(d_i \epsilon [L_i, U_i)\right) = F_\lambda(U_i) - F_\lambda(L_i) \\ &= e^{-\lambda L_i} - e^{-\lambda U_i} \end{aligned} \tag{3.24}$$

So, the likelihood of the data is,

$$\mathscr{L}(\lambda) = \prod_{i=1}^{n} \pi_i(\lambda),$$

Then, the log-likelihood is,

$$\begin{aligned} \ell(\lambda) &= \sum_{i=1}^{n} \log \pi_i(\lambda) \\ &= \sum_{i=1}^{n} \log \left(e^{-\lambda L_i} - e^{-\lambda U_i} \right). \end{aligned} \tag{3.25}$$

Differentiating the log-likelihood twice with respect to λ,

$$\begin{aligned} \frac{\partial \ell(\lambda)}{\partial \lambda} &= \sum_{i=1}^{n} \frac{-L_i e^{-\lambda L_i} + U_i e^{-\lambda U_i}}{e^{-\lambda L_i} - e^{-\lambda U_i}} \\ &= \sum_{i=1}^{n} \frac{U_i e^{-\lambda U_i} - L_i e^{-\lambda L_i}}{e^{-\lambda L_i} - e^{-\lambda U_i}}. \end{aligned} \tag{3.26}$$

$$\frac{\partial^2 \ell(\lambda)}{\partial \lambda^2} = -\sum_{i=1}^{n} \frac{e^{-\lambda(U_i+L_i)}(U_i - L_i)^2}{(e^{-\lambda L_i} - e^{-\lambda U_i})^2}. \tag{3.27}$$

Then using the Newton-Raphson method in R, an estimate $\hat{\lambda}$ of λ can be found.

The asymptotic variance of $\hat{\lambda}$ is

$$Var(\hat{\lambda}) = \frac{1}{\sum_{i=1}^{n} \frac{e^{-\lambda(U_i+L_i)}(U_i-L_i)^2}{(e^{-\lambda L_i} - e^{-\lambda U_i})^2}}. \tag{3.28}$$

And an estimate of this variance would be obtained by plugging in $\hat{\lambda}$ in (3.28).

3.7 Distribution of the distance between successive crossovers

Suppose that we want to somehow quantify the effect of genetic interference for different chromosomes against a group of individuals of a given genetic makeup or such an effect on the same chromosome between groups of different genetic backgrounds.

A direct comparison of the distribution of the distance between successive crossovers *a la* Kolmogorov-Smirnov may not be a good idea because the difference may occur due to the difference in the length of chromosomes of the individual basis or genetic group.

We could compare the expected value of the distance, but since we do not observe distance between successive crossovers, we need to find a way to estimate the mean of the distance. A more sensible alternative would be to compare the expected value of the distance between successive crossovers,

$$\hat{E}(d) = \int_0^\infty (1 - \hat{F}(t))dt,$$

where F is the distribution of the positive random variable d. In any case, a good starting point would be the estimation of the distribution of the distance between successive crossovers. However, there is a small complication. We do not observe such a distance, but the end points of an interval that covers such a distance i.e., our data consist of intervals such that, $d_i \, \epsilon \, [L_i, U_i)$.

This section deals with the problem of estimating the distribution of the interval censored distance between two successive crossovers.

3.7.1 The non-parametric maximum likelihood (NPML) estimator for interval censored data

We will start by finding the NPML estimator of the interval censored data. This problem can be reduced to one of maximizing the likelihood function subject to certain constraints. First the probability that a distance, d_i, lays in the observed interval $[L_i, U_i)$ is given by

$$Pr(L_i \leqslant d_i < U_i) = F(U_i) - F(L_i),$$

where F is the distribution that we want to estimate, non-parametrically. Furthermore, the probability of observing a particular set of intervals in a random sample of size n is given by the likelihood,

$$\mathscr{L}(F) = \prod_{i=1}^{n} (F(U_i) - F(L_i)). \tag{3.29}$$

Now, let $b_0 < b_1 < ...b_M$ be the grid of distances that include all L_i and U_i end points. We then get the set $\{[b_{j-1}, b_j),\ j = 1, 2, ..., M\}$ the observed end points,

with this notation,

$$\mathscr{L}(\mathbf{F}) = \prod_{i=1}^{n} \sum_{j=1}^{M} I_{[L_i,\infty)}(b_{j-1}) I_{(0,U_i]}(b_j)(F(b_i) - F(b_{j-i})). \tag{3.30}$$

So that the log-likelihood function, $\ell(\mathbf{F})$, is given by,

$$\begin{aligned} \ell(\mathbf{F}) &= \log \mathscr{L}(\mathbf{F}) \\ &= \sum_{i=1}^{n} \log \left(\sum_{j=1}^{M} I_{[L_i,\infty)}(b_{j-1}) I_{(0,U_i]}(b_j)(F(b_j) - F(b_{j-1})) \right). \end{aligned} \tag{3.31}$$

The problem can be further reduced by noticing that the sample data may not provide information to estimate each $F(b_j)$, $j = 1, 2, ..., M$ independently. For example, imagine that we have no data observations in a particular basic interval. Thus the data provides no evidence of a probability mass falling in such a basic interval. As a result, the estimator of $F(b_j)$ will be identical to that for $F(b_{j-1})$. Because of this, the M observed end points are aggregated, or consolidated, to include only the support of the non-parametric estimator of F. This set of intervals, say $\{[\ell_i, \nu_i)\}$ is found in such a way that they represent the maximum clique of the data [40, 11].

Thus, by construction, the maximal clique intervals are such that, $\ell_i \, \epsilon \, \{L_1, L_2, ..., L_n\}$ and $\nu_i \, \epsilon \, \{R_1, R_2, ..., R_n\}$ and such that no interval endpoint L_i and R_i occurs between ℓ_i and ν_i. The maximum clique intervals are completely ordered in the real line and we assign them an index from 1 to J. On the maximal clique intervals $\hat{F}$ is not decreasing, unless $\hat{F}(\nu_i) - \hat{F}(\ell_i -) = 0$. In any other interval $\hat{F}$ must be constant. Define,

$$\alpha_{ij} = I_{[L_i, \infty)}(\ell_j) I_{(0, U_i]}(\nu_j), \; j = 1, 2, ..., J.$$

Therefore the vector $\boldsymbol{\alpha}_i \, \epsilon \, \{0, 1\}^J$ indicates which maximal clique interval intersects with the ith observed interval $[L_i, U_i)$.

We know that $F(0) = F_0 = 0$ and our NPML have jumps only at each ℓ_j, $j = 1, 2, ..., J$. So, finding F consists of estimating the vector where the jumps occur. Define,

$$\Delta_j = F(\nu_j^*) - F(\ell_j), \; j = 1, 2, ..., J$$

where,

$$\nu_j^* = \begin{cases} \nu_j - & \text{if } \ell_j < \nu_j \\ \nu_j & \text{if } \ell_j = \nu_j. \end{cases}$$

Now, estimation of the distribution F is equivalent to the estimation of $\boldsymbol{\Delta}$. So this is a constrained optimization problem. Then the NPML problem can be rewritten as,

$$\hat{\boldsymbol{\Delta}} = \arg\max_{\boldsymbol{\Delta}} \sum_{i=1}^{N} \log \eta_i(\boldsymbol{\Delta}) \tag{3.32}$$

with

$$\eta_i(\boldsymbol{\Delta}) = \sum_{j=1}^{J} \alpha_{ij} \Delta_j.$$

Subject to,

$$\Delta_j \geq 0,\ j = 1, 2, ..., J$$

$$\sum_{j=1}^{J} \Delta_j = 1$$

so that $\hat{F}_j = \hat{F}_{j-1} + \hat{\Delta}_j$, $j = 1, 2, ..., J$, with $\hat{F}_0 = 0$ and $\hat{F}_j = \hat{F}(\ell_j)$.

Then the log-likelihood function of $\boldsymbol{\Delta}$,

$$\ell(\boldsymbol{\Delta}) = \sum_{j=1}^{n} \log \eta_i(\boldsymbol{\Delta}). \tag{3.33}$$

It is strictly concave if the Hessian of $\ell(\boldsymbol{\Delta})$ is negative definite.

First derivative w.r.t $\boldsymbol{\Delta}$,

$$\frac{\partial \ell}{\partial \boldsymbol{\Delta}} = \sum_{i=1}^{n} \frac{\boldsymbol{\alpha}_i}{\eta_i(\boldsymbol{\Delta})} = \boldsymbol{\beta}. \tag{3.34}$$

The Hessian matrix is given by

$$\frac{\partial^2 \ell}{\partial \boldsymbol{\Delta} \partial \boldsymbol{\Delta}'} = -\sum_{i=1}^{n} \frac{\boldsymbol{\alpha}_i \boldsymbol{\alpha}_i'}{{\eta_i}^2(\boldsymbol{\Delta})}. \tag{3.35}$$

Notice that, because of the way the maximum clique intervals were constructed, the Hessian is a simple quadratic form of full rank and negative definite.

Since the objective function is concave and the constraints are linear; the Lagranian function to be optimized can be written as [10]

$$\sum_{i=1}^{n} \log \eta_i(\boldsymbol{\Delta}) + \sum_{j=1}^{J} \mu_j \Delta_j - \nu(\sum_{j=1}^{J} \Delta_j - 1),$$

where ν and $\mu_1, \mu_2, ...\mu_J$ are Lagrange multipliers.

The estimate of $\boldsymbol{\Delta}$ that satisfies all four Kraus-Khun-Tucker first-order necessary conditions will be the unique NPML estimate of $\boldsymbol{\Delta}$, because our problem is a strictly convex one.

The Kraus-Khun-Tucker first-order necessary conditions for a maximum are:

i) Primal constraints

$$\Delta_j \geq 0,\ j = 1, 2, ..., J,$$

$$\sum_{j=1}^{J} \Delta_j = 1.$$

ii) Dual constraints

$$\mu_j \geq 0,\ j = 1, 2, ..., J.$$

iii) Complimentary slackness

$$\mu_j \Delta_j \geq 0.$$

iv) Stationarity

$$\beta_j + \mu_j - \nu = 0,\ j = 1, 2, ..., J,$$

where

$$\beta_j = \sum_{i=1}^{n} \frac{\alpha_{ij}}{\eta_i(\boldsymbol{\Delta})}.$$

The unique NPML estimate of $\boldsymbol{\Delta}$ satisfies the conditions (i)-(iv) simultaneously, if in application these conditions are not met, then we do not have an NPML estimate.

3.7.2 Turnbull's self-consistent algorithm

It is easy to see that for interval censored data no closed form to the NPML problem exists. In the survival literature there is an algorithm proposed by Turnbull (1976) that produces the so-called self-consistent estimates.

In our context, the stationarity conditions (iv) above imply that,

$$\boldsymbol{\beta}'\boldsymbol{\Delta} + \boldsymbol{\mu}'\boldsymbol{\Delta} = \nu\, \mathbb{1}'\boldsymbol{\Delta} \tag{3.36}$$

which together with the primary constraints (i) and the complementary slackness conditions (iii), we conclude that,

$$\boldsymbol{\beta}'\boldsymbol{\Delta} = \nu = N, \tag{3.37}$$

because,

$$\boldsymbol{\beta} = \sum_{i=1}^{n} \frac{\boldsymbol{\alpha}_i}{\eta_i(\boldsymbol{\Delta})} \text{ and } \eta_i(\boldsymbol{\Delta}) = \boldsymbol{\alpha}_i'\boldsymbol{\Delta}_i. \tag{3.38}$$

Expanding $\beta_j\Delta_j$ we have that

$$\beta_j\Delta_j = \sum_{i=1}^{n} \frac{\alpha_{ij}\Delta_j}{\eta_i(\Delta)} = \sum_{i=1}^{n} \pi_{ij}(\boldsymbol{\Delta}), \text{ for } j = 1, 2, ...J, \tag{3.39}$$

where $\pi()$ can be interpreted as the probability that the inter crossover distance for the ith observation lies in the jth maximum clique interval, given the observed data. Now,

equations (3.19)-(3.22) lead us to

$$\begin{aligned} \sum_{i=1}^{n} \pi_{ij}(\boldsymbol{\Delta}) + \mu_j \Delta_j &= n\Delta_j \\ \Rightarrow \Delta_j &= \frac{1}{n}\sum_{i=1}^{n} \pi_{ij}(\boldsymbol{\Delta}); \qquad j = 1, 2, ..., J. \end{aligned}$$

Based on this, the Turnbull's algorithm for our problem can be expressed as

1) Select an initial guess $\Delta^{(0)}$.

2) At the r-stage, compute the expected observation fractions in each maximum clique interval according to,

$$\pi_{ij}(\boldsymbol{\Delta}^{(r)}) = \frac{\alpha_{ij}\Delta_j^{(r)}}{\eta_i(\boldsymbol{\Delta}^{(r)})}$$

3) Obtain the update of $\boldsymbol{\Delta}^{(r+1)}$ as

$$\Delta_j^{(r+1)} = \frac{1}{n}\sum_{i=1}^{n} \pi_{ij}(\boldsymbol{\Delta}^{(r)}).$$

4) Whenever

$$\max_{j=1,2,...,J} |\Delta_j^{(r+1)} - \Delta_j^{(r)}| > \varepsilon$$

iterate again, i.e., increase r and go to step (2), otherwise stop.

The resulting vector $\hat{\boldsymbol{\Delta}}^{SC}$ holds the self-consistant estimator of $\boldsymbol{\Delta}$.

As Turnbull (1976) pointed out, this procedure can be seen as an instance of the EM algorithm with the E-step and M-step being step 2 and 3 above, respectively.

Furthermore, Nettleton (1999) showed that the sequence of iterations of Turnbull's algorithm converges to a fixed point which belongs to the set of self-consistant estimates. The question that remains to be answered is whether $\hat{\boldsymbol{\Delta}}^{SC}$ is also $\hat{\boldsymbol{\Delta}}^{NPML}$ or not.

To answer this, note that when deriving the self-consistent equations we did not use the dual constraints (ii) of the Kraus-Khun-Tucker necessary conditions for the NPML estimator. Consequently, although $\hat{\boldsymbol{\Delta}}^{NPML}$ is also self-consistant, the reciprocal does not hold, i.e., $\hat{\boldsymbol{\Delta}}^{SC}$ is not necessarily NPML estimator. For example, suppose that the data are the intervals $[0,1), [1,3), [1,3), [0,2), [0,2), [2,3)$. It is easy to see that the maximal clique intervals are $[0,1), [1,2), [2,3)$ the vector $(\frac{1}{2}, 0, \frac{1}{2})$ is a self-consistent estimate of $\boldsymbol{\Delta}$, so that $\boldsymbol{\Delta}^{SC} \geq \mathbb{0}$ and $\mathbb{1}'\boldsymbol{\Delta}^{SC} = 1$. However, the Kraus-Khun-Tucker conditions are violated at this point, since, $\boldsymbol{\beta} = (6,8,6)'$ implies $\mu_2 = -2 < 0$. The NMPL estimate of Δ is $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$.

Since Turnbull's algorithm did not enforce the dual constraints of the Lagrange multipliers $\mu_j \geq 0$, $j = 1, 2, ...J$, then there may exist points in the parameter space at which the estimates comply with the self-consistent equations, but one or more of the Langrange multipliers of the set $\{\mu_j\}$ are negative. Such self-consistent estimates cannot be the NPML estimate. Nonetheless, we can check if $\hat{\boldsymbol{\Delta}}^{SC}$ is $\hat{\boldsymbol{\Delta}}^{NPML}$ by verifying that $\hat{\mu}_j = N - \hat{\beta}_j^{SC} \geq 0,\ j = 1, 2, ..., J$.

3.7.3 Quadratic programming

Perhaps the most popular technique for solving constrained optimization problems like our NPML estimation problem, is quadratic programming. This approach consists of a battery of iterative algorithms on which a quadratic approximation of the constrained objective function is solved at each iteration.

For example, for our problem, we can take a second order Taylor series expansion

of the log-likelihood for the estimator at iteration $r+1$ about the estimate at iteration r, $\boldsymbol{\Delta}^{(r)}$, i.e.,

$$\begin{aligned}\ell(\boldsymbol{\Delta}^{(r+1)}) \approx \ell(\boldsymbol{\Delta}^{(r)}) &+ (\boldsymbol{\Delta}^{(r+1)} - \boldsymbol{\Delta}^{(r)})'\boldsymbol{\beta}^{(r)} \\ &+ \tfrac{1}{2}(\boldsymbol{\Delta}^{(r+1)} - \boldsymbol{\Delta}^{(r)})'\boldsymbol{H}^{(r)}(\boldsymbol{\Delta}^{(r+1)} - \boldsymbol{\Delta}^{(r)}),\end{aligned}$$

and maximize $\ell(\boldsymbol{\Delta}^{(\boldsymbol{r+1})})$ subject to $\boldsymbol{\Delta}^{(r+1)} \geq \mathbb{0}$ and $\mathbb{1}'\boldsymbol{\Delta}^{(r+1)} = 1$.

The first three Kraus-Khun-Tucker necessary conditions for this quadratic programming NPML problem are similar to before and together with the gradient of the Lagranian, these can be written as,

1) $\boldsymbol{\Delta}^{(r+1)} \geq \mathbb{0}$ and $\mathbb{1}'\boldsymbol{\Delta}^{(r+1)} = 1$.

2) $\boldsymbol{\mu} \geq \mathbb{0}$.

3) $\boldsymbol{\mu} \bullet \boldsymbol{\Delta}^{(r+1)} = \mathbb{0}$.

4) $\boldsymbol{H}^{(r)}\boldsymbol{\Delta}^{(r+1)}\boldsymbol{\mu} = \nu\mathbb{1} + \boldsymbol{H}^{(r)}\boldsymbol{\Delta}^{(r)} - \boldsymbol{\beta}^{(r)}$,

where $\boldsymbol{\beta}^{(r)}$ and $\boldsymbol{H}^{(r)}$ are the gradient of $\ell(\boldsymbol{\Delta})$, evaluated at $\boldsymbol{\Delta}^{(r)}$ respectively. i.e.,

$$\boldsymbol{\beta}^{(r)} = \sum_{i=1}^{N} \frac{\boldsymbol{\alpha}_i}{\eta_i(\boldsymbol{\Delta}^{(r)})}$$

and

$$\boldsymbol{H}^{(r)} = -\sum_{i=1}^{N} \frac{\boldsymbol{\alpha}_i\boldsymbol{\alpha}_i'}{\eta_i^2(\boldsymbol{\Delta}_i^{(r)})}.$$

An estimate of $var(\hat{\boldsymbol{\Delta}})$ is given by the negative of the inverse of the Hessian matrix evaluated at the NPML estimate, i.e.,

$$\hat{var}(\hat{\boldsymbol{\Delta}}) = \left(\sum_{i=1}^{N} \frac{\boldsymbol{\alpha}_i\boldsymbol{\alpha}_j'}{\eta_i^2(\hat{\boldsymbol{\Delta}}^{NPML})}\right)^{-1}.$$

3.7.4 Nonparametric analysis of censored time-to-event data

To keep the calculation simpler, we used the nonparametric analysis of censored time-to-event to find the estimate of $\boldsymbol{\lambda}$. The standard estimator of the survival function, proposed by Kaplan and Meier (1958), is known as the Product-Limit estimator. This estimator is defined as follows for all values of t in the range where there is data:

$$\hat{S}(t) = \begin{cases} 1 & if \;\; t < t_1, \\ \prod_{t_i \leq t}(1 - \frac{a_i}{v_i}), & if \;\; t_1 \leq t \end{cases} \tag{3.40}$$

We suppose that the events occur at n distinct times $t_1 < t_2 < ... < t_n$. In the above equation a_i and v_i are the number of events and the number at risk at time t_i respectively. The Product-Limit estimator is a step function with jumps at the observed event times. The size of these jumps depends on the number of events observed at each event time t_i and on the pattern of the censored observations prior to t_i.

To get the estimate of the survival function for interval censored data, a modified Product-Limit estimator has been suggested by Turnbull (1974). To construct the estimator, let $0 = \tau_0 < \tau_1 < ... < \tau_m$ be a grid of time which includes all the points L_i and U_i for $i = 1, ..., n$. For the ith observation, define a weight α_{ij} to be 1 if the interval (τ_{j-1}, τ_j) is contained in the interval $(L_i, U_i]$ and 0 otherwise. That is, α_{ij} indicates whether the event which occurs in the interval $(L_i, U_i]$ could have occurred at τ_j. Hence, α_{ij} can be written as

$$\alpha_{ij} = I_{[L_i, \infty)}(\tau_{j-1}) I_{(-\infty, U_i]}(\tau_j)$$

Starting with an initial guess at $\tilde{S}(\cdot)$ using Product-Limit estimator, as suggested

by Turnbull, the algorithm is as follows:

Step 1: Compute the probability of an event's occurring at time τ_j by

$p_j = \tilde{S}(\tau_{(j-1)}) - \tilde{S}(\tau_j), j = 1, ..., m.$

Step 2: Estimate the number of events which occurred at τ_j by

$$\tilde{d}_j = \sum_{i=1}^{n} \frac{\alpha_{ij} p_j}{\sum_k^m \alpha_{ij} p_k} = \frac{p_j \sum_{i=1}^{n} \alpha_{ij}}{\sum_k^m \alpha_{ij} p_k}. \tag{3.41}$$

Here, $\sum_k^m \alpha_{ij} p_k$ is the total probability assigned to possible event times in the interval $(L_i, U_i]$.

Step 3: Compute the estimated number at risk at time τ_j by

$$\tilde{Y}_j = \sum_{k=j}^{m} \tilde{a}_k. \tag{3.42}$$

Step 4: Compute the updated Product-Limit estimator using the pseudo data found in Steps 2 and 3. If the updated estimate of $\tilde{S}(\cdot)$ is close to the old version of $\tilde{S}(\cdot)$ for all $\tau_j's$, stop the iterative process, otherwise, repeat Steps $1 - 3$ using the updated estimate of $\tilde{S}(\cdot)$. [20]

Since, to our knowledge, no standard statistical package was found to estimate the survival function based on Turnbulls algorithm, we implemented the algorithm in R [12].

3.8 Ratio of cM and MB

The ratio of genetic length to physical length (cM/MB) on a given chromosome is a measure commonly used to describe "crossover intensity" in a broad and informal sense. It is well known that this ratio varies with species and chromosome. It is also known that the ratio may not be constant across a given chromosome and these variations can be taken as a local indicator of interference. In our study we tried to measure the rate of change of the genetic distance (cM) with respect to physical distance (MB) between the markers. Inter marker genetic distances were estimated to take the ratio with the corresponding inter marker physical distances. This ratio has been calculated for all the chromosomes from all five panels. Comparing them across all the panels would give an idea of the variability of crossover rates among the different panels [26].

The local ratio of estimated genetic distance and physical distance was calculated by

$$r_i = \frac{d_i - d_{i-1}}{m_i - m_{i-1}}, \tag{3.43}$$

where d_i is the estimated genetic distance at the ith locus and m_i is the physical distance at the ith locus.

As a reference an estimate of the overall ratio at each chromosome for every cross was calculated as,

$$\bar{r} = \frac{d_n - d_1}{m_n - m_1}; \text{ where, } d_1 \leq d_2 \leq ... \leq d_n \text{ and } m_1 \leq m_2 \leq ... \leq m_n \tag{3.44}$$

Chapter 4

Data Description

For this study, we used datasets from 5 different crosses. The crosses were cited using dams first, then sires for every data. Three of these datasets were generated at Jackson Laboratory, Bar Harbor, ME. The mouse strains were [(PERA X DDK)F_1 X B6]N_2 with chromosomes 1 to X and had 39 female individuals. [PERC X DDK)F_1X B6]N_2 with chromosomes 1 to X except chromosome 11 and had 36 females [15]. (SM X NZB)F_1 X NZB were used to produce female backcross progeny. Here we had 53 females with chromosomes from 1 to X [29].

The mice with backcross (B6-Pgk1^a X DDK)F_1 X B6 were used in two of the data sets. These two datasets originated from females that were genetically identical, but phenotypically different. In one group the X-inactivation pattern was more skewed towards the silencing of DDK X-linked genes (more skewed), while in the other group any of the two X chromosomes was randomly inactivated resulting in a similar number of cells with the parental phenotype (less skewed). The dataset that was less skewed had 81 females and the other one had 80. Both of these datasets had chromosomes from 1 to X [5].

4.1 Data Editing

Since we wanted to use standard software to estimate the genetic distance between consecutive markers which can be done through R/qtl, we formatted the file according to this particular data arrangement. As a starting point, we used the consensus genetic map as well as the physical distance from the appropriate sources. The data set was also checked for inconsistencies such as impossible or outrageous genotype and they were deleted.

We started our search for these positions from the following websites:

1.http://www.biolreprod.org/content/suppl/2006/10/18/biolreprod.106.056739.DC1/biolreprod.106.056739-1.pdf

2. http://www.informatics.jax.org/

3.http://www.shigen.nig.ac.jp/mouse/mmdbj/top.jsp

In some cases neither the genetic distances (cM) nor the physical distances (MB) were found from the above mentioned sources. In those cases the following formulae of interpolation and extrapolation were used to find the starting values needed for the iterative estimation algorithm.

***Linear Interpolation*:** Generally, linear interpolation takes two data points into account. Say we want to find the y of the point (x, y) which is in between (x_a, y_a) and (x_b, y_b), the interpolant is given by [3],

$$y = y_a + (y_b - y_a)\frac{(x - x_a)}{(x_b - x_a)}$$

To find the genetic position of a marker in cM, the physical map in MB units was

used as x's and in the same way the genetic distances in cM were used as x's to find the physical map in MB.

***Extrapolation*:** Suppose we have values for points (x_0, y_0), (x_1, y_1), $(x_2, y_2)...(x_n, y_n)$ and (x_{n+1}, y_{n+1}), where either y_0 or y_{n+1} is unknown. The following formula can be used to find the unknown values,

$$y_0 = y_1 - \lambda(x_1 - x_0)$$

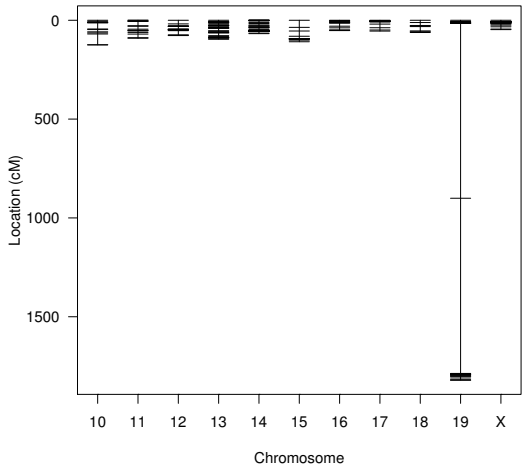
$$y_{n+1} = y_n + \lambda(x_{n+1} - x_n) \quad \text{where} \quad \lambda = \frac{y_n - y_1}{x_n - x_1}$$

Once a full set of initial genetic distances was completed, we estimated genetic distance by EM algorithm and plotted them. From this we found that between some of the markers there were too many recombination events in chromosome 19 of cross $\texttt{[(PERA X DDK)F}_1\texttt{ X B6]N}_2$ and in chromosomes 5 and 16 of cross $\texttt{[PERC X DDK)F}_1\texttt{ X B6]N}_2$. This fact has been shown in Figure 4.1 (a) and (b). This can happen because of an anomaly in the data, recombination hotspot or chiasma interference.

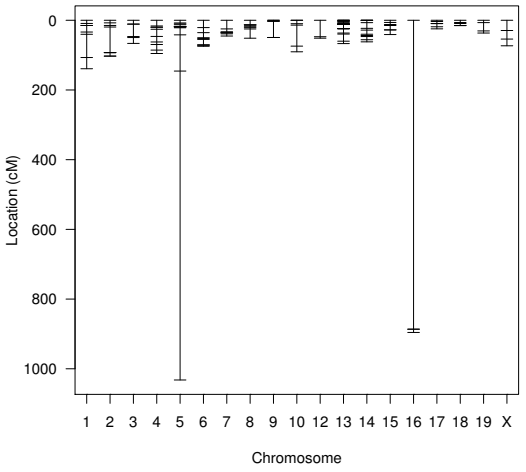
Further investigation of the data shows that it was happening as those markers were showing either 'H' or 'A' for all the individuals, which is highly unlikely. So those markers were deleted from the data and the genetic distance was estimated again with the edited data. Figure 4.1 (c) and (d) are showing the final estimate of the genetic distance.

Furthermore, at one point of the data analysis it was found that some individuals had double recombination in the same place, which seemed to be very interesting. However, further examination revealed that it happened because of the misplacement of some markers. So, those markers were also deleted.

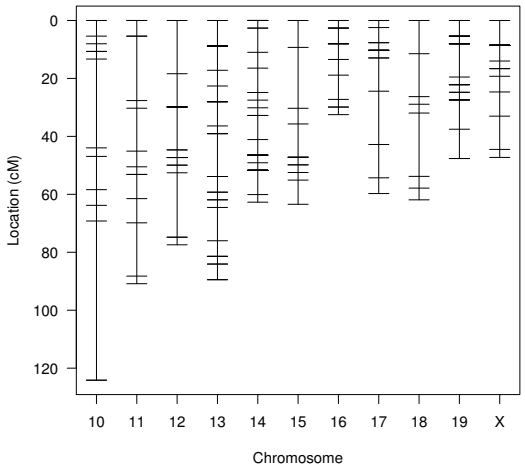
Once the data was massaged to remove gross errors, the analysis was undertaken.



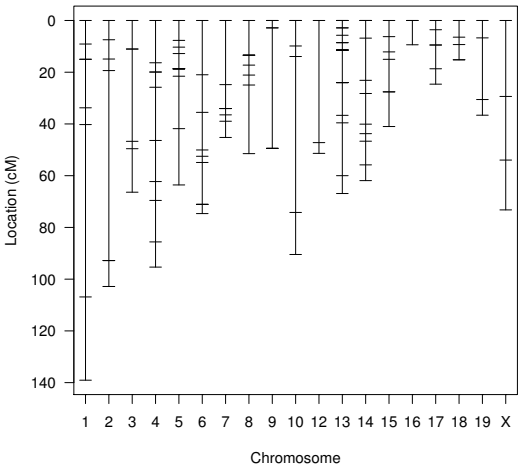
(a) Genetic Map of Cross `[(PERA X DDK)F`$_1$ `X B6]N`$_2$



(b) Genetic Map of Cross `[PERC X DDK)F`$_1$ `X B6]N`$_2$



(c) Genetic Map of Cross `[(PERA X DDK)F`$_1$ `X B6]N`$_2$ after deleting problem markers



(d) Genetic Map of Cross `[PERC X DDK)F`$_1$ `X B6]N`$_2$ after deleting problem markers

Figure 4.1: Genetic map

Chapter 5

Data Analysis and Results

5.1 Recombination Fraction

Once the genetic distance was estimated using an built-in function in R, it was used to obtain the estimated recombination fraction between adjacent markers using the Haldane function (eq. 3.23). For all the crosses, Tables 1 to 5 in Appendix-A give the chromosome number in the first column, marker name in the second column, consensus genetic distance in the third column and estimated genetic distance in the fourth column. Tables 6 to 10 provide the chromosome number in the first column, estimated genetic distance in the second column and recombination fraction for neighboring markers in the third column. In the table when the genetic distance is 0, the recombination fraction is also 0 as it is impossible to have any crossover. The probability of occurring crossover increases with the increase of genetic distances. The highest estimated genetic distance 79.98 in our data was found in chromosome 8 of [(SM X NZB)F_1 X NZB]N_2 and the corresponding recombination fraction was 0.39901135.

5.2 Distribution of genetic distance

From left to right, Tables 11 to 15 show chromosome number, individual ID and intercrossover distance for all crosses. Intercrossover distance was found using R code given in Appendix-B. Results from Tables 11 to 15 were used to find the results of Tables 16 to 20. The first and second column of Tables 16 to 20 give unique upper or lower bound of intercrossover distance and Product-Limit estimate (Section 3.7.4) respectively. R code to find Product-Limit estimate is given in Appendix-B. Results from Tables 16 to 20 were used to find distribution of genetic distance, F.

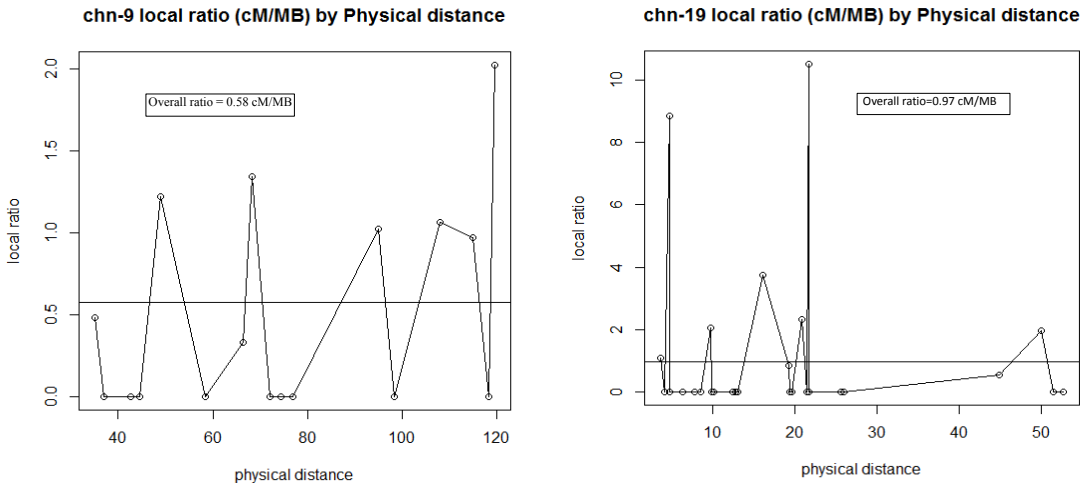
5.3 Crossover Rate

As described in Section 3.6 Crossover rate, its variance and standard deviation has been obtained by R code using the interval censored data for where there is double crossover as well as considering the data where there is only one crossover . R code is presented in Appendix-B. MLE was used to calculate the crossover rate. Estimated crossover rate for each chromosome with their variance and standard deviation are provided from Tables 21 to 25 in Appendix-A. The results in the tables show that variance or standard deviation is high for high crossover rate.

5.4 cM and MB Ratio

The ratio of genetic length to physical length (cM/MB) has been found to make a comment about the distribution of crossover rate and to find whether there is a sign of any kind of interference in the data. The results can be found in Appendix-A from Tables 26 to 30. From left to right the table depicts name of marker, ratio of cM and

MB, and overall ratio of corresponding chromosome. We could not get ratios for a few markers, as in those cases where both the estimated genetic distance and physical distance were zero. After plotting the ratios, the following kinds of graphs were found. Several graphs were made to comment on data; only two of them are presented here.



(a) [(PERA X DDK)F_1 X B6]N_2 (b) [PERA X DDK)F_1 X B6]N_2

Figure 5.1: Local ratio by physical distance

Sharp spikes in the graphs may occur either because there is hotspot of recombination or because of interference. In the above figure of [(PERA X DDK)F_1 X B6]N_2 for chromosome 9 we can see one sharp pick and for chromosome 19 we can see two sharp picks. This means there may be interference in these chromosomes. In the coming sections (Section 5.4 and 5.5) we aim to find more evidence of interference.

5.5 NPML of Lambda

To find the NPML of $\boldsymbol{\lambda}$ first the clique intervals needed to be found and for that we needed to plot the data. It was done for all the chromosomes of each cross, but only

chromosome 11 is shown here as an example.

The observed data for chromosome 11 of $\texttt{[(B6-Pgh1}^a\times\ \texttt{DDK)F}_1\times\ \texttt{B6]N}_2$ (more skewed) is:

$[35.00, 73.70),\ [35.00, 73.70),\ [25.94, 64.91),\ [25.94, 64.91),\ [6.59, 41.79),\ [54.35, 96.82),$
$[22.44, 54.35),\ [15.85, 35.00),\ [6.59, 41.79),\ [0.00, 42.47),\ [28.41, 54.35),\ [22.44, 54.35),$
$[54.35, 96.82)$

and in a plot this looks like:

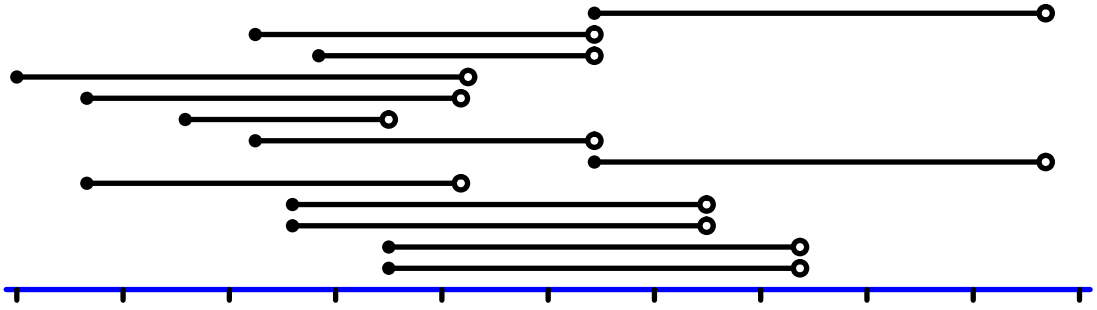


Figure 5.2: Observed data plot for chromosome 11

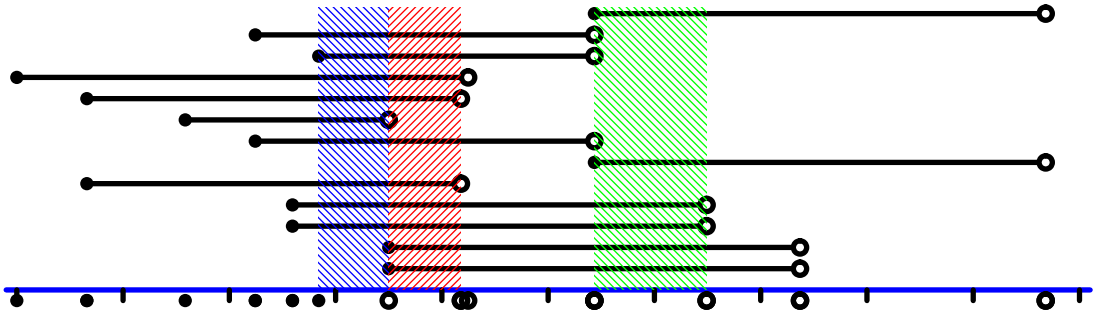


Figure 5.3: Clique intervals for chromosome 11

So the maximum clique intervals are:

$$\dot{\iota}_1 = [28.41, 35),\ \dot{\iota}_2 = [35, 41.79),\ \dot{\iota}_3 = [54.35, 64.91)$$

and the NPML of $\boldsymbol{\Delta}$ is

$$\hat{\boldsymbol{\Delta}} = \left(\frac{1}{3}, \frac{5}{12}, \frac{1}{4}\right)'$$

Then the NPML of $F(d)$ is

$$\hat{F}(d) = \begin{cases} 0 & \text{if } d < 28.41 \\ \frac{1}{3} & \text{if } 28.41 \leq d < 35 \\ \frac{3}{4} & \text{if } 35 \leq d < 54.35 \\ 1 & \text{if } d \geq 54.35 \end{cases}$$

and

$$\hat{\mathrm{E}}(d) = \int_0^\infty (1 - \hat{F}(t))\, dt$$

$$= 37.641$$

$\hat{F}(d)$ is found using nonparametric analysis of censored time-to-event method described in Section 3.7.4 and the result of that section is given in Tables 16 to 20 of Appendix-A. We found $\hat{\mathbf{\Delta}}$ for all the chromosomes of all the crosses in the above way and checked if they satisfy $\hat{\mu}_j = N - \hat{\beta}_j^{SC} \geq 0,\ j = 1, 2, ..., J$ described in Section 3.7.2. R program for checking the condition is given in Appendix-B.

5.6 Comparison between expected difference between consecutive crossovers

To find if there is interference, we compared the expected genetic distance $\hat{E}(d)$, found considering both chiasma and chromatid interference, that is explained in the previous section and the expected genetic distance, $1/\hat{\lambda}$, found without considering interference, that is presented from Tables 21 to 25.

Table 5.1: Expected difference between consecutive crossovers

$\hat{E}(g)$	$1/\hat{\lambda}$	Cross
Chromosome-2		
85.36	66.418	[PERC X DDK)F_1 X B6]N_2
56.84	56.138	[(SM X NZB)F_1 X NZB]N $_2$
Chromosome-3		
32.785	34.466	[(PERA X DDK)F_1 X B6]N_2
chromosome-5		
49.372	53.268	[(PERA X DDK)F_1 X B6]N_2
6.925	12.730	[PERC X DDK)F_1 X B6]N_2
25.926	33.198	[(SM X NZB)F_1 X NZB]N $_2$
Chromosome-6		
0	0.680	[PERC X DDK)F_1 X B6]N_2
44.08	50.680	[(B6-Pgk1^a X DDK)F_1X B6]N_2 (More skewed)
53.4	49.806	[(B6-Pgk1^a X DDK)F_1 X B6]N_2(Less skewed)
23.68	30.147	[(SM X NZB)F_1 X NZB]N $_2$
Chromosome-7		
40.405	47.174	[(B6-Pgk1^a X DDK)F_1X B6]N_2 (More skewed)
30.12	32.976	[(SM X NZB)F_1 X NZB]N $_2$
Chromosome-8		
52.52	56.640	[(B6-Pgk1^a X DDK)F_1X B6]N_2 (More skewed)
Chromosome-9		
41.16	41.162	[(B6-Pgk1^a X DDK)F_1X B6]N_2 (More skewed)
34.52	37.448	[(B6-Pgk1^a X DDK)F_1 X B6]N_2(Less skewed)
0	4.239	[(SM X NZB)F_1 X NZB]N $_2$

Continued...

$\hat{E}(g)$	$1/\hat{\lambda}$	Cross
Chromosome-10		
54.905	60.485	[(PERA X DDK)F_1 X B6]N_2
64.35	54.060	[PERC X DDK)F_1 X B6]N_2
45.746	55.417	[(B6-Pgk1^a X DDK)F_1X B6]N_2 (More skewed)
Chromosome-11		
43.035	55.741	[(PERA X DDK)F_1 X B6]N_2
37.641	40.743	[(B6-Pgk1^a X DDK)F_1X B6]N_2 (More skewed)
65.722	69.216	[(B6-Pgk1^a X DDK)F_1 X B6]N_2(Less skewed)
Chromsome-12		
34.18	47.295	[(PERA X DDK)F_1 X B6]N_2
27.21	41.677	[(B6-Pgk1^a X DDK)F_1X B6]N_2 (More skewed)
25.53	50.375	[(B6-Pgk1^a X DDK)F_1 X B6]N_2(Less skewed)
Chromosome-13		
36	43.298	[PERC X DDK)F_1 X B6]N_2
39.92	50.201	[(B6-Pgk1^a X DDK)F_1 X B6]N_2(Less skewed)
Chromosome-14		
24.555	30.298	[(PERA X DDK)F_1 X B6]N_2
26.84	43.610	[(B6-Pgk1^a X DDK)F_1X B6]N_2 (More skewed)
Chromosome-15		
28.57	55.692	[(B6-Pgk1^a X DDK)F_1X B6]N_2 (More skewed)
Chromosome-16		
25.18	29.890	[(B6-Pgk1^a X DDK)F_1X B6]N_2 (More skewed)
Chromosome-17		
20.685	23.785	[(PERA X DDK)F_1 X B6]N_2
35.74	41.058	[(B6-Pgk1^a X DDK)F_1 X B6]N_2(Less skewed)
Chromosome-18		
38.44	39.398	[(B6-Pgk1^a X DDK)F_1 X B6]N_2(Less skewed)
Chromosome-20		
24.59	27.593	[PERC X DDK)F_1 X B6]N_2
39.71	51.587	[(B6-Pgk1^a X DDK)F_1X B6]N_2 (More skewed)
35.1	57.916	[(B6-Pgk1^a X DDK)F_1 X B6]N_2(Less skewed)

Only the results of the crosses and chromosomes are presented here for which nonparametric estimate of the survival function was found and for which the condition $\hat{\mu}_j = N - \hat{\beta}_j^{SC} \geq 0,\ j = 1, 2, ..., J$ was satisfied.

High differences between $\hat{E}(g)$ and $1/\hat{\lambda}$ give us evidence that there may be interference in 20 chromosomes. To know if the difference is significant we have to do a test. However, in this work such a test was not developed, then it goes beyond the scope of the present thesis.

The chromosomes that show evidence of interference are chromosomes 11 and 12 of [(PERA X DDK)F_1 X B6]N_2, chromosomes $2, 5, 10$ and 13 of [PERC X DDK)F_1 X B6]N_2, chromosomes $6, 7, 10, 12, 14, 15$ and 20 of [(B6-Pgk1^a X DDK)F_1X B6]N_2 (More skewed), chromosomes $6, 12, 13$ and 20 of [(B6-Pgk1^a X DDK)F_1 X B6]N_2(Less skewed) and chromosomes $5, 6$ and 9 of [(SM X NZB)F_1 X NZB]N_2.

5.7 Comments on interference

Results from Sections 5.3 and 5.5 show evidence of interference in several chromosomes of different crosses. But since both of them showed evidence of interference in chromosome 13 of [PERC X DDK)F_1 X B6]N_2 and chromosome 10 of [(B6-Pgk1^a X DDK)F_1X B6]N_2 (More skewed), it is confirmed that they have interference.

Chapter 6

Conclusion

In mice we cannot recover the four products of meiosis, nor observe where the crossovers actually take place. For this reason, to infer about the crossover or chiasma process or check whether there is interference in the chromosomes, finding an appropriate model is necessary. Most statistical models use the genetic distance between successive crossovers as the basis for inference. However, since we do not know the exact location of crossovers, we cannot observe the distance between two successive crossovers and thus estimation of the distribution of the genetic distance is not possible. Fortunately we can find a set of flanking markers from the data, which means we then have a set of type-2 double censored observation.

The data sets that were used in this practicum consisted of five different crosses of female mice. Most of them were generated at Jackson Laboratory, Bar Harbor, Maine, US. Consensus genetic map and physical distance were found from some well known websites and interpolation and extrapolation method were used to find some missing cM and MB. After plotting the estimated genetic distance, some problem markers were identified and removed from the data. It is well-known that the chemistry of some

markers results in genotype detection that is not reliable. When the data was ready to use, type-2 censored data was obtained, presented and used with the techniques in this manuscript.

Expected distance between consecutive crossovers was found through a likelihood that did not consider the phenomenon of interference i.e. assuming that the distance between consecutive crossovers follows exponential distribution. However, since ignoring the effects of interference is not practical, we turned our attention to non-parametric alternatives. Although among several non-parametric methods the quadratic programming based on the second order Taylor series expansion should be preferred, as the estimate found in this method satisfies all the KKT conditions, in this thesis we used the method suggested by Turnbull (1976) depicted in Section 3.7.4 for time-to-event data. Afterwards, the estimates were checked to see if they had met the KKT conditions. Only the ones that satisfied the condition were taken for further analysis.

Expected distances under interference and no-interference modes were then compared in Table 5.1 . The result revealed that out of a total of 99 chromosomes of five crosses in this thesis, 20 chromosomes may have interference. These chromosomes were, chromosomes 11 and 12 of $\texttt{[(PERA X DDK)F}_1\texttt{ X B6]N}_2$, chromosomes $2, 5, 10$ and 13 of $\texttt{[PERC X DDK)F}_1\texttt{ X B6]N}_2$, chromosomes $6, 7, 10, 12, 14, 15$ and 20 of $\texttt{[(B6-Pgk1}^a\texttt{ X DDK)F}_1\texttt{X B6]N}_2$ (More skewed), chromosomes $6, 12, 13$ and 20 of $\texttt{[(B6-Pgk1}^a\texttt{ X DDK)F}_1\texttt{ X B6]N}_2$(Less skewed) and chromosomes $5, 6$ and 9 of $\texttt{[(SM X NZB)F}_1\texttt{ X NZB]N}_2$. For a visual confirmation, graphs were plotted for local ratio by physical distance for these 20 chromosomes to look for the evidence of interference. The graphs confirmed that only two of the chromosomes, chromosome 13 of cross $\texttt{[PERC X DDK)F}_1\texttt{ X B6]N}_2$ and chromosome 10 of cross $\texttt{[(B6-Pgk1}^a\texttt{ X DDK)F}_1\texttt{X B6]N}_2$ (More skewed) had interference.

Although we only found two chromosomes with interference, more chromosomes of this study may have interference. One reason for getting this result may be due to the use of the non-parametric analysis method suggested by Turnbull (1976). Most of the estimates that we found by using this method, did not meet the KKT conditions. So, this may not be a good method to find evidence of interference in the chromosome. Applying primal active set strategies for convex quadratic programming might lead us to find estimates that satisfy the KKT conditions. This means we could end up finding more chromosomes that have interference. Another reason for finding a very small number of chromosomes with interference might be the small data sets used in this study. In Section 5.3, we showed how a graphical presentation of the data was used to make a decision about the evidence of interference. However, because of the small dataset in many chromosomes, we could not obtain enough data points in the graphs to make any comment about the interfence.

At the beginning of this work, we considered different ways of modeling interference and made an extensive review. However, as time progressed we realized that inference on chromatid interference is a problem that requires of more exquisite models and larger data sets. We consider that physical models need to be taken into consideration to be able to assess the extent of chromatid interference. The magnitude of such endeavour falls beyond the scope of this thesis. Besides, fundamental models that may address some aspects of chromatid interference (like two pathway methods) require specialized data and such an information was not contained in our data set, so our contribution on this subject is a modest reviews that can be used as the basis of a comprehensive methodology.

Bibliography

[1] T. Billings, E. E. Sargent, J. P. Szatkiewicz, N. Leahy, I.-Y. Kwak, N. Bektassova, M. Walker, T. Hassold, J. H. Graber, K. W. Broman, et al. Patterns of recombination activity on mouse chromosome 11 revealed by high resolution mapping. *PLoS One*, 5(12):e15340, 2010.

[2] K. W. Broman and J. L. Weber. Characterization of human crossover interference. *The American Journal of Human Genetics*, 66(6):1911–1926, 2000.

[3] E. Cheney and D. Kincaid. *Numerical mathematics and computing*. Nelson Education, 2012.

[4] E. de Boer, P. Stam, A. J. Dietrich, A. Pastink, and C. Heyting. Two levels of interference in mouse meiotic recombination. *Proceedings of the National Academy of Sciences*, 103(25):9607–9612, 2006.

[5] E. de La Casa-Esperón, J. C. Loredo-Osti, F. P.-M. de Villena, T. L. Briscoe, J. M. Malette, J. E. Vaughan, K. Morgan, and C. Sapienza. X chromosome effect on maternal recombination and meiotic drive in the mouse. *Genetics*, 161(4):1651–1659, 2002.

[6] B. L. Dumont and B. A. Payseur. Evolution of the genomic rate of recombination in mammals. *Evolution*, 62(2):276–294, 2008.

[7] E. Foss, R. Lande, F. Stahl, and C. Steinberg. Chiasma interference as a function of genetic distance. *Genetics*, 133(3):681–691, 1993.

[8] E. J. Foss and F. W. Stahl. A test of a counting model for chiasma interference. *Genetics*, 139(3):1201–1209, 1995.

[9] F. Gauthier, O. C. Martin, and M. Falque. Coda (crossover distribution analyzer): quantitative characterization of crossover position patterns along chromosomes. *BMC bioinformatics*, 12(1):1, 2011.

[10] R. Gentleman and C. J. Geyer. Maximum likelihood for interval censored data: Consistency and computation. *Biometrika*, 81(3):618–623, 1994.

[11] R. Gentleman and A. C. Vandal. Computational algorithms for censored-data problems using intersection graphs. *Journal of Computational and Graphical Statistics*, 10(3):403–421, 2001.

[12] S. R. Giolo. Turnbull's nonparametric estimator for interval-censored data. *Department of Statistics, Federal University of Paraná*, pages 1–10, 2004.

[13] E. Housworth and F. Stahl. Crossover interference in humans. *The American Journal of Human Genetics*, 73(1):188–197, 2003.

[14] E. A. Housworth and F. W. Stahl. Is there variation in crossover interference levels among chromosomes from human males? *Genetics*, 183(1):403–405, 2009.

[15] F. Y. Ideraabdullah, K. Kim, D. Pomp, J. L. Moran, D. Beier, and F. P.-M. de Villena. Rescue of the mouse ddk syndrome by parent-of-origin-dependent modifiers. *Biology of reproduction*, 76(2):286–293, 2007.

[16] A. J. Jeffreys, L. Kauppi, and R. Neumann. Intensely punctate meiotic recombination in the class ii region of the major histocompatibility complex. *Nature genetics*, 29(2):217–222, 2001.

[17] H. Khalil, H. Tummala, and N. Zhelev. Atm in focus: A damage sensor and cancer target. *Biodiscovery*, 5, 2012.

[18] J. S. King and R. Mortimer. A polymerization model of chiasma interference and corresponding computer simulation. *Genetics*, 126(4):1127–1138, 1990.

[19] N. Kleckner, D. Zickler, G. H. Jones, J. Dekker, R. Padmore, J. Henle, and J. Hutchinson. A mechanical basis for chromosome function. *Proceedings of the National Academy of Sciences of the United States of America*, 101(34):12592–12597, 2004.

[20] J. P. Klein and M. L. Moeschberger. *Survival analysis: techniques for censored and truncated data*. Springer Science & Business Media, 2005.

[21] K. Lange. *Mathematical and statistical methods for genetic analysis*. Springer Science & Business Media, 2003.

[22] K. Lange and N. Risch. Comments on lack of interference in the four strand model of crossing over. *Journal of mathematical biology*, 5(1):55–59, 1977.

[23] J. Lian, Y. Yin, M. Oliver-Bonet, T. Liehr, E. Ko, P. Turek, F. Sun, and R. H. Martin. Variation in crossover interference levels on individual chromosomes from human males. *Human molecular genetics*, 17(17):2583–2594, 2008.

[24] L. Liu, H. Parekh-Olmedo, and E. B. Kmiec. The development and regulation of gene repair. *Nature Reviews Genetics*, 4(9):679–689, 2003.

[25] A. Lynn, S. Schrump, J. Cherry, T. Hassold, and P. Hunt. Sex, not genotype, determines recombination levels in mice. *The American Journal of Human Genetics*, 77(4):670–675, 2005.

[26] V. G. Nagabhushana. Computational analysis of variability of recombination rates in mice. Master's thesis, University of Texas, Arlington, August 2008.

[27] K. Paigen, J. P. Szatkiewicz, K. Sawyer, N. Leahy, E. D. Parvanov, S. H. Ng, J. H. Graber, K. W. Broman, and P. M. Petkov. The recombinational anatomy of a mouse chromosome. *PLoS Genet*, 4(7):e1000119, 2008.

[28] P. M. Petkov, K. W. Broman, J. P. Szatkiewicz, and K. Paigen. Crossover interference underlies sex differences in recombination rates. *Trends in Genetics*, 23(11):539–542, 2007.

[29] W. A. Pitman, R. Korstanje, G. A. Churchill, E. Nicodeme, J. J. Albers, M. C. Cheung, M. A. Staton, S. S. Sampson, S. Harris, and B. Paigen. Quantitative trait locus mapping of genes that regulate hdl cholesterol in sm/j and nzb/b1nj inbred mice. *Physiological genomics*, 9(2):93–102, 2002.

[30] K. Ross, P. Fransz, and G. Jones. A light microscopic atlas of meiosis inarabidopsis thaliana. *Chromosome research*, 4(7):507–516, 1996.

[31] S. M. Ross et al. *Stochastic processes*, volume 2. John Wiley & Sons New York, 1996.

[32] S. B. Roy. *Distribution of meiotic crossovers: heterogeneity, modeling interference, inter-pathway crosstalk*. PhD thesis, Université Paris Sud-Paris XI, 2014.

[33] T. P. Speed. Genetic map functions. In *Encyclopedia of Biostatistics*. John Wiley and Sons, Ltd, 2nd edition, 2005.

[34] F. Stahl. Defining and detecting crossover-interference mutants in yeast. *PloS one*, 7(6):e38476, 2012.

[35] F. W. Stahl and H. M. Foss. A two-pathway analysis of meiotic crossing over and gene conversion in saccharomyces cerevisiae. *Genetics*, 186(2):515–536, 2010.

[36] F. W. Stahl, H. M. Foss, L. S. Young, R. H. Borts, M. F. Abdullah, and G. P. Copenhaver. Does crossover interference count in saccharomyces cerevisiae? *Genetics*, 168(1):35–48, 2004.

[37] F. W. Stahl and E. A. Housworth. Methods for analysis of crossover interference in saccharomyces cerevisiae. *Meiosis: Volume 1, Molecular and Genetic Methods*, pages 35–53, 2009.

[38] R. Strich. Meiotic dna replication. *Current topics in developmental biology*, 61:29–60, 2004.

[39] F. Teuscher, G. Brockmann, P. Rudolph, H. Swalve, and V. Guiard. Models for chromatid interference with applications to recombination data. *Genetics*, 156(3):1449–1460, 2000.

[40] A. C. Vandal, R. Gentleman, and X. Liu. Constrained estimation and likelihood intervals for censored data. *The Canadian Journal of Statistics/La Revue Canadienne de Statistique*, pages 71–83, 2005.

[41] H. Zhao and T. P. Speed. Stochastic modeling of the crossover process during meiosis. *Communications in Statistics-Theory and Methods*, 27(6):1557–1580, 1998.

[42] H. Zhao, T. P. Speed, and M. S. McPeek. Statistical analysis of crossover interference using the chi-square model. *Genetics*, 139(2):1045–1056, 1995.

[43] D. Zickler and N. Kleckner. The leptotene-zygotene transition of meiosis. *Annual review of genetics*, 32(1):619–697, 1998.

Appendix A

Table 1: Estimated Genetic distance of Cross [(PERA X DDK)F_1 X B6]N_2

Chr.	Marker	Gen.	Est.
1	rs13475712	6.5	6.5
1	rs13475735	8.7	11.9
1	rs3711079	10	23.38
1	rs4222215	15	26.01
1	rs13475771	15.3	26.01
1	rs3677683	17	26.01
1	01.029.481	17.8	26.01
1	01.035.780	21	34.37
1	01.041.550	25.7	39.78
1	01.046.600	25.7	39.78
1	mCV23591750	25.7	39.78
1	rs3716105	32.8	42.41
1	rs6356603	41	47.82
1	01.076.110	47	51.07
1	01.087.170	54.5	61.77
1	rs13475982	58.5	61.77
1	rs13475988	58.7	61.77
1	rs13475989	58.7	61.77
1	rs13475991	58.7	61.77
1	rs6342650	58.7	61.77
1	rs6358447	58.7	61.77
1	CEL-1-98681809	58.7	61.77
1	CEL-1-98799654	58.7	61.77
1	rs3695980	58.7	61.77
1	rs13476012	58.7	61.77
1	rs3685663	59.5	64.41
1	rs3717264	59.5	64.41
1	rs3664662	59.5	64.41
1	rs3695581	63.1	72.76
1	rs13476089	63.1	72.76
1	rs6355835	67	75.39
1	01.135.010	70	75.39
1	rs13476147	73	78.02
1	rs6364156	73	78.02
1	rs13476259	78	113.98
1	01.178.925	100	113.98
1	CEL-1-181947877	101	113.98
1	rs13476290	102	119.39
1	rs3654705	106.3	119.39
1	rs6246360	112	119.39
2	rs13476318	2	2
2	rs13476330	4	7.4
2	02.016.175	9.5	18.9
2	rs6181760	17	24.33
2	02.041.990	29	29.74
2	02.054.160	30.5	33.07
2	02.065.760	37	43.86
2	rs6371268	38.3	43.86

Continued...

Chr.	Marker	Gen.	Est.
2	rs13476560	40.4	43.86
2	rs13476563	40.4	43.86
2	mCV25095764	44	46.5
2	CEL-2-79237503	47.9	46.5
2	rs3722345	48	49.13
2	02.079.300	48.1	49.13
2	02.083.650	48.1	54.55
2	rs4223268	48.1	59.97
2	02.093.700	50.3	62.6
2	rs13476663	50.3	65.23
2	rs6249987	50.3	65.23
2	rs13476667	50.3	65.23
2	rs3674721	50.3	65.23
2	mCV25337624	50.3	65.23
2	rs13476684	50.3	65.23
2	rs13476689	51.4	65.23
2	rs3022892	51.4	65.23
2	rs3693678	51.4	65.23
2	rs3701250	52.5	67.86
2	rs6276129	65.5	70.49
2	rs3723406	65.5	73.13
2	rs6340352	67	73.13
2	rs3697020	70	73.13
2	rs6411422	73	73.13
2	02.128.325	73	75.76
2	02.130.220	73.5	79.77
2	rs13476794	78	83.78
2	rs3710324	78	83.78
2	rs13476805	79.4	83.78
2	rs6360457	79.7	86.41
2	gnf02.141.261	80	86.41
2	rs6195594	81	89.05
2	rs3655895	81.7	89.05
2	rs3696870	82	89.05
2	02.146.685	82	89.05
2	rs3726342	92	97.4
2	rs6204920	92	100.03

Chr.	Marker	Gen.	Est.
2	CEL-2-163612103	96	102.66
2	rs6219107	99	108.07
2	rs3673248	100	113.47
3	rs6398851	4.6	4.6
3	rs13477046	16.5	22.98
3	D3Mit63	22	31.33
3	rs6239288	26.4	46.14
3	rs3696955	30	46.14
3	rs6226544	33.7	48.78
3	CEL-3-68001820	33.7	48.78
3	rs13477178	33.7	51.41
3	rs6198234	33.7	51.41
3	rs3698109	33.7	51.41
3	rs13477190	33.7	54.04
3	rs13477210	35.2	54.04
3	rs3715352	35.2	54.04
3	rs13477215	35.2	54.04
3	gnf03.079.138	38.3	59.45
3	rs6376008	39.7	62.08
3	rs13477244	39.7	62.08
3	rs3720007	45.2	64.71
3	rs6391963	49.7	67.35
3	rs3686473	49.7	67.35
3	03.100.150	49.7	67.35
3	03.119.365	55	69.98
3	rs6214597	55	69.98
3	03.152.282	79.4	88.37
3	rs13477498	83.5	88.37
3	rs6331755	84.9	88.37
4	rs13477534	0	0
4	rs13477546	1.9	0
4	04.011.950	5.2	0
4	rs13477592	5.2	2.63
4	rs13477599	6.3	2.63
4	04.021.985	6.3	8.16
4	04.029.760	7.5	26.21
4	rs13477662	12.1	29.3

Continued...

Chr.	Marker	Gen.	Est.
4	04.042.340	14.5	34.71
4	04.049.500	21.9	34.71
4	rs13477741	32.3	34.71
4	rs6258088	40	40.12
4	04.078.330	40	40.12
4	04.099.005	49.6	42.75
4	rs13477895	50.8	45.38
4	rs3670382	54.4	45.38
4	rs3696331	56	48.02
4	04.115.380	57	48.02
4	rs3671259	57.6	48.02
4	rs13477972	59.1	48.02
4	04.128.160	60	57.91
4	rs6268364	81	71.56
4	rs3693087	81	71.56
5	rs13478092	1	1
5	CEL-5-14611794	5	12.47
5	rs13478133	8	20.82
5	UT-5-19.849706	8	20.82
5	rs13478138	9	20.82
5	rs3706626	11	20.82
5	rs13478151	12	23.45
5	rs13478157	15	23.45
5	UT-5-30.642219	18	23.45
5	05.035.200	24	31.81
5	rs3716195	26	31.81
5	rs13478210	26	34.44
5	rs13478212	26	34.44
5	rs13478215	26	37.07
5	mCV27558149	39	51.88
5	gnf05.061.650	41	51.88
5	05.067.560	41	51.88
5	05.071.190	42	51.88
5	05.073.500	44	51.88
5	05.085.655	45	51.88
5	05.091.725	45	54.53
5	rs13478428	54	57.18

Chr.	Marker	Gen.	Est.
5	05.108.560	61	62.59
5	CEL-5-120064766	67	70.94
5	05.117.270	67	70.94
5	05.121.625	72	73.57
5	rs13478521	72	73.57
5	05.135.860	81	81.94
5	rs3668534	86	90.31
6	06.002.580	0.6	0.6
6	06.005.375	2.6	0.6
6	rs3655269	3.2	6
6	06.033.580	15.8	23.17
6	06.063.270	29.42	26.82
6	mhcCD8b4	30.5	26.82
6	rs13478841	33.5	29.45
6	06.085.360	36.5	34.87
6	06.097.530	43	34.87
6	06.102.675	46.5	37.51
6	rs3655148	46.5	45.86
6	rs6204829	51.5	48.49
6	06.118.265	51.5	51.12
6	rs3695724	58.8	51.12
6	CEL-6-122563022	59.3	51.12
6	06.125.555	61.2	59.41
6	rs3670851	62.3	67.69
6	rs6339546	63.9	70.32
6	06.140.060	68	75.76
6	rs6387265	71.4	78.4
7	07.000.385	0	0
7	rs13479163	8	5.4
7	mCV25220583	10.4	5.4
7	CEL-7-29429804	18	10.81
7	rs6313526	23.5	10.81
7	rs6295036	23.5	10.81
7	CEL-7-36545579	24.5	10.81
7	rs13479238	24.5	10.81
7	07.047.960	26.8	10.81
7	gnf07.050.858	27.8	10.81

Continued...

Chr.	Marker	Gen.	Est.
7	mCV23672419	27.8	10.81
7	rs3693038	27.8	10.81
7	rs13479274	27.8	10.81
7	rs13479276	27.8	10.81
7	rs6160140	27.8	13.44
7	rs3705155	37	16.08
7	rs13479317	37	16.08
7	rs3693876	37	16.08
7	rs13479321	37	18.71
7	rs13479334	37	18.71
7	rs13479338	37	18.71
7	07.077.280	40	31.88
7	rs13479427	50.3	45.06
7	rs3719258	65.6	63.45
7	D7Mit291	66	68.73
7	rs3663988	69	76.93
8	rs13479627	8	8
8	08.020.285	8	12.24
8	08.029.780	13	17.16
8	rs13479741	28	46.28
8	rs3726906	31.5	48.91
8	08.062.280	32	48.91
8	rs3712611	33	48.91
8	rs13479813	33	48.91
8	08.076.440	34	48.91
8	rs3690549	35.5	51.55
8	rs6296891	37	54.18
8	rs8236770	38.1	56.81
8	08.086.390	38.6	56.81
8	D8Mit322	61	83.34
8	rs13480026	67	94.91
8	rs4227456	67	94.91
9	09.029.420	17	17
9	rs3669224	17	17
9	rs13480160	25	17
9	09.038.640	26	17
9	09.043.100	27	22.41
9	09.052.500	31	22.41
9	09.060.410	35	25.05
9	rs6174757	38	27.68
9	rs13480267	41	27.68
9	rs13480277	42	27.68
9	rs13480285	43	27.68
9	rs3725272	49	46.07
9	09.094.990	52	46.07
9	09.104.660	60	56.39
9	rs13480436	61	63.01
9	rs6302293	67	63.01
9	gnf09.117.044	74	65.63
10	rs13480480	2	2
10	rs6192001	5	7.4
10	10.009.900	7	10.04
10	10.011.300	7	12.68
10	rs13480506	7	12.68
10	rs13480525	9	15.31
10	10.038.835	26.5	45.97
10	D10Mit40	29	48.91
10	rs3165937	40.7	60.39
10	rs13480703	50	65.8
10	rs3705990	56	71.21
10	rs13480804	69.5	126.15
10	10.127.600	70	126.15
11	rs13480847	1.1	1.1
11	rs13480869	2.4	1.1
11	rs13480889	8	6.51
11	rs3678321	11	6.51
11	rs3657760	17	28.75
11	rs13480997	20	31.38
11	mCV23044839	36	46.19
11	UT-11-68.607315	38	51.6
11	rs13481123	47	54.24
11	gnf11.093.966	47	54.24
11	rs3714299	53	62.59
11	rs3710148	56	70.94

Continued...

Chr.	Marker	Gen.	Est.
11	rs6384437	65	89.33
11	rs13481226	68	91.95
12	rs13481276	1	1
12	12.014.515	6	19.38
12	rs6187012	14	30.86
12	rs6223000	16	30.86
12	12.032.830	16	30.86
12	gnf12.033.545	16	30.86
12	rs6243157	17	30.86
12	rs3689063	25	45.67
12	rs13481465	25	45.67
12	rs3686891	28	48.3
12	rs3686378	29	48.3
12	rs13481531	29	48.3
12	mCV23169261	32	48.3
12	12.081.010	37	50.93
12	rs3696951	38	50.93
12	rs8259763	43	53.56
12	12.099.140	48	75.8
12	gnf12.101.501	50	75.8
12	CEL-12-101776500	52	78.42
13	D13Mit158	5	5
13	13.005.379	7	13.84
13	13.010.063	7	13.84
13	D13Mit172	7	13.84
13	13.010.368	7	13.84
13	13.011.447	7	13.84
13	13.013.030	7	13.84
13	13.013.605	7	13.84
13	rs3721858	7	22.19
13	13.017.553	7	27.61
13	13.020.344	8	33.03
13	13.019.050	10	33.03
13	D13Mit135	10	33.03
13	13.021.844	10	33.03
13	RS6158895	11	33.03
13	13.028.853	14	41.39

Chr.	Marker	Gen.	Est.
13	13.031.107	16	44.04
13	13.034.400	16	44.04
13	13.035.013	16	44.04
13	13.034.725	16	44.04
13	rs6259014	16	44.04
13	gnf13.045.330	30	58.85
13	rs6209128	34	58.85
13	rs3700819	36	64.26
13	mCV22624058	36	64.26
13	CEL-13-60831741	36	66.89
13	13.061.625	36	66.89
13	13.066.450	40	69.53
13	rs6179438	43	81.02
13	13.080.001	44	86.43
13	13.083.500	45	86.43
13	rs13481918	45	86.43
13	13.087.830	45	86.43
13	rs3655061	46	86.43
13	rs4230027	51	86.43
13	rs3705092	59	89.07
13	rs13481992	59	89.07
13	13.114.540	71	94.47
13	rs3657414	73	94.47
14	14.002.500	1.75	1.75
14	rs6340768	1.75	1.75
14	rs13482084	2.5	1.75
14	14.006.480	3	1.75
14	rs3719629	3	1.75
14	14.015.365	5.5	4.38
14	14.020.160	7.5	4.38
14	rs3722090	12.5	12.74
14	14.032.950	13.5	18.2
14	mCV23384307	17	26.61
14	rs13482179	17	26.61
14	rs6392664	22.5	29.24
14	14.055.010	22.5	31.89
14	rs13482214	27.5	34.54

Continued...

Chr.	Marker	Gen.	Est.
14	14.058.715	28	34.54
14	rs6179045	28.3	34.54
14	14.065.150	32.5	42.85
14	CEL-14-65598536	42	48.22
14	rs13482259	43.5	48.22
14	rs6298191	44	48.22
14	14.074.265	44.3	48.22
14	CEL-14-71690454	44.3	48.22
14	rs3023412	44.3	48.22
14	rs6325141	44.3	50.86
14	rs13482301	44.5	53.49
14	rs3725470	44.5	53.49
14	gnf14.085.610	45	53.49
14	rs6395984	45	53.49
14	rs6407863	45	53.49
14	rs3706761	45	53.49
14	rs3655019	45	53.49
14	rs6291434	45	53.49
14	rs6176735	45	53.49
14	rs6299927	45	53.49
14	rs4139735	45	53.49
14	CEL-14-85152539	45	53.49
14	rs13482311	45	53.49
14	rs13482312	45	53.49
14	rs13482313	45	53.49
14	rs13482314	45	53.49
14	14.093.815	45	53.49
14	rs3708779	54	61.84
14	rs3683221	54.5	64.46
15	rs13482418	6.7	6.7
15	15.017.570	14.5	15.99
15	rs13482490	14.5	37
15	15.030.400	14.8	42.41
15	15.053.380	26.2	53.89
15	rs3701449	26.2	53.89
15	rs13482618	29.6	53.89
15	15.066.375	39	56.53

Chr.	Marker	Gen.	Est.
15	15.078.600	46.7	56.53
15	rs3697744	48.5	59.16
15	rs3716673	57.8	61.8
15	rs3690173	63.4	70.14
16	rs4166445	20.9	20.9
16	rs4168890	20.9	23.52
16	16.029.450	22.2	23.52
16	rs4170974	23.5	23.52
16	16.031.880	24	28.93
16	rs4174174	27.8	28.93
16	16.035.790	27.8	28.93
16	rs4175353	27.8	28.93
16	16.055.570	38	34.35
16	rs4197416	54	39.77
16	16.075.770	54	48.14
16	rs3656592	57.7	50.78
16	rs4217061	58	50.78
16	rs3164088	70.65	53.4
17	rs3662575	7.6	7.6
17	17.013.500	9.32	10.04
17	rs6270865	9.32	15.27
17	rs3667748	9.33	15.27
17	rs4231344	12.6	17.9
17	17.022.870	12.6	17.9
17	rs3724223	18.15	17.9
17	UT-17-33.238924	18.15	17.9
17	rs8242408	18.7	17.9
17	rs3682923	20.9	17.9
17	17.041.250	21.65	17.9
17	mCV25197172	22.5	20.53
17	17.043.515	22.9	20.53
17	rs8273969	33.5	32.01
17	rs3675634	42	50.4
17	rs3687741	55.7	61.9
17	D17Mit123	56.7	67.33
18	rs13483210	4	4
18	rs3656185	9	15.47

Continued...

Chr.	Marker	Gen.	Est.
18	rs3691362	16	30.28
18	rs3714096	16	32.91
18	18.046.110	32	35.95
18	rs6350869	36	57.8
18	18.070.235	45	61.86
18	rs13483472	56	65.9
19	rs13483500	4	4
19	rs3671671	4	4
19	rs3713033	4	9.41
19	rs13483505	4	9.41
19	rs13483511	4	9.41
19	19.005.500	4	9.41
19	CEL-19-8529644	5	9.41
19	UT-19-10.709331	7	12.04
19	rs6163293	7	12.04
19	rs3700209	7	12.04
19	rs6237846	7	12.04
19	CEL-19-12911424	7	12.04
19	rs3692733	7	12.04
19	rs3669192	10.9	23.52
19	gnf19.017.711	15	26.15
19	rs3720318	15	26.15
19	rs13483557	15	26.15
19	19.018.140	15	26.15
19	rs13483563	15	28.79
19	rs6392565	15	28.79
19	rs3672759	16.4	31.42
19	rs3653630	16.4	31.42
19	rs13483577	20	31.42
19	rs3090325	20	31.42

Chr.	Marker	Gen.	Est.
19	19.043.320	43	41.53
19	19.048.500	48	51.64
19	rs13483669	48	51.64
19	rs6257938	52	51.64
X	DXMit166	15.5	15.5
X	rs13483834	30.7	24.08
X	CEL-X-68179178	30.7	24.08
X	CEL-X-68645226	30.7	24.08
X	rs13483838	30.7	24.08
X	CEL-X-71104123	31	24.08
X	CEL-X-72627341	31	24.08
X	CEL-X-73027245	31	24.08
X	gnfX.070.167	31	24.08
X	rs13483858	31	24.08
X	CEL-X-74073918	31	24.08
X	CEL-X-74272691	31	24.08
X	rs13483862	31	24.08
X	rs13483863	31	24.08
X	CEL-X-75125049	31	24.08
X	rs13483877	33.2	24.08
X	rs13483803	33.5	24.08
X	gnfX.076.619	34.6	24.08
X	rs13483888	37	29.49
X	DXMit16	37	32.13
X	CEL-X-91222960	37	32.13
X	CEL-X-94143306	41.5	34.76
X	rs13483935	48.4	40.17
X	rs13484004	55	48.52
X	rs13484094	63	60
X	DXMit29	73.3	62.77

NB: “Chr.” is the chromosome number; “Gen.” is the consensus genetic distance; “Est.” is the estimated genetic distance.

Table 2: Estimated Genetic distance of Cross [PERC X DDK)F_1 X B6]N_2

Chr.	Marker	Gen.	Est.
1	01.029.481	17.8	17.8
1	01.035.780	21	26.91
1	01.041.550	25.7	32.8
1	01.046.600	26.2	32.8
1	01.070.445	36.9	51.56
1	01.076.110	47	58.04
1	01.135.010	70	124.67
1	D1Mit270	92.3	156.86
2	02.041.990	29	29
2	02.054.160	30.5	36.45
2	02.065.760	37	43.91
2	02.079.300	48.1	48.37
2	02.125.700	70	121.81
2	02.128.325	73	121.81
2	02.130.220	73.5	121.81
2	02.151.240	81.7	131.84
2	02.146.685	82	131.84
3	03.014.785	4.6	4.6
3	03.047.215	23.3	15.64
3	03.054.150	29.5	15.64
3	03.100.150	49.2	51.32
3	03.119.365	55	54.17
3	03.152.282	79.4	70.99
4	04.011.950	5.2	5.2
4	04.021.985	6.3	21.55
4	04.029.760	7.5	25.13
4	04.042.340	14.5	25.13
4	04.049.500	21.9	31.02
4	04.078.330	40	51.58
4	04.099.005	49.6	67.47
4	04.115.380	57	74.76
4	04.128.160	60	90.8
4	04.131.640	63	100.54
5	05.010.335	5	5
5	05.035.200	24	12.7
5	05.038.350	26	15.31
5	05.046.885	28	17.79
5	05.067.560	41	23.69
5	05.071.190	42	23.69
5	05.073.500	44	23.69
5	05.085.655	45	23.69
5	05.091.725	45	23.69
5	05.108.560	61	26.55
5	05.117.270	67	46.82
5	05.135.860	81	68.58
6	06.002.580	0.6	0.6
6	06.014.800	3.2	21.58
6	06.033.580	15.8	36.12
6	06.063.270	29.42	50.66
6	06.070.455	30.5	53.1
6	06.074.995	31.5	55.49
6	06.097.530	43	71.69
6	06.102.675	46.5	71.69
6	06.108.730	46.5	71.69
6	06.118.265	51	75.25
7	07.000.385	0.67	0.67
7	07.047.960	26.8	25.46
7	07.087.220	46.4	34.71
7	07.091.695	49.9	34.71
7	07.096.985	50	34.71
7	07.110.535	52.6	37.16
7	07.112.290	52.8	39.62
7	07.115.675	53	45.89
8	08.067.625	33	33
8	08.086.390	38.6	46.4
8	08.090.187	41	46.4
8	08.092.425	43	46.4
8	08.096.955	45	50.25
8	08.101.010	45	54.11
8	08.111.015	53	57.95
8	08.124.650	67	84.47
9	09.029.420	17	17

Continued...

Chr.	Marker	Gen.	Est.
9	09.043.100	27	17
9	09.052.500	31	19.86
9	09.060.410	35	19.86
9	09.091.925	49	66.43
9	09.104.660	60	66.43
9	09.119.240-2	73	66.43
10	10.009.900	7	7
10	10.011.300	7	7
10	10.015.980	9	16.86
10	10.038.835	26.5	20.93
10	D10Mit95	51	81.21
10	10.127.600	70	97.46
12	D12Mit112	22	22
12	12.089.840	45	69.22
12	12.099.140	48	73.37
13	13.001.770	5	5
13	13.005.379	7	5
13	13.010.063	7	5
13	D13Mit172	7	7.86
13	13.010.368	7	7.86
13	13.011.447	7	10.72
13	13.013.030	7	13.58
13	13.013.605	7	13.58
13	13.017.553	7	16.44
13	13.020.344	8	16.44
13	13.019.050	10	16.44
13	D13Mit135	10	16.44
13	13.021.844	10	16.44
13	RS6158895	11	16.44
13	13.028.853	14	29
13	13.031.107	16	29
13	13.034.400	16	29
13	13.035.013	16	29
13	13.034.725	16	29
13	13.061.625	36	41.67
13	13.080.001	44	44.57
13	13.090.665	47	65
13	13.114.540	71	71.89
14	14.002.500	1.75	1.75
14	14.006.480	3	1.75
14	14.015.365	5.5	1.75
14	14.024.450	10	8.57
14	14.032.950	13.5	24.89
14	14.051.890	21	29.97
14	14.055.010	21.5	41.83
14	14.065.150	32.5	45.5
14	14.074.265	44.1	48.42
14	14.093.815	44.1	57.57
14	14.101.995	63	63.65
15	15.030.400	14.8	14.8
15	15.044.240	18.2	21.05
15	15.051.335	21.4	26.94
15	15.066.375	39	29.82
15	15.075.005	44.1	42.4
15	15.078.600	46.7	42.4
15	15.098.880	56.8	55.8
16	16.029.450	22.2	22.2
16	16.031.880	26.5	22.2
16	16.055.570	38	31.59
17	17.013.500	9.32	9.32
17	17.022.870	11.7	12.9
17	17.038.280	20.43	18.8
17	17.041.250	21.65	18.8
17	17.052.240	29.4	27.99
17	D17Mit39	45.3	33.94
18	18.025.640	16	16
18	18.042.400	22	22.45
18	18.051.990	26	25.31
18	18.061.750	37	31.2
18	18.072.500	47	31.2
19	19.003.270	4	4
19	19.018.140	15	10.73
19	19.043.320	43	34.56
19	19.048.500	51	40.62
X	DXMit166	15.5	15.5
X	DXMit16	37	44.87
X	DXMit234	58	69.46
X	DXMit29	73.3	88.72

NB: "Chr." is the chromosome number; "Gen." is the consensus genetic distance; "Est." is the estimated genetic distance.

Table 3: Estimated Genetic distance of Cross [(B6-Pgk1^a X DDK)F$_1$XB6]N$_2$ (More skewed)

Chr.	Marker	Gen.	Est.
1	D1Mit66	9	9
1	D1Mit318	18.5	24.84
1	D1Mit251	38.1	44.19
1	D1Mit132	43.1	50.78
1	D1Mit390	63.1	66.62
1	D1Mit424	81.6	89.75
1	D1Mit270	92.3	91
1	D1Mit293	109.6	118.21
2	D2Mit117	5	5
2	D2Mit83	16	19.17
2	D2Mit244	33	35.02
2	D2Mit37	45	49.19
2	D2Mit276	65	74.32
2	D2Mit285	86	95.52
2	D2Mit200	107	122.73
3	D3Mit164	2.4	2.4
3	D3Mit203	11.2	25.52
3	D3Mit63	22	43.09
3	D3Mit74	41	74.75
3	D3Mit254	64.1	116.71
3	D3Mit163	87.6	137.91
4	D4Mit227	3.2	3.2
4	D4Mit286	14.5	30.41
4	D4Mit164	28.6	38.42
4	D4Mit58	48.5	55.99
4	D4Mit37	56.5	68.56
4	D4Mit339	65.7	87.92
4	D4Mit256	82.7	117.3
5	D5Mit344	1	1
5	D5Mit79	26	45.88
5	D5Mit15	39	61.72
5	D5Mit314	59	88.93
5	D5Mit31	78	122.97
5	D5Mit143	86	130.98
6	D6Mit236	3.1	3.1
6	D6Mit384	27.5	28.22
6	D6Mit65	46	57.6
6	D6Mit25	65	87
6	D6Mit201	74.1	99.55
7	D7Mit178	0.5	0.5
7	D7Mit117	11	25.66
7	D7Mit310	18	32.27
7	D7Mit30	37	49.84
7	D7Mit323	50	56.42
7	D7Mit291	66	92.98
7	D7Mit223	72.4	102.45
8	D8Mit157	2	2
8	D8Mit191	21	38.53
8	D8Mit348	44	67.92
8	D8Mit166	56	91.05
8	D8Mit322	61	110.4
9	D9Mit126	6	6
9	D9Mit90	9	18.56
9	D9Mit97	29	25.14
9	D9Mit270	43	34.63
9	D9Mit212	61	66.3
9	D9Mit281	68	74.31
10	D10Mit298	3	3
10	D10Mit214	19	37.04
10	D10Mit40	29	49.6
10	D10Mit66	49	72.73
10	D10Mit233	62	97.87
10	D10Mit269	70	105.87
11	D11Mit71	1.1	1.1
11	D11Mit151	13	20.45
11	D11Mit20	20	33.01
11	D11Mit5	37	48.86
11	D11Mit66	47	55.45
11	D11Mit67	57	74.8
11	D11Mit168	71	97.92

Continued...

Chr.	Marker	Gen.	Est.
12	D12Mit182	2	2
12	D12Mit112	22	23.19
12	D12Mit260	45	50.4
12	D12Mit150	59	86.94
13	D13Mit16	10	10
13	D13Mit63	26	27.56
13	D13Mit99	40	46.92
13	D13Mit107	48	62.76
13	D13Mit78	75	73.76
14	D14Mit10	3	3
14	D14Mit54	12.5	39.54
14	D14Mit234	22.5	55.38
14	D14Mit162	44.3	66.38
14	D14Mit170	63	93.59
15	D15Mit12	4.7	4.7
15	D15Mit100	21	52.72
15	D15Mit234	34.2	63.72
15	D15Mit189	48.5	81.29
15	D15Mit16	61.7	108.51
15	D15Mit79	66.2	108.51
16	D16Mit182	3.4	3.4
16	D16Mit166	21	48.29
16	D16Mit140	42.8	62.47
16	D16Mit191	57.8	73.47
16	D16Mit106	71.45	81.48
17	D17Mit78	8.18	8.18
17	D17Mit175	17.7	16.19
17	D17Mit7	32.3	35.55
17	D17Mit39	45.3	49.72
17	D17Mit123	56.7	62.28
18	D18Mit67	4	4
18	D18Mit60	16	18.17
18	D18Mit123	31	41.3
18	D18Mit47	50	53.87
18	D18Mit144	57	60.45
19	D19Mit42	5	5
19	D19Mit111	15	22.55
19	D19Mit66	41	29.12
19	D19Mit137	55.7	48.47
20	DXMit124	2.8	2.8
20	DXMit166	15.5	32.18
20	DXMit210	29.5	48.03
20	DXPas29	42.15	63.87
20	DXMit117	50.8	71.89
20	DXMit135	69	91.24

NB: "Chr." is the chromosome number; "Gen." is the consensus genetic distance; "Est." is the estimated genetic distance.

Table 4: Estimated Genetic distance of Cross [(B6-Pgk1^a X DDK)F$_1$ X B6]N$_2$(Less skewed)

Chr.	Marker	Gen.	Est.
1	D1Mit66	9	9
1	D1Mit318	18.5	20.14
1	D1Mit251	38.1	62.94
1	D1Mit132	43.1	72.56
1	D1Mit390	63.1	88.63
1	D1Mit424	81.6	118.51
1	D1Mit270	92.3	126.63
1	D1Mit293	109.6	152.16
2	D2Mit117	5	5
2	D2Mit83	16	26.53
2	D2Mit244	33	46.18
2	D2Mit37	45	51.45
2	D2Mit276	65	79.13
2	D2Mit285	86	88.74
2	D2Mit200	107	125.94
3	D3Mit164	2.4	2.4
3	D3Mit203	11.2	18.47
3	D3Mit63	22	25.15
3	D3Mit74	41	41.22
3	D3Mit254	64.1	78.44
3	D3Mit163	87.6	108.33
4	D4Mit227	3.2	3.2
4	D4Mit286	14.5	28.73
4	D4Mit164	28.6	46.56
4	D4Mit58	48.5	56.17
4	D4Mit37	56.5	72.24
4	D4Mit339	65.7	95.74
4	D4Mit256	82.7	117.27
5	D5Mit344	1	1
5	D5Mit79	26	40.92
5	D5Mit15	39	70.81
5	D5Mit314	59	83.54
5	D5Mit31	78	143.75
5	D5Mit143	86	158.12
6	D6Mit236	3.1	3.1
6	D6Mit384	27.5	30.76
6	D6Mit65	46	54.26
6	D6Mit25	65	84.16
6	D6Mit201	74.1	93.76
7	D7Mit178	0.5	0.5
7	D7Mit117	11	23.99
7	D7Mit310	18	30.67
7	D7Mit30	37	58.34
7	D7Mit323	50	62.23
7	D7Mit291	66	108.07
7	D7Mit223	72.4	122.44
8	D8Mit157	2	2
8	D8Mit191	21	41.92
8	D8Mit348	44	65.41
8	D8Mit166	56	83.24
8	D8Mit322	61	99.3
9	D9Mit126	6	6
9	D9Mit90	9	20.38
9	D9Mit97	29	28.5
9	D9Mit270	43	44.58
9	D9Mit212	61	58.96
9	D9Mit281	68	73.34
10	D10Mit298	3	3
10	D10Mit214	19	14.14
10	D10Mit40	29	30.21
10	D10Mit66	49	49.86
10	D10Mit233	62	77.53
10	D10Mit269	70	95.36
11	D11Mit71	1.1	1.1
11	D11Mit151	13	12.25
11	D11Mit20	20	42.14
11	D11Mit5	37	59.98
11	D11Mit66	47	69.59
11	D11Mit67	57	95.14
11	D11Mit168	71	118.63

Continued...

Chr.	Marker	Gen.	Est.
12	D12Mit182	2	2
12	D12Mit112	22	47.81
12	D12Mit260	45	73.34
12	D12Mit150	59	101.01
13	D13Mit16	10	10
13	D13Mit63	26	44.65
13	D13Mit99	40	79.31
13	D13Mit107	48	84.57
13	D13Mit78	75	110.11
14	D14Mit10	3	3
14	D14Mit54	12.5	28.53
14	D14Mit234	22.5	48.19
14	D14Mit162	44.3	57.8
14	D14Mit170	63	68.95
15	D15Mit12	4.7	4.7
15	D15Mit100	21	30.23
15	D15Mit234	34.2	36.91
15	D15Mit189	48.5	48.07
15	D15Mit16	61.7	60.81
15	D15Mit79	66.2	60.81
16	D16Mit182	3.4	3.4
16	D16Mit166	21	43.32
16	D16Mit140	42.8	70.99

Chr.	Marker	Gen.	Est.
16	D16Mit191	57.8	76.26
16	D16Mit106	71.45	81.51
17	D17Mit78	8.18	8.18
17	D17Mit175	17.7	22.55
17	D17Mit7	32.3	42.21
17	D17Mit39	45.3	58.29
17	D17Mit123	56.7	74.36
18	D18Mit67	4	4
18	D18Mit60	16	18.37
18	D18Mit123	31	46.04
18	D18Mit47	50	67.58
18	D18Mit144	57	74.25
19	D19Mit42	5	5
19	D19Mit111	15	21.07
19	D19Mit66	41	40.72
19	D19Mit137	55.7	62.25
20	DXMit124	2.8	2.8
20	DXMit166	15.5	51.84
20	DXMit210	29.5	64.58
20	DXPas29	42.15	77.33
20	DXMit117	50.8	86.94
20	DXMit135	69	139.43

NB: “Chr.” is the chromosome number; “Gen.” is the consensus genetic distance; “Est.” is the estimated genetic distance.

Table 5: Estimated Genetic distance of Cross [(SM X NZB)F_1 X NZB]N_2

Chr.	Marker	Gen.	Est.
1	D1MIT20	18.5	18.5
1	D1MIT22	32.8	60.21
1	D1MIT92	64	106.46
1	D1MIT34	81.6	121.77
1	D1MIT17	106.3	168.03
2	D2MIT1	1	1
2	D2MIT88	30	42.72
2	D2MIT35	45	99.56
2	D2MIT49	95.5	150.84
3	D3MIT12	49	49
3	D3MIT11	49	50.91
3	D3MIT38	70.3	97.2
4	D4MIT2	6.5	6.5
4	D4MIT17	31.4	16.94
4	D4MIT9	44.5	32.28
4	D4MIT11	57.4	59.1
4	D4MIT312	69.8	74.43
5	D5MIT228	18	18
5	D5MIT114	44	65.78
5	D5MIT7	45	73.95
5	D5Mit10	54	86.79
5	D5MIT239	58	86.79
5	D5MIT316	59	86.79
5	D5MIT209	63	90.71
5	D5MIT65	68	90.71
5	D5MIT370	70	92.64
5	D5MIT29	72	96.53
5	D5MIT99	80	114.16
6	D6MIT50	3.3	3.3
6	D6MIT3	33.5	30.1
6	D6MIT44	51.5	53.78
6	D6MIT15	74	77.45
7	D7MIT25	16	16
7	D7MIT37	49.8	87.33
7	D7MIT71	65.2	117.45
7	D7Nds4	72.4	147.57
8	D8MIT4	14	14
8	D8MIT45	40	93.98
8	D8MIT56	73	121.26
9	D9MIT2	17	17
9	D9MIT21	31	50.7
9	D9MIT8	42	68.63
9	D9MIT15	61	89.96
10	D10Mit126	21	21
10	D10Mit11	50	78.09
10	D10MIT24	67	104.89
11	D11MIT19	13	13
11	D11MIT4	37	22.1
11	D11MIT41	49	45.78
11	D11MIT11	69	69.45
12	D12MIT2	19	19
12	D12MIT5	37	49.13
12	D12MIT7	50	72.8
12	D12MIT8	58	78.92
13	D13MIT34	30	30
13	D13Mit102	44	48.07
13	D13MIT73	55	71.88
13	D13MIT35	75	87.21
14	D14MIT15	11	11
14	D14MIT37	27.5	15
14	D14MIT7	44.5	34.29
14	D14MIT97	58	44.77
15	D15MIT18	18.7	18.7
15	D15MIT31	48.5	77.61
15	D15MIT39	56.6	107.38
16	D16MIT3	21	21
16	D16MIT5	38	58.36
16	D16MIT70	57	95.41
17	D17MIT50	23.2	23.2
17	D17MIT20	34.3	30.89

Continued...

Chr.	Marker	Gen.	Est.
17	D17MIT76	54.6	65.52
18	D18MIT34	12	12
18	D18MIT24	25	30.36
18	D18MIT9	42	46.03
19	D19MIT16	15	15

Chr.	Marker	Gen.	Est.
19	D19MIT27	43	74.48
19	D19MIT71	54	85.38
20	DXMIT89	3	3
20	DXMIT1	29.01	37.64

NB: "Chr." is the chromosome number; "Gen." is the consensus genetic distance; "Est." is the estimated genetic distance.

Table 6: Estimated recombination fraction of Cross [(PERA X DDK)F_1 X B6]N_2

Chr.	Gen. dis.	Recom. frac.	Chr.	Gen. dis.	Recom. frac.
1	5.4	0.0511862	1	5.41	0.05127596
1	11.48	0.10257426	1	0	0
1	2.63	0.02562028	1	0	0
1	0	0	2	5.4	0.0511862
1	0	0	2	11.5	0.1027332
1	0	0	2	5.43	0.05145541
1	8.36	0.07698481	2	5.41	0.05127596
1	5.41	0.05127596	2	3.33	0.03221532
1	0	0	2	10.79	0.09705177
1	0	0	2	0	0
1	2.63	0.02562028	2	0	0
1	5.41	0.05127596	2	0	0
1	3.25	0.03146627	2	2.64	0.02571515
1	10.7	0.09632581	2	0	0
1	0	0	2	2.63	0.02562028
1	0	0	2	0	0
1	0	0	2	5.42	0.05136569
1	0	0	2	5.42	0.05136569
1	0	0	2	2.63	0.02562028
1	0	0	2	2.63	0.02562028
1	0	0	2	0	0
1	0	0	2	0	0
1	0	0	2	0	0
1	0	0	2	0	0
1	2.64	0.02571515	2	0	0
1	0	0	2	0	0
1	0	0	2	0	0
1	8.35	0.07690019	2	0	0
1	0	0	2	2.63	0.02562028
1	2.63	0.02562028	2	2.63	0.02562028
1	0	0	2	2.64	0.02571515
1	2.63	0.02562028	2	0	0
1	0	0	2	0	0
1	35.96	0.25642909	2	0	0
1	0	0	2	2.63	0.02562028
1	0	0	2	4.01	0.03853413

Continued...

Chr.	Gen. dis.	Recom. frac.
2	4.01	0.03853413
2	0	0
2	0	0
2	2.63	0.02562028
2	0	0
2	2.64	0.02571515
2	0	0
2	0	0
2	0	0
2	8.35	0.07690019
2	2.63	0.02562028
2	2.63	0.02562028
2	5.41	0.05127596
2	5.4	0.0511862
3	18.38	0.15380296
3	8.35	0.07690019
3	14.81	0.12818066
3	0	0
3	2.64	0.02571515
3	0	0
3	2.63	0.02562028
3	0	0
3	0	0
3	2.63	0.02562028
3	0	0
3	0	0
3	0	0
3	5.41	0.05127596
3	2.63	0.02562028
3	0	0
3	2.63	0.02562028
3	2.64	0.02571515
3	0	0
3	0	0
3	2.63	0.02562028
3	0	0
3	18.39	0.15387219
3	0	0

Chr.	Gen. dis.	Recom. frac.
3	0	0
4	0	0
4	0	0
4	2.63	0.02562028
4	0	0
4	5.53	0.0523516
4	18.05	0.1515105
4	3.09	0.02996456
4	5.41	0.05127596
4	0	0
4	0	0
4	5.41	0.05127596
4	0	0
4	2.63	0.02562028
4	2.63	0.02562028
4	0	0
4	2.64	0.02571515
4	0	0
4	0	0
4	0	0
4	9.89	0.08973303
4	13.65	0.11945361
4	0	0
5	11.47	0.10249477
5	8.35	0.07690019
5	0	0
5	0	0
5	0	0
5	2.63	0.02562028
5	0	0
5	0	0
5	8.36	0.07698481
5	0	0
5	2.63	0.02562028
5	0	0
5	2.63	0.02562028
5	14.81	0.12818066
5	0	0

Continued...

Chr.	Gen. dis.	Recom. frac.
5	0	0
5	0	0
5	0	0
5	0	0
5	2.65	0.02580999
5	2.65	0.02580999
5	5.41	0.05127596
5	8.35	0.07690019
5	0	0
5	2.63	0.02562028
5	0	0
5	8.37	0.0770694
5	8.37	0.0770694
6	0	0
6	5.4	0.0511862
6	17.17	0.14532279
6	3.65	0.03519958
6	0	0
6	2.63	0.02562028
6	5.42	0.05136569
6	0	0
6	2.64	0.02571515
6	8.35	0.07690019
6	2.63	0.02562028
6	2.63	0.02562028
6	0	0
6	0	0
6	8.29	0.07639217
6	8.28	0.07630744
6	2.63	0.02562028
6	5.44	0.05154511
6	2.64	0.02571515
7	5.4	0.0511862
7	0	0
7	5.41	0.05127596
7	0	0
7	0	0
7	0	0

Chr.	Gen. dis.	Recom. frac.
7	0	0
7	0	0
7	0	0
7	0	0
7	0	0
7	0	0
7	0	0
7	2.63	0.02562028
7	2.64	0.02571515
7	0	0
7	0	0
7	2.63	0.02562028
7	0	0
7	0	0
7	13.17	0.11578277
7	13.18	0.1158596
7	18.39	0.15387219
7	5.28	0.05010776
7	8.2	0.07562899
8	4.24	0.040652
8	4.92	0.04685684
8	29.12	0.22072189
8	2.63	0.02562028
8	0	0
8	0	0
8	0	0
8	0	0
8	2.64	0.02571515
8	2.63	0.02562028
8	2.63	0.02562028
8	0	0
8	26.53	0.20587404
8	11.57	0.10328898
8	0	0
9	0	0
9	0	0
9	0	0
9	5.41	0.05127596

Continued...

Chr.	Gen. dis.	Recom. frac.
9	0	0
9	2.64	0.02571515
9	2.63	0.02562028
9	0	0
9	0	0
9	0	0
9	18.39	0.15387219
9	0	0
9	10.32	0.0932462
9	6.62	0.06200474
9	0	0
9	2.62	0.02552539
10	5.4	0.0511862
10	2.64	0.02571515
10	2.64	0.02571515
10	0	0
10	2.63	0.02562028
10	30.66	0.22919254
10	2.94	0.02855234
10	11.48	0.10257426
10	5.41	0.05127596
10	5.41	0.05127596
10	54.94	0.33336462
10	0	0
11	0	0
11	5.41	0.05127596
11	0	0
11	22.24	0.17952377
11	2.63	0.02562028
11	14.81	0.12818066
11	5.41	0.05127596
11	2.64	0.02571515
11	0	0
11	8.35	0.07690019
11	8.35	0.07690019
11	18.39	0.15387219
11	2.62	0.02552539
12	18.38	0.15380296

Chr.	Gen. dis.	Recom. frac.
12	11.48	0.10257426
12	0	0
12	0	0
12	0	0
12	0	0
12	14.81	0.12818066
12	0	0
12	2.63	0.02562028
12	0	0
12	0	0
12	0	0
12	2.63	0.02562028
12	0	0
12	2.63	0.02562028
12	22.24	0.17952377
12	0	0
12	2.62	0.02552539
13	8.84	0.08102632
13	0	0
13	0	0
13	0	0
13	0	0
13	0	0
13	0	0
13	8.35	0.07690019
13	5.42	0.05136569
13	5.42	0.05136569
13	0	0
13	0	0
13	0	0
13	0	0
13	8.36	0.07698481
13	2.65	0.02580999
13	0	0
13	0	0
13	0	0
13	0	0
13	14.81	0.12818066

Continued...

Chr.	Gen. dis.	Recom. frac.
13	0	0
13	5.41	0.05127596
13	0	0
13	2.63	0.02562028
13	0	0
13	2.64	0.02571515
13	11.49	0.10265374
13	5.41	0.05127596
13	0	0
13	0	0
13	0	0
13	0	0
13	0	0
13	2.64	0.02571515
13	0	0
13	5.4	0.0511862
13	0	0
14	0	0
14	0	0
14	0	0
14	0	0
14	2.63	0.02562028
14	0	0
14	8.36	0.07698481
14	5.46	0.05172446
14	8.41	0.07740761
14	0	0
14	2.63	0.02562028
14	2.65	0.02580999
14	2.65	0.02580999
14	0	0
14	0	0
14	8.31	0.07656158
14	5.37	0.05091683
14	0	0
14	0	0
14	0	0
14	0	0

Chr.	Gen. dis.	Recom. frac.
14	0	0
14	2.64	0.02571515
14	2.63	0.02562028
14	0	0
14	0	0
14	0	0
14	0	0
14	0	0
14	0	0
14	0	0
14	0	0
14	0	0
14	0	0
14	0	0
14	0	0
14	0	0
14	0	0
14	0	0
14	8.35	0.07690019
14	2.62	0.02552539
15	9.29	0.08478017
15	21.01	0.17154229
15	5.41	0.05127596
15	11.48	0.10257426
15	0	0
15	0	0
15	2.64	0.02571515
15	0	0
15	2.63	0.02562028
15	2.64	0.02571515
15	8.34	0.07681557
16	2.62	0.02552539
16	0	0
16	0	0
16	5.41	0.05127596
16	0	0
16	0	0

Continued...

Chr.	Gen. dis.	Recom. frac.
16	0	0
16	5.42	0.05136569
16	5.42	0.05136569
16	8.37	0.0770694
16	2.64	0.02571515
16	0	0
16	2.62	0.02552539
17	2.44	0.02381421
17	5.23	0.04965764
17	0	0
17	2.63	0.02562028
17	0	0
17	0	0
17	0	0
17	0	0
17	0	0
17	0	0
17	2.63	0.02562028
17	0	0
17	11.48	0.10257426
17	18.39	0.15387219
17	11.5	0.1027332
17	5.43	0.05145541
18	11.47	0.10249477
18	14.81	0.12818066
18	2.63	0.02562028
18	3.04	0.02949429
18	21.85	0.17701428
18	4.06	0.03899536
18	4.04	0.03881093
19	0	0
19	5.41	0.05127596
19	0	0
19	0	0
19	0	0
19	0	0
19	2.63	0.02562028
19	0	0

Chr.	Gen. dis.	Recom. frac.
19	0	0
19	0	0
19	0	0
19	0	0
19	11.48	0.10257426
19	2.63	0.02562028
19	0	0
19	0	0
19	0	0
19	2.64	0.02571515
19	0	0
19	2.63	0.02562028
19	0	0
19	0	0
19	0	0
19	0	0
19	10.11	0.09153424
19	10.11	0.09153424
19	0	0
19	0	0
X	8.58	0.07884198
X	0	0
X	0	0
X	0	0
X	0	0
X	0	0
X	0	0
X	0	0
X	0	0
X	0	0
X	0	0
X	0	0
X	0	0
X	0	0
X	0	0
X	5.41	0.05127596
X	2.64	0.02571515

Continued...

Chr.	Gen. dis.	Recom. frac.
X	0	0
X	2.63	0.02562028
X	5.41	0.05127596

Chr.	Gen. dis.	Recom. frac.
X	8.35	0.07690019
X	11.48	0.10257426
X	2.77	0.02694669

NB: “Chr.” is the chromosome number; “Gen. dis.” is the estimated genetic distance; “Recom. frac.” is the recombination fraction.

Table 7: Estimated recombination fraction of Cross [PERC X DDK)F_1 X B6]N_2

Chr.	Gen. dis.	Recom. frac.
1	9.11	0.08328268
1	5.89	0.0555631
1	0	0
1	18.76	0.15642408
1	6.48	0.06077663
1	66.63	0.36810474
1	32.19	0.23735374
2	7.45	0.06921544
2	7.46	0.06930159
2	4.46	0.04266869
2	73.44	0.38489922
2	0	0
2	0	0
2	10.03	0.09088017
2	0	0
3	11.04	0.09906148
3	0	0
3	35.68	0.25506127
3	2.85	0.02770297
3	16.82	0.14283134
4	16.35	0.13945813
4	3.58	0.03454841
4	0	0
4	5.89	0.0555631
4	20.56	0.16857283
4	15.89	0.13612584
4	7.29	0.06783472
4	16.04	0.13721582
4	9.74	0.08850038
5	7.7	0.07136399
5	2.61	0.02543049
5	2.48	0.024195
5	5.9	0.05565197
5	0	0
5	0	0
5	0	0
5	0	0
5	2.86	0.02779742
5	20.27	0.16664496
5	21.76	0.17643239
6	20.98	0.17134515
6	14.54	0.1261674
6	14.54	0.1261674
6	2.44	0.02381421
6	2.39	0.02333778
6	16.2	0.13837488
6	0	0
6	0	0
6	3.56	0.03436219
7	24.79	0.19545828
7	9.25	0.08444786
7	0	0
7	0	0
7	2.45	0.02390944
7	2.46	0.02400464
7	6.27	0.05892801
8	13.4	0.11754611
8	0	0
8	0	0
8	3.85	0.03705507
8	3.86	0.03714765
8	3.84	0.03696247
8	26.52	0.20581521
9	0	0
9	2.86	0.02779742
9	0	0
9	46.57	0.30299914
9	0	0
9	0	0
10	0	0
10	9.86	0.08948679
10	4.07	0.03908756

Continued...

Chr.	Gen. dis.	Recom. frac.
10	60.28	0.35024388
10	16.25	0.13873632
12	47.22	0.30554358
12	4.15	0.03982443
13	0	0
13	0	0
13	2.86	0.02779742
13	0	0
13	2.86	0.02779742
13	2.86	0.02779742
13	0	0
13	2.86	0.02779742
13	0	0
13	0	0
13	0	0
13	0	0
13	0	0
13	12.56	0.11106661
13	0	0
13	0	0
13	0	0
13	0	0
13	12.67	0.11192132
13	2.9	0.02817503
13	20.43	0.16770999
13	6.89	0.06436354
14	0	0
14	0	0
14	6.82	0.06375322
14	16.32	0.13924174
14	5.08	0.04830458
14	11.86	0.10558325
14	3.67	0.03538547
14	2.92	0.02836372
14	9.15	0.08361592
14	6.08	0.05724875
15	6.25	0.05875155
15	5.89	0.0555631
15	2.88	0.02798626
15	12.58	0.11122215
15	0	0
15	13.4	0.11754611
16	0	0
16	9.39	0.08560978
17	3.58	0.03454841
17	5.9	0.05565197
17	0	0
17	9.19	0.0839489
17	5.95	0.0560961
18	6.45	0.06051302
18	2.86	0.02779742
18	5.89	0.0555631
18	0	0
19	6.73	0.06296727
19	23.83	0.18955458
19	6.06	0.05707161
X	29.37	0.22211479
X	24.59	0.19423767
X	19.26	0.15984272

NB: “Chr.” is the chromosome number; “Gen. dis.” is the estimated genetic distance; “Recom. frac.” is the recombination fraction.

Table 8: Estimated Recombination Fraction of Cross [(B6-Pgk1^a X DDK)F_1XB6]N_2 (More skewed)

Chr.	Gen. dis.	Recom. frac.	Chr.	Gen. dis.	Recom. frac.
1	15.84	0.13576178	7	17.57	0.14814889
1	19.35	0.16045445	7	6.58	0.0616542
1	6.59	0.06174186	7	36.56	0.25933448
1	15.84	0.13576178	7	9.47	0.08627227
1	23.13	0.18517778	8	36.53	0.25919003
1	1.25	0.01234504	8	29.39	0.22222593
1	27.21	0.20984708	8	23.13	0.18517778
2	14.17	0.12339078	8	19.35	0.16045445
2	15.85	0.13583462	9	12.56	0.11106661
2	14.17	0.12339078	9	6.58	0.0616542
2	25.13	0.19752214	9	9.49	0.08643773
2	21.2	0.17278806	9	31.67	0.23460797
2	27.21	0.20984708	9	8.01	0.07401331
3	23.12	0.18511481	10	34.04	0.24689407
3	17.57	0.14814889	10	12.56	0.11106661
3	31.66	0.23455488	10	23.13	0.18517778
3	41.96	0.28397198	10	25.14	0.19758263
3	21.2	0.17278806	10	8	0.07392811
4	27.21	0.20984708	11	19.35	0.16045445
4	8.01	0.07401331	11	12.56	0.11106661
4	17.57	0.14814889	11	15.85	0.13583462
4	12.57	0.11114439	11	6.59	0.06174186
4	19.36	0.16052235	11	19.35	0.16045445
4	29.38	0.22217037	11	23.12	0.18511481
5	44.88	0.2962267	12	21.19	0.17272261
5	15.84	0.13576178	12	27.21	0.20984708
5	27.21	0.20984708	12	36.54	0.25923819
5	34.04	0.24689407	13	17.56	0.14807851
5	8.01	0.07401331	13	19.36	0.16052235
6	25.12	0.19746163	13	15.84	0.13576178
6	29.38	0.22217037	13	11	0.0987406
6	29.4	0.22228148	14	36.54	0.25923819
6	12.55	0.11098881	14	15.84	0.13576178
7	25.16	0.19770357	14	11	0.0987406
7	6.61	0.06191713	14	27.21	0.20984708

Continued...

Chr.	Gen. dis.	Recom. frac.
15	48.02	0.30863012
15	11	0.0987406
15	17.57	0.14814889
15	27.22	0.2099051
15	0	0
16	44.89	0.29626745
16	14.18	0.12346609
16	11	0.0987406
16	8.01	0.07401331
17	8.01	0.07401331
17	19.36	0.16052235
17	14.17	0.12339078
17	12.56	0.11106661

Chr.	Gen. dis.	Recom. frac.
18	14.17	0.12339078
18	23.13	0.18517778
18	12.57	0.11114439
18	6.58	0.0616542
19	17.55	0.14800812
19	6.57	0.06156652
19	19.35	0.16045445
20	29.38	0.22217037
20	15.85	0.13583462
20	15.84	0.13576178
20	8.02	0.0740985
20	19.35	0.16045445

NB: “Chr.” is the chromosome number; “Gen dis.” is the estimated genetic distance; “Recom. frac.” is the recombination fraction.

Table 9: Estimated Recombination Fraction of Cross [(B6-Pgk1^a X DDK)F_1 X B6]N_2 (Less skewed)

Chr.	Gen. dis.	Recom. frac.
1	11.14	0.09986256
1	42.8	0.28757094
1	9.62	0.08751159
1	16.07	0.13743343
1	29.88	0.22493482
1	8.12	0.07494945
1	25.53	0.19993231
2	21.53	0.17494055
2	19.65	0.16248563
2	5.27	0.05001777
2	27.68	0.21256174
2	9.61	0.08742909
2	37.2	0.26239536
3	16.07	0.13743343
3	6.68	0.06253001
3	16.07	0.13743343
3	37.22	0.26249039
3	29.89	0.22498982
4	25.53	0.19993231
4	17.83	0.14997377
4	9.61	0.08742909
4	16.07	0.13743343
4	23.5	0.18749887
4	21.53	0.17494055
5	39.92	0.27497577
5	29.89	0.22498982
5	12.73	0.11238674
5	60.21	0.35003408
5	14.37	0.12489421
6	27.66	0.21244674
6	23.5	0.18749887
6	29.9	0.22504482
6	9.6	0.08734657
7	23.49	0.18743636
7	6.68	0.06253001

Chr.	Gen. dis.	Recom. frac.
7	27.67	0.21250424
7	3.89	0.03742528
7	45.84	0.30010183
7	14.37	0.12489421
8	39.92	0.27497577
8	23.49	0.18743636
8	17.83	0.14997377
8	16.06	0.13736091
9	14.38	0.12496922
9	8.12	0.07494945
9	16.08	0.13750594
9	14.38	0.12496922
9	14.38	0.12496922
10	11.14	0.09986256
10	16.07	0.13743343
10	19.65	0.16248563
10	27.67	0.21250424
10	17.83	0.14997377
11	11.15	0.09994258
11	29.89	0.22498982
11	17.84	0.15004377
11	9.61	0.08742909
11	25.55	0.20005231
11	23.49	0.18743636
12	45.81	0.29998185
12	25.53	0.19993231
12	27.67	0.21250424
13	34.65	0.2499632
13	34.66	0.2500132
13	5.26	0.04992776
13	25.54	0.19999231
14	25.53	0.19993231
14	19.66	0.16255312
14	9.61	0.08742909
14	11.15	0.09994258

Continued...

Chr.	Gen. dis.	Recom. frac.
15	25.53	0.19993231
15	6.68	0.06253001
15	11.16	0.10002258
15	12.74	0.11246425
15	0	0
16	39.92	0.27497577
16	27.67	0.21250424
16	5.27	0.05001777
16	5.25	0.04983774
17	14.37	0.12489421
17	19.66	0.16255312
17	16.08	0.13750594
17	16.07	0.13743343

Chr.	Gen. dis.	Recom. frac.
18	14.37	0.12489421
18	27.67	0.21250424
18	21.54	0.17500555
18	6.67	0.06244251
19	16.07	0.13743343
19	19.65	0.16248563
19	21.53	0.17494055
20	49.04	0.31249451
20	12.74	0.11246425
20	12.75	0.11254175
20	9.61	0.08742909
20	52.49	0.32499613

NB: “Chr.” is the chromosome number; “Gen dis.” is the estimated genetic distance; “Recom. frac.” is the recombination fraction.

Table 10: Estimated Recombination Fraction of Cross [(SM X NZB)F_1 X NZB]N_2

Chr.	Gen. dis.	Recom. frac.	Chr.	Gen. dis.	Recom. frac.
1	41.71	0.28288914	8	27.28	0.21025301
1	46.25	0.30173429	9	33.7	0.24516708
1	15.31	0.13188032	9	17.93	0.15067312
1	46.26	0.30177394	9	21.33	0.17363771
2	41.72	0.28293256	10	57.09	0.34037807
2	56.84	0.33957796	10	26.8	0.20745804
2	51.28	0.32070937	11	9.1	0.08319933
3	1.91	0.01873979	11	23.68	0.18862185
3	46.29	0.30189284	11	23.67	0.18855957
4	10.44	0.09422123	12	30.13	0.22630671
4	15.34	0.13210113	12	23.67	0.18855957
4	26.82	0.20757504	12	6.12	0.05760281
4	15.33	0.13202754	13	18.07	0.15164987
5	47.78	0.30770934	13	23.81	0.18943038
5	8.17	0.07537429	13	15.33	0.13202754
5	12.84	0.11323855	14	4	0.03844183
5	0	0	14	19.29	0.16004675
5	0	0	14	10.48	0.09454573
5	3.92	0.03770274	15	58.91	0.34608383
5	0	0	15	29.77	0.22432901
5	1.93	0.01893226	16	37.36	0.26315448
5	3.89	0.03742528	16	37.05	0.26168148
5	17.63	0.14857086	17	7.69	0.07127825
6	26.8	0.20745804	17	34.63	0.24986317
6	23.68	0.18862185	18	18.36	0.15366445
6	23.67	0.18855957	18	15.67	0.13452126
7	71.33	0.37993802	19	59.48	0.34782851
7	30.12	0.22625197	19	10.9	0.09793728
7	30.12	0.22625197	20	34.64	0.24991319
8	79.98	0.39901135			

NB: "Chr." is the chromosome number; "Gen dis." is the estimated genetic distance; "Recom. frac." is the recombination fraction.

Table 11: Inter Crossover Distance of Cross [(PERA X DDK)F_1 X B6]N_2

Chr.	ID	Inter Cross. Dis.
Chr.-1	1	(27.69 , 49.87]
	2	(74.2 , 85.02]
	5	(54.64 , 102.08]
	10	(35.61 , 74.2]
	28	(24.44 , 39.17]
	28	(0 , 13.95]
	28	(0 , 13.34]
	33	(66.12 , 107.48]
	34	(16.25 , 62.91]
	39	(27.57 , 35.61]
Chr.-2	1	(56.86 , 76.38]
	10	(42.55 , 53.54]
	14	(25.67 , 41.86]
	14	(45.19 , 64.33]
	15	(38.27 , 46.33]
	15	(42.84 , 50.87]
	26	(64.88 , 79.01]
	27	(48.1 , 58.92]
	28	(51.59 , 65.73]
	32	(30.64 , 37.28]
	33	(35.49 , 43.53]
	38	(48.11 , 58.94]
Chr.-3	9	(10.53 , 34.33]
	17	(7.9 , 28.92]
	27	(38.65 , 65.39]
	28	(8.35 , 41.54]
	29	(41.73 , 62.75]
	38	(36.02 , 47]
	39	(23.84 , 57.04]
Chr.-4	25	(34.59 , 42.75]
Chr.-5	7	(50.13 , 66.86]
	15	(69.47 , 89.31]
	16	(33.73 , 41.77]
	25	(21.69 , 44.87]
	29	(50.12 , 69.94]
Chr.-6	5	(16.25 , 29.96]
	5	(0 , 16.57]
	7	(13.61 , 24.54]
	7	(0 , 16.57]
	18	(47.15 , 69.76]
	21	(25.32 , 45.12]
	23	(0 , 16.57]
	32	(14.34 , 39.86]
Chr.-7	1	(26.35 , 47.37]
	27	(21.07 , 39.66]
	34	(0 , 13.48]
	37	(58.05 , 68.73]
	37	(0 , 13.48]
	38	(39.66 , 63.45]
Chr.-8	7	(66.18 , 86.91]
	28	(7.9 , 37.06]
	38	(10.53 , 66.18]
Chr.-10	1	(63.81 , 124.15]
	16	(25.24 , 110.84]
	18	(25.24 , 110.84]
	19	(10.82 , 77.24]
	25	(61.17 , 118.75]
	26	(58.53 , 116.11]
	28	(50.49 , 58.53]
	29	(25.24 , 110.84]
	32	(41.51 , 58.39]
	35	(14.42 , 50.49]
Chr.-11	3	(60.58 , 85.44]
	11	(25.49 , 56.08]
Chr.-12	27	(28.92 , 49.93]
	35	(34.18 , 74.8]
	39	(34.18 , 74.8]
Chr.-13	1	(0 , 8.84]
	1	(0 , 0.002]
	1	(55.69 , 67.18]
	23	(61.46 , 72.28]
	34	(0 , 8.84]

Continued...

Chr.	ID	Inter Cross. Dis.
Chr.-13	34	(0 , 0.002]
	34	(0 , 8.35]
	34	(66.88 , 80.63]
	37	(0 , 8.84]
	37	(0 , 0.002]
	37	(67.18 , 72.59]
	38	(45.04 , 56.04]
Chr.-14	3	(0 , 13.68]
	38	(49.11 , 60.09]
Chr.-15	37	(24.8 , 54.15]
Chr.-17	23	(0 , 7.67]
	34	(41.37 , 49.43]
	37	(32.5 , 46.63]
	37	(0 , 16.93]
Chr.-18	8	(5.67 , 42.33]
Chr.-X	6	(16.09 , 33.02]
	39	(0 , 0.002]

NB: "Chr." is the chromosome number; "ID" is the individual ID; "Inter Cross. Dis." is the intercrossover distance.

Table 12: Inter Crossover Distance of Cross $\texttt{[PERC X DDK)F}_1\texttt{ X B6]N}_2$

Chr.	ID	Inter Cross. Dis.
chr.-1	2	(97.76 , 139.06]
	9	(66.63 , 124.06]
	11	(6.48 , 124.06]
	18	(0 , 105.3]
	24	(5.89 , 33.76]
	24	(0 , 25.24]
	27	(0 , 91.87]
	33	(25.24 , 97.76]
chr.-2	21	(4.46 , 92.81]
	30	(85.36 , 102.84]
chr.-4	9	(9.47 , 46.38]
	11	(65.67 , 78.99]
	14	(43.74 , 65.67]
	36	(39.22 , 69.52]
chr.-5	10	(0 , 5.09]
	12	(13.85 , 41.82]
chr.-6	21	(0 , 4.83]
	21	(0 , 18.59]
chr.-7	26	(0 , 34.04]
	26	(4.91 , 20.43]
	36	(9.25 , 36.49]
	36	(0 , 4.91]
chr.-10	12	(4.07 , 74.21]
	31	(64.35 , 90.46]
chr.-13	13	(36 , 55.45]
	14	(30.99 , 54.28]
chr.-14	33	(33.26 , 43.75]
chr.-15	10	(0 , 25.98]
chn-X	8	(0 , 53.96]
	10	(0 , 53.96]
	10	(0 , 43.85]
	23	(24.59 , 73.22]
	31	(24.59 , 73.22]

NB: "Chr." is the chromosome number; "ID" is the individual ID; "Inter Cross. Dis." is the intercrossover distance.

Table 13: Inter Crossover Distance of Cross [(B6-Pgk1^a X DDK)F_1XB6]N_2(More skewed)

Chr.	ID	Inter Cross. Dis.
chn.-1	4	(41.78 , 80.75]
	4	(1.25 , 51.59]
	16	(66.16 , 109.21]
	18	(41.78 , 80.75]
	25	(24.38 , 67.43]
	27	(23.13 , 40.22]
	31	(46.81 , 93.37]
	36	(22.43 , 64.91]
	42	(24.38 , 67.43]
	48	(1.25 , 51.59]
	51	(0 , 38.97]
	52	(19.35 , 41.78]
	52	(40.22 , 74.02]
	54	(66.16 , 109.21]
	60	(24.38 , 67.43]
	62	(22.43 , 64.91]
	63	(40.22 , 74.02]
	68	(22.43 , 64.91]
	72	(46.81 , 93.37]
	73	(66.16 , 109.21]
chn.-2	11	(39.3 , 76.35]
	14	(46.33 , 87.71]
	20	(39.3 , 76.35]
	27	(21.2 , 73.54]
	28	(21.2 , 73.54]
	29	(60.5 , 103.56]
	39	(76.35 , 117.73]
	54	(30.02 , 69.32]
	54	(0 , 46.33]
	55	(30.02 , 69.32]
	61	(14.17 , 55.15]
	61	(21.2 , 73.54]
	62	(30.02 , 69.32]
	64	(55.15 , 90.52]
	66	(39.3 , 76.35]
	78	(55.15 , 90.52]
	80	(0 , 48.41]
chr..-3	2	(41.96 , 94.82]
	3	(91.19 , 135.51]
	4	(73.62 , 112.39]
	27	(41.96 , 94.82]
	36	(0 , 73.62]
	41	(41.96 , 94.82]
	45	(49.23 , 114.31]
	48	(91.19 , 135.51]
	50	(91.19 , 135.51]
	52	(0 , 73.62]
	53	(41.96 , 94.82]
	61	(31.66 , 91.19]
	63	(17.57 , 72.35]
	69	(31.66 , 91.19]
	71	(49.23 , 114.31]
	72	(31.66 , 91.19]
	73	(31.66 , 91.19]
	75	(91.19 , 135.51]
	77	(31.66 , 91.19]
	78	(31.66 , 91.19]
	80	(31.66 , 91.19]
chr.-4	3	(38.15 , 84.72]
	6	(8.01 , 52.79]
	6	(12.57 , 49.5]
	7	(31.93 , 78.88]
	10	(17.57 , 38.15]
	10	(19.36 , 61.31]
	22	(49.5 , 86.89]
	26	(49.5 , 86.89]
	53	(38.15 , 84.72]
	56	(38.15 , 84.72]
	58	(57.51 , 114.1]
	60	(38.15 , 84.72]

Continued...

Chr.	ID	Inter Cross. Dis.
	61	(57.51 , 114.1]
	71	(31.93 , 78.88]
	78	(25.58 , 65.36]
chr.-5	1	(27.21 , 77.09]
	3	(43.05 , 121.97]
	6	(0 , 60.72]
	6	(27.21 , 77.09]
	7	(77.09 , 129.98]
	9	(0 , 43.05]
	10	(15.84 , 87.93]
	19	(43.05 , 121.97]
	21	(15.84 , 87.93]
	24	(27.21 , 77.09]
	26	(61.25 , 85.1]
	34	(0 , 61.25]
	39	(15.84 , 87.93]
	42	(27.21 , 77.09]
	44	(34.04 , 69.26]
	45	(0 , 60.72]
	49	(27.21 , 77.09]
	62	(77.09 , 129.98]
	64	(77.09 , 129.98]
	65	(43.05 , 121.97]
	66	(43.05 , 121.97]
	74	(43.05 , 121.97]
chr.-6	6	(0 , 54.5]
	20	(0 , 58.78]
	35	(29.38 , 83.9]
	38	(58.78 , 96.45]
	46	(29.38 , 83.9]
	60	(29.38 , 83.9]
	61	(29.38 , 83.9]
	62	(58.78 , 96.45]
	80	(29.38 , 83.9]
chr.-7	2	(6.61 , 49.34]
	3	(67.32 , 101.95]
	4	(24.15 , 67.32]
	6	(24.18 , 55.92]

Chr.	ID	Inter Cross. Dis.
	6	(36.56 , 52.61]
	17	(6.58 , 60.71]
	19	(30.76 , 92.48]
	23	(24.15 , 67.32]
	28	(30.76 , 92.48]
	54	(30.76 , 92.48]
	65	(24.15 , 67.32]
	71	(30.76 , 92.48]
chr.-8	8	(52.52 , 108.4]
	21	(29.39 , 89.05]
	32	(52.52 , 108.4]
	34	(29.39 , 89.05]
	39	(52.52 , 108.4]
	48	(29.39 , 89.05]
	51	(29.39 , 89.05]
	55	(29.39 , 89.05]
	60	(0 , 65.92]
	67	(23.13 , 71.87]
	76	(29.39 , 89.05]
chr.-9	28	(41.16 , 55.75]
	42	(41.16 , 55.75]
	49	(16.07 , 60.3]
	65	(16.07 , 60.3]
chr.-10	6	(35.69 , 94.87]
	10	(35.69 , 94.87]
	27	(25.14 , 56.27]
	31	(35.69 , 94.87]
	45	(35.69 , 94.87]
	51	(23.13 , 60.83]
	55	(60.83 , 102.87]
	56	(60.83 , 102.87]
	59	(23.13 , 60.83]
	60	(12.56 , 69.73]
chr.-11	2	(35 , 73.7]
	11	(35 , 73.7]
	15	(25.94 , 64.91]
	18	(25.94 , 64.91]
	19	(6.59 , 41.79]

Continued...

Chr.	ID	Inter Cross. Dis.
	23	(54.35 , 96.82]
	30	(22.44 , 54.35]
	41	(15.85 , 35]
	53	(6.59 , 41.79]
	53	(0 , 42.47]
	70	(28.41 , 54.35]
	78	(22.44 , 54.35]
	80	(54.35 , 96.82]
chr.-12	6	(27.21 , 84.94]
	20	(27.21 , 84.94]
	21	(0 , 48.4]
	33	(27.21 , 84.94]
	49	(27.21 , 84.94]
	50	(0 , 63.75]
	56	(27.21 , 84.94]
chr.-13	60	(19.36 , 52.76]
chr.-14	13	(26.84 , 90.59]
	43	(26.84 , 90.59]
	44	(15.84 , 63.38]
	61	(26.84 , 90.59]
	67	(11 , 54.05]
chr.-15	5	(11 , 76.59]

Chr.	ID	Inter Cross. Dis.
	6	(28.57 , 103.81]
	11	(28.57 , 103.81]
	33	(28.57 , 103.81]
	34	(28.57 , 103.81]
	36	(28.57 , 103.81]
	49	(28.57 , 103.81]
	61	(28.57 , 103.81]
	63	(28.57 , 103.81]
chr.-16	28	(11 , 33.19]
	39	(11 , 33.19]
	40	(14.18 , 70.07]
	44	(25.18 , 78.08]
chr.-17	30	(33.53 , 54.1]
chr.-18	74	(12.57 , 42.28]
chr.-19	19	(0 , 24.12]
chr.-20	19	(39.71 , 88.44]
	26	(39.71 , 88.44]
	56	(31.69 , 69.09]
	60	(8.02 , 43.21]
	65	(31.69 , 69.09]
	69	(39.71 , 88.44]
	71	(39.71 , 88.44]

NB: “Chr.” is the chromosome number; “ID” is the individual ID; “Inter Cross. Dis.” is the intercrossover distance.

Table 14: Inter Crossover Distance of Cross [(B6-Pgk1^a X DDK)F_1 X B6]N_2(Less skewed)

Chr.	ID	Inter Cross. Dis.
chr.-1	6	(54.07 , 89.22]
	7	(9.62 , 68.49]
	9	(68.49 , 109.51]
	10	(0 , 45.95]
	13	(55.57 , 106.49]
	14	(106.49 , 143.16]
	17	(45.95 , 63.69]
	27	(63.69 , 132.02]
	31	(55.57 , 106.49]
	33	(63.69 , 132.02]
	34	(0 , 45.95]
	36	(25.69 , 98.37]
	37	(25.69 , 98.37]
	38	(29.88 , 54.07]
	41	(25.69 , 98.37]
	41	(8.12 , 63.53]
	42	(38 , 79.6]
	47	(63.69 , 132.02]
	50	(25.69 , 98.37]
	51	(9.62 , 68.49]
	51	(38 , 79.6]
	54	(16.07 , 55.57]
	54	(8.12 , 63.53]
	56	(25.69 , 98.37]
	68	(54.07 , 89.22]
	74	(0 , 45.95]
chr.-2	1	(62.21 , 120.94]
	6	(5.27 , 52.6]
	15	(37.29 , 79.76]
	17	(42.56 , 99.41]
	18	(9.61 , 74.49]
	20	(9.61 , 74.49]
	25	(42.56 , 99.41]
	31	(42.56 , 99.41]
	32	(0 , 46.81]
	41	(9.61 , 74.49]
	42	(9.61 , 74.49]
	43	(42.56 , 99.41]
	44	(42.56 , 99.41]
	55	(62.21 , 120.94]
	60	(62.21 , 120.94]
	62	(24.92 , 74.13]
	64	(52.6 , 83.74]
	65	(9.61 , 74.49]
	67	(32.95 , 62.21]
	75	(52.6 , 83.74]
chr.-3	16	(0 , 53.29]
	17	(16.07 , 59.97]
	19	(53.29 , 89.86]
	21	(22.75 , 76.04]
	25	(59.97 , 105.93]
	27	(53.29 , 89.86]
	29	(37.22 , 83.18]
	53	(59.97 , 105.93]
	54	(22.75 , 76.04]
	70	(59.97 , 105.93]
	75	(22.75 , 76.04]
chr.-4	6	(43.51 , 92.54]
	12	(0 , 39.57]
	17	(43.51 , 92.54]
	18	(39.57 , 70.71]
	23	(67.01 , 114.07]
	39	(9.61 , 43.51]
	48	(49.18 , 88.54]
	54	(39.57 , 70.71]
	55	(27.44 , 69.04]
	62	(43.51 , 92.54]
	65	(43.51 , 92.54]
	66	(0 , 27.44]
	72	(17.83 , 52.97]

Continued...

Chr.	ID	Inter Cross. Dis.
	73	(49.18 , 88.54]
	76	(25.68 , 67.01]
chr.-5	5	(0 , 72.94]
	8	(102.83 , 157.12]
	10	(42.62 , 142.75]
	11	(42.62 , 142.75]
	12	(42.62 , 142.75]
	16	(0 , 69.81]
	18	(102.83 , 157.12]
	19	(0 , 72.94]
	21	(60.21 , 87.31]
	29	(42.62 , 142.75]
	30	(12.73 , 102.83]
	33	(72.94 , 117.2]
	36	(12.73 , 102.83]
	39	(12.73 , 102.83]
	40	(72.94 , 117.2]
	42	(12.73 , 102.83]
	48	(12.73 , 102.83]
	51	(72.94 , 117.2]
	56	(12.73 , 102.83]
	64	(12.73 , 102.83]
	66	(29.89 , 82.54]
	67	(12.73 , 102.83]
	69	(42.62 , 142.75]
	75	(12.73 , 102.83]
	78	(42.62 , 142.75]
chr.-6	26	(23.5 , 81.06]
	31	(23.5 , 81.06]
	34	(23.5 , 81.06]
	40	(53.4 , 90.66]
	73	(23.5 , 81.06]
	75	(29.9 , 63]
	78	(23.5 , 81.06]
chr.-7	1	(3.89 , 77.4]
	17	(38.24 , 107.57]
	18	(3.89 , 77.4]
	22	(34.35 , 61.73]

Chr.	ID	Inter Cross. Dis.
	23	(31.56 , 84.08]
	28	(84.08 , 121.94]
	32	(49.73 , 91.77]
	37	(31.56 , 84.08]
	60	(34.35 , 61.73]
	66	(3.89 , 77.4]
	68	(38.24 , 107.57]
	70	(38.24 , 107.57]
	71	(49.73 , 91.77]
	72	(38.24 , 107.57]
	74	(77.4 , 98.45]
	78	(38.24 , 107.57]
chr.-8	1	(41.32 , 97.3]
	3	(0 , 33.89]
	13	(23.49 , 81.24]
	20	(0 , 63.41]
	24	(41.32 , 97.3]
	32	(41.32 , 97.3]
	53	(41.32 , 97.3]
	60	(41.32 , 97.3]
	71	(23.49 , 81.24]
chr.-9	53	(24.2 , 52.96]
	56	(8.12 , 38.58]
	68	(30.46 , 52.96]
	69	(38.58 , 67.34]
chr.-10	6	(47.32 , 81.22]
	8	(35.72 , 74.53]
	12	(35.72 , 74.53]
	18	(47.32 , 81.22]
	19	(35.72 , 74.53]
	21	(27.67 , 65.15]
	29	(19.65 , 63.39]
	35	(63.39 , 92.36]
	44	(47.32 , 81.22]
	55	(19.65 , 63.39]
	57	(0 , 27.21]
	64	(35.72 , 74.53]
	78	(19.65 , 63.39]

Continued...

Chr.	ID	Inter Cross. Dis.
chr.-11	17	(53 , 106.38]
	29	(57.34 , 94.04]
	36	(35.16 , 76.49]
	42	(53 , 106.38]
	43	(47.73 , 68.49]
	46	(82.89 , 117.53]
	48	(53 , 106.38]
	54	(9.61 , 53]
	56	(57.34 , 94.04]
	61	(53 , 106.38]
	64	(17.84 , 57.34]
	67	(57.34 , 94.04]
	68	(53 , 106.38]
	74	(82.89 , 117.53]
	78	(53 , 106.38]
	80	(35.16 , 76.49]
chr.-12	2	(25.53 , 99.01]
	7	(25.53 , 99.01]
	31	(25.53 , 99.01]
	54	(25.53 , 99.01]
	55	(25.53 , 99.01]
	59	(25.53 , 99.01]
	63	(25.53 , 99.01]
	65	(0 , 71.34]
chr.-13	7	(39.92 , 100.11]
	13	(34.66 , 74.57]
	29	(39.92 , 100.11]
	30	(39.92 , 100.11]
	40	(39.92 , 100.11]
	44	(5.26 , 65.46]
	47	(39.92 , 100.11]
	60	(5.26 , 65.46]
	68	(5.26 , 65.46]
	77	(39.92 , 100.11]
	79	(0 , 69.31]
chr.-14	10	(29.27 , 65.95]
	20	(29.27 , 65.95]
	66	(29.27 , 65.95]
chr.-16	5	(27.67 , 72.86]
chr.-17	6	(35.74 , 66.18]
	75	(19.66 , 50.11]
chr.-18	15	(21.54 , 55.88]
	21	(27.67 , 63.58]
	47	(49.21 , 70.25]
	63	(0 , 49.21]
chr.-20	6	(35.1 , 136.63]
	12	(35.1 , 136.63]
	15	(9.61 , 74.85]
	19	(35.1 , 136.63]
	20	(12.74 , 74.53]
	21	(35.1 , 136.63]
	23	(12.74 , 74.53]
	27	(35.1 , 136.63]
	28	(35.1 , 136.63]
	31	(35.1 , 136.63]
	32	(22.36 , 87.59]
	35	(35.1 , 136.63]
	42	(22.36 , 87.59]
	45	(12.74 , 74.53]
	56	(22.36 , 87.59]
	66	(25.49 , 84.14]
	72	(22.36 , 87.59]
	76	(35.1 , 136.63]

NB: “Chr.” is the chromosome number; “ID” is the individual ID; “Inter Cross. Dis.” is the intercrossover distance.

Table 15: Inter Crossover Distance of Cross [(SM X NZB)F_1 X NZB]N_2

Chr.	ID	Inter Cross. Dis.
chr.-1	1	(0 , 87.96]
	2	(0 , 87.96]
	9	(61.56 , 149.53]
	20	(0 , 87.96]
	25	(15.31 , 107.82]
	27	(0 , 61.56]
	29	(0 , 87.96]
	32	(61.56 , 149.53]
	34	(46.25 , 103.27]
	34	(0 , 61.57]
	36	(0 , 87.96]
	37	(0 , 61.57]
	38	(0 , 87.96]
	40	(0 , 61.56]
	40	(0 , 61.57]
	41	(0 , 87.96]
	41	(15.31 , 107.82]
	44	(0 , 61.56]
	45	(0 , 87.96]
	49	(0 , 87.96]
	49	(15.31 , 107.82]
chr.-2	5	(56.84 , 149.84]
	6	(56.84 , 149.84]
	19	(0 , 108.12]
	20	(0 , 108.12]
	22	(0 , 98.56]
	25	(0 , 108.12]
	27	(0 , 98.56]
	27	(0 , 108.12]
	31	(0 , 98.56]
	38	(56.84 , 149.84]
	42	(56.84 , 149.84]
	44	(0 , 98.56]
	44	(0 , 108.12]
	53	(56.84 , 149.84]
chr.-4	51	(26.82 , 57.49]
chr.-5	2	(18.69 , 30.75]
	2	(0 , 21.52]
	4	(26.86 , 78.53]
	10	(8.17 , 68.79]
	10	(9.74 , 40.21]
	14	(30.75 , 96.16]
	21	(8.17 , 68.79]
	23	(30.75 , 96.16]
	37	(8.17 , 68.79]
chr.-6	10	(0 , 50.48]
	12	(0 , 47.35]
	19	(23.68 , 74.15]
	47	(23.68 , 74.15]
chr.-7	4	(30.12 , 131.57]
	6	(30.12 , 131.57]
	7	(0 , 101.45]
	7	(0 , 60.24]
	11	(0 , 101.45]
	15	(0 , 60.24]
	28	(0 , 101.45]
	28	(0 , 60.24]
	34	(0 , 101.45]
	35	(0 , 101.45]
	37	(30.12 , 131.57]
	44	(0 , 101.45]
	44	(0 , 60.24]
chr.-8	2	(0 , 107.26]
	11	(0 , 107.26]
	18	(0 , 107.26]
	36	(0 , 107.26]
	42	(0 , 107.26]
chr.-9	7	(0 , 39.26]
	24	(0 , 51.63]
	24	(0 , 39.26]
	35	(0 , 51.63]
	35	(0 , 39.26]

Continued...

Chr.	ID	Inter Cross. Dis.
	37	(0 , 51.63]
chr.-10	36	(0 , 83.89]
	47	(0 , 83.89]
chr.-11	3	(0 , 47.35]
	16	(23.68 , 56.45]
	36	(0 , 47.35]
chr.-12	5	(0 , 53.8]
	20	(0 , 53.8]
	50	(0 , 53.8]
chr.-14	5	(0 , 29.77]

Chr.	ID	Inter Cross. Dis.
chr.-15	1	(0 , 88.68]
	9	(0 , 88.68]
	15	(0 , 88.68]
	36	(0 , 88.68]
	41	(0 , 88.68]
chr.-16	7	(0 , 74.41]
	48	(0 , 74.41]
	52	(0 , 74.41]
chr.-17	49	(0 , 42.32]
chr.-18	19	(0 , 34.03]

NB: “Chr.” is the chromosome number; “ID” is the individual ID; “Inter Cross. Dis.” is the intercrossover distance.

Table 16: Survival Functions of Cross [(PERA X DDK)F_1 X B6]N_2

Dis.	Nonpara. surv.
Chr.-1	
0	1
13.34	0.8
13.95	0.8
16.25	0.8
24.44	0.8
27.57	0.8
27.69	0.8
35.61	0.53333
39.17	0.32
49.87	0.32
54.64	0.32
62.91	0.32
66.12	0.32
74.2	0.26667
85.02	0
102.08	0
107.48	0
Chr.-2	
25.67	1
30.64	1
35.49	1
37.28	0.79265
38.27	0.79265
41.86	0.79265
42.55	0.79265
42.84	0.79265
43.53	0.57524
45.19	0.57524
46.33	0.57524
48.1	0.57524
48.11	0.57524
50.87	0.39654
51.59	0.39654

Dis.	Nonpara. surv.
53.54	0.39611
56.86	0.39611
58.92	0.14381
58.94	0.14381
64.33	0.14381
64.88	0.14381
65.73	0
76.38	0
79.01	0
Chr.-3	
7.9	1
8.35	1
10.53	1
23.84	1
28.92	0.50012
34.33	0.50012
36.02	0.50012
38.65	0.50012
41.54	0.49976
41.73	0.49976
47	0
57.04	0
62.75	0
65.39	0
Chr.-5	
21.69	1
33.73	1
41.77	0.6
44.87	0.6
50.12	0.6
50.13	0.6
66.86	0.3
69.47	0.3

Continued...

Dis.	Nonpara. surv.	Weibull surv.
69.94	0	
89.31	0	
Chr.-6		
13.61	1	
14.34	1	
16.25	1	
16.57	0.3	
24.54	0.3	
25.32	0.3	
29.96	0.125	
39.86	0.125	
45.12	0.125	
47.15	0.125	
69.76	0	
Chr.-7		
0	1	
13.48	0.66667	
21.07	0.66667	
26.35	0.66667	
39.66	0.33346	
47.37	0.3332	
58.05	0.3332	
63.45	0	
68.73	0	
Chr.-8		
7.9	1	
10.53	1	
37.06	0.33333	
66.18	0.33333	
86.91	0	
Chr.-10		
10.82	1	
14.42	1	
25.24	1	

Dis.	Nonpara. surv.
41.51	1
50.49	0.75
58.39	0.5
58.53	0.5
61.17	0.5
63.81	0.5
77.24	0
110.84	0
116.11	0
118.75	0
124.15	0
Chr.-11	
25.49	1
56.08	0.5
60.58	0.5
85.44	0
Chr.-12	
28.92	1
34.18	1
49.93	0
74.8	0
Chr.-13	
0	1
8.35	0.55556
8.84	0.55556
45.04	0.55556
55.69	0.55556
56.04	0.37037
61.46	0.37037
66.88	0.37037
67.18	0.27778
72.28	0
72.59	0
80.63	0

Continued...

Dis.	Nonpara. surv.
Chr.-14	
0	1
13.68	0.5
49.11	0.5
60.09	0

Dis.	Nonpara. surv.
Chr.-17	
0	1
7.67	0.5
16.93	0.5
32.5	0.5
41.37	0.5
46.63	0
49.43	0

NB: “Chr.” is the chromosome number; “Dis.” is the unique upper or lower bound of intercrossover distance; “Nonpara. surv.” is the Product-Limit estimate.

Table 17: Survival Functions of Cross [PERC X DDK)F_1 X B6]N_2

Dis.	Nonpara. surv.
chr-1	
0	1
5.89	1
6.48	1
25.24	0.625
33.76	0.5
66.63	0.5
91.87	0.25
97.76	0.25
105.3	0
124.06	0
139.06	0
chr.-2	
4.46	1
85.36	1
92.81	0
102.84	0
chr.-4	
9.47	1
39.22	1
43.74	1
46.38	0.33333
65.67	0.33333
69.52	0
78.99	0
chr.-5	
0	1
5.09	0.5
13.85	0.5
41.82	0

Dis.	Nonpara. surv.
chr.-6	
0	1
4.83	0
18.59	0
chr.-7	
0	1
4.91	0.66667
9.25	0.66667
20.43	0
34.04	0
36.49	0
chr.-10	
4.07	1
64.35	1
74.21	0
90.46	0
chr.-13	
30.99	1
36	1
54.28	0
55.45	0
chr.-X	
0	1
24.59	1
43.85	0
53.96	0
73.22	0

NB: “Chr.” is the chromosome number; “Dis.” is the unique upper or lower bound of intercrossover distance; “Nonpara. surv.” is the Product-Limit estimate.

Table 18: Survival Functions of Cross [(B6-Pgk1^a X DDK)F_1XB6]N_2 (More skewed)

Dis.	Nonpara. surv.
chr.-1	
0	1
1.25	1
19.35	1
22.43	1
23.13	1
24.38	1
38.97	0.66667
40.22	0.66667
41.78	0.66667
46.81	0.66667
51.59	0.375
64.91	0.375
66.16	0.375
67.43	0
74.02	0
80.75	0
93.37	0
109.21	0
chr.-2	
0	1
14.17	1
21.2	1
30.02	1
39.3	1
46.33	0.57143
48.41	0.57143
55.15	0.57143
60.5	0.57143
69.32	0.1
73.54	0.1
76.35	0.1
87.71	0
90.52	0

Dis.	Nonpara. surv.
103.56	0
117.73	0
chr.-3	
0	1
17.57	1
31.66	1
41.96	1
49.23	1
72.35	0.33333
73.62	0.33333
91.19	0.33333
94.82	0
112.39	0
114.31	0
135.51	0
chr.-4	
8.01	1
12.57	1
17.57	1
19.36	1
25.58	1
31.93	1
38.15	0.8
49.5	0.6669
52.79	0.66643
57.51	0.66643
61.31	0
65.36	0
78.88	0
84.72	0
86.89	0
114.1	0

Continued...

Dis.	Nonpara. surv.
chr.-5	
0	1
15.84	1
27.21	1
34.04	1
43.05	0.83333
60.72	0.30769
61.25	0.30769
69.26	0.30769
77.09	0.30769
85.1	0
87.93	0
121.97	0
129.98	0
chr.-6	
0	1
29.38	1
54.5	0.5
58.78	0.5
83.9	0
96.45	0
chr.-7	
6.58	1
6.61	1
24.15	1
24.18	1
30.76	1
36.56	1
49.34	0.125
52.61	0.125
55.92	0.125
60.71	0.125
67.32	0.125
92.48	0
101.95	0

Dis.	Nonpara. surv.
chr.-8	
0	1
23.13	1
29.39	1
52.52	1
65.92	0
71.87	0
89.05	0
108.4	0
chr.-9	
16.07	1
41.16	1
55.75	0
60.3	0
chr.-10	
12.56	1
23.13	1
25.14	1
35.69	1
56.27	0.4
60.83	0.4
69.73	0
94.87	0
102.87	0
chr.-11	
0	1
6.59	1
15.85	1
22.44	1
25.94	1
28.41	1
35	0.66667
41.79	0.25
42.47	0.25

Continued...

Dis.	Nonpara. surv.
54.35	0.25
64.91	0
73.7	0
96.82	0
chr.-12	
0	1
27.21	1
48.4	0
63.75	0
84.94	0
chr.-14	
11	1
15.84	1
26.84	1
54.05	0
63.38	0
90.59	0

Dis.	Nonpara. surv.
chr.-15	
11	1
28.57	1
76.59	0
103.81	0
chr.-16	
11	1
14.18	1
25.18	1
33.19	0
70.07	0
78.08	0
chr.-20	
8.02	1
31.69	1
39.71	1
43.21	0
69.09	0
88.44	0

NB: “Chr.” is the chromosome number; “Dis.” is the unique upper or lower bound of intercrossover distance; “Nonpara. surv.” is the Product-Limit estimate.

Table 19: Survival Functions of Cross [(B6-Pgk1^a X DDK)F_1 X B6]N_2(Less skewed)

Dis.	Nonpara. surv.
chr.-1	
0	1
8.12	1
9.62	1
16.07	1
25.69	1
29.88	1
38	1
45.95	0.59286
54.07	0.59286
55.57	0.59286
63.53	0.3794
63.69	0.3794
68.49	0.33333
79.6	0.06626
89.22	0.06626
98.37	0.06626
106.49	0.06626
109.51	0
132.02	0
143.16	0
chr.-2	
0	1
5.27	1
9.61	1
24.92	1
32.95	1
37.29	1
42.56	1
46.81	0.625
52.6	0.625
62.21	0.625
74.13	0
74.49	0
79.76	0
83.74	0
99.41	0
120.94	0
chr.-3	
0	1
16.07	1
22.75	1
37.22	1
53.29	0.71429
59.97	0.71429
76.04	0
83.18	0
89.86	0
105.93	0
chr.-4	
0	1
9.61	1
17.83	1
25.68	1
27.44	0.75
39.57	0.75
43.51	0.75
49.18	0.75
52.97	0.33333
67.01	0.33333
69.04	0
70.71	0
88.54	0
92.54	0
114.07	0

Continued...

Dis.	Nonpara. surv.
chr.-5	
0	1
12.73	1
29.89	1
42.62	1
60.21	1
69.81	0.5
72.94	0.5
82.54	0.15385
87.31	0.15385
102.83	0.15385
117.2	0
142.75	0
157.12	0
chr.-6	
23.5	1
29.9	1
53.4	1
63	0
81.06	0
90.66	0
chr.-7	
3.89	1
31.56	1
34.35	1
38.24	1
49.73	1
61.73	0.22222
77.4	0.22222
84.08	0.22222
91.77	0
98.45	0
107.57	0
121.94	0

Dis.	Nonpara. surv.
chr.-8	
0	1
23.49	1
33.89	0.83333
41.32	0.83333
63.41	0
81.24	0
97.3	0
chr.-9	
8.12	1
24.2	1
30.46	1
38.58	0.5
52.96	0
67.34	0
chr.-10	
0	1
19.65	1
27.21	0.88889
27.67	0.88889
35.72	0.88889
47.32	0.88889
63.39	0.25
65.15	0
74.53	0
81.22	0
92.36	0
chr.-11	
9.61	1
17.84	1
35.16	1
47.73	1
53	0.85714
57.34	0.75

Continued...

Dis.	Nonpara. surv.
68.49	0.4
76.49	0.4
82.89	0.4
94.04	0
106.38	0
117.53	0
chr.-12	
0	1
25.53	1
99.01	0
71.34	0
chr.-13	
0	1
5.26	1
34.66	1
39.92	1
65.46	0
69.31	0
74.57	0
100.11	0
chr.-17	
19.66	1

Dis.	Nonpara. surv.
35.74	1
50.11	0
66.18	0
chr.-18	
0	1
21.54	1
27.67	1
49.21	0.5
55.88	0
63.58	0
70.25	0
chr.-20	
9.61	1
12.74	1
22.36	1
25.49	1
35.1	1
74.53	0
74.85	0
84.14	0
87.59	0
136.63	0

NB: “Chr.” is the chromosome number; “Dis.” is the unique upper or lower bound of intercrossover distance; “Nonpara. surv.” is the Product-Limit estimate.

Table 20: Survival Functions of Cross [(SM X NZB)F_1 X NZB]N_2

Dis.	Nonpara. surv.
chr.-1	
0	1
15.31	1
46.25	1
61.56	0.4
61.57	0
87.96	0
103.27	0
107.82	0
149.53	0
chr.-2	
0	1
56.84	1
98.56	0
108.12	0
149.84	0
chr.-5	
0	1
8.17	1
9.74	1
18.69	1
21.52	0.6
26.86	0.6
30.75	0.6
40.21	0
68.79	0

Dis.	Nonpara. surv.
78.53	0
96.16	0
chr.-6	
0	1
23.68	1
47.35	0
50.48	0
74.15	0
chr.-7	
0	1
30.12	1
60.24	0
101.45	0
131.57	0
chr.-9	
0	1
39.26	0
51.63	0
chr.-11	
0	1
23.68	1
47.35	0
56.45	0

NB: "Chr." is the chromosome number; "Dis." is the unique upper or lower bound of intercrossover distance; "Nonpara. surv." is the Product-Limit estimate.

Table 21: Estimated Crossover Rate of Cross [(PERA X DDK)F_1 X B6]N_2

Chr.	Lambda	Var.	Std.
1	0.022809	0.000054	0.007350
2	0.019996	0.000034	0.005788
3	0.029014	0.000126	0.011213
4	0.025956	0.000676	0.026005
5	0.018773	0.000071	0.008437
6	0.044974	0.000267	0.016350
7	0.031433	0.000169	0.012993
8	0.023242	0.000192	0.013852
10	0.016533	0.000030	0.005500

Chr.	Lambda	Var.	Std.
11	0.017940	0.000164	0.012818
12	0.021144	0.000156	0.012491
13	0.035578	0.000106	0.010311
14	0.033005	0.000552	0.023501
15	0.026607	0.000745	0.027288
17	0.042043	0.000452	0.021257
18	0.054837	0.004166	0.064548
20	0.088851	0.004312	0.065662

NB: "Chr." is the chromosome number; "Lambda" is the estimated crossover rate; "Var." is the variance, "Std." is the standard deviation.

Table 22: Estimated Crossover Rate of Cross [PERC X DDK)F_1 X B6]N_2

Chr.	Lambda	Var.	Std.
1	0.019952	0.000059	0.007682
2	0.015056	0.000122	0.011043
4	0.019553	0.000098	0.009891
5	0.078557	0.003705	0.060868
6	1.469909	51.738410	7.192942

Chr.	Lambda	Var.	Std.
7	0.103153	0.003913	0.062551
10	0.018498	0.000185	0.013594
13	0.023096	0.000272	0.016499
14	0.026133	0.000687	0.026215
15	0.364614	19.245570	4.386978
X	0.036242	0.000342	0.018505

NB: "Chr." is the chromosome number; "Lambda" is the estimated crossover rate; "Var." is the variance, "Std." is the standard deviation.

Table 23: Estimated Crossover Rate of Cross [(B6-Pgk1^a X DDK)F_1XB6]N_2 (More skewed)

Chr.	Lambda	Var.	Std.	Chr.	Lambda	Var.	Std.
1	0.020257	0.000022	0.004660	11	0.024544	0.000049	0.007032
2	0.019107	0.000023	0.004763	12	0.023994	0.000096	0.009802
3	0.014468	0.000011	0.003246	13	0.030016	0.000979	0.031289
4	0.018894	0.000025	0.005021	14	0.022930	0.000121	0.010997
5	0.017459	0.000015	0.003899	15	0.017956	0.000041	0.006438
6	0.019732	0.000047	0.006866	16	0.033456	0.000325	0.018040
7	0.021198	0.000041	0.006387	17	0.023257	0.000551	0.023479
8	0.017656	0.000031	0.005562	18	0.040828	0.001882	0.043378
9	0.024294	0.000155	0.012467	19	0.389653	20.729680	4.552986
10	0.018045	0.000035	0.005898	20	0.019385	0.000057	0.007550

NB: "Chr." is the chromosome number; "Lambda" is the estimated crossover rate; "Var." is the variance, "Std." is the standard deviation.

Table 24: Estimated Crossover Rate of Cross [(B6-Pgk1^a X DDK)F_1 X B6]N_2(Less skewed)

Chr.	Lambda	Var.	Std.	Chr.	Lambda	Var.	Std.
1	0.017542	0.000013	0.003570	10	0.020139	0.000033	0.005718
2	0.018246	0.000018	0.004248	11	0.014448	0.000013	0.003673
3	0.017462	0.000029	0.005414	12	0.019851	0.000059	0.007652
4	0.019736	0.000027	0.005234	13	0.019920	0.000040	0.006363
5	0.015068	0.000010	0.003201	14	0.022147	0.000173	0.013141
6	0.020078	0.000063	0.007938	16	0.021425	0.000496	0.022272
7	0.016688	0.000019	0.004331	17	0.024356	0.000310	0.017619
8	0.019756	0.000048	0.006916	18	0.025382	0.000173	0.013145
9	0.026704	0.000187	0.013660	20	0.017266	0.000020	0.004434

NB: "Chr." is the chromosome number; "Lambda" is the estimated crossover rate; "Var." is the variance, "Std." is the standard deviation.

Table 25: Estimated Crossover Rate of Cross [(SM X NZB)F_1 X NZB]N_2

Chr.	Lambda	Var.	Std.
1	0.027563	0.000053	0.007303
2	0.017813	0.000029	0.005417
4	0.024860	0.000649	0.025467
5	0.030122	0.000122	0.011030
6	0.033171	0.000343	0.018529
7	0.030325	0.000125	0.011171
8	0.101530	0.932314	0.965564
9	0.235885	2.079055	1.441893

Chr.	Lambda	Var.	Std.
10	0.126885	2.980091	1.726294
11	0.043946	0.000856	0.029265
12	0.189596	3.097847	1.760070
14	0.322766	16.795430	4.098223
15	0.120657	1.127649	1.061908
16	0.141439	2.239838	1.496609
17	0.235358	11.814740	3.437258
18	0.286289	14.692920	3.833134

NB: “Chr.” is the chromosome number; “Lambda” is the estimated crossover rate; “Var.” is the variance, “Std.” is the standard deviation.

Table 26: Ratio Information of Cross [(PERA X DDK)F_1 X B6]N_2

Marker	Ratio	Overall
chr.-1		
rs13475712	0.88	0.602
rs13475735	0.70	
rs3711079	1.64	
rs4222215	1.27	
rs13475771	0.00	
rs3677683	0.00	
01.029.481	0.00	
01.035.780	1.32	
01.041.550	1.03	
01.046.600	0.00	
mCV23591750	0.00	
rs3716105	0.25	
rs6356603	0.40	
01.076.110	0.86	
01.087.170	1.35	
rs13475982	0.00	
rs13475988	0.00	
rs13475989	0.00	
rs13475991	0.00	
rs6342650	0.00	
rs6358447	0.00	
CEL-1-98681809	0.00	
CEL-1-98799654	0.00	
rs3695980	0.00	
rs13476012	0.00	
rs3685663	5.62	
rs3717264	0.00	
rs3664662	0.00	
rs3695581	0.60	
rs13476089	0.00	
rs6355835	0.33	
01.135.010	0.00	
rs13476147	0.36	
rs6364156	0.00	

Marker	Ratio	Overall
rs13476259	4.45	
01.178.925	0.00	
CEL-1-181947877	0.00	
rs13476290	1.11	
rs3654705	0.00	
rs6246360	0.00	
Chr.-2		
rs13476318	0.65	0.657
rs13476330	1.99	
02.016.175	0.85	
rs6181760	0.64	
02.041.990	0.35	
02.054.160	0.27	
02.065.760	0.92	
rs6371268	0.00	
rs13476560	0.00	
rs13476563	0.00	
mCV25095764	0.38	
CEL-2-79237503	0.00	
rs3722345	2.07	
02.079.300	0.00	
02.083.650	1.23	
rs4223268	0.65	
02.093.700	1.38	
rs13476663	0.56	
rs6249987	0.00	
rs13476667	0.00	
rs3674721	0.00	
mCV25337624	0.00	
rs13476684	0.00	
rs13476689	0.00	
rs3022892	0.00	
rs3693678	0.00	
rs3701250	1.15	

Continued...

Marker	Ratio	Overall
rs6276129	1.81	
rs3723406	1.71	
rs6340352	0.00	
rs3697020	0.00	
rs6411422	0.00	
02.128.325	1.52	
02.130.220	2.12	
rs13476794	0.84	
rs3710324	0.00	
rs13476805	0.00	
rs6360457	1.07	
gnf02.141.261	0.00	
rs6195594	5.18	
rs3655895	0.00	
rs3696870	0.00	
02.146.685	0.00	
rs3726342	0.63	
rs6204920	9.39	
CEL-2-163612103	1.45	
rs6219107	1.38	
rs3673248	1.03	
chr.-3		
rs6398851	0.32	0.589
rs13477046	1.02	
D3Mit63	1.01	
rs6239288	0.74	
rs3696955	0.00	
rs6226544	0.58	
CEL-3-68001820	0.00	
rs13477178	1.26	
rs6198234	0.00	
rs3698109	0.00	
rs13477190	1.21	
rs13477210	0.00	
rs3715352	0.00	
rs13477215	NaN	

Marker	Ratio	Overall
gnf03.079.138	1.39	
rs6376008	0.57	
rs13477244	0.00	
rs3720007	0.32	
rs6391963	0.97	
rs3686473	0.00	
03.100.150	0.00	
03.119.365	0.17	
rs6214597	0.00	
03.152.282	0.57	
rs13477498	0.00	
rs6331755	0.00	
chr.-4		
rs13477534	0.00	0.477
rs13477546	0.00	
04.011.950	0.00	
rs13477592	0.57	
rs13477599	0.00	
04.021.985	1.55	
04.029.760	2.24	
rs13477662	0.47	
04.042.340	1.16	
04.049.500	0.00	
rs13477741	0.00	
rs6258088	0.30	
04.078.330	0.00	
04.099.005	0.13	
rs13477895	0.87	
rs3670382	0.00	
rs3696331	0.40	
04.115.380	0.00	
rs3671259	0.00	
rs13477972	0.00	
04.128.160	2.05	
rs6268364	0.74	
rs3693087	0.00	

Continued...

Marker	Ratio	Overall
chr.-5		
rs13478092	0.28	0.628
CEL-5-14611794	1.04	
rs13478133	1.55	
UT-5-19.849706	0.00	
rs13478138	0.00	
rs3706626	0.00	
rs13478151	2.46	
rs13478157	0.00	
UT-5-30.642219	0.00	
05.035.200	0.97	
rs3716195	0.00	
rs13478210	3.60	
rs13478212	0.00	
rs13478215	3.55	
mCV27558149	0.68	
gnf05.061.650	0.00	
05.067.560	0.00	
05.071.190	0.00	
05.073.500	0.00	
05.085.655	0.00	
05.091.725	0.49	
rs13478428	0.50	
05.108.560	0.46	
CEL-5-120064766	1.02	
05.117.270	0.00	
05.121.625	0.59	
rs13478521	0.00	
05.135.860	0.63	
rs3668534	1.27	
Chr.-6		
06.002.580	0.12	0.553
06.005.375	0.00	
rs3655269	0.55	
06.033.580	0.91	
06.063.270	0.13	

Marker	Ratio	Overall
mhcCD8b4	0.00	
rs13478841	0.38	
06.085.360	0.66	
06.097.530	0.00	
06.102.675	0.51	
rs3655148	2.09	
rs6204829	0.32	
06.118.265	0.75	
rs3695724	0.00	
CEL-6-122563022	0.00	
06.125.555	1.67	
rs3670851	4.81	
rs6339546	0.55	
06.140.060	0.89	
rs6387265	0.46	
Chr.-7		
07.000.385	0.00	0.606
rs13479163	0.34	
mCV25220583	0.00	
CEL-7-29429804	0.47	
rs6313526	0.00	
rs6295036	0.00	
CEL-7-36545579	0.00	
rs13479238	0.00	
07.047.960	0.00	
gnf07.050.858	0.00	
mCV23672419	0.00	
rs3693038	0.00	
rs13479274	0.00	
rs13479276	0.00	
rs6160140	0.36	
rs3705155	1.17	
rs13479317	0.00	
rs3693876	0.00	
rs13479321	3.76	
rs13479334	0.00	

Continued...

Marker	Ratio	Overall
rs13479338	0.00	
07.077.280	1.48	
rs13479427	0.76	
rs3719258	0.53	
D7Mit291	2.81	
rs3663988	2.95	
Chr.-8		
rs13479627	0.43	0.825
08.020.285	2.24	
08.029.780	0.51	
rs13479741	1.82	
rs3726906	0.22	
08.062.280	0.00	
rs3712611	0.00	
rs13479813	0.00	
08.076.440	0.00	
rs3690549	1.36	
rs6296891	0.87	
rs8236770	0.69	
08.086.390	0.00	
D8Mit322	0.79	
rs13480026	2.34	
rs4227456	0.00	
Chr.-9		
09.029.420	0.48	0.576
rs3669224	0.00	
rs13480160	0.00	
09.038.640	0.00	
09.043.100	1.22	
09.052.500	0.00	
09.060.410	0.33	
rs6174757	1.34	
rs13480267	0.00	
rs13480277	0.00	
rs13480285	0.00	

Marker	Ratio	Overall
rs3725272	1.02	
09.094.990	0.00	
09.104.660	1.06	
rs13480436	0.97	
rs6302293	0.00	
gnf09.117.044	2.02	
Chr.-10		
rs13480480	0.27	1.027
rs6192001	1.09	
10.009.900	8.00	
10.011.300	1.89	
rs13480506	0.00	
rs13480525	0.64	
10.038.835	1.40	
D10Mit40	0.39	
rs3165937	0.42	
rs13480703	0.35	
rs3705990	0.68	
rs13480804	2.29	
10.127.600	0.00	
Chr.-11		
rs13480847	0.18	0.876
rs13480869	0.00	
rs13480889	0.84	
rs3678321	0.00	
rs3657760	1.67	
rs13480997	0.25	
mCV23044839	0.67	
UT-11-68.607315	1.71	
rs13481123	0.20	
gnf11.093.966	0.00	
rs3714299	1.45	
rs3710148	2.41	
rs6384437	2.41	
rs13481226	0.44	

Continued...

Marker	Ratio	Overall	Diff.
Chr.-12			
rs13481276	0.27	0.789	
12.014.515	1.34		
rs6187012	1.02		
rs6223000	0.00		
12.032.830	0.00		
gnf12.033.545	0.00		
rs6243157	0.00		
rs3689063	0.83		
rs13481465	0.00		
rs3686891	0.26		
rs3686378	0.00		
rs13481531	0.00		
mCV23169261	0.00		
12.081.010	0.47		
rs3696951	0.00		
rs8259763	0.58		
12.099.140	1.70		
gnf12.101.501	0.00		
CEL-12-101776500	0.69		
Chr.-13			
D13Mit158	1.06	0.822	
13.005.379	13.60		
13.010.063	0.00		
D13Mit172	0.00		
13.010.368	0.00		
13.011.447	0.00		
13.013.030	0.00		
13.013.605	0.00		
rs3721858	2.59		
13.017.553	7.13		
13.020.344	1.95		
13.019.050	0.00		
D13Mit135	0.00		
13.021.844	0.00		
RS6158895	0.00		

Marker	Ratio	Overall	Diff.
13.028.853	1.99		
13.031.107	1.18		
13.034.400	0.00		
13.035.013	0.00		
13.034.725	0.00		
rs6259014	0.00		
gnf13.045.330	1.47		
rs6209128	0.00		
rs3700819	1.19		
mCV22624058	0.00		
CEL-13-60831741	1.04		
13.061.625	0.00		
13.066.450	0.74		
rs6179438	3.59		
13.080.001	0.58		
13.083.500	0.00		
rs13481918	0.00		
13.087.830	0.00		
rs3655061	0.00		
rs4230027	0.00		
rs3705092	0.94		
rs13481992	0.00		
13.114.540	0.70		
rs3657414	0.00		
Chr.-14			
14.002.500	0.29	0.634	
rs6340768	0.00		
rs13482084	0.00		
14.006.480	0.00		
rs3719629	0.00		
14.015.365	0.49		
14.020.160	0.00		
rs3722090	0.87		
14.032.950	1.69		
mCV23384307	1.41		
rs13482179	0.00		

Continued...

Marker	Ratio	Overall
rs6392664	0.72	
14.055.010	0.55	
rs13482214	1.93	
14.058.715	0.00	
rs6179045	0.00	
14.065.150	1.55	
CEL-14-65598536	1.67	
rs13482259	0.00	
rs6298191	0.00	
14.074.265	0.00	
CEL-14-71690454	0.00	
rs3023412	0.00	
rs6325141	1.49	
rs13482301	0.33	
rs3725470	0.00	
gnf14.085.610	0.00	
rs6395984	0.00	
rs6407863	0.00	
rs3706761	0.00	
rs3655019	NaN	
rs6291434	0.00	
rs6176735	0.00	
rs6299927	0.00	
rs4139735	0.00	
CEL-14-85152539	0.00	
rs13482311	0.00	
rs13482312	0.00	
rs13482313	0.00	
rs13482314	0.00	
14.093.815	0.00	
rs3708779	0.65	
rs3683221	1.46	
Chr.-15		
rs13482418	1.91	0.648
15.017.570	0.57	
rs13482490	2.99	

Marker	Ratio	Overall
15.030.400	0.91	
15.053.380	0.52	
rs3701449	0.00	
rs13482618	0.00	
15.066.375	1.33	
15.078.600	0.00	
rs3697744	0.36	
rs3716673	0.48	
rs3690173	0.94	
Chr.-16		
rs4166445	0.74	0.471
rs4168890	0.88	
16.029.450	0.00	
rs4170974	0.00	
16.031.880	3.73	
rs4174174	0.00	
16.035.790	0.00	
rs4175353	0.00	
16.055.570	0.28	
rs4197416	0.67	
16.075.770	0.68	
rs3656592	0.33	
rs4217061	0.00	
rs3164088	0.39	
Chr.-17		
rs3662575	0.81	0.724
17.013.500	0.47	
rs6270865	6.71	
rs3667748	0.00	
rs4231344	0.39	
17.022.870	0.00	
rs3724223	0.00	
UT-17-33.238924	0.00	
rs8242408	0.00	
rs3682923	0.00	

Continued...

Marker	Ratio	Overall
17.041.250	0.00	
mCV25197172	1.26	
17.043.515	0.00	
rs8273969	1.06	
rs3675634	1.04	
rs3687741	0.70	
D17Mit123	0.94	
Chr.-18		
rs13483210	0.33	0.858
rs3656185	1.47	
rs3691362	1.46	
rs3714096	0.43	
18.046.110	0.25	
rs6350869	1.70	
18.070.235	0.36	
rs13483472	0.34	
Chr.-19		
rs13483500	1.07	0.972
rs3671671	0.00	
rs3713033	8.87	
rs13483505	0.00	
rs13483511	0.00	
19.005.500	0.00	
CEL-19-8529644	0.00	
UT-19-10.709331	2.05	
rs6163293	0.00	
rs3700209	0.00	
rs6237846	0.00	
CEL-19-12911424	0.00	
rs3692733	0.00	
rs3669192	3.75	
gnf19.017.711	0.83	
rs3720318	0.00	
rs13483557	0.00	
19.018.140	0.00	
rs13483563	2.32	

Marker	Ratio	Overall
rs6392565	0.00	
rs3672759	10.52	
rs3653630	0.00	
rs13483577	0.00	
rs3090325	0.00	
19.043.320	0.54	
19.048.500	1.95	
rs13483669	0.00	
rs6257938	0.00	
Chr.-X		
DXMit166	0.35	0.415
rs13483834	0.36	
CEL-X-68179178	0.00	
CEL-X-68645226	0.00	
rs13483838	0.00	
CEL-X-71104123	0.00	
CEL-X-72627341	0.00	
CEL-X-73027245	0.00	
gnfX.070.167	0.00	
rs13483858	0.00	
CEL-X-74073918	0.00	
CEL-X-74272691	0.00	
rs13483862	0.00	
rs13483863	0.00	
CEL-X-75125049	0.00	
rs13483877	0.00	
rs13483803	0.00	
gnfX.076.619	0.00	
rs13483888	2.16	
DXMit16	0.66	
CEL-X-91222960	0.00	
CEL-X-94143306	0.90	
rs13483935	0.78	
rs13484004	0.37	
rs13484094	0.49	
DXMit29	0.26	

NB: “Marker” is the name of marker; “Ratio” is the ratio of cM and MB; “Overall” is the overall ratio of corresponding chromosome.

Table 27: Ratio Information of Cross [PERC X DDK)F_1 X B6]N_2

Marker	Ratio	Overall
Chr.-1		
01.029.481	0.54	1.008
01.035.780	1.44	
01.041.550	1.12	
01.046.600	0	
01.070.445	0.79	
01.076.110	1.11	
01.135.010	1.19	
D1Mit270	0.9	
Chr.-2		
02.041.990	0.67	0.980
02.054.160	0.62	
02.065.760	0.64	
02.079.300	0.33	
02.125.700	1.57	
02.128.325	0	
02.130.220	0	
02.151.240	0.77	
02.146.685	0	
Chr.-3		
03.014.785	0.26	0.503
03.047.215	0.33	
03.054.150	0	
03.100.150	0.82	
03.119.365	0.18	
03.152.282	0.51	
Chr.-4		
04.011.950	0.35	0.796
04.021.985	1.62	
04.029.760	0.44	
04.042.340	0	
04.049.500	0.82	
04.078.330	0.7	
04.099.005	0.76	
04.115.380	0.46	
04.128.160	1.19	
04.131.640	2.82	
Chr.-5		
05.010.335	0.34	0.511
05.035.200	0.31	
05.038.350	0.81	
05.046.885	0.29	
05.067.560	0.28	
05.071.190	0	
05.073.500	0	
05.085.655	0	
05.091.725	0	
05.108.560	0.17	
05.117.270	2.3	
05.135.860	1.17	
Chr.-6		
06.002.580	0.12	0.651
06.014.800	1.7	
06.033.580	0.76	
06.063.270	0.53	
06.070.455	0.33	
06.074.995	0.52	
06.097.530	0.71	
06.102.675	0	
06.108.730	0	
06.118.265	0.38	
Chr.-7		
07.000.385	33.5	0.427
07.047.960	0.6	

Continued...

Marker	Ratio	Overall
07.087.220	0.24	
07.091.695	0	
07.096.985	0	
07.110.535	0.18	
07.112.290	1.38	
07.115.675	1.84	
Chr.-8		
08.067.625	0.49	0.923
08.086.390	0.78	
08.090.187	0	
08.092.425	0	
08.096.955	0.84	
08.101.010	0.94	
08.111.015	0.38	
08.124.650	1.93	
Chr.-9		
09.029.420	0.48	0.923
09.043.100	0	
09.052.500	0.3	
09.060.410	0	
09.091.925	1.62	
09.104.660	0	
09.119.240-2	0	
Chr.-10		
10.009.900	0.55	0.571
10.011.300	0	
10.015.980	2.12	
10.038.835	0.18	
D10Mit95	1.17	
10.127.600	0.45	
Chr.-12		
D12Mit112	0.52	0.782
12.089.840	1.07	

Marker	Ratio	Overall
12.099.140	0.44	
Chr.-13		
13.001.770	1.09	0.958
13.005.379	0	
13.010.063	0	
D13Mit172	143	
13.010.368	0	
13.011.447	2.65	
13.013.030	1.77	
13.013.605	0	
13.017.553	0.72	
13.020.344	0	
13.019.050	0	
D13Mit135	0	
13.021.844	0	
RS6158895	0	
13.028.853	2.98	
13.031.107	0	
13.034.400	0	
13.035.013	0	
13.034.725	0	
13.061.625	0.5	
13.080.001	0.18	
13.090.665	2.18	
13.114.540	0.29	
Chr.-14		
14.002.500	0.29	0.633
14.006.480	0	
14.015.365	0	
14.024.450	0.76	
14.032.950	1.9	
14.051.890	0.38	
14.055.010	3.79	
14.065.150	0.36	
14.074.265	0.32	

Continued...

Marker	Ratio	Overall
14.093.815	0.49	
14.101.995	0.72	
Chr.-15		
15.030.400	0.45	0.668
15.044.240	0.44	
15.051.335	0.98	
15.066.375	0.19	
15.075.005	1.44	
15.078.600	0	
15.098.880	0.66	
Chr.-16		
16.029.450	0.7	0.608
16.031.880	0	
16.055.570	0.39	
Chr.-17		
17.013.500	0.64	0.352
17.022.870	0.38	
17.038.280	0.47	
17.041.250	0	

Marker	Ratio	Overall
17.052.240	0.84	
D17Mit39	0.26	
Chr.-18		
18.025.640	0.56	0.423
18.042.400	0.41	
18.051.990	0.29	
18.061.750	0.6	
18.072.500	0	
Chr.-19		
19.003.270	0.65	0.329
19.018.140	0.5	
19.043.320	0.95	
19.048.500	1.17	
Chr.-X		
DXMit166	0.35	0.642
DXMit16	0.63	
DXMit234	0.6	
DXMit29	0.74	

NB: “Marker” is the name of marker; “Ratio” is the ratio of cM and MB; “Overall” is the overall ratio of corresponding chromosome.

Table 28: Ratio Information of Cross [(B6-Pgk1^a X DDK)F_1XB6]N_2 (More skewed)

Marker	Ratio	Overall
chr.-1		
D1Mit66	0.68	0.789
D1Mit318	0.78	
D1Mit251	0.53	
D1Mit132	0.98	
D1Mit390	0.34	
D1Mit424	0.58	
D1Mit270	0.12	
D1Mit293	1.24	
Chr.-2		
D2Mit117	0.58	0.707
D2Mit83	0.70	
D2Mit244	0.59	
D2Mit37	0.75	
D2Mit276	0.58	
D2Mit285	0.62	
D2Mit200	1.00	
Chr.-3		
D3Mit164	0.32	0.711
D3Mit203	1.18	
D3Mit63	1.25	
D3Mit74	0.67	
D3Mit254	0.96	
D3Mit163	0.86	
Chr.-4		
D4Mit227	0.32	0.793
D4Mit286	0.82	
D4Mit164	0.49	
D4Mit58	0.45	
D4Mit37	0.70	
D4Mit339	1.14	
D4Mit256	1.46	

Marker	Ratio	Overall
Chr.-5		
D5Mit344	0.19	1.077
D5Mit79	1.15	
D5Mit15	0.74	
D5Mit314	0.62	
D5Mit31	1.18	
D5Mit143	0.64	
Chr.-6		
D6Mit236	0.17	0.706
D6Mit384	0.68	
D6Mit65	0.64	
D6Mit25	0.96	
D6Mit201	0.86	
Chr.-7		
D7Mit178	0.14	0.865
D7Mit117	0.96	
D7Mit310	0.47	
D7Mit30	0.46	
D7Mit323	0.34	
D7Mit291	1.03	
D7Mit223	1.18	
Chr.-8		
D8Mit157	0.22	0.865
D8Mit191	1.33	
D8Mit348	0.55	
D8Mit166	0.97	
D8Mit322	2.48	
Chr.-9		
D9Mit126	0.26	0.702
D9Mit90	1.34	
D9Mit97	0.36	

Continued...

Marker	Ratio	Overall
D9Mit270	0.37	
D9Mit212	0.97	
D9Mit281	0.77	
Chr.-10		
D10Mit298	0.35	0.771
D10Mit214	2.02	
D10Mit40	0.54	
D10Mit66	0.58	
D10Mit233	0.98	
D10Mit269	0.55	
Chr.-11		
D11Mit71	0.16	1.106
D11Mit151	1.07	
D11Mit20	0.64	
D11Mit5	0.71	
D11Mit66	0.42	
D11Mit67	1.35	
D11Mit168	1.42	
Chr.-12		
D12Mit182	0.18	0.939
D12Mit112	0.58	
D12Mit260	0.64	
D12Mit150	1.40	
Chr.-13		
D13Mit16	0.49	1.005
D13Mit63	0.79	
D13Mit99	0.56	
D13Mit107	0.81	
D13Mit78	0.47	
Chr.-14		
D14Mit10	0.27	0.703
D14Mit54	1.60	

Marker	Ratio	Overall
D14Mit234	0.65	
D14Mit162	0.51	
D14Mit170	1.10	
Chr.-15		
D15Mit12	1.49	0.560
D15Mit100	1.00	
D15Mit234	0.79	
D15Mit189	1.00	
D15Mit16	1.34	
D15Mit79	0.00	
Chr.-16		
D16Mit182	0.61	0.845
D16Mit166	1.76	
D16Mit140	0.36	
D16Mit191	0.65	
D16Mit106	0.77	
Chr.-17		
D17Mit78	0.55	0.841
D17Mit175	0.48	
D17Mit7	0.88	
D17Mit39	0.67	
D17Mit123	0.67	
Chr.-18		
D18Mit67	0.33	0.956
D18Mit60	0.69	
D18Mit123	0.99	
D18Mit47	0.57	
D18Mit144	0.86	
Chr.-19		
D19Mit42	0.49	1.167
D19Mit111	1.47	
D19Mit66	0.34	

Continued...

Marker	Ratio	Overall
D19Mit137	1.09	
Chr.-20		
DXMit124	0.33	0.904
DXMit166	0.74	

Marker	Ratio	Overall
DXMit210	0.76	
DXPas29	0.50	
DXMit117	0.31	
DXMit135	0.58	

NB: “Marker” is the name of marker; “Ratio” is the ratio of cM and MB; “Overall” is the overall ratio of corresponding chromosome.

Table 29: Ratio Information of Cross [(B6-Pgk1^a X DDK)F_1 X B6]N_2(Less skewed)

Marker	Ratio	Overall
Chr.-1		
D1Mit66	0.68	0.789
D1Mit318	0.55	
D1Mit251	1.17	
D1Mit132	1.43	
D1Mit390	0.35	
D1Mit424	0.75	
D1Mit270	0.80	
D1Mit293	1.16	
Chr.-2		
D2Mit117	0.58	0.707
D2Mit83	1.07	
D2Mit244	0.73	
D2Mit37	0.28	
D2Mit276	0.63	
D2Mit285	0.28	
D2Mit200	1.37	
Chr.-3		
D3Mit164	0.32	0.711
D3Mit203	0.82	
D3Mit63	0.48	
D3Mit74	0.34	
D3Mit254	0.85	
D3Mit163	1.22	
Chr.-4		
D4Mit227	0.32	0.793
D4Mit286	0.77	
D4Mit164	1.10	
D4Mit58	0.24	
D4Mit37	0.90	
D4Mit339	1.39	
D4Mit256	1.07	

Marker	Ratio	Overall
Chr.-5		
D5Mit344	0.19	1.077
D5Mit79	1.02	
D5Mit15	1.40	
D5Mit314	0.29	
D5Mit31	2.09	
D5Mit143	1.15	
Chr.-6		
D6Mit236	0.17	0.706
D6Mit384	0.75	
D6Mit65	0.51	
D6Mit25	0.98	
D6Mit201	0.66	
Chr.-7		
D7Mit178	0.14	0.865
D7Mit117	0.90	
D7Mit310	0.48	
D7Mit30	0.73	
D7Mit323	0.20	
D7Mit291	1.29	
D7Mit223	1.79	
Chr.-8		
D8Mit157	0.22	0.865
D8Mit191	1.45	
D8Mit348	0.44	
D8Mit166	0.75	
D8Mit322	2.06	
Chr.-9		
D9Mit126	0.26	0.702
D9Mit90	1.54	
D9Mit97	0.45	

Continued...

Marker	Ratio	Overall
D9Mit270	0.63	
D9Mit212	0.44	
D9Mit281	1.39	
Chr.-10		
D10Mit298	0.35	0.771
D10Mit214	0.66	
D10Mit40	0.70	
D10Mit66	0.50	
D10Mit233	1.07	
D10Mit269	1.22	
Chr.-11		
D11Mit71	0.16	1.106
D11Mit151	0.62	
D11Mit20	1.51	
D11Mit5	0.80	
D11Mit66	0.62	
D11Mit67	1.79	
D11Mit168	1.45	
Chr.-12		
D12Mit182	0.18	0.939
D12Mit112	1.25	
D12Mit260	0.60	
D12Mit150	1.06	
Chr.-13		
D13Mit16	0.49	1.005
D13Mit63	1.55	
D13Mit99	1.00	
D13Mit107	0.27	
D13Mit78	1.10	
Chr.-14		
D14Mit10	0.27	0.703
D14Mit54	1.12	

Marker	Ratio	Overall
D14Mit234	0.80	
D14Mit162	0.48	
D14Mit170	0.42	
Chr.-15		
D15Mit12	1.49	0.560
D15Mit100	0.53	
D15Mit234	0.48	
D15Mit189	0.64	
D15Mit16	0.63	
D15Mit79	0.00	
Chr.-16		
D16Mit182	0.61	0.845
D16Mit166	1.56	
D16Mit140	0.70	
D16Mit191	0.31	
D16Mit106	0.51	
Chr.-17		
D17Mit78	0.55	0.841
D17Mit175	0.83	
D17Mit7	0.91	
D17Mit39	0.78	
D17Mit123	0.83	
Chr.-18		
D18Mit67	0.33	0.956
D18Mit60	0.70	
D18Mit123	1.18	
D18Mit47	0.98	
D18Mit144	0.87	
Chr.-19		
D19Mit42	0.49	1.167
D19Mit111	1.35	
D19Mit66	1.01	

Continued...

Marker	Ratio	Overall
D19Mit137	1.22	
Chr.-20		
DXMit124	0.31	0.904
DXMit166	1.26	

Marker	Ratio	Overall
DXMit210	0.61	
DXPas29	0.40	
DXMit117	0.37	
DXMit135	1.57	

NB: “Marker” is the name of marker; “Ratio” is the ratio of cM and MB; “Overall” is the overall ratio of corresponding chromosome.

Table 30: Ratio Information of Cross [(SM X NZB)F_1 X NZB]N_2

Marker	Ratio	Overall
Chr.-1		
D1MIT20	0.58	0.938
D1MIT22	1.49	
D1MIT92	0.69	
D1MIT34	0.5	
D1MIT17	1.37	
Chr.-2		
D2MIT1	0.26	0.944
D2MIT88	1.01	
D2MIT35	1.95	
D2MIT49	0.58	
Chr.-3		
D3MIT12	0.49	1.143
D3MIT11	19.1	
D3MIT38	1.1	
Chr.-4		
D4MIT2	0.25	0.589
D4MIT17	0.28	
D4MIT9	0.48	
D4MIT11	0.95	
D4MIT312	0.83	
Chr.-5		
D5MIT228	0.56	0.881
D5MIT114	0.95	
D5MIT7	0.68	
D5Mit10	1.27	
D5MIT239	0	
D5MIT316	0	
D5MIT209	0.53	
D5MIT65	0	
D5MIT370	0.57	

Marker	Ratio	Overall
D5MIT29	0.96	
D5MIT99	1.75	
Chr.-6		
D6MIT50	0.19	0.574
D6MIT3	0.44	
D6MIT44	0.63	
D6MIT15	0.78	
Chr.-7		
D7MIT25	0.42	1.150
D7MIT37	1.08	
D7MIT71	1.11	
D7Nds4	1.42	
Chr.-8		
D8MIT4	0.42	1.090
D8MIT45	1.41	
D8MIT56	0.65	
Chr.-9		
D9MIT2	0.46	1.060
D9MIT21	1.66	
D9MIT8	0.96	
D9MIT15	0.72	
Chr.-10		
D10Mit126	0.79	0.921
D10Mit11	0.88	
D10MIT24	1.03	
Chr.-11		
D11MIT19	0.51	0.653
D11MIT4	0.21	
D11MIT41	1.16	

Continued...

Marker	Ratio	Overall
D11MIT11	1.03	
Chr.-12		
D12MIT2	0.45	0.840
D12MIT5	0.77	
D12MIT7	1.19	
D12MIT8	0.5	
Chr.-13		
D13MIT34	0.65	0.767
D13Mit102	0.51	
D13MIT73	0.95	
D13MIT35	1.09	
Chr.-14		
D14MIT15	0.34	0.398
D14MIT37	0.14	
D14MIT7	0.69	
D14MIT97	0.37	
Chr.-15		
D15MIT18	0.66	1.311
D15MIT31	1.13	
D15MIT39	1.93	

Marker	Ratio	Overall
Chr.-16		
D16MIT3	0.73	1.302
D16MIT5	1.38	
D16MIT70	1.23	
Chr.-17		
D17MIT50	0.52	1.039
D17MIT20	0.64	
D17MIT76	1.21	
Chr.-18		
D18MIT34	0.33	1.036
D18MIT24	1.29	
D18MIT9	0.84	
Chr.-19		
D19MIT16	0.73	1.794
D19MIT27	2.43	
D19MIT71	0.74	
Chr.-20		
DXMIT89	0.31	0.580
DXMIT1	0.58	

NB: “Marker” is the name of marker; “Ratio” is the ratio of cM and MB; “Overall” is the overall ratio of corresponding chromosome.

Appendix B

1.1) R code for table 11 to 15

```
data1=read.table("B2_4_estgen_code.txt")
nc=ncol(data1)
nr=nrow(data1)

#-------------------------------------

chn=matrix(0,nc,2)
kk=0
jj=1
  for (i in 3:nc)
    {
          if(data1[1,i]==data1[1,i-1])
           {
           jj=jj+1
           }
          else
           {
           kk=kk+1
           chn[kk,1]=data1[1,i-1]
           chn[kk,2]=jj
           jj=1
            }
    }
           chn[kk+1,1]=data1[1,i-1]
           chn[kk+1,2]=jj

#-------------------------------------

mv=max(chn[,2])
mr=kk+1
nal=sum(chn[1:mr,2])

nrb=matrix(0:0,nr-2,mr)
y=array(,c(nr-2,mv,mr))
z=array(,c(nr-2,mv,mr))

ub=array(,c(mv,nr-2,mr))
lb=array(,c(mv,nr-2,mr))
```

```
    jj=0
for (l in 1:mr)
 {
  for (j in 3:nr)
   {
          k=0
         for (ii in 3:(chn[l,2]+1))
           {
                 i= jj+ii
            if (data1[j,i-1]!=9)
                 {
            if (data1[j,i]!=9)
             {
                  if(data1[j,i]!=data1[j,i-1])
                   {
                         k=k+1
                         y[j-2,k,l]=data1[2,i-1]
                         z[j-2,k,l]=data1[2,i]
                   }
             }
                 else
                 {
            if (i < nc && ii < (chn[l,2]+1))
                 {
            if (data1[j,i+1]!=9)
                 {
                  if(data1[j,i+1]!=data1[j,i-1])
                   {
                         k=k+1
                         y[j-2,k,l]=data1[2,i-1]
                         z[j-2,k,l]=data1[2,i+1]
                   }
                 }
                 }
                 }
                 }
         }
           if(k >1)
         {
          for (m in 2:k)
               {
                    ub[m-1,j-2,l]=z[j-2,m,l]-y[j-2,m-1,l]
                    lb[m-1,j-2,l]=y[j-2,m,l]-z[j-2,m-1,l]
              nrb[j-2,l]=k-1
                    }
         }
  }
             jj=jj+chn[l,2]
 }
```

```
#-------------------------------------------------

nid = nr-2
mx = matrix(,nid,1)
for (j in 1:nid)
  {
          mx[j,1]=max(nrb[j,])
  }
          nrow = sum(mx[,1])

 ncol = (2*mr) + 1
tdat = matrix(,nrow,ncol)

 k=0
for (j in 1:nid)
 {
          if(mx[j,1] != 0)
{
    for (i in 1:mx[j,1])
    {
      k = k + 1
      tdat[k,1]=data1[j+2,1]
   for (l in 1:mr)
   {
      ll = (2*l)-1
           tdat[k,ll+1]= lb[i,j,l]
           tdat[k,ll+2]= ub[i,j,l]
    }
   }
}
 }

tdat

#matrix(c(lb[1:4,1:5,1],ub[1:4,1:5,1]),ncol=2)
# This command will show the interval for the 1st individual and Choromosom no 1.
```

1.2) R code for table 21 to 25

```
data=read.table("LiUi.txt",header=TRUE)
nr=nrow(data)

li=matrix(,nr,1)
li[,1]=data[,1]
ui=matrix(,nr,1)
ui[,1]=data[,2]

diff=ui-li
#-----------------------------------------
exnd=matrix(,nr,1)

ld=0.01
z <- 999
k <- 0

#while(k < 5)
while(z > 0.000001)
{
for (j in 1:nr){
exnd[j]=exp(-ld*diff[j])
}
#function-------------------------------------------------
sum1=0.0
for (j in 1:nr)
        {
sum1=sum1+(((ui[j]*exnd[j])-li[j])/(1-exnd[j]))
        }

#derivative of function-----------------------------------------

sum2=0.0
for (j in 1:nr)
        {
sum2=sum2+((diff[j]^2)*exnd[j]/((1-exnd[j])^2))
        }
#interation-----------------------------------------------------
ld.new <- ld + (sum1/sum2)
z <- abs (ld - ld.new)
ld <- ld.new
k <- k+1
}
#while loop ends here-------------------
var=1/sum2
var
std=var^(1/2)
std
```

1.3) R code for table 16 to 20

```
require(survival)
dat <- read.table('data_nonparametric.txt',header=T)

cria.tau <- function(data){
l <- data$left
r <- data$right
tau <- sort(unique(c(l,r[is.finite(r)])))
return(tau)
}

tau=cria.tau(dat)

S.ini <- function(tau){
m<-length(tau)
ekm<-survfit(Surv(tau[1:m-1],rep(1,m-1))~1)
So<-c(1,ekm$surv)
p <- -diff(So)
return(p)
}
p=S.ini(tau)

cria.A <- function(dat,tau){
tau12 <- cbind(tau[-length(tau)],tau[-1])
interv <- function(x,inf,sup) ifelse(x[1]>=inf & x[2]<=sup,1,0)
A <- apply(tau12,1,interv,inf=dat$left,sup=dat$right)
id.lin.zero <- which(apply(A==0, 1, all))
if(length(id.lin.zero)>0) A <- A[-id.lin.zero, ]
return(A)
}
A=cria.A(dat,tau)

Turnbull <- function(p, A, dat, eps=1e-7, iter.max=5000, verbose=FALSE){
n<-nrow(A)
m<-ncol(A)
Q<-matrix(1,m)
iter <- 0
repeat {
iter <- iter + 1
diff<- (Q-p)
maxdiff<-max(abs(as.vector(diff)))
if (verbose)
print(maxdiff)
if (maxdiff<eps | iter>=iter.max)
break
Q<-p
C<-A%*%p
p<-p*((t(A)%*%(1/C))/n)
```

```
}
cat("Iterations = ", iter,"\n")
cat("Max difference = ", maxdiff,"\n")
cat("Convergence criteria: Max difference < 1e-7","\n")
dimnames(p)<-list(NULL,c("P Estimate"))
surv<-round(c(1,1-cumsum(p)),digits=5)
right <- dat$right
if(any(!(is.finite(right)))){
t <- max(right[is.finite(right)])
return(list(time=tau[tau<t],surv=surv[tau<t]))
}
else
return(list(time=tau,surv=surv))
}
Turnbull(p,A,dat)
```

1.4) R code for checking condition:

```
data <- as.matrix(read.table('data11_3.txt'))
a<-data
d<- as.matrix(read.table('del11_3.txt'))
t=nrow(a)
n=ncol(a)
e=matrix(0,n,1)
e=a%*%d
eta=c(e)
div=a/eta
sss<-colSums(div)
m<-t-sss
m
```