KARYOTYPE EVOLUTION IN VIVO AND IN VITRO

by

Bengt Olle Bengtsson

the Queen's College

Thesis submitted for the degree of

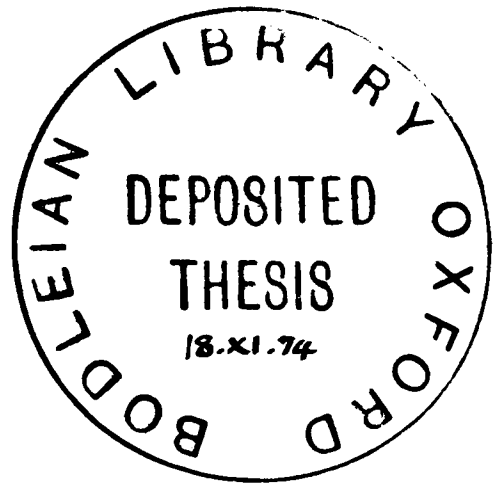
Doctor of Philosophy

in the University of Oxford

Genetics Laboratory

Department of Biochemistry

September 1974



"On peut discuter avec tout
le monde, sauf avec les
chromosomes." P. Jardin

ACKNOWLEDGEMENTS

I would like to thank:

- The Fellows of the Queen's College who gave me a Florey Studentship and thereby made it possible for me to come to Oxford;

- All the friends and colleagues who for shorter or longer periods have worked at the Genetics Laboratory when I was there. Here I also include the technical staff which made the research possible. Without the friendliness and the encouragement that met me in the lab, this thesis would never have been written;

- S. Povey and D. Swallow from the Galton Laboratory in London and R. Kennett and M. Nabholz from the Genetics Laboratory in Oxford for allowing me to use their results in Chapter 5; I am also grateful to D. Mayers from the Oxford University Computing Laboratory who calculated Table 2 in Chapter 1;

- My friends at l'Institut National d'Etudes Démographiques, Paris, and the Institute of Genetics, Lund, who provided me with facilities to write this thesis and with whom I have had many fruitful discussions;

- Everyone who has patiently corrected my English;

- M. Sergo for typing the thesis;

- Finally, Walter Bodmer who was my supervisor. Everything I understand in the field of genetics, I have learnt from him. But most of all I am grateful to Walter for his confidence in my work and for never questioning the way in which I did it.

TABLE OF CONTENTS

		Page
Abstract		I
PART I: Theoretical Considerations on Karyotype Evolution in Mammals		
Chapter 1	The Increase of a Balanced Chromosome Mutation	
	1.1 Introduction	1
	1.2 Basic model	4
	1.3 Random genetic drift	8
	1.4 Segregation distortion	10
	1.5 Viability differences	12
	1.5.1 Chromosome mutations with effects on viabilities	12
	1.5.2 Chromosome mutations decreasing the linkage between two segregating loci	14
	1.6 Final comments	20
	1.7 Appendix	21
	1.8 References	24
Chapter 2	The Fertility of Heterozygotes for a Balanced Chromosome Mutation	
	2.1 Introduction	27
	2.2 The frequency of unbalanced gametes formed by a chromosomal heterozygote	27
	2.3 The fertility of chromosomal heterozygotes	29
	2.4 The selective disadvantage of chromosomal heterozygotes in humans	31
	2.5 References	36
Chapter 3	Chromosome Mutations, Gene-Flow and Speciation	
	3.1 Introduction	37
	3.2 The creation of the difference in karyotype	38
	3.3 The effect of a karyotype difference on the gene-flow between two populations	40
	3.4 Chromosome mutations and speciation	44
	3.5 References	46
Chapter 4	Mammalian Chromosomes Similar in Length are also Similar in Shape	
	4.1 Introduction	47
	4.2 Method	48
	4.3 Results	49
	4.4 Discussion	50
	4.5 References	53

PART II: Karyotype Evolution in Cell Lines

Chapter 5 Karyotypic and Genetic Analysis of Human Intraspecific Somatic Cell Hybrids
5.1 Introduction 54
5.2 Materials and Methods 55
5.3 Results 59
5.3.1 Characterization of the fusion products 59
5.3.2 Karyotype evolution during continued growth of the hybrids 63
5.3.3 Culture conditions and chromosome loss 65
5.3.4 Thioguanine resistant lines 66
5.3.5 Spontaneous segregant lines 68
5.4 Discussion 70
5.5 References 72
5.6 Tables and Figures 75

Chapter 6 Karyotype Selection in Cell Lines
6.1 Introduction 76
6.2 A model of karyotype selection 77
6.2.1 Preliminaries 77
6.2.2 The model 78
6.2.3 Equilibrium behaviour 79
6.2.4 Numerical example 81
6.3 Discussion 82
6.4 Appendix 84
6.5 References 85

The tables and the figures are included in the text immediately after they have been referred to, except in Chapter 5 where they are collected in a special section (5.6). Tables and figures are not given page numbers.

Throughout the thesis expressions of type 1/2 xyz shall be read as $\frac{1}{2} \cdot xyz$.

ABSTRACT

KARYOTYPE EVOLUTION IN VIVO AND IN VITRO

The present thesis is divided into two parts; the first part, comprising chapters 1 to 4, called "Theoretical considerations on karyotype evolution in mammals", and the second part, chapters 5 and 6, called "Karyotype evolution in cell lines". All the chapters are selfcontained and have their own lists of references.

The first three chapters deal with different population genetical aspects arising from the reduced fertility of chromosomal heterozygotes relative to chromosomal homozygotes.

In chapter 1, after an introduction about karyotypes and chromosome mutations, various factors which can influence the behaviour of a chromosome mutation in a population are considered. The probabilities are given for a chromosome mutation to become fixed in a small population due to random genetic drift, and the conditions are specified under which segregation distortion, viability advantage and recombination modification can help a chromosome mutation to increase in frequency in a large, random mating population.

Chapter 2 brings up to discussion the factors which determine the fertility disadvantage of chromosomal heterozygotes. The role of the reproductive system in influencing the effective fertility of individuals which produce lethal but functional gametes is stressed. In the last part of chapter 2 we describe how the fitness disadvantage associated with human chromosome mutations can be estimated.

The importance of chromosome mutations in the speciation process is debated in chapter 3. According to a common idea, chromosome mutations are important for speciation because they can help to decrease

the effective gene-flow between two partially separated populations. However, there are a number of problems associated with this view. The main obstacle is the rather small effect of karyotype differences on gene-flow between two populations; this is shown in chapter 3 with a simple mathematical model. The chapter ends with us turning the question around through arguing that a chromosome mutation may have a much better chance to spread in a speciation situation, which implies that speciation is important for chromosome mutations rather than the reverse.

Chapter 4 is a short chapter introducing the use of nonparametric statistics to karyotype research. We here show that there exists a significantly high proportion of mammalian karyotypes containing acrocentric and bi-armed chromosomes in a non-random mixture. The result is perhaps not very surprising, but the chapter should indicate that this approach to karyotype research can be fruitfully continued, given more data and the help of a computer.

The results obtained from an investigation of a series of human intra-specific hybrids made between lymphocytes and D98/AH-2 cells are described in chapter 5. The author of this thesis did most of the tissue culture work involved and the karyotype analysis. Special effort was put into determining and describing karyotypic and genetic - enzyme and immunological - markers of the hybrid cells. All the hybrid lines were karyotypically very stable, and the karyotype evolution of one line, DM, was closely followed during more than 500 days of growth in culture. A number of 6-thioguanine resistent lines were isolated, and examples are given which show how such segregant lines can be used for genetic analysis of the human genome.

Established cell lines grown under constant conditions normally reach a kind of equilibrium in respect to the karyotypes of the cells in the cell line. The equilibrium is characterized by almost all cells having different but, at the same time, similar karyotypes. In chapter 6 is proposed a model which we hope can be used to describe such situations of variability and

stability. Based on the idea that there is one type of cells in the cell line which is ideal and that there are many ways by which cells can differ from the ideal type of cells, this model is in certain aspects very crude. It is, however, the first model of its kind, as far as we know, and other models which try to be interpretable must probably be based on premises very similar to the ones used in chapter 6.

CHAPTER 1

THE INCREASE OF A BALANCED CHROMOSOME MUTATION

In the first part of this thesis we shall consider various population genetical aspects of chromosome mutations. Most chromosome mutations, as will be discussed below, give a reduced fertility and, thereby, fitness to chromosomal heterozygotes. They thus behave as heterozygote disadvantageous genes. This is the simple fact on which the discussion in the first three chapters is based. Chapter 2 looks at what determines the size of the fitness decrease in heterozygotes and how the decrease can be measured, and in chapter 3 is discussed the effect which differences in karyotypes may have on the rate of gene-flow between two populations. In the present chapter, after a short introduction on karyotypes and chromosome mutations, we shall review the ways by which a chromosome mutation can spread in a population despite its selective disadvantage. We refer in the discussions and in the empirical examples almost exclusively to chromosome mutations involving autosomes in mammals, but the arguments can of course be applied to any genetic factor in any outbreeding species which behaves in a similar way.

1.1 Introduction

Most mammals have chromosome complements, karyotypes, which are specific for the species to which they belong. Even with such a crude measure of karyotype similarity as the diploid chromosome number, one frequently finds differences between related species, for example between humans with 46 chromosomes per cell and chimpanzees with 48 chromosomes. The comparison of karyotypes of different species and populations has become an important tool for studies on evolutionary relationships

(Benirschke, ed. 1969), and new staining techniques (see Caspersson and Zech, eds. 1973) have made it possible to compare karyotypes in great detail. Thus, the human karyotype is now said to differ from the karyotype of the chimpanzee by "six pericentric inversions, one telomeric fusion-translocation, 4 intercalar deletions or insertions, 16 deletions or additions of terminal Q-bands and an important variation of heterochromatin" (Lejeune et al. 1973).

There exist, as we can see from this quotation, many types of chromosome mutations which can be involved in differentiation and divergence of karyotypes (Hsu and Mead 1969, Ford 1970, Meylan 1970, White 1973). We shall in this chapter limit our discussion to balanced structural mutations, i.e. those changes in the karyotype which arise through breakage and reunion of chromosomes and which leave the genes and their dose-relations intact. The simplest kinds of such changes are translocations and inversions. Translocations are sometimes divided into three groups: 1) interchanges (reciprocal translocations) in which chromosome segments on different chromosomes change places, 2) centric fusions (Robertsonian translocations) where two one-armed chromosomes join to create a bi-armed chromosome, and 3) centric fissions where one bi-armed chromosome splits into two one-armed chromosomes. Inversions, on the other hand, are divided into pericentric and paracentric, depending on whether the inverted segment contains the centromere or not.

Heterozygotes and homozygotes for inversions and translocations have, if the changes are balanced, all the genes that other animals have, and are therefore usually phenotypically normal. (See however Jacobs (1974), where it is shown that the frequency of chromosomal heterozygotes is higher among mentally retarded individuals than in the general population, which indicates that chromosome mutations sometimes can have an effect on the phenotype.) The problem arises in the gamete production by chromosomal heterozygotes, where the pairing of chromosomes in meiosis followed by their separation leads to the creation of occasional gametes with

an unbalanced haploid chromosome set. (In Lewis and John (1963), there is a very thorough discussion of the meiosis in chromosomal heterozygotes.) The frequency of gametes which are unbalanced due to effects of a chromosome mutation has been found to vary with different mutations from negligible to more than half of the gametic production. Mammalian unbalanced gametes are functional even when they contain a very abnormal set of chromosomes, as has been shown genetically by Snell (1946) and cytologically by Ford (1972). Matings between animals where at least one parent produces unbalanced gametes, therefore lead to the formation of chromosomally abnormal zygotes which almost always abort, or if not, lead to young that die early. This leads to the effect that chromosomally heterozygote animals, on average, have lower fertility than other animals of the same species.

Such decreases in fertility have been found, directly or indirectly, to exist, for example, in heterozygotes for X-ray induced reciprocal translocations in mice (Koller 1944, Snell 1946, Carter et al. 1955), in heterozygotes for a Robertsonian translocation in cattle (Gustavsson 1969, 1971), and in human translocation heterozygotes of different kinds (Jacobs et al. 1970). There exist, unfortunately, very little empirical data on the fertility of inversion heterozygotes since only a few well established instances of inversions have been studied in mammals.

Already in one of the first papers on population genetics (Fisher 1922), it was stated that

> 'If selection favours the homozygotes, no stable equilibrium will be be possible, and selection will then tend to eliminate whichever gene is below its equilibrium proportion; such factors will therefore not commonly be found in nature: ..."

This argument applies equally to the chromosome mutations discussed above, since selection favours the chromosomal homozygotes by disfavouring the heterozygotes. There is thus an apparent opposition between the fertility disadvantage of chromosomal heterozygotes and the occasional

spread to fixation of structural chromosome mutations which constitutes one of the major causes of karyotype evolution.

This problem was recognized a long time ago but has, apart from a few exceptions, not attracted much attention from theoretical population geneticists. Wright (1941) wrote an article on the probability of fixation of a chromosome mutation by random drift (this problem is further discussed below), and Haldane discussed the effect of heterozygote disadvantage in connection with the finding of the Rh-polymorphism in the same year (1941). Aspects of the problem have later been taken up by entomologists studying the effect on the over-all fertility of an insect population when chromosomally changed mosquitos are released (see for example Whitten 1971, Curtis and Hill 1971). The theoretical work done on inversions in Drosophila (a review of most of that literature is given in Charlesworth and Charlesworth 1973), and in Moraba scurra (Lewontin and White 1960, White et al. 1963, Turner 1972) is, unfortunately, not relevant to the present problem since the inversion heterozygotes in these species have normal fertility.

1.2 Basic model

We shall in this chapter look at the possible effects of 1) random genetic drift, 2) segregation distortion, 3) association with a viability advantage, and 4) linking of two interacting genes on the behaviour of a new chromosome mutation in nature. These evolutionary factors are probably the most important ones which may influence the fate of a new chromosome mutation.

To describe the situation in a population which has a new balanced chromosome mutation introduced, we shall throughout the chapter use a simple multiplicative fertility model and assume random mating. The fertilities given to animals with different karyotypes will be according to table 1, where the standard karyotype is denoted by SS, the heterozygote karyotype by ST, and the karyotype which is homozygote for the chromosome

mutation by TT. The model is designed to describe heterozygote infertility, so s_f and s_m will always be positive. A fertility model is said to be multiplicative if the relative offspring size of a mating is equal to the product of the parental fertilities. Thus, in our model, a mating between two homozygote animals produces on average 1 "offspring unit", while a mating between two heterozygotes produces only $(1-s_m)(1-s_f)$ "units". Bodmer (1965) showed that, in random mating populations, multiplicative fertility models are equivalent to models assuming differential viabilities (with different viability parameters in the two sexes).

Both Wright (1941) and Curtis and Hill (1971) discuss the following fact which may make multiplicative fertility models less suited to describe the behaviour of a chromosome mutation in a population: Two chromosomal heterozygotes can at a mating produce unbalanced gametes which are such that they complement each other. The fertility of matings between two chromosomal heterozygotes is thus somewhat greater than would be expected from a multiplicative fertility model. The effect of this phenomenon on the population is, however, minute, as pointed out by Wright (1941), and the argument that the fertility of a mating stands in a direct linear relationship to the frequency of balanced zygotes formed, is anyway not generally tenable (see Chapter 2). It is therefore reasonable to use the present, simple model and postpone the discussion of more complex fertility models until some empirical data on the differential fertilities of different matings exist.

The heterozygotes of the two sexes are given different fertilities in table 1. This is done because there exist data that, for example, X-ray induced translocations can have different effects on the fertilities of males and females in mice (Snell 1946, Carter et al. 1955). Some human Robertsonian translocations do also differ in the frequencies with which heterozygotes in the two sexes produce offspring with unbalanced karyotypes (Stene 1970). When the distinction between the sexes is not made, the fertility of heterozygotes is called 1-s.

	Karyotype		
	SS	ST	TT
Males	1	$1-s_m$	1
Females	1	$1-s_f$	1

Table 1. Postulated fertilities of animals with different karyotypes.

If assortative mating is to be discussed in association with the fertility model described in table 1, then the evolution of the population cannot be expressed only by the change in the frequency of the chromosome mutation. A mating matrix must be constructed to deduce the change of mating-type and karyotype frequencies from one generation to the next. Such models become very similar to those which take the effect of complementary gametes into account (Curtis and Hill 1971).

Let us, before we discuss the effects that other evolutionary factors may have on the behaviour of a chromosome mutation, first look at some of the properties of the simple model outlined in table 1. Regard a very big, random mating population, and call the proportion of SS animals in generation n for P, the proportion of ST animals Q, and the proportion of TT animals for R. Using q to stand for the frequency of the chromosome mutation in the population (q is thus equal to R + 1/2Q), we find the karyotype frequencies in the next generation to be:

$$P' = (P + 1/2(1-s_f)Q) \cdot (P + 1/2(1-s_m)Q) / W = (p - 1/2s_fQ)(p - 1/2s_mQ) / W,$$

and similarly

$$Q' = \left[(p - 1/2s_fQ)(q - 1/2s_mQ) + (p - 1/2s_mQ)(q - 1/2s_fQ)\right] / W$$

$$R' = (q - 1/2s_fQ)(q - 1/2s_mQ) / W,$$

where

$$p = 1 - q = P + 1/2Q$$

and

$$W = (p - 1/2s_fQ)(p - 1/2s_mQ) + (p - 1/2s_fQ)(q - 1/2s_mQ) +$$

$$+ (p - 1/2s_mQ)(q - 1/2s_fQ) + (q - 1/2s_fQ)(q - 1/2s_mQ) =$$

$$= (1 - s_fQ)(1 - s_mQ).$$

The frequency of the chromosome mutation in generation n+1 is:

$$q' = R' + 1/2Q' = \left[(q - 1/2s_f Q)(q - 1/2s_m Q) + 1/2(p - 1/2s_f Q)(q - 1/2s_m Q) + \right.$$

$$\left. + 1/2(p - 1/2s_m Q)(q - 1/2s_f Q)\right] / W =$$

$$= q - (1/2 - q) \cdot Q \cdot \left[\frac{s_f}{2(1-s_f Q)} + \frac{s_m}{2(1-s_m Q)}\right].$$

When q is small, the change in frequency between two generations, q' - q is approximately:

$$- 1/2 \cdot 2q\left[\frac{s_f}{2} + \frac{s_m}{2}\right] = - q \cdot \frac{s_f + s_m}{2} .$$

The frequency of the chromosome mutation will thus always decrease in the population if the initial frequency is less than a half. The mutation disappears at a geometric rate determined by the mean of the fertility decreases in the two sexes.

It is worth noticing a peculiarity in the equations of the karyotype frequencies, namely that the population is not in Hardy-Weinberg proportions if s_m and s_f are unequal. The deviation is such that there are relatively more heterozygotes than expected from Hardy-Weinberg considerations, although the heterozygotes are the animals which have the selective disadvantage. The observation is, however, mainly of academic interest, since the surplus of heterozygotes due to the effect of different gametic outputs in the two sexes (Purser 1966) is so small that it cannot for example be detected in nature, and so would not be confused with heterozygote fitness advantage.

Let us now introduce other evolutionary factors into this model to see what importance they may have for the fate of a chromosome mutation in a random mating population.

1.3 Random genetic drift

Many writers on evolution, for example Mayr (1969), have suggested that karyotype differentiation normally occurs in small, isolated populations where chromosome mutations can spread due to random genetic drift. Wright (1941) discussed the possible effect of drift on reciprocal translocations and found that, in a population of N animals, the formula $(2e/N)(3/4)^{2N}$ approximates the probability that a semi-sterile chromosome mutation increases from a single mutation until fixation due to genetic drift. Since 1941 population genetics has been given new tools to solve problems of this kind. Kimura (1962) derived a general formula, based on diffusion approximation theory, for the probability of increase until fixation in a finite population of any mutation for which the selective values of the genotypes involved were known. When discontinuous generations are considered, the probability of increase until fixation, assuming that the mutation at first only exists in one heterozygote, is:

$$u = \int_0^{(2N)^{-1}} W(q)^{-2N}\,dq \Big/ \int_0^{1} W(q)^{-2N}\,dq \ ,$$

where N is the effective population size, and W(q) is the fitness of a population with the mutation present at frequency q (Kimura 1970).

Other authors have discussed the probability of fixation of genes with heterozygote advantage (Robertson 1962), but nobody, to our knowledge, has so far used this method to study the behaviour of heterozygote disadvantageous factors. If no difference between male and female heterozygotes is assumed, then in our case W(q) is equal to $1 - 2sq(1-q)$. The probability of fixation of a chromosome mutation due to random drift in a population of effective size N, is thus:

$$u = \int_0^{(2N)^{-1}} \left[1-2sq(1-q)\right]^{-2N} dq \Big/ \int_0^{1} \left[1-2sq(1-q)\right]^{-2N} dq \ .$$

Numerical methods to estimate this formula have been worked out by Dr. Mayers at the Computing Laboratory at Oxford University, and he

has also calculated u for some different values of s and N. The results are given in table 2. It is interesting to notice how well the probabilities given by Wright's formula correspond to the here presented values for s = 1/2. Thus, according to Wright, $u = 0.35 \cdot 10^{-13}$ for N = 50 and $u = 0.56 \cdot 10^{-26}$ for N = 100, while the values given in table 2 are $0.38 \cdot 10^{-13}$ and $0.87 \cdot 10^{-26}$.

The information given in table 2 is easier to analyse if it is transformed into the average number of chromosome mutations which will spread through a population during a certain time due to genetic drift. In table 3 is given the expected number of chromosome mutations that will become fixed in a population of size N during 10^6 generations, assuming - rather arbitrarily - that the mutation frequency of chromosome mutations with heterozygote fertility decrease s is 10^{-5}.

The size of a population is, of course, important if it is to evolve karyotypically due to drift. A chromosome mutation causing only a relative fertility decrease in heterozygotes of 0.001 has practically no chance of spreading in a population of 10 000 animals, while mutations reducing the heterozygote's fertility by 25 % would have a chance to become fixed in populations of 10 animals. If we assume that hardly any mammalian population is of smaller effective size than 100 animals over a long time, then only chromosome mutations with fertility effects less than 0.1 have reasonable chances to increase due to drift. In most species this value is probably many orders of magnitude smaller.

It is important to notice how the probability of fixation in a population of a given size depends on s. A chromosome mutation behaves like a neutral gene if s is smaller than 1/N, and it is heavily selected against if s is much greater than 1/N. Small changes of s can however cause great changes in the probability of fixation, particularly when s is around 10/N. The difference is thus 17 orders of magnitude between the probabilities of fixation for a chromosome mutation with s = 0.01 and a mutation with s = 0.05 in a population consisting of 1 000 animals. Factors which influence the

N	s = 0.50	0.25	0.10	0.05	0.01	0.001
10	$0.82 \cdot 10^{-3}$	$0.88 \cdot 10^{-2}$	$0.27 \cdot 10^{-1}$	$0.37 \cdot 10^{-1}$	$0.47 \cdot 10^{-1}$	$0.50 \cdot 10^{-1}$
50	$0.38 \cdot 10^{-13}$	$0.90 \cdot 10^{-7}$	$0.17 \cdot 10^{-3}$	$0.16 \cdot 10^{-2}$	$0.72 \cdot 10^{-2}$	$0.97 \cdot 10^{-2}$
100	$0.87 \cdot 10^{-26}$	$0.10 \cdot 10^{-12}$	$0.72 \cdot 10^{-6}$	$0.85 \cdot 10^{-4}$	$0.25 \cdot 10^{-2}$	$0.47 \cdot 10^{-2}$
1 000	$0.2 \cdot 10^{-251}$	$0.1 \cdot 10^{-117}$	$0.17 \cdot 10^{-46}$	$0.4 \cdot 10^{-24}$	$0.80 \cdot 10^{-7}$	$0.2 \cdot 10^{-3}$
10 000	$0.8 \cdot 10^{-2\,501}$	$0.4 \cdot 10^{-1\,162}$	$0.5 \cdot 10^{-448}$	$0.1 \cdot 10^{-222}$	$0.16 \cdot 10^{-46}$	$0.8 \cdot 10^{-8}$
1 000 000	$0.1 \cdot 10^{-249\,880}$	$0.4 \cdot 10^{-115\,987}$	$0.3 \cdot 10^{-44\,556}$	$0.2 \cdot 10^{-21\,994}$	$0.8 \cdot 10^{-4\,358}$	$0.7 \cdot 10^{-439}$

Table 2. The probability of fixation of a chromosome mutation with fertility decrease s in a population of size N.
(Computed by Dr. Mayers.)

N	s = 0.50	0.25	0.10	0.05	0.01	0.001
10	$0.82 \cdot 10^{-1}$	0.88	2.7	3.7	4.7	5.0
50	$0.19 \cdot 10^{-10}$	$0.45 \cdot 10^{-4}$	$0.85 \cdot 10^{-1}$	0.80	3.6	4.9
100	$0.87 \cdot 10^{-23}$	$0.10 \cdot 10^{-9}$	$0.72 \cdot 10^{-3}$	$0.85 \cdot 10^{-1}$	2.5	4.7
1 000	$0.2 \cdot 10^{-247}$	$0.1 \cdot 10^{-113}$	$0.17 \cdot 10^{-42}$	$0.4 \cdot 10^{-20}$	$0.80 \cdot 10^{-3}$	2
10 000	$0.8 \cdot 10^{-2\,496}$	$0.4 \cdot 10^{-1\,157}$	$0.5 \cdot 10^{-443}$	$0.1 \cdot 10^{-217}$	$0.16 \cdot 10^{-41}$	$0.8 \cdot 10^{-3}$
1 000 000	$0.1 \cdot 10^{-249\,873}$	$0.4 \cdot 10^{-115\,980}$	$0.3 \cdot 10^{-44\,549}$	$0.2 \cdot 10^{-21\,987}$	$0.8 \cdot 10^{-4\,351}$	$0.7 \cdot 10^{-432}$

Table 3. Mean number of chromosome mutations fixed during 10^6 generations assuming a mutation rate of 10^{-5}.

fertility of chromosomal heterozygotes, for example the functioning of the reproductive system (see chapter 2), can thus in some situations be of great importance in determining the rate of karyotype evolution by drift.

1.4 Segregation distortion

Another evolutionary factor which has been suggested (Sandler and Novitski 1957) to be associated with chromosome mutations is segregation distortion. This is a name for everything which may influence the segregation of the two factors in a heterozygote so that they do not get represented in equal proportions in the offspring. In table 4, we have named the segregation ratio of a chromosomal heterozygote with respect to the chromosome mutation t_m/n_m and t_f/n_f for males and females respectively. Letting P, Q and R stand for the same karyotype frequencies as before, we find that the frequency of the chromosome mutation in the next generation now becomes:

$$q' = R' + 1/2Q' = (R + t_f(1-s_f)Q) \cdot (R + t_m(1-s_m)Q) \; / \; W +$$

$$+ 1/2\left[(P+ n_f(1-s_f)Q)(R + t_m(1-s_m)Q) + (P+ n_m(1-s_m)Q)(R + t_f(1-s_f)Q) \right] / \; W =$$

$$= \frac{(R + (1-s_f)t_fQ)(1-s_mQ)}{2W} + \frac{(R + (1-s_m)t_mQ)(1-s_fQ)}{2W} \quad ,$$

where

$$W = (1-s_fQ)\,(1-s_mQ)$$

as before.

When the chromosome mutation is rare, we can express the frequency as

$$q' \sim \frac{(1-s_f)t_f2q}{2} + \frac{(1-s_m)t_m2q}{2} = q(t_f(1-s_f) + t_m(1-s_m)) \; .$$

The condition for increase when the chromosome mutation is rare, is thus

$$t_f(1-s_f) + t_m(1-s_m) > 1.$$

	Chromosome type:	
	S	T
ST males	n_m	t_m
ST females	n_f	t_f

Table 4. The frequency with which male and female chromosomal heterozygotes give the different chromosome types to their offspring. $n_m + t_m = 1$, and $n_f + t_f = 1$.

One can rather easily show that under this condition the frequency of the chromosome mutation will continue to increase, i.e. q' is always bigger than q. Segregation distortion can thus never give rise to a stable chromosomal polymorphism in the present system.

Let us give a numerical example of how segregation distortion may help a chromosome mutation to spread: If male heterozygotes for a chromosome mutation have relative fertility 0.95 and no segregation distortion, while the females have fertility 0.91 and a deviation from the normal segregation pattern so that 62 % of her offspring have the chromosome mutation, then the mutation will increase in the population until it has become fixed.

The only well known case of segregation distortion in relation to chromosomes in mammals is the segregation distortion of the X chromosome in XO female mice. (These mice are fertile unlike their human equivalents.) Much more than half of their daughters have the normal XX chromosome complement, which seems to be due to meiotic drive favouring the formation of gametes with X chromosomes as opposed to these without (Kaufman 1972). No convincing instance of deviation from the normal segregation ratio has so far been found when various translocations have been studied: for example, X-ray induced reciprocal translocations in mouse (Snell 1946, Carter et al. 1955), a spontaneously derived centric fusion translocation also in mouse (Evans et al. 1967), and a centric fusion translocation in cattle (Gustavsson 1969). There were reports at one time that segregation distortion was involved in the transmission of some human translocations (D/G), but later, more careful statistical investigations have found no evidence for this (Stene 1970, Jacobs et al. 1970). This finding is important because the studies involve pooling data on many chromosome mutations. If there existed variation between individual chromosome mutations in segregation ratios, then those with ratios favouring transmission of the chromosome mutation would be over-represented in the studied group of mutations and it might have been possible to detect

a deviation of the mean segregation ratio from one half. The absence of segregation distortion in data from pooled human translocations is thus evidence against the existence of a high variation of segregation ratios between different mutation.

1.5 Viability differences

In the preceding sections it has been assumed that the lower fertility of heterozygotes is the only fitness difference of importance in a chromosomally polymorphic population. (We use the word fitness to mean "the overall measure of natural selection ... /which/ includes both the effects of viability and fertility" (Cavalli-Sforza and Bodmer 1971, p. 72).) It is, however, likely that rearrangement of genes which is a concomitant of every chromosome mutation, also can be associated with changes in viability. Here we shall discuss two different cases. In the first it is assumed that chromosome mutations themselves not only have effects on the fertility of animals but also on their viabilities. In the second case we shall consider chromosome mutations which do not by themselves influence viabilities, but which change the degree of linkage between two loci, each having two alleles with different effects on fitness.

1.5.1 Chromosome mutations with effects on viabilities

We assume that the viability effects of the chromosome mutation are identical in the two sexes and denote by $1 + d$ the viability of TT animals, by $1 + hd$ the viability of ST animals and by 1 the viability of animals with the standard genotype, SS. The most likely event will not be considered here since it is not of interest in the present context, namely that d and/or hd are negative, for example due to damage of important genetic material by the breaking of chromosomes involved in the chromosome mutation. Only if the heterozygotes have higher viability than animals with the standard karyotype, i.e. if $hd>0$, can the viability differences play a role in the karyotype evolution.

The combined effects of viability and fertility give rise to a fitness model, illustrated in table 5, where male and female heterozygotes have different fitness values. Such models of sex-differences in selection have been thoroughly studied, mainbly by Owen (1953), Bodmer (1965), Mandel (1971) and Karlin (1972).

The necessary condition for an increase in frequency of a chromosome mutation which is very rare in the population, is simple to give (from Parsons 1961):

$$\frac{(1-s_f)(1+hd)}{2} + \frac{(1-s_m)(1+hd)}{2} > 1 .$$

In this situation, unlike the case of segregation distortion, initial increase does not automatically imply fixation. A chromosome mutation which initially increases in frequency may not go to fixation but may move instead to an internal equilibrium. Owen (1952) showed that models of differential fitness in the two sexes can have up to three internal equilibria, two of which are stable. We shall now see when there may exist a stable internal equilibrium in the model described by table 5.

To simplify the discussion we say that the SS equilibrium is unstable when we mean that, if the chromosome mutation is rare, then it will increase in frequency. The condition for this was just given. Similarly, the TT equilibrium is unstable if and only if

$$\frac{(1-s_f)(1+hd)}{2} + \frac{(1-s_m)(1+hd)}{2} > 1 + d .$$

We are only interested in the situations when the SS equilibrium is unstable and we can then distinguish two cases depending on the stability of the TT equilibrium:

1. Both the SS and the TT equilibria are unstable. In this situation Mandel (1971) has shown that there will exist at least one stable internal equilibrium, as would be expected.

	Karyotype		
	SS	ST	TT
Males	1	$(1+hd)(1-s_m)$	1+d
Females	1	$(1+hd)(1-s_f)$	1+d

Table 5. General fitness model showing the combined effect of differential viabilities and fertilities in animals with different karyotypes.

2. The TT equilibrium is stable and the SS equilibrium is unstable. One can then show, as has been done in the appendix, that there will not exist any internal equilibria if both heterozygotes are more fit than the standard homozygote SS, i.e. if

$$\text{both } (1-s_f)(1+hd) > 1 \quad \text{and} \quad (1-s_m)(1+hd) > 1.$$

If one heterozygote is less fit than the standard homozygote (both of them cannot be because the SS equilibrium is unstable), then there may exist a stable internal equilibrium. It is, however, unlikely that a population will stay a long time at an internal equilibrium of this type, where a considerable proportion of the population has a decrease of fertility. Genetic factors which influence the meiosis of heterozygotes so that they achieve a more normal fertility will in such situations become important and will be favoured by selection. The fitness values will thereby change and the population will probably move to fixation of the chromosome mutation, unless the heterozygotes develop higher fitness than both types of homozygotes.

No chromosome mutation in a mammal is known to carry an advantage in viability relative to other animals. Nor is it certain that any of the chromosome polymorphism found in mammalian populations are due to the existence of a stable polymorphism which can be investigated by the present type of models. Our ignorance is mainly due to the great difficulties associated with determining a chromosome mutation's effects on fertility and viability. This problem and how it can be solved at least partially for human translocations will be further discussed in chapter 2.

1.5.2 Chromosome mutations decreasing the linkage between two segregating loci

Fisher (1930) was the first to point out that selection may favour closer linkage between two advantageous alleles. The behaviour of two loci when the recombination value between them is changed, has in later years become the subject of many papers in population genetics. Recombination

modifier genes with no direct effects on fitness have then normally been discussed (see Feldman 1972 for the most important results and a review of the literature). The term secondary selection has been introduced to denote selection favouring an allele, for example a neutral recombination modifier allele, which does not by itself change the fitness of individuals but which can influence the fitness of the population. Discussing the difference between primary and secondary selection, Feldman and Balkau (1973) wrote: "we may even conceive of a gene which decreased the fitness of its carriers but still increased by virtue of secondary selection, i.e. its effect on population fitness". It is this type of situation which we shall now investigate. Chromosome mutations can behave just like the gene mentioned in the quotation, since chromosome mutations decrease the fitness of their carriers but may also change the linkage between genes and thereby the fitness of the population.

We shall consider two loci, A and B, each with two alleles, $\underline{A}$, $\underline{a}$ and $\underline{B}$, $\underline{b}$, where two of the alleles, say $\underline{A}$ and $\underline{B}$, get closely linked together by a chromosome mutation. It will be assumed that recombinations cannot occur between these two alleles in chromosomal heterozygotes, so that gametes which carry the chromosome mutation will always also carry the alleles $\underline{A}$ and $\underline{B}$. This is not a very severe restriction; most chromosome mutations, even reciprocal translocations (Roberts 1970), are known to prevent recombination over big regions of the chromosomes.

Let the recombination fraction between the loci A and B be **r** in individuals who do not carry the chromosome mutation, and let the fitness parameters of the different genotypes in chromosomal homozygotes be according to table 6. Chromosomal heterozygotes have fitness 1-s times the value given by table 6. So is, for example, the fitness of an individual with genotype $\underline{AB}/\underline{aB}$ and karyotype ST equal to $(1-s) \cdot w_{13}$. Let the frequency of gametes with genotype $\underline{AB}$ (without the chromosome mutation), $\underline{Ab}$, $\underline{aB}$, $\underline{ab}$ and $\underline{AB}$ (with the chromosome mutation) be x_1, x_2, x_3, x_4 and x_T. The frequencies of these gametes are then in the next generation:

	Fitnesses of genotypes:		
	BB	Bb	bb
AA	w_{11}	w_{12}	w_{22}
Aa	w_{13}	$w_{23}=w_{14}=1$	w_{24}
aa	w_{33}	w_{34}	w_{44}

Table 6. The two locus fitness models used in the text.

$$x_1' = (w_1x_1 - rD + (1-s)w_{11}x_1x_T) / W_T$$

$$x_2' = (w_2x_2 + rD + (1-s)w_{12}x_2x_T) / W_T$$

$$x_3' = (w_3x_3 + rD + (1-s)w_{13}x_3x_T) / W_T$$

$$x_4' = (w_4x_4 - rD + (1-s)x_4x_T) / W_T$$

$$x_T' = (w_{11}x_T^2 + (1-s)w_1x_T) / W_T$$

where

$$D = x_1x_4 - x_2x_3$$

$$WT = w_1x_1 + w_2x_2 + w_3x_3 + w_4x_4 + 2(1-s)w_1x_T + w_{11}x_T^2$$

and

$$w_i = \sum_{j=1}^{4} x_jw_{ij} \quad ; \quad i = 1,2,3,4 \quad ; \quad w_{14} = w_{23} = 1 \quad ; \quad w_{ij} = w_{ji} .$$

(See for example Bodmer and Felsenstein (1967) for an introduction to the theoretical analysis of the two locus models.)

If the frequency of the chromosome mutation is rare, then one can make the following approximations:

$$x_1' = (w_1x_1 - rD) / W$$

$$x_2' = (w_2x_2 + rD) / W$$

$$x_3' = (w_3x_3 + rD) / W$$

$$x_4' = (w_4x_4 - rD) / W$$

$$x_T' = (1-s)w_1x_T / W$$

where

$$W = \sum_{i=1}^{4} w_ix_i .$$

The chromosome mutation will thus increase in the population, despite its selective disadvantage, if and only if

$$\frac{(1-s) \cdot w_1}{W} > 1 ,$$

which is equivalent to

$$s < \frac{w_1 - W}{w_1} .$$

Let us now assume that the population existed at a stable equilibrium at the two loci A and B when the chromosome mutation was introduced, and that the equilibrium frequencies of the different gametes were $\hat{x}_1$, $\hat{x}_2$, $\hat{x}_3$ and $\hat{x}_4$. At equilibrium

$$\hat{x}_1 = (\hat{w}_1 \hat{x}_1 - r\hat{D}) / W ,$$

which implies that

$$\hat{w}_1 - \hat{W} = \frac{r \hat{D}}{\hat{x}_1} .$$

The chromosome mutation will thus increase in this situation if and only if

$$s < \frac{r \hat{D}}{\hat{x}_1 \hat{w}_1} = \frac{r \hat{D}}{\hat{x}_1 \hat{W} + r\hat{D}} .$$

The factors that determine the fate of a chromosome mutation which closely links alleles from two segregating loci at a stable equilibrium, are thus, 1) the recombination fraction between the segregating loci in individuals with the normal karyotype; 2) the linkage disequilibrium $\hat{D}$ existing between the alleles at the equilibrium. Feldman (1972) has earlier shown that a neutral recombination modifier factor will increase in a population only if it decreases the linkage between two segregating loci with linkage disequilibrium; 3) the alleles which the chromosome mutation links together. There are two combinations of alleles, AB and ab, which are associated with the positive linkage disequilibrium. If the chromosome mutation had linked together the a and b alleles instead, then the condition for increase would be the same as given above, except that $\hat{x}_1$ would be changed to $\hat{x}_4$. It is thus advantageous for the chromosome mutation if it

links together the most rare combination of alleles with positive linkage disequilibrium.

A special situation arises when the population is at an equilibrium where at one of the loci, say A, there is only one allele present. Bodmer and Felsenstein (1967) have investigated this case and have given the conditions under which a second allele at the initially fixed locus will increase. A particularly interesting case occurs when the new allele increases only if there is close linkage between the two loci. We shall now see how such a situation can favour the spread of a chromosome mutation.

Assume that the population is polymorphic with the gametes $\underline{aB}$ and $\underline{ab}$ present at an equilibrium due to heterozygote advantage at the B locus. The same fitness scheme as above is used (w_{34} is thus greater than w_{33} and w_{44}), except that $w_{14} = w_{23}$ is not necessarily equal to 1. We assume, arbitrarily, that

$$w^{+}_{1} > w^{+}_{2} \quad ,$$

where

$$w^{+}_{1} = uw_{13} + vw_{14} \quad \text{and}$$

$$w^{+}_{2} = uw_{23} + vw_{24} \quad .$$

u is the equilibrium frequency of the gamete $\underline{aB}$ and v the equilibrium frequency of the gamete $\underline{ab}$, i.e.

$$u = \frac{w_{34} - w_{44}}{2w_{34} - w_{33} - w_{44}} \quad \text{and}$$

$$v = 1 - u$$

For reasons which will become obvious later, we are only interested in those cases where $w^{+}_{1} > w$, defining

$$w = u^2 w_{33} + 2uvw_{34} + v^2 w_{44}$$

as the mean fitness of the population at equilibrium.

Under these conditions the allele A will not increase in frequency when introduced in the population (Bodmer and Felsenstein 1967) if

$$(w - w^{+}_{2}) \cdot v > (w^{+}_{1} - w) \cdot u$$

and

$$r > \frac{(w - w^{+}_{2})(w^{+}_{1} - w)}{((w - w^{+}_{2})v - (w^{+}_{1} - w)u)w_{14}} ,$$

where r is the recombination fraction between the two loci.

If we under these conditions introduce a new gamete with the alleles A and B plus a chromosome mutation which prevents crossing over between these alleles in the chromosomal heterozygotes - but which also causes a relative fitness decrease of 1-s in all genotypes, then this new genetic combination will behave like a third allele at the B locus. It is therefore easy to give the condition for increase of this new "allele", which turns out to be

$$1 < \frac{(1-s)\, w^{+}_{1}}{w} ,$$

which is equivalent to

$$s < \frac{w^{+}_{1} - w}{w^{+}_{1}} = \frac{(w_{13}-w_{33})(w_{34}-w_{44}) + (w_{14}-w_{34})(w_{34}-w_{33})}{w_{13}(w_{34}-w_{44}) + w_{14}(w_{34}-w_{33})} .$$

This is thus the condition under which the chromosome mutation will increase in the described situation. One should observe that this is not a simple "hitch-hike" model, because the allele A will not increase by itself.

Let us now give a numerical example. Assume that there exists a polymorphism at the B locus because the fitness of aaBB animals is 0.98, of aaBb animals 1.00 and of aabb animals 0.99. The allele A is introduced by mutation at the first locus but does not increase in frequency even though the fitness of AaBB and AaBb animals is 1.06, because the fitness of Aabb animals is only 0.88. The two loci are unlinked, making r = 1/2.

If a chromosome mutation now appears in the population which links together the two alleles A and B so that recombinations between these alleles do not occur in chromosomal heterozygotes, then this chromosome mutation will increase in frequency in the population, even if all chromosomal heterozygotes have a fertility of only 95 % relative to the chromosomally homozygote animals with identical genetic constitution.

This example is of course highly simplified, but it describes how a chromosome mutation, despite negative effects on fitness, may increase in a population due to effects of the associated reorganization of the genome that it causes.

1.6 Final comments

We have now reviewed how a chromosome mutation may increase in frequency in a population despite the fitness disadvantage of the chromosomal heterozygotes. One would have preferred having more information about the existence of some of the evolutionary factors discussed. Empirical data which is sadly lacking would have been needed for this. The data is missing, both because karyologists have not always collected the information relevant to evolutionary considerations and because many of the important parameters are very difficult to estimate.

The mathematical models are nevertheless of interest, even if the relevant empirical information does not exist. The models tell us how other evolutionary factors must be related quantitatively to the fertility of the chromosomal heterozygotes, if they are going to influence the behaviour of the chromosome mutations in nature. Knowledge gained in this way not only helps us to better understand the connections between various evolutionary factors, but can also help field workers to decide which biological parameters are important to study.

1.7 Appendix

On the existence of internal equilibria in the model given by table 5.

To simplify table 5, we define a as 1+d, b as $(1+hd)(1-s_m)$ and c as $(1+hd)(1-s_f)$. Arbitrarily we shall throughout assume that $c < b$. To correspond to the model in the text, we assume that a is greater than 1, and that $b+c > 2$, i.e. the SS equilibrium is unstable (Parsons 1961).

We shall show that if the TT equilibrium is stable, then there can only exist internal equilibria if $c < 1$. There are, for example, never any internal equilibria when $1 < c < a < b$ and the TT equilibrium is stable. Bodmer (1965) has already shown that there are no internal equilibria when $1 < c < b < a$. Karlin (1972) gave a description of the internal equilibria in the general two sex differential fitness model, but the analysis is, unfortunately, not complete.

The frequency of S carrying gametes in the gametic output of males is denoted by p and the frequency in females by P. q is 1-p and Q is 1-P. Define $u_1 = p/q$ and $u_2 = P/Q$, and $t = u_2/u_1$. All these definitions are taken from Bodmer (1965). Algebra gives (see for example Bodmer 1965):

$$u_1' = \frac{u_1u_2 + 1/2b(u_1+u_2)}{a + 1/2b(u_1+u_2)}$$

$$u_2' = \frac{u_1u_2 + 1/2c(u_1+u_2)}{a + 1/2c(u_1+u_2)}$$

and

$$\Delta u_1 = u_1' - u_1 = \frac{u_1u_2 + 1/2b(u_1+u_2) - au_1 - 1/2bu_1(u_1+u_2)}{a + 1/2b(u_1+u_2)}$$

$$\Delta u_2 = u_2' - u_2 = \frac{u_1u_2 + 1/2c(u_1+u_2) - au_2 - 1/2cu_2(u_1+u_2)}{a + 1/2c(u_1+u_2)}$$

At an equilibrium, when $\bar{t} = \bar{u}_2/\bar{u}_1$, one has

$$\bar{u}_1 = \frac{b(1+\bar{t}) - 2a}{b(1+t) - 2t}$$

$$\bar{u}_2 = \frac{c(1+\bar{t}) - 2a\bar{t}}{c(1+\bar{t}) - 2}$$

From

$$\bar{u}_1 = \frac{\bar{u}_1\bar{u}_2 + 1/2b(\bar{u}_1+\bar{u}_2)}{a + 1/2b(\bar{u}_1+\bar{u}_2)}$$

it is obvious that

$$\bar{u}_1 > 1 \Leftrightarrow \bar{u}_1\bar{u}_2 > a.$$

Similarly one can show that

$$\bar{u}_2 > 1 \Leftrightarrow \bar{u}_1\bar{u}_2 > a.$$

Since

$$\bar{t} > 1 \Leftrightarrow \frac{(\bar{u}_1\bar{u}_2 + 1/2c(\bar{u}_1+\bar{u}_2))(a + 1/2b(\bar{u}_1+\bar{u}_2))}{(a + 1/2c(\bar{u}_1+\bar{u}_2))(\bar{u}_1\bar{u}_2 + 1/2b(\bar{u}_1+\bar{u}_2))} > 1 \Leftrightarrow \bar{u}_1\bar{u}_2 - a > 0$$

when $b > c$ (as assumed), we have thus shown that

$$\bar{t} > 1 \Leftrightarrow \bar{u}_1 > 1 \text{ and } \bar{u}_2 > 1$$

$$\bar{t} < 1 \Leftrightarrow \bar{u}_1 < 1 \text{ and } \bar{u}_2 < 1 \quad ,$$

under the conditions considered.

t must, of course, be positive to have a biological interpretation.

We shall now show that if $c > 1$, then there is no equilibrium with a corresponding $\bar{t}$ value greater than 1.

Assume that there existed such an equilibrium. The corresponding $\bar{u}_1$ value should then also be greater than 1. But from (3) follows:

$$\bar{u}_1 > 1 \Rightarrow \bar{u}_2 / \bar{t} > 1 \Rightarrow c\bar{t}(1+\bar{t}) - 2\bar{t} < c(1+\bar{t}) - 2a\bar{t} \Rightarrow a < 1 - \frac{c(\bar{t}-1)(\bar{t}+1)}{\bar{t}}$$

This result implies that $a < 1$, which contradicts our original assumption that $a > 1$. Thus, there are no equilibria with t values greater than 1 when $c > 1$.

Finally it will be shown that if $\frac{b+c}{2} < a$, i.e. if the border equilibrium $p = P = 0$ is stable, then there are never any internal equilibria with corresponding $\bar{t}$ values smaller than 1.

Suppose there existed such an equilibrium. Both $\Delta\bar{u}_1$ and $\Delta\bar{u}_2$ should then be equal to 0. But if $\Delta\bar{u}_1 = 0$ and $\Delta\bar{u}_2 = 0$, then $A\Delta\bar{u}_1 + B\Delta\bar{u}_2$ is also equal to 0, where A and B are arbitrary constants. Put $A = a + 1/2\, b(\bar{u}_1+\bar{u}_2)$ and $B = a + 1/2c(\bar{u}_1+\bar{u}_2)$. Then we have:

$$A\Delta\bar{u}_1 + B\Delta\bar{u}_2 = 1/2((b+c-2a)(\bar{u}_1+\bar{u}_2) + 4\bar{u}_1\bar{u}_2 - (\bar{u}_1+\bar{u}_2)(b\bar{u}_1+c\bar{u}_2)) <$$
$$< 1/2(4\bar{u}_1\bar{u}_2 - (\bar{u}_1+\bar{u}_2)(b\bar{u}_1+c\bar{u}_2)) = 1/2\bar{u}_1{}^2(\bar{t}(2-b-c) + 2\bar{t}-b-c\bar{t}^2) <$$
$$< 1/2\bar{u}_1{}^2(2\bar{t} - b - c\bar{t}^2).$$

It is simple to show that this expression is always negative for $\bar{t} < 1$. This leads thus to a contradiction since $A\Delta\bar{u}_1 + B\Delta\bar{u}_2$ should be equal to 0. There can thus be no equilibrium under these conditions for which $\bar{t} < 1$.

The trivial case $\bar{t} = 1$ can be excluded because

$$\bar{t} = 1 \Rightarrow \bar{u}_1 = \bar{u}_2 \Rightarrow \frac{b(1+\bar{t}) - 2a}{b(1+\bar{t}) - 2\bar{t}} = \frac{c(1+\bar{t}) - 2a\bar{t}}{c(1+\bar{t}) - 2} \Rightarrow (a - 1)(b - c) = 0,$$

which contradicts the original assumptions.

Hereby we have shown that the model given by table 5 can have no internal equilibria when the SS equilibrium is unstable, the TT equilibrium is stable and the chromosomal heterozygotes of both sexes have higher fitness than animals with the standard karyotype.

1.8 References

Benirschke, K. ed. 1969. Comparative Mammalian Cytogenetics. Springer-Verlag, Berlin-Heidelberg-New York.

Bodmer, W.F. 1965. Differential fertility in population genetics models. Genetics 51: 411-424.

Bodmer, W.F. and Felsenstein, J. 1967. Linkage and selection: Theoretical analysis of the deterministic two locus random mating model. Genetics 57: 237-265.

Carter, T.C., Lyon, M.F. and Phillips, R.J.S. 1955. Gene-tagged chromosome translocations in eleven stocks of mice. J. Genet. 53: 154-166.

Caspersson, T. and Zech, L. eds. 1973. Chromosome Identification. Technique and applications in biology and medicine. Nobel Symposium 23. Academic Press, New York and London.

Cavalli-Sforza, L.L. and Bodmer, W.F. 1971. The Genetics of Human Populations. W.H. Freeman and Company, San Francisco.

Charlesworth, B. and Charlesworth, D. 1973. Selection of new inversions in multi-locus genetic systems. Genet. Res. 21: 167-183.

Curtis, C.F. and Hill, W.G. 1971. Theoretical studies on the use of translocations for the control of Tsetse flies and other disease vectors. Theor. Pop. Biol. 2: 71-90.

Evans, E.P., Lyon, M.F. and Daglish, M. 1967. A mouse translocation giving a metacentric marker chromosome. Cytogenetics 6: 105-119.

Feldman, M.W. 1972. Selection for linkage modification: I. Random mating populations. Theor. Pop. Biol. 3: 324-346.

Feldman, M.W. and Balkau, B. 1973. Selection for linkage modification: II. A recombination balance for neutral modifiers. Genetics 74: 713-726.

Fisher, R.A. 1922. On the dominance ratio. Proc. Roy. Soc. Edinburgh 42: 321-341.

Fisher, R.A. 1930. The Genetical Theory of Natural Selection. Clarendon Press, Oxford.

Ford, C.E. 1970. The population cytogenetics of other mammalian species. In Human Population Cytogenetics. Ed. by Jacobs, P.A., Price, W.H. and Law, L. University Press, Edinburgh.

Ford, C.E. 1972. Gross genome unbalance in mouse spermatozoa: Does it influence the capacity to fertilize? In The Genetics of the Spermatozoon. Ed. by Beatty, R.A. and Glueckson-Waelsch, S. Edinburgh.

Gustavsson, I. 1969. Cytogenetics, distribution and phenotypic effects of a translocation in Swedish cattle. Hereditas 63: 68-169.

Gustavsson, I. 1971. Culling rates in daughters of sires with a translocation of centric fusion type. Hereditas 67: 65-74.

Haldane, J.B.S. 1941. Selection against heterozygosis in man. Ann. Eugen. 11: 333-340.

Hsu, T.C. and Mead, R.A. 1969. Mechanisms of chromosomal changes in mammalian speciation. In Benirschke, ed. 1969.

Jacobs, P.A. 1974. Correlation between euploid structural chromosome rearrangements and mental subnormality in humans. Nature 249: 164-165.

Jacobs, P.A., Aitken, J., Franckiewicz, A., Law, P., Newton, M.S. and Smith, P.G. 1970. The inheritance of translocations in man: data from families ascertained through a balanced heterozygote. Ann. Hum. Genet. 34: 119-131.

Karlin, S. 1972. Some mathematical models of population genetics. Amer. Math. Monthly 79: 699-739.

Kaufman, M.H. 1972. Non-random segregation during mammalian oogenesis. Nature 238: 465-466.

Kimura, M. 1962. On the probability of fixation of mutant genes in a population. Genetics 47: 713-719.

Kimura, M. 1970. Stochastic processes in population genetics, with special reference to distribution of genefrequencies and probability of gene fixation. In Mathematical Topics in Population Genetics. Ed. by Kojima, K. Springer-Verlag, Berlin-Heidelberg-New York.

Koller, P.C. 1944. Segmental interchange in mice. Genetics 29: 242-263.

Lejeune, J., Dutrillaux, B., Rethoré, M.O. and Prieur, M. 1973. Comparaison de la structure fine des chromatides d'Homo sapiens et de Pan troglodytes. Chromosoma 43: 423-444.

Lewis, K.R. and John, B. 1963. Chromosome marker. J. & A. Churchill Ltd., London.

Lewontin, R.C. and White, M.J.D. 1960. Interaction between inversion polymorphisms of two chromosome pairs in the grasshopper Moraba scurra. Evolution 14: 116-129.

Mandel, S.P.H. 1971. Owen's model of a genetical system with differential viability between sexes. Heredity 26: 49-63.

Mayr, E. 1969. Species, speciation and chromosomes. In Benirschke, K., ed. 1969.

Meylan, A. 1970. Chromosomal polymorphism in mammals. Symp. zool. Soc. Lond. 26: 211-222.

Owen, A.R.G. 1952. A genetical system admitting of two stable equilibria. Nature 170: 1127.

Owen, A.R.G. 1953. A genetical system admitting of two distinct stable equilibria under natural selection. Heredity 7: 97-102.

Parsons, P.A. 1961. The initial progress of new genes with viability differences between sexes with sex linkage. Heredity 16: 103-107.

Purser, A.F. 1966. Increase in heterozygote frequency with differential fertility. Heredity 21: 322-327.

Roberts, P.A. 1970. Screening for X-ray-induced suppressors in Drosophila melanogaster: Prevalence and effectiveness of translocations. Genetics 65: 429-448.

Robertson, A. 1962. Selection for heterozygotes in small populations. Genetics 47: 1291-1300.

Sandler, L. and Novitski, E. 1957. Meiotic drive as an evolutionary force. Amer. Natur. 41: 105-110.

Snell, G.D. 1946. An analysis of translocations in the mouse. Genetics 31: 157-180.

Stene, J. 1970. Statistical inference on segregation ratios for D/G-translocations, when the families are ascertained in different ways. Ann. Hum. Genet. 34: 93-115.

Turner, J.R.G. 1972. Selection and stability in the complex polymorphism of Moraba scurra. Evolution 26: 334-343.

White, M.J.D. 1973. Animal Cytology and Evolution. Third Edition. Cambridge University Press, Cambridge.

White, M.J.D., Lewontin, R.C. and Andrew, L.E. 1963. Cytogenetics of the grasshopper Moraba scurra. VII. Geographic variation of the adaptive properties of inversions. Evolution 17: 147-162.

Whitten, M.J. 1971. Insect control by genetic manipulation of natural populations. Science 171: 682-684.

Wright, S. 1941. On the probability of fixation of reciprocal translocations. Amer. Natur. 75: 513-522.

CHAPTER 2

THE FERTILITY OF HETEROZYGOTES FOR A BALANCED CHROMOSOME MUTATION

2.1 Introduction

Snell (1933, 1935, 1941) described and studied the first ascertained cases of balanced chromosome mutations in mammals. Through radiation he had obtained a number of male mice which gave smaller litters than normal. He called them "semi-sterile" and went on to show that these animals were heterozygotes for reciprocal translocations. This connection between decreased fertility and chromosome mutations has since then been demonstrated in many other instances in mammals (see Chapter 1 for some of the more important studies). In humans, for example, it has been shown that individuals in the general population which are heterozygotes for a balanced translocation have higher rates of fertility failures (stillbirths, abortions) than normal (Jacobs et al. 1970), and also that the frequency of balanced translocation carriers is significantly higher in families with recurrent pregnancy failures than in the general population (Carr 1971). The purpose of the present chapter is to discuss this connection between chromosome mutations and the lower fertility of heterozygotes for these mutations, and to indicate how this decrease in fertility can be estimated in humans.

2.2 The frequency of unbalanced gametes formed by a chromosomal heterozygote

The lower fertility of translocation heterozygotes and other heterozygotes for balanced chromosome mutations, is caused by their production of gametes with an unbalanced haploid genome. In animals as opposed to plants such gametes are functional, but zygotes formed by fertilizations involving

unbalanced gametes will have an abnormal genotype that usually will cause the embryo to abort. The fertility decrease of chromosomal heterozygotes is thus due to zygotic loss.

The frequency of unbalanced gametes formed by a chromosomal heterozygote depends on many factors; first of all on the type of chromosome mutations involved. Reciprocal translocations are, for example, normally associated with a higher frequency of unbalanced gametes than are centric fusions and fissions. Nothing definite can, unfortunately, be said about the production of unbalanced gametes by mammalian inversion heterozygotes of different types, since information is very scarce about such heterozygotes.

But even within the same type of chromosome rearrangement there can be great differences between individual mutations. For example, each of the seven centric fusions that constitute the karyotype difference between the tobacco mouse and the normal mouse has a different tendency to cause the production of unbalanced gametes (Ford and Evans 1973, Cattanach and Mosley 1973). Also genetic factors not directly associated with the chromosome mutation, such as sex, the degree of inbreeding and the general genetic background, can influence the frequency with which unbalanced gametes are formed by a chromosomal heterozygote.

The frequency of unbalanced gametes produced by heterozygotes for a certain chromosome mutation is an interesting biological parameter. However, it is from an evolutionary point of view more important to know what effect this fraction of abnormal gametes has on the fertility of the chromosomal heterozygotes, since this is what determines the expected behaviour of the chromosome mutation in a population. Some writers on the population genetics of chromosome mutations have equated the frequency of unbalanced gametes (or rather the frequency of unbalanced zygotes formed) with the heterozygotes' fertility decrease (see, for example, Wright 1941, or Curtis and Hill 1971), but a factor is then ignored which seems to be of great importance for determining the rate of karyotype evolution in nature, namely

the ability of some chromosomal heterozygotes to "buffer" their fertility against a certain amount of zygotic loss.

2.3 The fertility of chromosomal heterozygotes

The difference between fertility on the one hand, and the fraction of normal zygotes formed by a mating on the other, is well illustrated by an early experiment in genetics performed by Florence Durham in 1911. Durham made crosses between mice which were heterozygous for the gene yellow, which is a recessive lethal, and found that the litters produced by these matings were not much smaller than the litters produced by control matings - despite the fact that a quarter of the zygotes formed by the heterozygote x heterozygote matings should be homozygotes for the gene yellow and therefore die at some stage of development. This is thus a case where there is no direct relationship between the frequency of lethal zygotes formed and the decrease in fertility. Such an insensitivity of an animal's fertility to the genetic constitution of the fertilization products, can be due to many mechanisms, two of which we shall now discuss.

The first possibility, which is only relevant for multiparous animals, is within-litter selection. The disappearance of a fraction of the zygotes or embryos due to their abnormal genotypes, may well give the surviving normal embryos and young a better chance to survive. The competition ability of zygotes with balanced genomes is thus improved when there are fewer surviving zygotes in the litter. Selection favouring individual survival of the balanced fertilization products can occur at all stages of development, from the time before implantation in the uterus to the end of the period of parental expenditure. The loss of some zygotes after a mating due to their genetic unbalance, can then, at least partly, be compensated by an improved chance for the remaining normal zygotes to survive and give rise to adult animals.

The second factor, which is relevant for multiparous animals but also for monotocous animals with more than one oestrus cycle per breeding season, is the rapidity with which the unbalanced zygotes abort. It is important for multiparous animals since an early death of the unbalanced zygotes is advantageous for the survival of the normal zygotes. The later the abortion of the unbalanced zygotes occurs, the shorter time there is for the decrased within-litter selection to have any effects. Also for animals carrying one young at a time, it is important that this embryo abort early if it has an abnormal genotype. If this happens, and a new fertilization rapidly follows, then the total reproductive effort lost by carrying the genetically dead embryo is relatively small. This argument is only applicable to situations, such as in humans, where an abortion can quickly be followed by a new fertilization.

This point leads to the rather surprising result that of two chromosome mutations with similar effects on meiosis, the one which produces the most severe effects in unbalanced fertilization products will be the least disadvantageous mutant. It is, for exampel, possible that from this viewpoint human D/D translocations are less severely selected against than are D/G translocations, since a translocation heterozygote mother of the last type may waste almost one year of her reproductive time carrying a genetically dead child to birth, while mothers with D/D translocations only rarely carry an abnormal offspring for a noticeable length of time.

The conclusions which can be drawn from these considerations on an animal's production of unbalanced gametes and the fertility of the animal, are simple but important. One aspect is the following: A chromosome mutation in one species which causes a certain fraction of unbalanced gametes to be formed, may lead to a strongly reduced fertility of the chromosomal heterozygotes, while an identical mutation in an other species with a different reproductive system may cause only a slight decrease of the fertility of the chromosomal heterozygotes. In the first case only extraordinary circumstances can help the chromosome mutation to spread

while in the second species the probability of fixation of the chromosome mutation due to either drift or association with other evolutionary factors, is much greater. This implies that karyotype evolution will probably be very slow in the first species and more rapid in the second species. It may well be that differences in reproductive systems are more important for the explanation of differences between groups of animals in rates of karyotypic evolution, than are differences in population parameters, such as effective population size and degree of inbreeding. Some of the aspects of different rates of karyotype evolution in different mammalian orders have been discussed by Árnason (1972).

2.4 The selective disadvantage of chromosomal heterozygotes in humans

How is one to know the fertility of a certain type of chromosomal heterozygote? If the decrease in fertility is equal to the fraction of unbalanced gametes formed then one can, at least sometimes, determine this frequency by studying the meiosis of the heterozygote. But it is more or less impossible to determine the effective fertility of a group of animals in nature, especially since the effects can be smaller than 1 % and still be of evolutionary importance, as in all other similar situations. Laboratory studies of the fertility of the animals concerned are clearly unsatisfactory in answering the question about what effective fertility these animals would have in the wild.

The only possibility to obtain valid estimates of the selective effects associated with chromosome mutations is in humans, which offer an almost unique opportunity of studying whole families. However, what can be investigated in humans is not the selective values that one mutation of unique origin has, but the (average) forces that work on chromosome mutations of a certain type. Before discussing how this can be done, we must mention two problems. The first is that the empirical foundation for studying the selective effects of chromosome mutations is very weak but will im-

prove with further population surveys of karyotypes. The second is that we have so far in this chapter given too simplified a picture by assuming that the only selective effects of chromosome mutations are the associated fertility decreases. Chromosome mutations may, however, have deleterious effects besides changing the karyotype. Heterozygotes for chromosome mutations will therefore sometimes have lower viability than chromosomal homozygotes. Such phenotypic differences have been indicated by Jacobs (1974) who has shown that the frequency of carriers of balanced chromosome mutations is higher among the mentally retarded than in the general population. We shall therefore not restrict our attention only to fertility but consider the general fitness of chromosomal heterozygotes.

Let us regard a large random mating population in which chromosome mutations of a certain type are induced with frequency μ. Double chromosomal heterozygotes, or homozygotes for one particular chromosome mutation, are so rare that they can be neglected. To simplify, we define the selective effect of a chromosome mutation as the decrease in fitness of heterozygotes for this mutation. A mutation carried by a heterozygote will be called new if neither of the heterozygote's parents carried the mutation.

In the first case we assume that all chromosome mutations of the specified type have an identical selective effect on the chromosomal heterozygotes, giving them a fitness 1-s instead of the normal fitness, which we put equal to 1. A stable mutation-selection balance, making the frequency of chromosomal heterozygotes equal to $2\frac{\mu}{s}$, will then be established in the population. In every generation a fraction 2μ of the population carries new mutations. The fraction of chromosomal heterozygotes carrying new mutations to the total of chromosomal heterozygotes is thus $2\mu / 2\frac{\mu}{s} = s$. We have here found a method for studying the selection against chromosomal heterozygotes: A number of chromosomal heterozygotes are ascertained in the general population, for example by a newborn survey, and then the fraction of heterozygotes carrying new mutations in this group is

determined by family studies. The ratio thus obtained is an estimate of the selection against the chromosomal heterozygotes.

This method of determining the selection against a dominant factor, corresponds well to what intuition tells us: If a mendelian factor is heavily selected against in the heterozygote, then it will only rarely be transmitted in many generations, and almost all heterozygotes that are found in one particular generation will therefore be carrying new mutations. Thus, if the adverse selection is strong, then the ratio of new mutants among the heterozygotes is high. The argument can then be reversed for mutations that are only weakly selected against.

However, a problem arises when one regards a group of heterogenous mutations, such as for example reciprocal translocations or D/D centric fusions. Here there is no reason to assume that all mutations will have the same selective effect. We shall therefore now, in this more realistic situation, see what the ratio of chromosomal heterozygotes carrying new mutations to the total of chromosomal heterozygotes estimates:

We are still studying a large random mating population, where chromosome mutations of the relevant type are induced with frequency μ. The normal fitness in the population is 1, but all heterozygotes for chromosome mutations have lower fitness. A distinction will be made between newly introduced mutations and mutations found in the population; the second group includes mutations transmitted by carriers from earlier generations. Chromosome mutations found in the population thus represent a biased sample of newly introduced chromosome mutations.

Let the distribution of the selective effects of newly introduced mutations be given by f(s), where $0 < s < 1$, and $\int_0^1 f(s)\,ds = 1$. Then, the (arithmetic) mean effect of newly introduced mutations is $\int_0^1 sf(s)\,ds$.

The mutation-selection balance in the population is now such, that the frequency of chromosomal heterozygotes is $2\int_0^1 \frac{\mu f(s)}{s}\,ds$. The mean

selective effect of the chromosome mutations in the population is therefore $\int_0^1 \left[t \cdot \frac{\mu f(t)}{t} \cdot \left(1 \Big/ \int_0^1 \frac{\mu f(s)}{s}\, ds \right) \right] dt = \left(\int_0^1 \frac{f(s)}{s}\, ds \right)^{-1}$. We see that the mean selective effect of chromosome mutations in the population is different, as expected, from the mean selective effect of newly introduced chromosome mutations.

If a large random sample is taken from the population and a number of chromosomal heterozygotes are found by karyotyping, then, in correspondence with the earlier situation, a fraction $2\mu \Big/ 2\int_0^1 \frac{\mu f(s)}{s}\, ds = \left(\int_0^1 \frac{f(s)}{s}\, ds \right)^{-1}$ of these will be heterozygotes carrying new mutations. This is the value we are interested in. It can be given two different interpretations:

The first is simply that it estimates the mean selective effect of the chromosome mutations in the population. This result corresponds well with the earlier situation.

The second is that the value estimates a weighted average of the selective effects of newly introduced chromosome mutations where every mutation with effect s is given weight $\left(s \cdot \int_0^1 \frac{f(s)}{s}\, ds \right)^{-1}$. Chromosome mutations with small effects are thus of greater importance in determining the average than mutations with strong effects.

That the ratio studied gives an estimate of a weighted average of the effects of newly introduced chromosome mutations, follows from the fact that the mutation-selection balance in the population is determined by the average effect of the chromosome mutations in the populations, where chromosome mutations with smaller effects are over-represented relative to chromosome mutations with greater effects.

Applying the described method on the data given in table 1 (from a review of the relevant literature by Jacobs et al. 1974) we get the results that the estimated fitness effects of chromosomal heterozygotes for D/D translocations is $1/16 = 0.06$, for D/G translocations is $1/6 = 0.17$, and for reciprocal translocations is $8/24 = 0.33$. These estimates are based on

Type of translocation:	Total No. studied:	One parent affected:	Both parents normal:
D/D	16	15	1
D/G	6	5	1
Reciprocal translocation	24	16	8

Table 1. Chromosome constitutions of parents to translocation heterozygote babies found in surveys of consecutive liveborn babies. Data taken from Appendix Table 14 in Jacobs et al. 1974. The total number of babies examined in these surveys is 43 558.

only a few individuals, but table 1 shows the effort which was necessary to collect the data.

Of particular interest is the fact that heterozygotes for reciprocal translocations have an estimated fitness of 0.67. This implies that there probably exists a group of translocation carriers with much greater fitness than 0.50, which is the greatest value translocation heterozygotes would have if they, as is commonly assumed, had only half of the normal fertility. A more careful analysis of the estimated selection values, for example of the existence of any selective differences between D/D and D/G translocation carriers, will have to wait until more information is available.

2.5 References

Árnason, Ú. 1972. The role of chromosomal rearrangement in mammalian speciation with special reference to Cetacea and Pinnipedia. Hereditas 70: 113-118.

Carr, D.H. 1971. Chromosomes and abortion. In Advances in Human Genetics 2. Ed. by Harris, H. and Hirschhorn, K. Plenum Press, New York - London.

Cattanach, B.M. and Moseley, H. 1973. Nondisjunction and reduced fertility caused by the tobacco mouse metacentric chromosomes. Cytogenet. Cell Genet. 12: 264-287.

Curtis, C.F. and Hill, W.G. 1971. Theoretical studies on the use of translocations for the control of Tsetse flies and other disease vectors. Theor. Pop. Biol. 2: 71-91.

Durham, F. 1911. Further experiments on the inheritance of coat colour in mice. J. Genet. 1: 159-178.

Ford, C.E. and Evans, E.P. 1973. Robertsonian translocations in mice: Segregation irregularities in male heterozygotes and zygotic unbalance. In Chromosomes Today, Volume Four. Ed. Wahrman, J. and Lewis, K.R. John Wiley & Sons, Chichester.

Jacobs, P.A. 1974. Correlation between euploid structural chromosome rearrangements and mental subnormality in humans. Nature 249: 164-165.

Jacobs, P.A., Aitken, J., Franckiewicz, A., Law, P., Newton, M.S. and Smith, P.G. 1970. The inheritance of translocations in man: Data from families ascertained through a balanced heterozygotes. Ann. Hum. Genet. 34: 119-131.

Jacobs, P.A., Melville, M., Ratcliffe, S., Keay, A.J. and Syme, J. 1974. A cytogenetic survey of 11,680 newborn infants. Ann. Hum. Genet. 37: 359-376.

Snell, G.D. 1933. Genetic changes induced in mice by X-rays. Amer. Natur. 67: 81.

Snell, G.D. 1935. The induction by X-rays of hereditary changes in mice. Genetics 20: 545-567.

Snell, G.D. 1941. Linkage studies with induced translocations in mice. Genetics 26: 169.

Wright, S. 1941. On the probability of fixation of reciprocal translocations. Amer. Natur. 75: 513-522.

CHAPTER 3
CHROMOSOME MUTATIONS, GENE-FLOW AND SPECIATION

3.1 Introduction

Some writers on karyotype evolution in animals, for example White (1969) and Todd (1970), stress the role that chromosome mutations play in the speciation process. Allopatric speciation, i.e. when one interbreeding gene pool splits into two independent, geographically isolated gene pools which then diverge genetically, can easily be explained without any reference to chromosome mutations. If, however, one wants to imagine a process of speciation that is not completely allopatric, then an explanation must be given as to why contact between the populations representing an old and an incipient species does not lead to destruction of the genetic differences between the populations by a flow of genes into the incipient species. This is where differences in karyotypes between the two populations have been considered important. Differences in karyotypes must be the result of fixation of separate chromosome mutations in the two populations, and since heterozygotes for most chromosome mutations have a reduced fertility (Chapter 1), this implies that matings of animals from the two populations will result in offspring with lower fertility than normal. The effective flow of genes into the incipient species in a situation like this, is thus lower than the rate of interbreeding would suggest. It has been proposed that with such a barrier to gene flow, evolution would proceed independently in the two populations. Other factors could then appear which finally would prevent all inter-population mating completely. In this way, a new species would have been created.

This argument is expressed differently by different authors. We shall here only quote one example:

> "If an alteration of this nature /a considerable structural alteration in a particular chromosome/ can persist and spread in a population it may play a large part in making the chromosomes of the population structurally incompatible with those of another, so that if the two populations are able to increase their ranges until they meet, they mey have the greatest difficulty in producing hybrids. It can be expected, then, that isolated populations may gradually diverge through differential incidence of mutations." (Cain 1971, pp. 143-144.)

There are two difficulties involved in this type of explanation of how genetic divergence can come about between populations which interbreed for at least some time during the speciation process. The first relates to how the populations become different in karyotypes; the second arises from the fact that a single difference in karyotypes between two populations acts as a poor barrier to gene-flow between them. We will discuss these problems, and finally briefly indicate a different view of the relationship between chromosome mutations and speciation.

3.2 The creation of the difference in karyotypes

Nothing is normally said by the authors using the argument outlined above about how the karyotypes of the two populations have come to be different. It is as if it is normal that chromosome mutations, for which the heterozygotes are less fertile, occasionally become fixed in most populations. But how such chromosome mutations ever become fixed in any population, is an intriguing evolutionary problem, discussed extensively in Chapter 1, because these chromosome mutations function as dominant deleterious genes when they are introduced in a population. Two different kinds of explanations can, as discussed before, be given for the increase: Either the chromosome mutation not only has a fertility disadvantage but also an evolutionary advantage, for example through its linking two favourable genes closer together or being favoured by segregation distortion. In this case

one would expect the increase in the chromosome mutation to happen, not only in one population, but also in the other population, once contact has been established between them. If, however, the chromosome mutation is favoured in both populations, then it cannot function as a gene-flow barrier.

The other possibility is that the chromosome mutation has become fixed in one population (probably the incipient species) by random drift. This can only happen if either the selective disadvantage of the chromosomal heterozygotes is very small, or the effective population size is very small (Wright 1941, Chapter 1). Since we are interested in the cases where the heterozygote fertility decrease is appreciable (this disadvantage is the cause of the decrease in gene-flow between the populations), we are left only with those situations where the chromosome mutation has become fixed due to the very small effective population size of the group of animals which constitutes the incipient species. The scheme of speciation presented in the introduction to this chapter would then rely on evolution due to founder effects, something which probably is not intended by many of the proponents of the argument.

The problem of increase of chromosome mutations can be overcome if very different selective regimes are assumed to rule in the two populations. But if strong niche effects are incorporated into the model, then very little remain of the original argument which puts the evolutionary emphasis on selection against hybrid animals resulting from inter-population matings.

3.3 The effect of a karyotype difference on the gene-flow between two populations

Let us now disregard the problem of how the two populations came to differ in karyotype and instead examine the effect that this has on the gene-flow between the populations. We shall here show by the use of a simple mathematical model that a single chromosome mutation difference between two populations has a very moderate effect on the gene flow between them, even if the chromosomal heterozygotes have a low fertility. There will, of course, be a stronger effect if the populations differ by many chromosome mutations, but then we are back to the problem discussed above. It may be possible to understand how the populations have come to differ by one chromosome mutation, but it is very unlikely that they will differ by many mutations unless the populations have been separated for a long time. There then would probably have appeared behaviour and other differences between the animals which would make inter-population matings much less frequent.

For the investigation of the effect of a karyotype difference between two populations on the gene-flow between them we shall assume a simple model of two populations, M and R. We interprete M as the main population and R as the incipient species. M and R have identical origins, but they have been separated for some time, for example due to a geographic change. During the time of separation, a chromosome mutation has become fixed in R making all animals there of karyotype TT, while the animals in M have the standard karyotype SS. Both SS and TT animals have normal fertility, while hybrids between them are chromosomal heterozygotes, ST, with fertility reduced to 1-s of the normal. If T is a translocation or an inversion, then s is a positive number unlikely to be much greater than 1/2.

The following system will be used to estimate the gene-flow that occurs after the two populations have come in contact with each other again:

Suppose that the populations differ not only in karyotypes, but also at a gene locus X (not linked to any of T's breakpoints), so that all animals in population M are homozygotes for an allele $\underline{x}_m$, while this allele is absent in R. All alleles at the locus X are assumed to be selectively neutral. The rate with which the allele $\underline{x}_m$ increases in frequency in R after contact between R and M is re-established, will be our measure of how well population R is "buffered" by the difference in karyotype, against inflow of genes from M.

The re-established contact between M and R is interpreted in the model as migration of animals from M into R, starting in generation t. It is assumed that the frequency of animals in R that were born in M, in every generation after t constitutes 1/2 m of the size of R. There is random mating in R and successive generations are thought of as separate. Assume m sufficiently small so that terms of order m^2 can be disregarded. Nothing explicit is said about the migration of animals from R to M, except that it is so small that it does not significantly change the genetic constitution of either M or R. The frequency of the allele $\underline{x}_m$ in R in generation T, will be denoted by f(T).

We are now ready to study the effect of migration on the incipient species R, by determining what happens to the allele $\underline{x}_m$ which flows into R, carried by the migrating animals. If the first wave of migration from M to R occurs before mating in generation t, then f(t+1), i.e. the frequency of allele $\underline{x}_m$ at birth of generation t+1, will be m. All the $\underline{x}_m$ alleles exist in hybrid animals which are chromosomal heterozygotes. If the migration in generation t+1 is the same as in generation t, then at the birth of generation t+2 we find that $f(t+2) = m + (1-s)m$, where in the second term we have taken into account the lower fertility of the chromosomal heterozygotes. The X locus and the breakpoints of T segregate independently, so half of the "second generation hybrids" carrying the $\underline{x}_m$ allele are heterozygotes for the chromosome mutation, while the other half has chromosome complement TT, which is the normal in R. Therefore, we

find that $f(t+3) = m + (1-s)m + 1/2(1-s)^2 m$, which is equal to $f(t+2) +$ $+ 1/2(1-s)m + 1/2(1-s)^2 m$. Going even one generation further we get $f(t+4) = f(t+3) + 1/2(1-s)m + 1/4(1-s)^2 m + 1/4(1-s)^3 m$.

If migration continues in the same way, then the frequency in generation T becomes $f(T-1) + \sum_{k=1}^{T-t-2} (1/2(1-s))^k m + 2 \cdot (1/2(1-s))^{T-t-1} m$, from which follows that the change in frequency of the $\underline{x}_m$ allele between the generations T-1 and T can be approximated by

$$f(T) - f(T-1) \sim \frac{1-s}{1+s} m .$$

This is the rate at which, at least in the beginning, the allele $\underline{x}_m$ increases every generation in R, under the assumption that migration in and out of R is small. If this is not the case or if migration goes on for a long time, then the proportion of animals in R with the SS karyotype cannot be disregarded and the rate of increase of $\underline{x}_m$ becomes greater than the estimate given by the formula.

If there were no difference in karyotype between animals in R and M, then the allele $\underline{x}_m$ would steadily increase every generation in R with the rate m. The karyotype difference has thus caused a decrease in gene-flow, as measured by our system, to a fraction 1-s / 1+s of what it otherwise would be. This value is, for example, 0.33 when s is 0.50, and 0.74 when s is 0.15 (which is a reasonable estimate of the fertility decrease in heterozygotes for centric fusions in mammals).

In a similar way, one can show that if there are many karyotype differences between the two populations and if the X-locus is unlinked to the breakpoints of the chromosome mutations, then the flow of the allele $\underline{x}_m$ into population R will be decreased to $\prod_i \frac{1-s_i}{1+s_i}$ of the value if would be if there were no differences in karyotype between animals in R and M. s_i is the fertility decrease in animals which are heterozygotes for the chromosome mutation T_i. A difference between two populations of three centric fusions, causing a fertility decrease in heterozygotes of res-

pectively 15 %, 10 % and 5 %, will thus decrease the effective flow of genes between the populations to $\frac{0.85}{1.15} \cdot \frac{0.90}{1.10} \cdot \frac{0.95}{1.05} = 55\ \%$ of what it would be if there were no differences in karyotypes.

This is thus the second problem of the argument about chromosome mutations and speciation that was given in the introduction: One, or even a few, chromosome mutations function poorly as a barrier against inflowing genes which are not linked to the chromosome mutation(s).

To overcome this problem, one must either postulate the existence of many karyotype differences between the two populations which would involve all the problems discussed above, or find a reason as to why a gene-flow decrease of perhaps much less than one order of magnitude can be of major importance for speciation. The point of the argument under discussion is that speciation will not occur in the given biological situation, unless the effective gene-flow into the incipient species is less than would be expected from the frequency of inter-population matings. It is hard to see that the modest change in effective gene-flow which follows from one - or even a few - karyotype differences, will dramatically change the chances of successful speciation in such a situation.

Other authors, for example Maynard Smith (1966), discussing different models of partial sympatric speciation have reached the same conclusion, namely, that genetic factors which cause a decrease of heterozygotes' fitness are not very good barriers to gene-flow. It is, however, important to observe that this conclusion does not imply that heterozygote disadvantageous factors will easily spread between populations with migration. An illustration to this is found in the model just described, where the frequency of chromosomal heterozygotes in population R rapidly will become $2 \cdot \frac{m}{s}$ and will remain at this value. (The equilibrium is found by regarding the situation as a mutation-selection balance, with migration taking the place of mutation.) This frequency of chromosomal heterozygotes in R is small if, as assumed, migration is small and the effect of the chromosome mutation on the fitness of heterozygotes is appreciable.

There is a possibility that in a real situation one may be misled by a low frequency of chromosomal heterozygotes in a population, to conclude that the inflow of genes from a chromosomally distinct, neighbouring population is negligible. This is however, as we have seen, sometimes false. A narrow zone of hybridization between two chromosomally distinct populations, judged from the low frequency with which chromosomal heterozygotes are found, can never prove that the populations are genetically isolated from each other. This consideration must be remembered, for example when one wants to explain the found genetic homogeneity between chromosomally different populations of burrowing rodents. (For a review of the relevant data, see Nevo et al. 1974.)

3.4 Chromosome mutations and speciation

In the introduction we described a frequently used argument which links karyotype change to the process of speciation, and then tried to show that this argument is not very convincing. In a paper discussing the same question, Wallace (1959) presented some slightly different ideas about the connection between chromosome mutations and speciation. Describing how a new species can be created at or just outside the boundary of an already existing species, Wallace, like the proponents of the argument discussed above, stressed the role that chromosome mutations may play in the birth process of a new species. He did, however, make two important points in his discussion:

The first was that chromosome mutations can create blocks of genes which are favourable at the boundary of a species, and then protect these from being dispersed by crossovers with other gene complexes adapted to the ecological situation in the centre of the species. The chromosome mutations are thus thought to protect the local population from inflowing gene complexes from other parts of the species' territory. But here, unlike the case in the argument discussed above, chromosome mutations

are considered to protect only the genes linked to the mutations. A model of how the linking of two genes together may cause the spread of a chromosome mutation, has been discussed in Chapter 1.

The second point is simply that every animal has, effectively, zero fertility at the species boundary and just beyond it. There, it does not matter if an animal has lower than normal fertility. The models used, for example in Chapter 1, to discuss the increase of chromosome mutations based on the fertility disadvantage of chromosomal heterozygotes, are therefore not valid in these situations. The important question close to the species boundary is whether an animal is able to survive, mate and have offspring which can be protected and fed until the end of the period of parental expenditure. A lower number of offspring may even in such situations be of advantage.

The biological situation favourable for the creation of a new species, may thus, as we have seen, also be favourable for the spreading of chromosome mutations. This is, in some way, the initial argument turned around: Instead of explaining speciation with the help of chromosome mutations, it may be more fruitful to regard situations of speciation as favourable for karyotypic change.

It is very likely that only some speciation situations, typical for certain animals, give possibilities for chromosome mutations to spread, while different speciation situations, common in other animals, may not favour karyotypic change. We have here found that not only differences in reproductive biology, as discussed in Chapter 2, but also differences in the way new speices are created, can cause separate rates of karyotype evolution in different groups of animals.

3.5 References

Cain, A.J. 1971. Animal Species and Their Evolution. Hutchinson & Co. London.

Maynard Smith, J. 1966. Sympatric speciation. Amer. Natur. 100: 637-650.

Nevo, E., Kim, Y.J., Shaw, C.R. and Thaeler, Jr., C.S. 1974. Genetic variation, selection and speciation in Thomomys talpoides pocket gephers. Evolution 28: 1-23.

Todd, N.B. 1970. Karyotypic fissioning and canid phylogeny. J. theor. Biol. 26: 445-480.

Wallace, B. 1959. The influences of genetic systems on geographical distribution. Cold Spring Harbor Symp. Qant. Biol. 24: 193-204.

White, M.J.D. 1969. Chromosomal rearrangements and speciation in animals. Ann. Rev. Genet. 3: 75-98.

Wright, S. 1941. On the probability of fixation of reciprocal translocations. Amer. Natur. 75: 513-522.

CHAPTER 4

MAMMALIAN CHROMOSOMES SIMILAR IN LENGTH ARE ALSO SIMILAR IN SHAPE

4.1 Introduction

If the autosomal chromosomes of the human haploid genome are ordered by their length, one will find that chromosomes of similar shapes will occur next to each other. Chromosomes with no arms, or with only a very small short arm, take for example the places 13-14-15 and 21-22. The haploid autosomal karyotype can thus be written 22222222222211122222211, where the twos stand for bi-armed chromosomes and the ones for acrocentric chromosomes. The number of runs in this row of ones and twos is only four, if by a run is meant an uninterupted succession of identical symbols. It thus looks as if the one- and bi-armed chromosomes were "clumped together" in the haploid karyotype.

This idea can be tested statistically by using a table published by Swed and Eisenhart (1943) (reprinted in Siegel), which for every pair of integers s_1 and s_2 gives the number of runs r_o, such that if s_1 and s_2 elements were mixed at random many times, then in only 2.5 percent of the cases this should lead to a number of runs smaller than or equal to r_o. The critical number for $s_1=5$ and $s_2=17$ is $r_o=6$, which means that if the one-armed and the bi-armed chromosomes of the human autosomal haploid karyotype were ordered randomly, then the probability that this would result in a row with only 6 or fewer runs, is smaller than 2.5 %. Since the observed number of runs is 4, it is reasonable to suggest that one- and bi-armed chromosomes are not in a random mixture in the human haploid autosomal karyotype ordered by length, but that chromosomes similar in length are also similar in shape. In this chapter we will show that this finding can be extended to many other mammalian species, i.e. if the chromosomes of

the haploid autosomal genotype of a species are ordered according to length, then frequently the one- and bi-armed chromosomes are spread in the row in a way which suggests a non-random mixture of the two types of chromosomes. A brief discussion of how the observed situation can be explained is given in the last section of the chapter.

4.2 Method

The 300 karyotypes published in the first six volumes of "An Atlas of Mammalian Chromosomes" (Hsu and Benirschke, 1967, 68, 69, 70, 71a and 71b) were used for the investigation. The number of acrocentric and telocentric chromosomes (which we call one-armed chromosomes) in every haploid autosomal karyotype was denoted by s_1, and the number of other chromosomes (the bi-armed) was denoted by s_2. Most chromosomes fall naturally into one of these subgroups; in cases of doubt, the published descriptions of the chromosomes were strictly adhered to.

The number of runs of similar chromosomes in the haploid autosomal karyotype ordered by length, was then recorded and denoted by **r**. It was sometimes only possible to get an upper estimate of **r**. By using the table mentioned above, we could then select the karyotypes with such a low number of runs that would only be the outcome in 2.5 % or less of the cases if s_1 and s_2 one- and bi-armed chromosomes were mixed at random. The selected karyotypes are given in table 1 with their corresponding s_1, s_2 and **r** values. Some of the selected karyotypes are identical to karyotypes of other species from the same order, as indicated in table 1. Altogether 34 karyotypes were selected by the described procedure, to which correspond 27 different karyotype patterns if the similarity between some karyotypes are taken into account.

In table 2 we give the number of karyotypes from different orders which so far have been published in the Atlas, and how the selected karyotypes are distributed between the orders.

ORDER	SPECIES (Number of folio in the Chromosome Atlas)	No	Karyotype identical to	s_1	s_2	r
Insectivora	Sorex caecutiens (205)	1		10	10	4
Chiroptera	Myotis velifer incautus (3)	2		17	4	4
	Myotis grisescens (57)	3	2			
	Lasiurus intermedius (158)	4		3	9	2
	Nycticeius humeralis (210)	5		20	2	2
Lagomorpha	Lepus alleni alleni (6)	6		2	21	2
	Lepus californicus eremicus (7)	7	6			
Rodentia	Perognathus penicillatus (58)	8		16	6	4
	Perognathus spinatus (165)	9		15	6	2
	Mus poschiavinus (177)	10		5	7	2
	Acomys cahirinus (276)	11		2	16	2
	Capromys pilorides (282)	12		6	13	4
Carnivora	Genetta genetta neumanni (126)	13		2	23	2
	Nyctereutes procyonoides (230)	14		7	13	4
Perrosodactyla	Equus zebra hartmannae (37)	15		2	23	2
	Equus grevyi (290)	16		6	16	4
Artodactyla	Ovis aries (45)	17		23	3	2
	Antidorcas marsupialis (89)	18		25	2	2
	Oryx gazella (94)	19	18			
	Tayassu tajacu (132)	20		6	8	2
	Bubalus bubalis (139)+	21		18	5	3
	Hemitragus jemlahicus (141)	22		17	6	2
	Kobus kob (143)	23		19	5	2
	Ovis ammon nigrimontana (144)	24	18			
	Syncerus caffer caffer (145)	25		21	4	2
	Bubalus bubalis (189)+	26	23			
	Oreamnos americanus (190)	27		11	9	2
	Damaliscus dorcas (293)	28		7	11	2
	Ovis orientalis (294)	29	17			
Primates	Callithrix jacchus (46)	30		7	15	6
	Hylobates concolor leucogenys (149)	31		3	22	2
	Callithrix humeralifer (196)	32		5	16	3
	Saguinus oedipus (198)	33	30			
	Lemur fulvus fulvus (244)	34		15	8	2

+ Intra-species chromosome polymorphism

Table 1. Species represented in Hsu and Benirschke (1967, 68, 69, 70, 71a and 71b) where one-armed and bi-armed chromosomes are non-randomly distributed in the karyotype. See the text for the method used to select the karyotypes and for the definitions of s_1, s_2 and r.

ORDER	A	B	C
Marsupialia	16	0	0
Insectivora	9	1	1
Chiroptera	23	4	3
Edentata	5	0	0
Lagomorpha	8	2	1
Rodentia	93	5	5
Cetacea	1	0	0
Carnivora	53	2	2
Pinnipedia	1	0	0
Tubulidentata	1	0	0
Proboscidea	2	0	0
Hyracoidea	1	0	0
Perrosodactyla	7	2	2
Artiodactyla	48	13	9
Primates	32	5	4
SUM	300	34	27

Table 2. The distribution over the mammalian orders of A, the number of karyotypes published in Hsu and Benirschke (1967, 68, 69, 70, 71a and 71b) and B, the number of karyotypes selected as being non-randomly organized of one-armed and bi-armed chromosomes. C gives the number of different karyotype patterns selected from each order, taking into account the similarity between some species' karyotypes.

If the chromosomes were randomly distributed in all the karyotypes, one would expect to select $0.025 \cdot 300 = 7.5$ karyotypes from the material by the method of selection used. The number found, 34, is of course significantly higher than the expected value. So is also 27, which takes into account the identity between some of the karyotypes.

Most of the selected karyotypes come from species of the order Artiodactyla, as shown in table 2, but even if this order is excluded from the data, the number of selected karyotypes, 21 resp. 18, is still much higher than the number expected from random mixing, $(0.025 \cdot 252 =)$ 6.3.

4.3 Results

The method used to select these karyotypes where the chromosomes similar in length are also similar in shape, is very crude. Only karyotypes with the number of runs certainly below the critical value, r_o, have been included in table 1. Many karyotypes have therefore been excluded since the published pictures do not give clear information about the length relationships between certain chromosomes. This is, for example, the case with Homo sapiens, whose karyotype shows significant clumping of the chromosome types when the chromosomes are ordered according to length (Paris Conference, 1971), as discussed in the introduction. Homo is, however, not included in table 1, since the karyotype published in the Atlas is not good enough to certify that the arrangement of chromosomes used in the Atlas, actually corresponds to an ordering strictly according to length.

Much interesting information is also excluded when only the distinction between one- and bi-armed chromosomes is made. A more careful investigation taking into account the exact position of the centromere on every chromosome would probably detect more non-randomly organized karyotypes, than has been found by using the simple dichotomy. Such an investigation is however difficult to do.

Some karyotypes have similar chromosomes next to each other in the ordered row, but have not been selected since the total chromosome number is low. This is, for example, the explanation to why there is no karyotype from Marsupiala represented in table 1. A test taking into account the general tendency of the whole material would therefore give much stronger evidence for non-random organization of the chromosomes in mammalian karyotypes than what is possible with the method used here.

Thus, it is clear that the present method has severely underestimated the number of karyotypes where chromosomes of similar lengths also have similar shapes. That a high number of karyotypes nevertheless has been found, relative to the number expected from the selection procedure if chromosomes were randomly mixed in all karyotypes, therefore indicates that the phenomenon is very common in mammalian karyotypes.

4.4 Discussion

How is this non-random organization of the one-armed and bi-armed chromosomes in mammalian karyotypes to be explained? The first mechanism which may give a non-random appearance to a karyotype is tetraploidy. In a tetraploid every chromosome in the autosomal haploid genome has a partner chromosome of identical length and shape. Commings (1972), following an idea of Ohno's (1970), has suggested that the banding patterns of the human chromosomes carry evidence of ancient tetraploidy. It is, however, very unlikely that traces of polyploidy can be detected by the crude test used here, particularly since that polyploidization must have occured in the early phases of mammalian evolution and all karyotypes have changed and evolved in different directions afterwards. A brief look at the selected karyotypes is also enough to tell that this mechanism cannot be the main, nor even a significant cause of the non-randomness found. That partial tetraploidy (multiple tetrasomy)

should be the explanation to the karyotypic regularity found, can be excluded by the same arguments.

A second type of explanation is the following: If a genome consists of a number of chromosome arms that randomly get involved in centric fusions (and fissions), then most of the longest chromosomes will be bi-armed and most of the shortest will be one-armed. The test used here may detect this as a similarity in shape of chromosomes with similar length. Almost all of the 27 different karyotypes represented in table 1 have, in accordance with this explanation, a bi-armed chromosome as the longest autosome and a one-armed chromosome as the shortest autosome. Centric fusions of chromosome arms taken at random from the karyotype may then, at least partly, explain the ordered appearance of many mammalian karyotypes. Other types of chromosome mutations such as reciprocal translocations, pericentric inversions and quantitative changes in the amount of heterochromatin, will tend to destroy the relationship between length and shape of chromosomes caused by centric fusions. The regularity of mammalian karyotypes described in this chapter, can thus be interpreted as a sign of the importance of centric fusions in the evolution of many mammalian karyotypes, relative to other mechanisms involved in the process.

However, there are some indications that random fusions of chromosome arms, which is a rather trivial cause of the reported ordered appearances of mammalian karyotypes, cannot be the sole explanation of this finding. First of all, there exist karyotypes whose regularity hardly can have been created by centric fusions. This is the case, for example, with many karyotypes in the genus Rattus, where most of the smallest autosomes are bi-armed and most of the longest are one-armed (Yosida 1973). Secondly, one can show that in a number of karyotypes, the arms of the bi-armed chromosomes are longer than the arms of the one-armed chromosomes. (The Mann-Whitney U-test can be used for this purpose, see Siegel.) Among the karyotypes represented in table 1, this is the

case, for example, in Nycticeius humeralis, Perognathus spinatus, Acomys cahirinus, Ovis aries and Lemur fulvus fulvus. The correspondence in length of chromosome arms of bi-armed chromosomes in some karyotypes, makes it less likely that centric fusions of randomly chosen chromosome arms should have produced the ordered appearances between one- and bi-armed chromosomes.

The alternative explanation to this regularity is that selection is involved. Chromosome mutations which help to give a well-ordered organization of the karyotype may be favoured relative to chromosome mutations creating chromosomes of shapes and sizes different from those of the already existing chromosomes. The biological basis for such a selection may, for example, involve the stability of the spindle apparatus at meiosis or mitosis.

From the data presented here one cannot exactly determine to what extent the described regularity of mammalian karyotypes is due to selective forces working more or less directly on this quality and to what extent it is a product of unrelated evolutionary processes. With more data on karyotypes and with better statistical methods, it should, however, be possible to further develop the new approach to karyotype research which we have introduced in this chapter, to successfully investigate this and other similar problems.

4.5 References

Comings, D.E. 1972. Evidence for ancient tetraploidy and conservation of linkage groups in mammalian chromosomes. Nature 238:455-457.

Hsu, T.C. and Benirschke, K. 1967, 68, 69, 70, 71a and 71b. An atlas of mammalian chromosomes. Volumes 1, 2, 3, 4, 5 and 6. Springer-Verlag, Berlin-Heidelberg-New York.

Ohno, S. 1970. Evolution by gene duplication. Springer-Verlag, Berlin-Heidelberg-New York.

Paris Conference (1971). 1972. Standardization in human cytogenetics. The National Foundation, Vol. 8, No. 7.

Siegel, S. Nonparametric statistics for the behavioral sciences. McGraw-Hill Kogakusha Ltd. No date or place of printing given.

Swed, F.S. and Eisenhart, C. 1943. Tables for testing randomness of grouping in a sequence of alternatives. Ann. Math. Statist. 14: 66-87.

Yosida, T.H. 1973. Evolution of karyotypes and differentiation in 13 Rattus species. Chromosoma 40: 285-297.

CHAPTER 5

KARYOTYPIC AND GENETIC ANALYSIS OF HUMAN INTRASPECIFIC SOMATIC CELL HYBRIDS

In the second part of this thesis we turn our attention to karyotype evolution in cell lines. Two recent findings are of great importance for the study of karyotype variability in vitro. The first is the development of new chromosome staining techniques with which one can analyse the chromosomes of individual cells in greatest detail. The second is the method of making somatic cell hybrids. From parent cells with known karyotypes one can now produce cell lines with "artificial" initial karyotypes. Through following the evolution of these karyotypes, information can be obtained about the selective forces which work on karyotypes.

In the present chapter we shall describe how it is possible to follow the karyotype evolution of human hybrid cell lines with chromosomal and genetical markers. Although the investigation was made over a long time, the hybrid lines studied never reached a stable state in respect to the karyotype evolution as most cell lines do, given enough time. At such a stable state there is both karyotype variability within the cell line and a population stability so that the "average karyotype" of the cell line remains the same over a long time. An attempt to analyse such situations mathematically is given in chapter 6, where we outline a model of karyotype selection which leads the cell population towards a stable state.

5.1 Introduction

Genetic analysis with somatic cell hybrids has so far been based predominantly on the use of interspecific hybrids following the discovery of extensive preferential loss of human chromosomes from human-mouse hybrids (1) and making use of species differences as a source of genetic markers (see for example 2, 3). Most studies, so far, using human

intraspecific hybrid cell lines and heterokaryonts, have been concerned with the biochemical status of cells derived from fusions between two mutant cells, or between a mutant and a normal cell (4, 5, 6, 7, 8). Human intraspecific hybrid lines have not been used for genetic analysis due to 1) the much lower rate of chromosome loss in most such hybrid lines (4, 9, 10, 11) relative to the loss in interspecific hybrids, and 2) the more limited availability of genetic marker differences. That considerable chromosome loss does however eventually occur in intraspecific hybrids during prolonged growth was reported by Engel et al. (12) for the mouse, and such segregation has been used to study malignancy (13, 14, 15, 16) and the expression of differentiated functions (17, 18).

We describe here the behaviour of human intraspecific hybrids between the 8-azaguanine resistant HeLa derivate, D98/AH-2, and PHA stimulated lymphocytes from a normal male. Special emphasis is placed on establishing the range of markers which can be studied in these hybrids and on characterizing the karyotypic behaviour of the hybrid lines.

5.2 Materials and Methods

Cells. D98/AH-2 is an 8-azaguanine (8AZG), and 6-thioguanine (6TG) resistant human cell line (19) now presumed to be of HeLa origin (20), which is deficient in the enzyme hypoxanthine-guanine phosphoribosyltransferase (HGPRT) and is therefore killed by HAT medium (21). Lymphocytes were obtained from peripheral blood taken from a normal male, MAR, and separated under sterile conditions on a Ficoll-Triosil gradient using standard techniques (22). The lymphocytes were cultured with PHA for five days before use for fusion. A careful check on the purity of the cell population was not made, but the frequency of cells other than lymphocytes is presumed to have been low.

Fusion. Most of the hybrids described were derived from a fusion done as described by Pontecorvo (23).

The day before the hybridization 1.5×10^6 D98/AH-2 cells were added to each of two petri dishes. Immediately before fusion the attached cells were washed using medium without serum, and cooled for 3 minutes at $4^{o}C$. Inactivated sendai virus (880 or 480 HAU/plate) was then added and incubation continued for 10 minutes at $4^{o}C$ after which the excess virus was washed away. 5×10^5 washed and resuspended lymphocytes were then added to each plate, which was incubated first for 15 minutes at $4^{o}C$ and then 20 minutes at $37^{o}C$. Finally unattached lymphocytes were washed off. One day later the medium was changed to selective HAT medium (21) containing 10^{-4}M hypoxanthine, 1.6×10^{-5}M thymidine and 5×10^{-5}M amethopterin. Nine days later colonies of growing cells had appeared and eight were isolated using micropipettes. Five of these were carried as cell lines called DM 1, DM 3, DM 4, DM 6, and DM 8, the first three originating from the first dish, and the last two from the second dish. A mass culture, derived from the combination of many different colonies was also maintained and called DM.

The clonal line DM 17 originated from a later hybridization experiment made in a similar way, but with both D98/AH-2 and lymphycytes in suspension.

The information on the fusions and the derived cell lines is collected in table 1.

Cell culture conditions. Cells were maintained routinely in RPMI 1640 medium with 10 % foetal calf serum and, generally, penicillin and streptomycin. Cells were transferred by trypsinization using standard techniques. Hybrid cells, presumed to have HGPRT activity and selected in HAT medium, were grown in this medium, while the DMR cell lines (6-thioguanine resistant derivatives of DM - see below) were normally maintained in medium containing 2×10^{-5}M 6-thioguanine.

For storage cells were frozen in a mixture of culture medium and 10 % DMSO (later changed to foetal calf serum with 5 % DMSO) using standard techniques (24), and were kept in liquid nitrogen.

All the cell lines described here grow firmly attached to glass. The cells round up at mitosis, and hybrid cells are then especially easy to shake loose due to their comparatively large volume.

Many different morphologies of growth are seen in the different lines. Some grow in dense patterns, while others grow like a loose net or in a spider form. The growth behaviour is, however, dependent on minor changes in the culture conditions, and so will not be discussed further here.

Selection of 6-thioguanine resistant cells. After 83 days in culture following fusion a part of the mass culture DM was transferred into non-selective culture medium, i.e. medium without the HAT ingredients. After a further 11 days, 5×10^5 cells were seeded into each of 8 bottles. Two days later the medium was changed to a medium containing thioguanine at a final concentration of either 10^{-4}M or 2×10^{-5}M. Almost all cells died and detached within a few days. Eventually, however, a number of colonies of growing cells appeared in all bottles. Some of these colonies were picked, and gave rise to the DMR cell lines (see table 1). The selection procedure did not guarantee that any two lines were of independent origin, since they may be derived from variants that arose before the cells were separated into different bottles.

DMO. A bottle of DM cells, set up 297 days after the fusion was kept in the incubator for one month without further transfers. (Normal maintenance involved transfers at intervals of at most 5 to 7 days by which time a culture was usually confluent.) The culture overgrew after a week, following which most cells died. New cell growth started, however, when the medium was changed after two weeks. The culture was transferred at the end of the month, and the same procedure repeated. The cells were then transferred normally and the resulting cell population was named DMO. The chromosome analysis reported in the text was made two weeks after these cycles of abnormal growth, namely after a total of 372 days in culture.

Chromosome studies. Standard procedure for making air-dried chromosome preparations was used (25). Cells were treated with colcemid usually at a concentration of 0.5 μg/ml for periods ranging from 15 to 60 minutes.

Following harvest by trypsinization the cells were subjected to hypotonic treatment (in 0.075 M KCl) for 20 minutes to as long as 45 minutes. They were then fixed in 3:1 methanol-acetic acid and air-dried in the usual way.

Normal Giemsa-stained preparations were used to determine chromosome numbers. C-banding was produced by the method described by Sumner (26) and G11 banding by the method described by Bobrow et al. (27). Both trypsin digestion (28) and heating in salt solution, the ASG method (29), were used for G-banding. The presence of the Y chromosome in cells was established by C-banding or by quinacrine staining (30).

Serological tests. HL-A typing was carried out using a fluorochromatic microcytotoxicity assay either with cells suspended in droplets, as in the usual lymphocytoxic assay (31), or growing as a monolayer in the wells of Falcon micro-test plates (24). The specificities of the sera used were established by their cytotoxic reactions on a panel of well characterised lymphocytes. HeLa and its derivatives are, in general, much more sensitive than peripheral blood lymphocytes to killing due to the naturally-occurring heterophile antibodies in the normal rabbit serum used as a source of complement. This phenomenon we call 'complement sensitivity'. Rabbit sera to be used as complement were either first screened on D98/AH-2 to select for low levels of heterophile antibody or appropriately absorbed at 4^{o}C. Further details of these techniques and the results of HL-A typing of the hybrids are described in Kennett et al. (32).

The technique used for Xg^{a} typing and a discussion of the results obtained have been published elsewhere (46).

Enzyme analysis. Cells were trypsinized as usual, washed at least three times in PBS, spun down in Beckman microfuge tubes and stored in liquid nitrogen after removal of the final wash supernatant. Horizontal starch-gel enzyme electrophoretic analysis on frozen and thawed extracts was carried out using standard, previously published techniques. For the enzymes specifically used in this study these were as follows:

Glucose 6-phosphate dehydrogenase (G6PD): Meera Khan (33), Motulsky and Yoshida (34).

Hypoxanthine-guanine phosphorybosyltransferase (HGPRT): Watson et al. (35).

Peptidase C (Pep C): Povey et al. (36).

Phosphoglucomutase$_3$ (PGM$_3$): Hopkinson and Harris (37).

Red cell acid phosphatase (ACP$_1$): Swallow et al. (38).

Alkaline phosphatase: for electrophoresis - Poulik (39); for staining - Boyer (40).

5.3 Results

5.3.1 Characterization of the fusion products

Chromosomes. The karyotype of the lymphocyte donor is normal. No uncommon characteristics are shown by C, G 11 or Q banding.

The D98/AH-2 cells that were used for the fusion experiments had a unimodal distribution of chromosome numbers, with a sharp peak at 63 chromosomes (Fig. 1).

Description of karyotypes of this cell line by Miller et al. (41) and by Francke et al. (11) agree well with our results. A typical karyotype of the cell line, as seen after ASG banding, is shown in figure 2 using the nomenclature of Francke et al. (11) for marker chromosomes. Our cell

line's karyotype does not have their marker chromosomes m 5, m 9, m 11 and m 12, while two other markers, m 3A and m 5A, are present in our cells but not those of Francke et al. (11). m 3A is a metacentric chromosome that consists of one long arm of a number nine chromosome with a big heterochromatic region close to the centromere, attached to another long arm of a number nine that has less G 11 banding heterochromatin. This marker is thus a kind of iso-chromosome for the long arm of number nine but with asymmetry in the amount of G 11 staining material. The origin of the subtelocentric m 5A chromosome is not known, but it resembles a number twelve chromosome deleted just beneath the strongly staining band on the long arm and with an unusual small or rearranged short arm. The possible origins of the other marker chromosomes are discussed by Francke et al. (11).

In agreement with Miller et al. (41) and Francke et al. (11) we normally find two X chromosomes in D98/AH-2. Neither of these X chromosomes is late labelling (see below). They do look slightly different and it is not clear whether they are structurally normal.

Chromosome 1, of which there are normally three per cell, is easily recognised when C banded while no chromosome in D98/AH-2 stains like the human Y chromosome. G 11 banding detects one normal chromosome 9 and the m 3A marker chromosome in almost every cell. The marker chromosome m 3 is very distinct and can be picked out from the rest of the chromosomes by its overall shape, without any particular staining method.

Most hybrid lines were studied for the distribution of chromosome numbers at a time earlier than or around two months after the fusion and before the cells were first frozen and stored in liquid nitrogen. Chromosome number distributions are shown in figure 3. Cells that looked as if they were of higher than usual ploidy were not counted.

Almost all the hybrid lines, at this early stage of growth, have a modal chromosome number between 105 and 110, which is very close to the expected full chromosome complement of about 109 chromosomes (46 from the lymphocytes and about 63 from D98/AH-2). Whether the apparent bimodality shown by DM 1, counted after 68 days in culture, really corresponds to the existence of two different cell populations in the cell line is now known. The mass culture, DM, showed no obvious difference from the clones, nor is there a difference between the lines that were isolated from the different fusion experiments (DM 1, 3, 4, 6, v. DM 17). The hybrid lines isolated from fusions between human lymphocytes and D98/AH-2 cells thus have a chromosome stability at least as great as that normally found in intraspecific hybrid cell lines. Figure 4 shows the karyotype of a typical cell from the DM mass culture. It corresponds closely to what is obtained if a normal male karyotype is added to the karyotype shown in figure 2.

The clones were all checked specifically for the presence of the Y chromosome and the m 3A marker chromosome, both of which were always found.

Enzymes. The zymogram phenotypes of red cells and leukocytes from the lymphocyte donor and of D98/AH-2 cells were determined for 22 enzymes in order to detect differences that could be used as genetic markers in the derived hybrids. The enzyme systems tested for were: phosphoglucomutases 1, 2 and 3 (PGM), phosphohexose isomerase, pyrophosphatase, the soluble and the mitochondrial forms of glutamate oxoaloacetate transaminase, indophenol oxidase A, 6-phosphogluconate dehydrogenase, peptidases A, B, C and D (Pep), adenosine deaminase, adenylate kinase, soluble malate dehydrogenase, lactate dehydrogenase, isocitrate dehydrogenase, acid (ACP_1) and alkaline phosphatases, glucose 6-phosphate dehydrogenase (G6PD), and hypoxanthine-guanine phosphoribosyltransferase (HGPRT).

D98/AH-2 cells lack HGPRT as expected since they were selected for 8-azaguanine resistance and die in HAT medium. They also contain a form

of alkaline phosphatase, which both electrophoretically and immunologically (42) appears to be similar to placental alkaline phosphatase. This enzyme is, of course, not found in lymphocytes. Differences in zymogram patterns between D98/AH-2 cells and cells from the lymphocyte donor were found for the enzymes Pep C, PGM_3, G6PD (as expected since HeLa is known to have a G6PD A or A-like variant) and the red cell acid phosphatase (ACP_1). The patterns found in the parent cells and a typical hybrid (DM 4 tested soon after the fusion) are summarized in table 2.

The zymograms for PGM_3 and ACP_1 of DM 4 showed the results expected from the dosage relationships between the different alleles created by the fusion. The same is also true for Pep C, where the Pep C3 band is much weaker than that for Pep C1.

The three G6PD bands on the gels, both on starch and cellogel, showed a clear 1:2:1 relationship, whereas an excess of component A from the D98/AH-2 parent might have been expected. Silagi et al. (4) got a similar result, as judged from their published picture, when they tested a hybrid cell line (e-DAU) derived from a fusion of D98/AH-2 cells with diploid male fibroblasts. Since there seem to be two X chromosomes in all D98/AH-2 cells, this suggests either that the two G6PD alleles have relatively low activities (HeLa lines are known to vary in their G6PD expression (43)) or that only one allele is functionally active.

An experiment with tritiated thymidine showed no obvious late replicating X chromosome in D98/AH-2 cells, and they are also known to have lacked Barr-bodies for a long time (44). The unexpectedly low level of G6PD A in the hybrids therefore does not seem to be explained by X inactivation, unless this is not necessarily accompanied by late replication. If D98/AH-2 has an inactive X chromosome, it should be possible to obtain clones of hybrid cells that do not have any G6PD A but which still retain one of the X chromosomes of D98/AH-2.

No alkaline phosphatase could be detected in the hybrid lines tested nor in any of the thioguanine resistant derivatives of DM. The only instance where it appeared in a hybrid line, DM 4 CS, will be discussed later.

Antigens. D98/AH-2 cells cannot be typed for HL-A antigens in a normal way (24). They occasionally react with high-titred sera that are specific for HL-A 2 and W 28, though other results indicate that D98/AH-2 expresses W 28 (and some other antigens - see 32) and that the HL-A 2 reactions are due to cross-reaction. As discussed earlier, D98/AH-2 is also 'comp-ment sensitive' (CS) relative to peripheral blood lymphocytes.

The HL-A type of the lymphocyte donor is HL-A3, 7, 9 and W5. The mass culture and all the hybrid clones when tested at an early stage of growth behaved in essentially the same way as lymphocytes from the donor. Thus they were not complement sensitive and clearly expressed the same HL-A type as the lymphocyte parent. Similar results were found independently with a D98/AH-2 fibroblast hybrid (HDAU) by Miggiano and co-workers (personal communication). Some of the hybrids reacted also with anti W 28 sera. A more detailed analysis of these and other similar reactions, will be published elsewhere (32).

A summary of the CS, HL-A and Xg types (46) of the parent cells and derived hybrids is given in table 3.

5.3.2 Karyotype evolution during continued growth of the hybrids

The mass culture DM, consisting of cells derived from different fusion events and so presumably harbouring more genetic variability than any of the clones, was used for studies on the stability of the hybrid karyotype. DM was kept growing under normal culture conditions in HAT medium for a total of more than 500 days. The results of five determinations of the distribution of chromosome numbers at different times are given in figure 5. It is clear that the modal chromosome number, even after 100 days in culture, remained very close to 109, the number expected in the

initial fusion products. The chromosome number then fell slowly with an average loss of one chromosome for every fifty days in culture. The cell population, however, apparently split into two sub-populations perhaps even before 200 days of growth. After 500 days one of the sub-populations had a modal chromosome number of 102 while the other had its mode at 95-96. That the bimodality of the chromosome number distribution really represents two different cell populations in the culture, is strongly suggested by the presence of a minute marker chromosome in most of the cells with a chromosome number above 96. This chromosome fragment is not included in the total chromosome count. Only once was a fragment found in a cell with a chromosome number below 95 (see figure 5).

Easily identifiable chromosomes were used as representative indicators of karyotype evolution. Such chromosomes include those readily detected by C or G 11 staining, as illustrated in figures 6a and 6b, namely chromosomes 1, 9, the Y chromosome and the marker chromosomes m 3A and m3. This sample originally constituted about 10 % of the total chromosome number (5+3+1+1+1/109 = 10.1 %).

After 111 days in culture, all these chromosomes were present, as initially, in most cells, but after 518 days in culture changes had occured (table 4). The frequency of rearranged chromosomes is comparatively high, and most cells now have only four number 1 chromosomes. The Y chromosome is still present in every cell. The two sub-populations with different chromosome numbers do not differ with respect to the presence of any of these chromosomes. However, a number of cells, probably representing a new subline, were found with one more iso-9 marker chromosome similar to m 3A but with more and equal amounts of heterochromatin on both arms.

The gradual loss of chromosomes, though slow, poses a problem: Is the loss random with respect to (1) the different chromosome types, and (2) the origin of the chromosomes, namely from D98/AH-2 or the lymphocyte?

It appears that the loss of chromosomes cannot be very non-random. This follows from the fact that an estimate of the mean and variance of the total chromosome number distribution in the population based on the behaviour of the chromosomes sample given in table 4, gives figures that correspond closely to the observed values for the total chromosome set (see table 5). This good agreement implies that the chromosomes, chosen only because of their staining properties, are a representative sample of the whole chromosome set with regard to their behaviour in this hybrid mass culture (apart from the split into two sub-populations). The relatively high frequency of rearranged chromosomes 9 suggests, however, that there may well exist differences in the rate of rearrangement between different chromosomes.

5.3.3 Culture conditions and chromosome loss

Every evolutionary process is limited by the rate of creation of new variants and, if it is not solely dependent on random fluctuations, by the strength of the selective forces acting on these variants. The karyotype stability shown by these hybrid lines can thus either be explained by a slow rate of creation of new variants or by the fact that new karyotypes have only small selective advantages, if any. Both the rate of mutation and the selective values may depend more or less on the environment, namely the culture conditions. If there were little dependence on the environment then changes in the culture conditions should not cause substantial changes in the rate of karyotype evolution.

Weiss and Chaplain (17) have previously reported on the use of abnormal growth conditions as a method for obtaining segregants with substantially reduced chromosome complements from intraspecific hybrid cell lines.

In DMO, the sub-population derived from DM by "maltreatment" as described in Materials and Methods, the modal chromosome number fell sharply compared to that of the DM cells kept under normal conditions during the same time (fig. 7). Whether this drop in chromosome number

was due to selection of cells from an already existing sub-population with a lower chromosome number or whether it was due to creation and selection of new karyotypes is not known. The stability of the karyotype in intraspecific hybrids thus depends at least to some extent on the cell culture conditions.

5.3.4 Thioguanine resistant lines

D98/AH-2 cells can grow in 6-thioguanine while normal hybrid cells, which have HGPRT activity, cannot. Hybrid cells that lack HGPRT can thus be selected for by growth in appropriate concentrations of 6-thioguanine, as described in Materials and Methods. The most likely origin of such 6-thioguanine resistant variants is through loss of the X-chromosome coming from the lymphocyte parent which carried the normal HGPRT gene. It was anticipated that such an enforced loss of one chromosome might be accompanied by concomitant loss of other chromosomes, and so give rise to segregants that could be used for genetic analysis. Thus, in an experiment involving selection for 6-thioguanine resistant cells in intraspecific hybrid hamster lines, Marin and Littlefield (45) obtained cell lines with markedly reduced chromosome numbers. However, resistance can also arise in other ways, specifically by a new mutation of the functioning HGPRT gene or by some change in the genes controlling thioguanine uptake.

The normal X chromosome and X-linked markers. Data on three X-linked markers: HGPRT activity, G6PD phenotype and the Xg^a antigen in a number of thioguanine resistant subclones of DM are given in table 6. All the clones tested lack HGPRT activity, indicating that the resistance to thioguanine is actually due to the loss of this enzyme activity. DMR 41 and DMR 10 both die in HAT medium, as expected, while DMR 41, which was isolated in 2×10^{-5} M 6-thioguanine, can still grow in a concentration of 10^{-4} M 6-thioguanine.

The four lines DMR 10, 36, 63 and 76 have lost the G6PDB allele activity but retain the A activity. DMR 10 and 36 have also lost the X-linked Xg^a cell surface antigen (46). This is consistent with a complete loss (or major deletion) of the normal X chromosome coming from the lymphocytes in these four lines. DMR 41 behaves quite differently, in that it retains G6PD A and B as well as the Xg^a antigen. Its normal karyotype also contains three chromosomes similar to the human X chromosome (see figure 8), while only two X chromosomes were found in the DMR 10 and DMR 63 cells tested. The loss of the enzyme HGPRT in this clone is thus probably due, either to a mutation in the gene coding for the enzyme, or to a very small deletion of the corresponding chromosome region. No colony of cells able to grow in HAT medium was obtained from 10^6 DMR 41 cells. The mutation at the HGPRT locus is thus not readily revertible.

Autosomal enzymes and antigens. The information on autosomal genetic markers in these thioguanine resistant lines is given in table 7. None of the lines have become complement sensitive or show any alkaline phosphatase activity. Only DMR 10, of all the lines studied, shows a difference in one respect from the zymogram patterns that were found in the initial hybrid lines. It retains only the Pep C1 band and appears to have lost Pep C3 activity.

Autosomes. Consistent with the low frequency of changes in the expression of autosomal markers in the 6-TG resistant lines their chromosome numbers remained quite high (table 8). All the lines tested have a mean chromosome number greater than 100. The lowest means are found in the two lines, DMR 10 and DMR 63, that had been in culture longest before they were tested. The mean chromosome numbers in most of these lines are, however, smaller than can be explained by loss of just one chromosome (the lymphocyte donor's X) plus the expected loss over time. This may be due to the more severe culture conditions involved in cloning with selection which may increase the rate of chromosome loss, as happened with DMO.

The marker chromosomes, studied as before (see table 4) are also, apart from a few exceptions, the same as in DM (see table 9).

DMR 10 and DMR 41 have both lost one chromosome 1, and a third line, DMR 36, has a chromosome 1 with a part of the long arm deleted. The data on the Pep C phenotypes, which are known to be controlled by a gene on chromosome 1 (3), shows that the chromosomes 1 lost in DMR 10 and 41 cannot be the same. This follows from the fact that DMR 41 has the same Pep C phenotype as DM and therefore retains the chromosome from the lymphocyte parent that carries the Pep C3 allele, while this is the chromosome that DMR 10 must have lost. Whether DMR 41 has lost a chromosome 1 from D98/AH-2 or from the lymphocyte donor cannot be established with this data. DMR 10 has also lost the very distinctive m 3 marker chromosome. Hors-Cayla et al. (9) report the loss of this chromosome in one of their cell lines, derived from a fusion between D98/AH-2 cells and human fibroblasts, while Francke et al. (11) describe other cases of loss of D98/AH-2 marker chromosomes from similar hybrids. There is thus good cytological evidence that D98/AH-2 as well as normal diploid parent chromosomes are lost from these human intraspecific hybrids.

DMR 63 cells have lost a normal chromosome 9 of unknown origin, while all the lines tested retain the Y chromosome.

5.3.5 Spontaneous segregant lines

Two cell populations that showed interesting differences from the rest of the lines are described here because they illustrate further how cell lines derived from fusions between D98/AH-2 and diploid cells can be used for human genetic analysis.

DM 17 A. A DM 17 culture grown for several transfers after storage in liquid nitrogen was found to differ from other DM 17 cells presumably as a result of chromosome loss during further growth.

This cell line, called DM 17 A, had lost the PGM_3 2 activity and also concomitantly the HL-A antigens 3 and 7. Neither of these losses has been observed so far in any other of the hybrid lines discussed here (table 10). All other genetic markers are the same in DM 17A as in DM 17. The mean chromosome number in DM 17A was 101.6 (14 metaphases counted; range: 100-105; no counts excluded), which is lower than the number DM 17 had when it was tested soon after fusion (see figure 3).

It is now fairly well established that the genes for HL-A and PGM_3 are on the same chromosome (47). Data from the lymphocyte donor's family show that his genotype is HL-A 3, 7 / HL-A 9, W5 and PGM_3 1/2, but do not establish the linkage phase of PGM_3 1 and 2 relative to the two HL-A haplotypes. The information from DM 17A confirms the HL-A genotype of the lymphocyte donor and suggests that the PGM_3 2 allele and the HL-A 3,7 haplotype are carried in coupling on the same chromosome.

DM 4CS. Another variant derived from DM 4 in an exactly analogous way (DM 4CS) has regained the complement sensitivity characteristic of D98/AH-2. These cells also express the placental type of alkaline phosphatase. With respect to the other enzymes and HL-A, DM 4CS behaves like DM 4 (table 10). This table also gives the markers found in HDAU, a hybrid line derived from a fusion between D98/AH-2 and cells from a human fibroblast strain (4). No information is available concerning the presence of these markers during its early stages of growth, but it is now similar to DM 4CS in being both complement sensitive and alkaline phosphatase positive.

The most likely interpretation of the reappearance in DM 4CS and HDAU of alkaline phosphatase activity and the surface antigens against which the naturally occuring rabbit heterophile antibodies are directed, is that these cells have lost one or more genetic factors contributed to the hybrid

cells by the lymphocyte which switch off these D98/AH-2 properties. Consistent with this argument is the fact that DM 4CS and HDAU have lost a substantial number of chromosomes (DM 4CS: 10 metaphases counted, range 88-102; HDAU: all chromosome counts < 100). More data on similar segregants are needed to establish whether there is any significance to the combined reappearance of alkaline phosphatase activity and complement sensitivity in these two lines.

5.4 Discussion

We have confirmed that intraspecific hybrids are chromosomally not completely stable and have shown that systematic isolation of chromosome segregants is feasible. In our studies segregants were produced in a variety of ways. Thus,

a) some segregants appeared spontaneously during extended periods of culture,
b) selection imposed non-specifically by growth in relatively poor nutritional environment seemed to increase the rate of appearance of segregants, and
c) selection for 6-thioguanine resistant cells in the sensitive hybrid have rise to lines, most of which had lost not only the X chromosome carrying the normal HGPRT gene but also other chromosomes.

Furthermore, there is no evidence for non-random chromosome loss under non-selective conditions and the losses appear to involve chromosomes from both parents.

Preliminary results (Kennett et al., unpublished observations) have clearly indicated that hybrids made between D98/AH-2 and lymphocytes are karyotypically more stable than those made with fibroblasts. Thus, whereas the DM hybrids initially and for some time retain average chromosome numbers between 100 and 105, crosses with fibroblasts fairly rapidly produce

clones with chromosome numbers well below 100. The relative instability of the initial clones is not a result of the senescence of the fibroblasts because the same result is seen when fibroblasts from a fresh explant are used in the hybridization. It may prove interesting to analyse the effect of using cells from different tissues in intraspecific hybrids to determine the effects of different cell types on hybrid stability. Karyotypic analysis may indicate that there is preferential loss or retention of specific chromosomes depending on the tissue origin of the cell type hybridized.

Intraspecific hybrids have so far been considered unsuitable for genetic segregation analysis because of their greater karyotypic stability and the relative scarcity of usable markers compared to interspecific hybrids. Our studies have, however, shown that genetic analysis using intraspecific hybrids faces no particular obstacles. Our approach to genetic analysis using intraspecific hybrids is closely analogous to Pontecorvo and coworkers' (48) classical genetic analysis of Aspergillus using the parasexual cycle. In this case stable diploids were first established using selective techniques, and then chromosome segregation and crossing over was forced by further selection. In our case the initial hybrids between D98/AH-2 and lymphocytes correspond to the stable diploid, and chromosome segregants are produced from this by selection or adventitiously. The use of such segregant lines for the analysis of genetic markers, either on a chromosome carrying a selective marker, the X, or on other chromosomes, is illustrated by the segregation of the HL-A 3,7 haplotype with the PGM_3-2 allele and the use of the Pep C3 activity to identify the parental origin of a lost chromosome 1.

5.5 References

1. Weiss, M.C. and Green, H. 1967. Proc. Nat. Acad. Sci. U.S.A. 58: 1104-1111.
2. Bodmer, W.F. 1971. Proc. Fourth Int. Congr. Hum. Genet. pp. 365-373.
3. Ruddle, F.H. 1973. Nature 242: 165-169.
4. Silagi, S., Darlington, G. and Bruce, S.A. 1969. Proc. Nat. Acad. Sci. U.S.A. 62: 1085-1092.
5. Siniscalco, M., Klinger, H.P., Eagle, H., Koprowski, H., Fujimoto, W.Y. and Seegmiller, J.E. 1969. Proc. Nat. Acad. Sci. U.S.A. 62: 793-799.
6. Nadler, H.L., Chacko, C.M. and Rachmeler, M. 1970. Proc. Nat. Acad. Sci. U.S.A. 67: 976-982.
7. de Weerd-Kastelein, E.A., Keijzer, W. and Bootsma, D. 1972. Nature New Biol. 238: 80-83.
8. Schneider, J.A., Francke, U., Hammond, D.S., Pellett, O.L. and Becker, F.L.A. 1973. Nature 244: 289-292.
9. Hors-Cayla, M.C., Trébuchet, C. and Heuertz, S. 1972. Ann. Génét. 15: 153-158.
10. Glaser, R. and O'Neill, F.J. 1972. Science 176: 1245-1247.
11. Francke, U. Hammond, D.S. and Schneider, J.A. 1973. Chromosoma 41: 111-121.
12. Engel, E., McGee, B.J. and Harris, H. 1969. Nature 223: 152-155.
13. Klein, G., Bregula, U., Wiener, F. and Harris, H. 1971. J. Cell Sci. 8: 659-672.
14. Bregula, U., Klein, G. and Harris, H. 1971. J. Cell Sci. 8: 673-680.
15. Wiener, F., Klein, G. and Harris, H. 1971. J. Cell Sci. 8: 681-692.
16. Wiener, F., Klein, G. and Harris, H. 1973. J. Cell Sci. 12: 253-261.
17. Weiss, M.C. and Chaplain, M. 1971. Proc. Nat. Acad. Sci. U.S.A. 68: 3026-3030.
18. Minna, J., Glazer, D. and Nirenberg, M. 1972. Nature New Biol. 235: 225-231.
19. Szybalski, W., Szybalska, E.H. and Ragni, G. 1962. Nat. Cancer Inst. Monogr. 7: 75-89.

20. Gartler, S.M. 1968. Nature 217: 750-751.

21. Littlefield, J.W. 1964. Cold Spring Harbor Symp. Quant. Biol. 29: 161-166.

22. Böyum, A. 1968. Scand. J. Clin. Lab. Invest. 21: Suppl. 97.

23. Pontecorvo, G. 1971. Nature 230: 367-369.

24. Miggiano, V., Nabholz, M. and Bodmer, W.F. 1970. in Histocompatibility Testing 1970. Munksgaard, Copenhagen. pp. 623-629.

25. Hamerton, J.L. 1971. Human Cytogenetics, Volume 1. Academic Press, New York and London.

26. Sumner, A.T. 1972. Exptl. Cell Res. 75: 304-306.

27. Bobrow, M., Kamlesh, M. and Pearson, P.L. 1972. Nature New Biol. 238: 122-124.

28. Seabright, M. 1971. Lancet ii: 971-972.

29. Evans, H.J., Buckton, K.E. and Sumner, A.T. 1971. Chromosoma 35: 310-325.

30. Bobrow, M., Pearson, P.L., Pike, M.C. and El-Alfi, O.S. 1971. Cytogenetics 10: 190-198.

31. Bodmer, W., Tripp, M. and Bodmer, J. 1967. in Histocompatibility Testing 1967. Munksgaard, Copenhagen. pp. 341-350.

32. Kennett, R., Bengtsson, B.O. and Bodmer, W.F. 1974. Tissue Antigens (in press).

33. Meera Khan, P. 1971. Arch. Biochem. Biophys. 145: 470-483.

34. Motulsky, A.G. and Yosida, A. 1969. in Biochemichal Methods in Red Cell Genetics, ed. Yunis, J.J. Academic Press, London. p. 51.

35. Watson, B., Gormley, I.P., Gardiner, S.E., Evans, H.J. and Harris, H. 1972. Exptl. Cell Res. 75: 401-409.

36. Povey, S., Corney, G., Lewis, W.H.P., Robson, E.B., Parrington, J.M. and Harris, H. 1972. Ann. Hum. Genet. 35: 455-465.

37. Hopkinson, D.A. and Harris, H. 1968. Ann. Hum. Genet. 31: 359-367.

38. Swallow, D.M., Povey, S. and Harris, H. 1973. Ann. Hum. Genet. 37: 31-38.

39. Poulik, M.D. 1957. Nature 180: 1477-1479.

40. Boyer, S.H. 1961. Science 134: 1002-1004.

41. Miller, O.J., Miller, D.A., Allderdice, P.W., Dev, V.G. and Grewal, M.S. 1971. Cytogenetics 10: 338-346.

42. Tan, K.K. and Aw, S.E. 1973. Immunochemistry 10: 209-212.

43. Steele, M.W. 1970. Biochem. Genet. 4: 25-40.

44. Dewitt, S.H., Rabson, A.S., Legallais, F.Y., Del Veccio, P.R. and Malmgren, R.A. 1959. J. Nat. Cancer Inst. 23: 1089-1095.

45. Marin, G. and Littlefield, J.W. 1968. J. Virol. 2: 69-77.

46. Fellous, M., Bengtsson, B., Finnegan, D. and Bodmer, W.F. 1974. Ann. Hum. Genet. 37: 421-430.

47. Edwards, J.H., Allen, F.H., Glenn, K.P., Lamm, L.U. and Robson, E.B. 1973. in Histocompatibility Testing 1972. Munksgaard, Copenhagen. pp. 745-751.

48. Pontecorvo, G. 1959. Trends in Genetic Analysis. Oxford University Press, Oxford.

5.6 Tables and Figures

		Name	Derived lines
Fusion 1	Massculture:	DM	All DMR lines; DMO
	Clones:	DM 1	
		DM 3	
		DM 4	DM 4 CS
		DM 6	
		DM 8	
Fusion 2[a]	Clone:	DM 17	DM 17A

[a] Other clones were isolated, but only DM 17 has been used for the work reported in this chapter.

Table 1. Cell lines produced by fusions between lymphocytes from a male donor, MAR, and D 98/AH-2 cells with selection in HAT-medium.

Enzymes	Parent cells		Typical hybrids
	D98/AH-2	MAR	DM 4
Pep C	1	3-1	3-1
PGM_3	1	2/1	2/1
G6PD	A	B	AB
ACP_1	BA	B	BA
Alk. phosphatase[a]	+	-	-
HGPRT	-	+	+

[a] Of placental type.

Table 2. Enzyme activity and zymogram patterns in DM 4 and its parent cells. Only those enzymes for which D 98/AH-2 cells are different from cells of the lymphocyte donor are included in this table.

See text for the other enzymes investigated and for the abbreviations used. DM 4 and other similar hybrids have considerably more of the enzyme activity corresponding to the alleles Pep C1, PGM_3 1 and ACP_1 B than to the alleles Pep C3, PGM_3 2 and ACP_1 A.

		Complement sensitivity	Xg^a type	HL-A
Parents	D98/AH-2	+	-	(W28)
	MAR	-	+	3,7,9,W5
Hybrids	DM	-	NT	3,7,9,W5,(W28)
	DM 4	-	NT	3,7,9,W5,(W28)
	DM 6	-	NT	3,7,9,W5,(W28)
	DM 8	-	+	3,7,9,W5,(W28)
	DM 17	-	+	3,7,9,W5,(W28)

Table 3. Immunological reactions of the parent cells and some of the derived hybrid cell lines.

Chromosome type	Cells with rearranged chromosomes	Number of chromosome copies per cell					
		0	1	2	3	4	5
1	1				2	42	5
9 + m 3A	4[a]				2	43	1
Y	0	1	48	1			
m 3	1		49				

[a] 2 of these cells had the same new marker chromosome.

Table 4. Chromosome variation in DM after 518 days in culture. Every chromosome type was scored for in 50 metaphases.

Table 5. The distribution of chromosome numbers in DM after 518 days.

	Mean	Variance
Estimated from the chromosome sample[a]:	99.5	2.5
Found in the cell population with		
Low chromosome numbers:	94.9	3.5
High chromosome numbers:	101.7	1.8

[a] The following methods have been used to calculate the table: The distribution of chromosome numbers in DM after 518 days in culture, given in figure 5, is divided into two, one with low chromosome numbers (91-98) and one with high (99-104). The means and variances of these two new distributions are then computed in the usual way to give the 4 lower figures in the table. For the estimates, information from table 4 is used to calculate the average number of copies of the different chromosome types per cell and the corresponding variances. There are, for example, on average 4.06 copies of chromosome 1 per cell with a variance of 0.142. These values are then added, assuming that the presence or absence of the different chromosomes are independent, to give an average number of chromosomes per cell from the sample of 10.4, with a variance of 0.242. If the sample is representative of all the chromosomes and if the initial hybrid cells had a total of 109 chromosomes with 11 from the sample refered to in table 4, then multiplying the sample values by 109/11, gives the estimate of 99.48 chromosomes per cell with a variance of 2.459 as shown in the first line of the table.

Lines	Conc. of 6-thioguanine used for isolation	G6PD	Xg^a	HGPRT
DMR 41	$2\ 10^{-5}$M	AB	+	-
DMR 76	$2\ 10^{-5}$M	A	NT	-
DMR 36	$2\ 10^{-5}$M	A	-	-
DMR 10	10^{-4}M	A	-	-
DMR 63	10^{-4}M	A	NT	-

Table 6. X linked markers in the DMR lines tested. The data on Xg^a is taken from Fellous et al. (46).

Lines	Pep C	PGM_3	Acid phos-phatase	Alk. phos-phatase	Complement sensitivity	HL-A[a]
DMR 41	3-1	2/1	BA	-	-	3,7,9,W5
DMR 76	3-1	2/1	BA	-	-	3,7,9,W5
DMR 36	3-1	2/1	NT	-	-	3,7,9,W5
DMR 10	1	2/1	BA	-	-	3,7,9,W5
DMR 63	3-1	2/1	BA	-	-	3,7,9,W5

[a] Presumed to express W28 though not specifically tested for this.

Table 7. Autosomal genetic markers in the DMR cell lines.

Lines	Days grown after fusion	Metaphases counted	Counts excluded[a]	Range	Mean
DMR 41	159	12	95	103-106	105.0
DMR 42	160	8	-	99-107	102.9
DMR 43	160	8	-	98-105	101.4
DMR 36	213	8	-	102-107	104.3
DMR 76	220	9	95	104-107	105.5
DMR 10	240	9	110	98-105	100.9
DMR 63	249	8	-	98-102	100.3

[a] Chromosome counts which clearly differed from the rest are not included in the description of the range and the calculation of the mean.

Table 8. Chromosome numbers in DMR lines.

Cell lines	Chromosome types				
	1	9	m 3A	Y	m 3
DMR 41	4	3	1	1	1
DMR 76	5	3	1	1	1
DMR 36	4[a]	3	1	1	1
DMR 10	4	NT	NT	1	0
DMR 63	5	2	1	1	1

[a] Plus one chromosome 1 with a part of the long arm deleted.

Table 9. The occurence of selected chromosomes in the DMR cell lines.

Cells	Pep C	PGM_3	G6PD	ACP_1	Alk. phosphatase	Complement sensitivity	HL-A[b]
MAR	3-1	2/1	B	B	-	-	3,7,9,W5
D98/AH-2	1	1	A	BA	+	+	
DM 4 & DM 17	3-1	2/1	AB	BA	-	-	3,7,9,W5
DM 17A	3-1	1	AB	BA	-	-	9,W5
DM 4CS	3-1	2/1	AB	BA	+	+	3,7,9,W5
HDAU[a]	1	2/1	AB	(B)A[c]	+	+	(2,9)

[a] The other parent of this hybrid line (4) was not available for study at the time these data were obtained.

[b] Only the HL-A antigens from the lymphocyte donor are listed, except for HDAU for which the fibroblast parent types are not known.

[c] ACP_1 B very much weaker than A.

Table 10. Genetic markers in DM 17A, DM 4CS and HDAU compared with the original hybrid lines from which they were derived and with the parent cells.

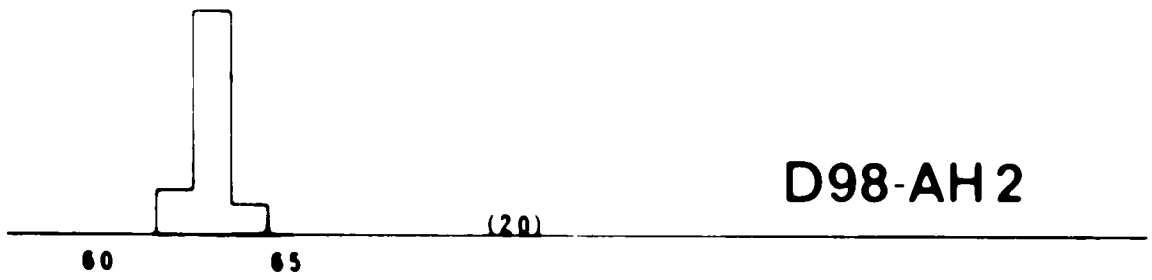


Figure 1. Chromosome distribution of D 98/AH-2. 20 metaphases counted.

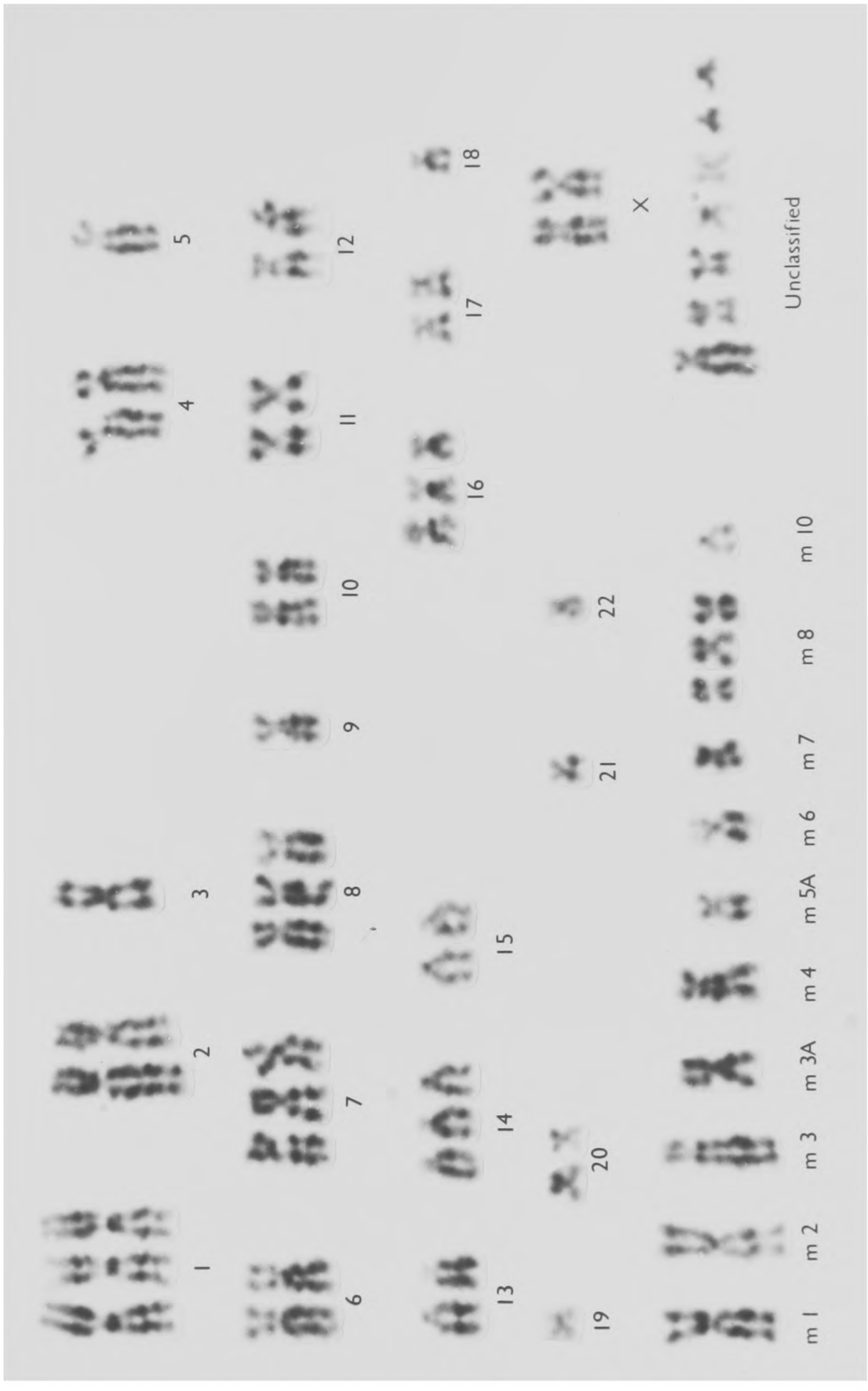


Figure 2. Karyotype of a D 98/AH-2 cell. G banding by the ASG method. The nomenclature for the marker chromosomes is essentially the same as used by Francke et al. 1973.

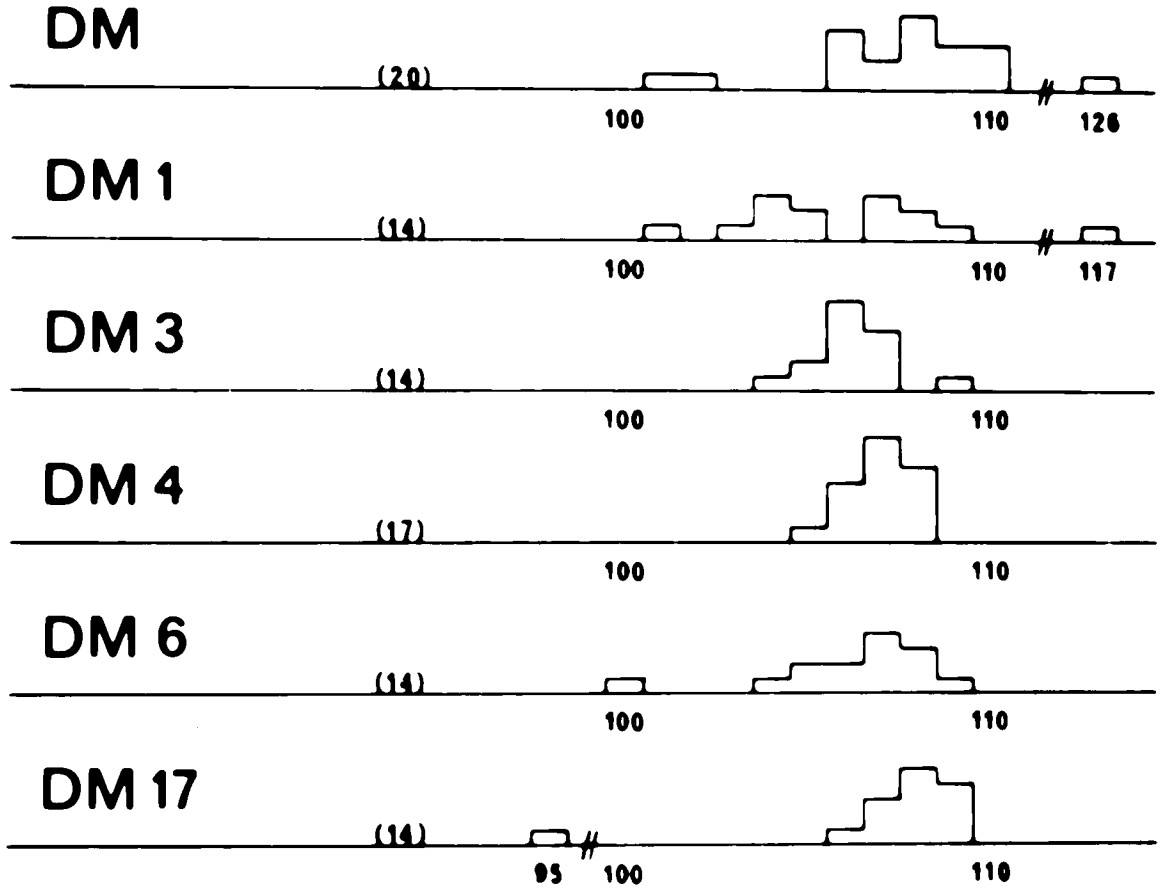


Figure 3. Chromosome number distributions of the massculture DM and six independent clones DM 1, 3, 4, 6 and 17, tested around or before two months in culture. Cells of higher ploidy are excluded. The number of counted metaphases is given in parenthesis to the left of the distributions. The means of the distributions are: DM 107.3 (126 excluded); DM 1 105.5 (117 excluded); DM 3 106.2; DM 4 106.9; DM 6 106.2; DM 17 107.9.

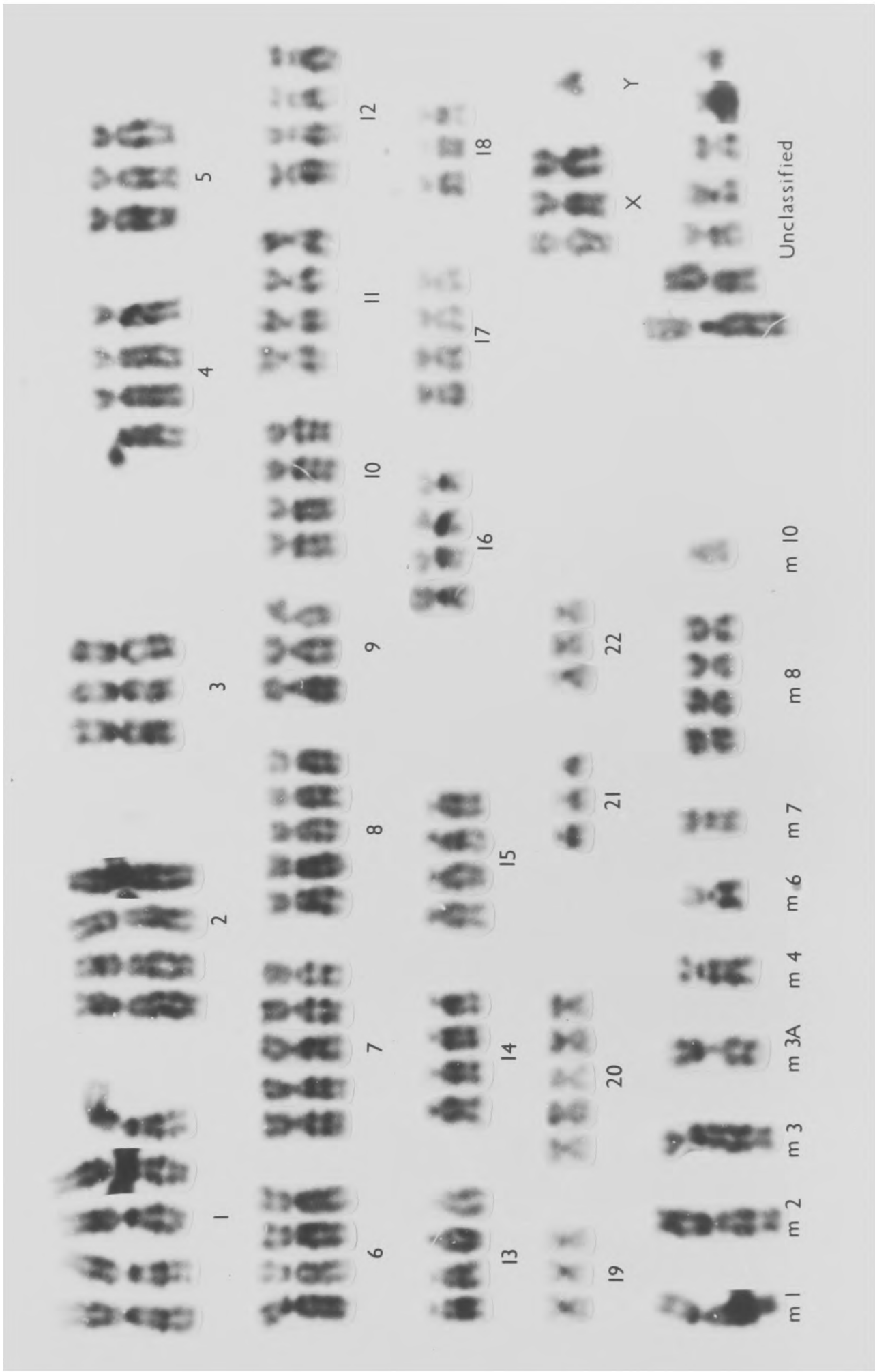


Figure 4. Karyotype of a DM cell after 111 days in culture. G banding by the ASG method. Compare with figure 2.

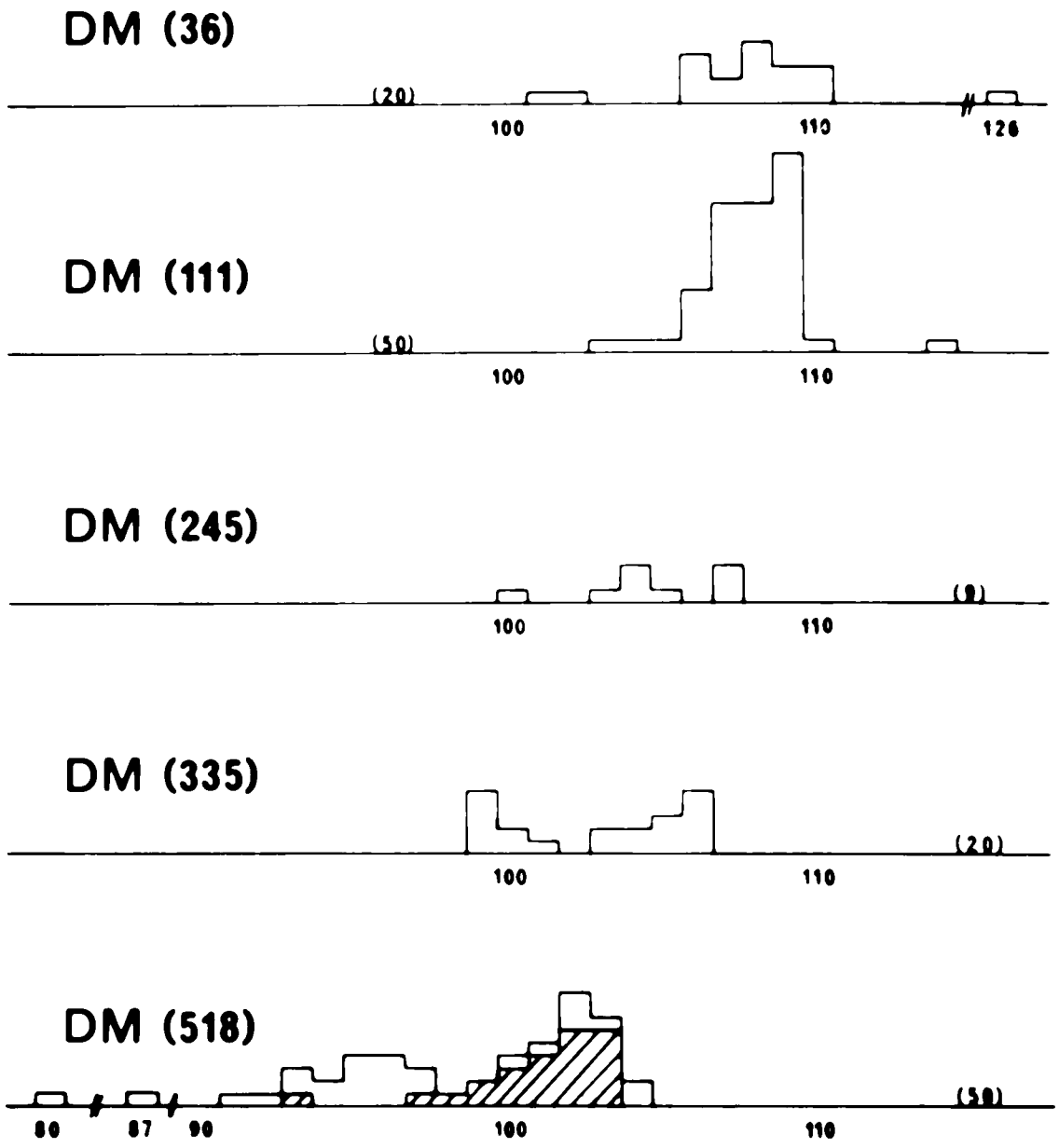


Figure 5. Chromosome distributions of the massculture DM after 36, 111, 245, 335 and 518 days in culture. The number of counted metaphases is given in parenthesis. Cells with a very small chromosome, recorded as a fragment, are shaded in the last distribution. The means of the distributions are: DM(36) 107.3 (126 excluded); DM(111) 109.8; DM(245) 104.7; DM(335) 102.8; DM(518) 99.0 (80 and 87 excluded).

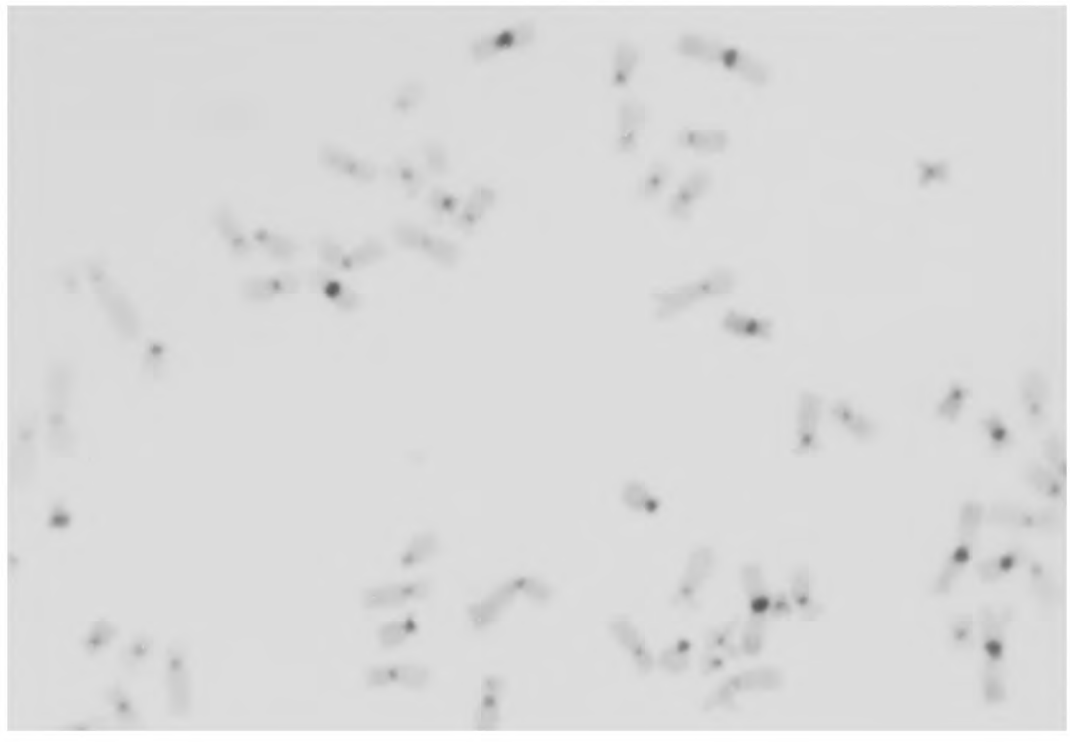


Figure 6a. Part of a C banded DM metaphase. 4 number one chromosomes are easy to find by their length and distinct heterochromatin staining. The Y chromosome is also readily recognized (far left).

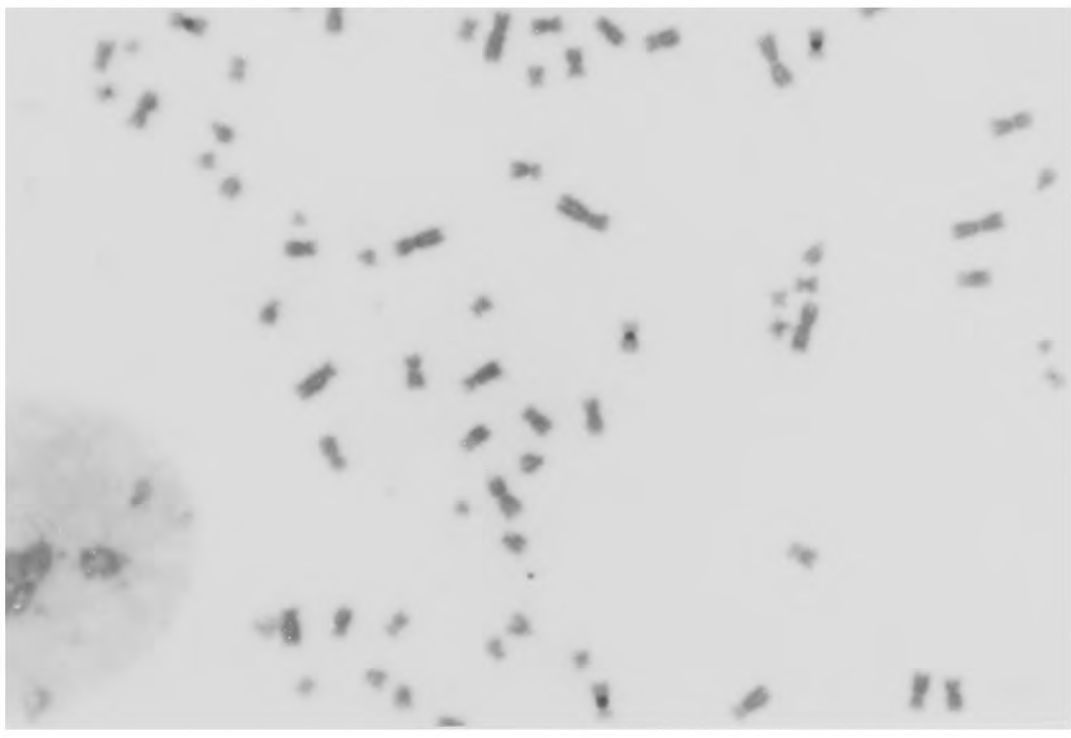


Figure 6b. Part of a G 11 banded DM metaphase. Three normal number nine chromosomes are seen (bottom, middle and top right), and the m 3A marker chromosome (top left).

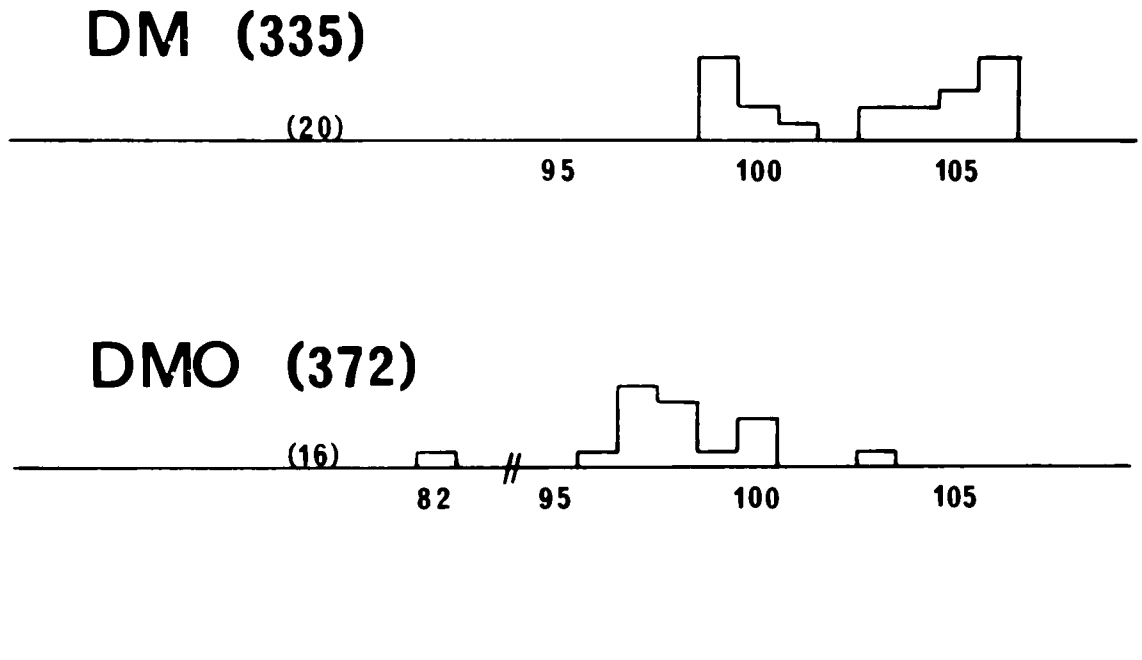


Figure 7. Chromosome distribution of DMO after 372 days in culture, compared with the normal DM after 335 days, showing the effect of different culture conditions. The mean numbers are for DM 102.8 and for DMO 98.3 (82 excluded).

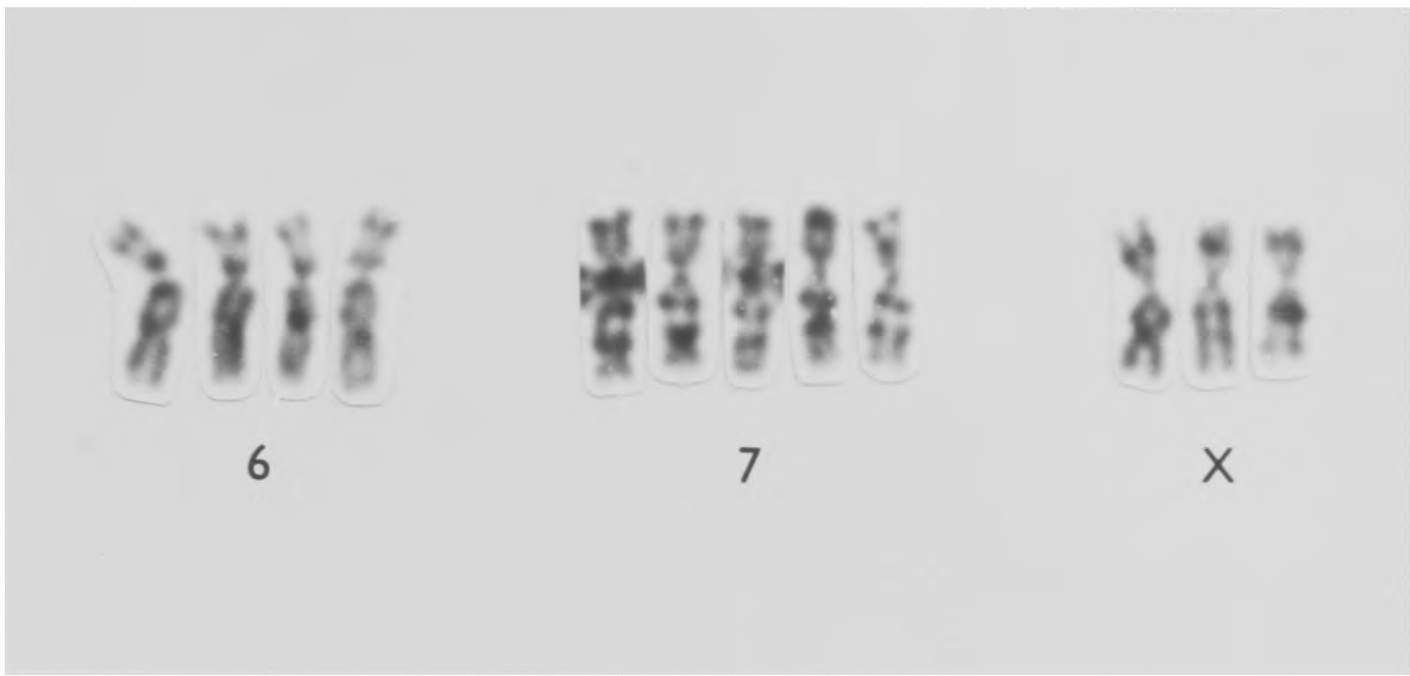


Figure 8. Part of the karyotype of a DMR 41 cell. G banding by trypsin treatment. Three chromosomes looking like the human X chromosome are present. Chromosomes 6 and 7 are included for comparison.

CHAPTER 6
KARYOTYPE SELECTION IN CELL LINES

6.1 Introduction

In this the last chapter we shall outline a model which describes the balance between mutation and selection of karyotypes in an established cell line. It is a well known fact that two cells in a cell line only rarely have the same karyotype. But it is also true that most cells in a cell line have karyotypes which are very similar and typical for the line in question; and even over long periods of growth, the "average" karyotype of a cell line normally remains the same.

We try to describe the combination of variation and stability in a mathematical model by assuming the existence of cells with an ideal karyotype. This means that no other cells in the cell line in question are better adapted than these to the in vitro life. Furthermore, it is assumed that cells which have almost the same karyotype as the ideal are almost as fit as the cells with the ideal karyotype. If finally it is assumed that there is a chance at every cell division that the mitotic apparatus fails, thereby making the karyotypes of the daughter cells different from the karyotype of the parent cell, then a dynamic system has been described which can lead to a stable equilibrium of karyotypes in the cell line.

The model is described in its most simplified form, leaving the discussion of its worst limitations to the end of the chapter.

6.2 A model of karyotype selection

6.2.1 Preliminaries

Cells. We denote the karyotype of a cell, c_x, by a vector $(m_{x1}, m_{x2}, \ldots, m_{xk})$. This means that the cell c_x has m_{x1} copies of chromosome type one, m_{x2} copies of type two, and so on. A chromosome type is, for example, the number one or the Y human chromosome, but it can also be a marker chromosome characteristic for the cell line.

$S(c_x)$ is defined as the total number of chromosomes in the cell c_x, i.e.

$$S(c_x) = \sum_{i=1}^{k} m_{xi} .$$

We introduce a measure of the difference between the karyotypes of two cells, c_x and c_y, by

$$D(c_x - c_y) = \sum_{i=1}^{k} \left| m_{xi} - m_{yi} \right| .$$

This is the number of chromosomes which has to be added and taken away from c_x's karyotype to obtain c_y's karyotype, and vice versa.

Cell cycle and fitness. To interpret the fitness value which every cell will be given, we imagine that all newly divided cells will divide again after exactly the same length of time. The cell cycle, if completed, is thus equally long for all cells. The difference between cells is that not all cells survive the cell cycle. We define $F(c_x)$ as the probability that the cell c_x survives the cell cycle and divides, and this is our measure of cell fitness. A different interpretation of cell fitness could have been given; we could have thought, for example, that all cells always survive the cell cycle but complete it in different lengths of time.

Mutation. If a cell has survived the cell cycle it will then divide. The mitosis normally makes the karyotypes in the two daughter cells identical to the karyotype of the parent cell. But a mutation can occur, either during

the parent cell's interphase or during the mitosis. The probability that a daughter cell has a karyotype which is different from its parent cell's initial karyotype, we call a. For reasons which will become obvious later, we define 1-a as a_0.

6.2.2 The model

We assume that there exists one ideal karyotype in the cell-line and we express the fitness of all cells in the cell-line as a function of how different the cells' karyotypes are from the ideal one. One of the simplest ways to do this is to say that for every chromosome in a cell which differs from the ideal karyotype, the cell has a certain fixed decrease in fitness. In mathematical terms this is expressed:

$$F(c_x) = 1 - b \cdot D(c_x - c_I) \ ,$$

where c_I is a cell with the ideal karyotype and b is a coefficient of selection. By this definition of fitness, only cells with the ideal karyotype have fitness 1. To exclude negative fitness values we have to add to the definition that if $1 - b \cdot D(c_x - c_I)$ is negative, then $F(c_x)$ is 0. With this definition of fitness, we can divide a population of cells into a (finite) number of fitness classes: cells with fitness 1 (cells with the ideal karyotype), cells with fitness 1-b (cells with the ideal karyotype except for one single chromosome too much or too few), cells with fitness 1-2b, and so on to cells with fitness 0.

We assume that chromosome mutations, caused by non-disjunctions for example, change the vectors describing the karyotypes by adding and/or subtracting one or more chromosomes of different types. It is also assumed that the number of chromosome types is rather high. From the preceding section we can then expect that cells with many chromosome types present in deviant numbers will be rare in the cell-line because they will be strongly selected against. Most chromosome mutations will therefore be to the detriment of the cells concerned, since only rarely

will it happen that, say, a non-disjunction will change a karyotype towards a better state, rather than towards a worse. We shall thus not consider backmutations on the fitness scale in our model of mutations; this we have to do to make the model solvable.

The probability that a cell has the same karyotype as the parent cell we defined before as a_0. We now define a_1 as the probability that the daughter cell has one chromosome more differing from the ideal karyotype than the parent cell. This implies that the daughter cell belongs to a fitness class one step below the parent cell. In the same way we define a_i as the probability that a daughter cell has a karyotype which differs by i more chromosomes from the ideal karyotype than the parent cell does. Obviously, $\sum_{i=1}^{\infty} a_i$ is equal to a, the total mutation rate, and $\sum_{j=0}^{\infty} a_j = 1$. We shall assume that the mutation rate is the same for all cells, independent of their fitnesses.

6.2.3 Equilibrium behaviour

We consider an infinitely large population of cells. It will be described by the proportions of cells in the different fitness classes. The proportions are determined at different points in time, each separated by the length of time it takes for the cells to complete a cell cycle. For convenience, we call our discontinuous points in time different generations.

Call the frequency of cells with fitness $1 - b \cdot i$ in generation k, by $z_i(k)$. We can then represent the cell population in generation k by the vector $Z(k) = (z_0(k), z_1(k), \ldots, z_i(k), \ldots)$. Let $z_0(0)$ be a positive number, since otherwise there will never be any cells with the ideal karyotype in the population. Let L be the column vector with all elements equal to 1 and with the same number of elements as Z, i.e. the number of fitness classes.

Now we introduce the matrix A which describes the joint effects of selection and mutation:

$$A = \begin{pmatrix} a_0 & a_1 & a_2 \ldots\ldots & a_i \ldots\ldots \\ 0 & a_0(1-b) & a_1(1-b) \ldots & a_{i-1}(1-b) \ldots \\ 0 & 0 & a_0(1-2b) \ldots & a_{i-2}(1-2b) \ldots \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 \ldots\ldots & a_0(1-ib) \ldots \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

With this matrix we can express the connection between two different generations by:

$$Z(k) = Z(k-1) \cdot A \;/\; (Z(k-1) \cdot A \cdot L) \; .$$

The dividing factor, $Z(k-1) \cdot A \cdot L$, is a normalizing constant making $Z(k)$ a vector of sum 1, as required.

From standard matrix theory we know that when k increases Z(k) will go towards the left eigenvector of sum 1 corresponding to A's biggest eigenvalue, which in this case is a_0. (For the exact conditions, see the appendix.)

The equilibrium population, $Z = (z_0, z_1, \ldots, z_i, \ldots)$ can thus be found by solving the relevant equation system:

$$a_0 z_0 = a_0 z_0$$
$$a_1 z_0 + a_0(1-b) z_1 = a_0 z_1$$
$$\vdots$$
$$a_i z_0 + a_{i-1}(1-b) z_1 + \ldots + a_0(1-ib) z_i = a_0 z_i$$
$$\vdots$$

plus $z_0 + z_1 + z_2 + \ldots + z_i + \ldots = 1$,

which simplified becomes:

$$z_0 = h$$
$$z_1 = a_1 h \,/\, a_0 b$$
$$z_2 = (a_2 h + a_1(1-b)z_1) \,/\, 2a_0 b$$
$$\cdot$$
$$\cdot$$
$$z_{i+1} = (a_i h + a_{i-1}(i-b)z_1 + \ldots + a_i(1-ib)z_i) \,/\, ia_0 b$$
$$\cdot$$
$$\cdot$$

where h is such that $\sum_{i=0} z_i = 1$.

We have thus obtained a series of equations which makes it possible to exactly determine the equilibrium the cell population will reach due to the opposing forces of mutation and selection.

6.2.4 Numerical example

We shall now give an example to illustrate the model: Suppose that every loss or gain of a chromosome gives a disadvantage of 3 % to the cell concerned, i.e. b = 0.03. 85 % of all cells are assumed to have the same karyotype as their parent cells; 5 % will have lost or gained one chromosome, 4 % two chromosomes, and so on to 1 % which have lost or gained altogether 5 chromosomes. Expressed in the terms used above this means that $a_0 = 0.85$, $a_1 = 0.05$, $a_2 = 0.04$, $a_3 = 0.03$, $a_4 = 0.02$ and $a_5 = 0.01$. If the cell-line started with at least a few cells with the ideal karyotype, then, after a number of cell divisions, the cell population will reach an equilibrium which can be described by the system of equations given above. In table 1 is given the equilibrium distribution of the cell-line, described by fitness classes. However, we are probably more interested in the distribution of chromosome numbers in the equilibrium population. This can be given, if we assume that loss and gains of chromosomes by mutation are equally likely and if we assume that the chromosome number of the ideal karyotype is m. The distribution of chromosome numbers that this

i =	0	1	2	3	4	5	6	7	8	9	10
% of the population	4.4	8.7	11.8	13.3	13.6	12.8	11.0	9.0	6.9	5.0	3.5

Table 1. The percentage of cells in the different fitness classes $1 - b \cdot i$ at equilibrium.

	Number of chromosomes per cell									
	m	$m \pm 1$	$m \pm 2$	$m \pm 3$	$m \pm 4$	$m \pm 5$	$m \pm 6$	$m \pm 7$	$m \pm 8$	$m \pm 9$
% of the population	21.6	17.0	11.2	6.0	3.1	1.2	0.5	0.1	0.1	0.0

Table 2. The distribution of cells with different number of chromosomes in the equilibrium population. The assumptions involved in the calculation of this table are described in the text.

will give is presented in table 2 and illustrated in figure 1. (In this example we have only calculated the 11 first fitness classes, since the remaining ones are of very little importance.)

6.3 Discussion

This model is, as far as we know, the first mathematical model which tries to describe the selective forces working on cells in culture. However, other ideas expressed in different forms have been discussed in connection with karyotype evolution in vitro. The present model resembles partly the common theories about stem lines in tumors and established cell lines (Makino 1952, 1956): It is assumed that there is only one karyotype which is ideal in the cell line. But to this idea we have added two other features. First, that cells with almost the ideal karyotype are also almost as fit as the cells of the stem line, and secondly, that there are very many different karyotypes which are similar to the ideal karyotype. These factors taken together lead to situations, such as the one described by figure 1, where only a very small proportion of the cells in the cell line belongs to the stem line and where most cells have different karyotypes at the same time as they all are similar.

To assess the quality of the model it should be checked with empirical data. The new chromosome banding techniques have made it possible to obtain the kind of information which is relevant, but the task of analysing the karyologists' results and comparing them with theoretical models such as the present one, is difficult and could (and may) become the subject of a second thesis.

We shall here restrict ourselves to mentioning the three main limitations of the model described:

1) It was assumed in the model that all chromosome mutations are deleterious for the cells concerned. This is, of course, unrealistic even if

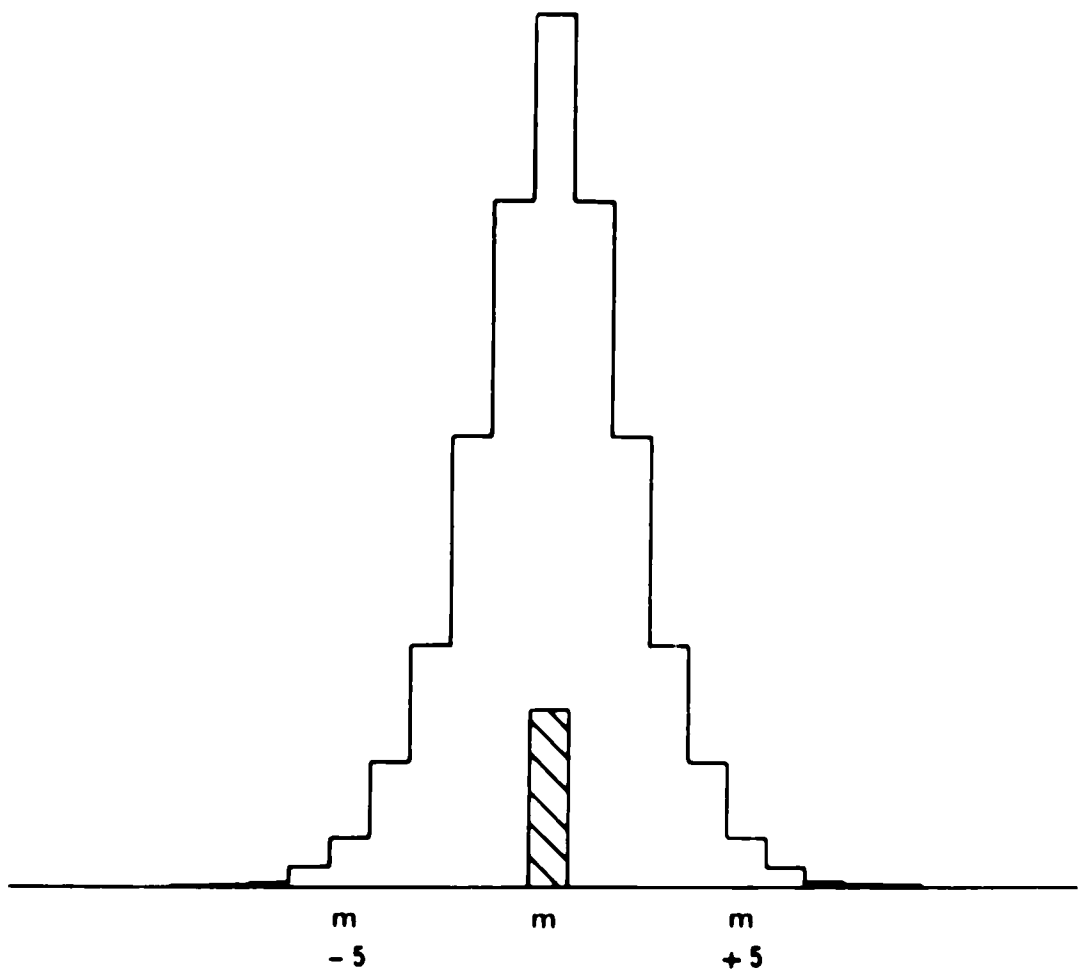


Figure 1. The distribution in the equilibrium population of cells with different chromosome numbers. Cells with the ideal karyotype are shaded.

the results will change only slightly if positive mutations are included. The matrix A is not triangular when positive mutations are considered, which makes it more difficult to determine its greatest eigenvalue. This problem can, however, easily be solved numerically by a computer. The limitation to deleterious mutations is thus not an important one.

2) A loss and a gain of a chromosome in respect to the ideal karyotype has in the model the same effect; It is also assumed that there is no difference in effects between the different chromosome types. Both these assumptions are crude. The problem can be circumvented if the connection between karyotypes and fitness is not specified. The population can always be regarded as divided into different fitness classes, and mutations can be seen as the changing of cells from one fitness class to another. The equilibrium population can thus be described by the distribution of cells in the different fitness classes, exactly as was done previously. However, if no connection between karyotypes and fitnesses is postulated, then the karyotype variability of the equilibrium population cannot be described.

3) Finally, it must be remembered that the model describes the chromosomal situation of a cell population close to an equilibrium. Nothing in the model can account for the appearance of a new ideal karyotype or a change from one ploidity level to another. The chromosome information discussed in the preceding chapter can thus not be interpreted in terms of the present model.

6.4 Appendix

A matrix theorem frequently used in population genetics

Define L as the column vector with m elements all equal to 1. If u is a row vector with m elements, then $u \cdot L$ is the sum of u's elements.

We say "the sum of a vector", when we mean the sum of the vector's elements.

A vector, or matrix, is non-negative, if all the elements of the vector or the matrix are zero or positive.

Theorem: If 1) A is a non-negative m x m matrix which can be diagonalized and has exactly one greatest eigenvalue to which corresponds a column eigenvector V_1 and a row eigenvector W_1;

2) z(0) is a non-negative row vector of sum 1;

3) $z(0) \cdot V_1 \neq 0$; and

4) z(i), i = 1,2,3, ... , is defined by

$$z(i) = z(i-1) \cdot A \ / \ z(i-1) \cdot A \cdot L \ ,$$

then

$$\lim_{i \to \infty} z(i) = W_1 \ / \ W_1 \cdot L \ ;$$

that is, the process described by assumption 4 goes towards a state given by the row eigenvector of sum 1 which corresponds to A's greatest eigenvalue.

The proof of this theorem can be given by methods found in any standard text on linear algebra.

6.5 References

Makino, S. 1952. A cytological study of the Yoshida sarcoma, an ascites tumor of white rats. Chromosoma 4: 649-674.

Makino, S. 1956. Further evidence favoring the concept of the stem cell in ascites tumors of rats. Ann. N.Y. Acad. Sci. 63: 818-830.