The Genetics of Recombination

D. G. Catcheside

The Genetics of Recombination

D. G. Catcheside, F.R.S.

Formerly Professor of Genetics, Australian National University, Canberra

Edward Arnold

First published 1977
by Edward Arnold (Publishers) Limited
25 Hill Street, London W1X 8LL

Boards edition ISBN: 0 7131 2612 4
Paper edition ISBN: 0 7131 2613 2

Filmset by Keyspools Ltd, Golborne, Lancs
Printed in Great Britain by William Clowes & Sons, Limited
London, Beccles and Colchester

Preface

Recombination is genetic analysis, the fourth and distinctive mode of biological study, the others being morphological description, chemical and physical methods of analysis. Functionally, recombination is commonly concerned in ensuring the accuracy of chromosome segregation during meiosis. It is implicated in the molecular parasitism of temperate bacteriophages initiated by the insertion of their DNA into those of their hosts. It is a mechanism for generating and conserving genetic variability so that new combinations are subjected to the tests of natural selection. It is the hallmark of sexual reproduction.

The purpose of this book is to examine the process itself. To do so fully at the molecular level proves to be difficult with present knowledge. Nevertheless the shape of the solution to the problem seems reasonably clear. In spite of the diversity of the derivative functions, which are the products of evolutionary advantage, it appears that the basic mechanisms of recombination have been conservative in evolution. Thus the mechanisms can be investigated, by a wide variety of means, by analysing the manifestations and properties of recombination wherever and however they are revealed. However, to do so requires consideration for each case of the reproductive context and means of analysis and so separate treatment.

Homologous DNA molecules, or chromosomes, given propinquity in an appropriate biological environment, will interact to produce recombinant versions of themselves. They do so by annealing of complementary polynucleotide chains contributed by each parental molecule to form local hybrid sections and the complexes of homologues can be resolved to yield intact molecules or chromosomes. In the course of resolution, a switch of parentage may occur at or near the hybrid section or the original parental molecules may be reformed substantially. In either case, the hybrid section is usually, if not always, retained in the molecules that emerge from the interaction. The hybrid section, at least, is subject to repair which involves local destruction and resynthesis of DNA. These events may occur *pari passu* with resolution. Recombination is an orderly sequence of events conducted, like other biological processes, under the supervision of a battery of enzymes, themselves determined by genes and subject to regulation. In higher organisms, the regulation extends to restriction on the extent of possible recombination by a mechanism, the synaptinemal

complex, which acts both to bring homologues into proximity and to prevent all but a minor fraction of each from interacting at the molecular level.

Molecular interaction ranges from the minimum of a fragment of single stranded DNA, contributed by a donor molecule to a competent recipient in bacterial transformation, to the maximum in which the interaction of parts of two complete molecules results in the formation of one or two molecules of joint parentage. The latter is best exemplified in the classical reciprocal crossing over and the conversion as seen in eucaryotes. Procaryotes provide examples over virtually the whole range.

This book attempts to describe the different aspects of recombination exhibited by various organisms and to relate the evidence each provides to the formulation of a revised theory of the mechanism. The areas of conflicting information about the processes, most marked among lower eucaryotes, are currently those most worthy of the effort to resolve.

The gene symbols used in practice for different organisms differ in a number of respects, most of which are relatively trivial. These include the use of numbers or letters to distinguish loci with names otherwise similar (e.g. *rec1*, *rec2* in yeast and *recA*, *recB* in *Escherichia coli*) and the presence or absence of a hyphen between locus name and number (e.g. *his-1* in *Neurospora crassa* and *his1* in yeast). In maize, it has been a frequent practice to place numbers, distinguishing different loci, as subscripts. However, it seems more convenient to write these on the line as *gl7*, rather than gl_7, for *glossy7*. It is usual to identify the 'wild' or common type allele at a locus by a superscript '+' (e.g. *his-1*$^+$ or *his1*$^+$); the wild type may be that of an arbitrarily chosen reference strain. Other mutant alleles are distinguished by letters or numbers or combinations thereof. In *Drosophila melanogaster* and maize these are written as superscripts (e.g. ry^8, wx^{90} and wx^{Coe}) but in most other organisms are written on the same line (e.g. *arg4-1* and *his1-7* in yeast). It would be defensible to enforce uniformity within the covers of one book, but this could cause trouble in going from the book to original literature. In consequence, it has been deemed wise to adhere to the usage for each organism as generally applied. In one respect only has a common practice been avoided. The book does not use a superscript '−' to designate mutants which lack a function present in the wild type; this is because the superscript '+' signifies the wild type rather than the presence of a function. The usage of '−' has arisen by confusion between gene symbol and phenotypic expression.

Acknowledgments

I gratefully acknowledge a Visiting Fellowship held in the Research School of Biological Sciences, Australian National University, during the tenure of which this book was composed and work on the control of recombination furthered. I owe very much to the encouragement of Professor Bernard John. Others have advised on particular sections,

notably Professors W. Hayes and C. J. Driscoll. However, the treatment, opinions and conclusions are wholly my own responsibility. I am greatly indebted to Mrs Erica Lockwood who kindly typed the whole book and to Miss Cathy Porter who drew the illustrations. The book is greatly enhanced by the photographs, for which I am grateful to Professor John and Professor Ditter von Wettstein and to Miss Diana Combes, who helped with Fig. 1.3.

Adelaide 1976 D.G.C.

Contents

Introduction

Genetics is autonomous and must not be mixed up with physico-chemical conceptions. M. Delbrück 1935.

Segregation and recombination are the tools of genetic analysis. Their occurrence in any organism constitutes evidence of sexual reproduction in that organism. The observation of recombination in bacteria and their viruses provided the first evidence of sexuality in these procaryotes.

Mendel was the first to enunciate these two properties of organisms. His first law stated that different heritable factors concerned with a given character segregate from one another in the formation of gametes and that in doing so each is unaltered by the other. These alternative genetic factors were later called allelomorphs or allelic genes. Mendel's second law stated that the segregation of different pairs of factors is independent, so that all combinations of non-allelic genes will be formed with equal frequency.

Within a few years of the rediscovery at the beginning of the twentieth century of Mendel's work and the confirmation of the general applicability of the laws to a wide range of plants and animals, exceptions to the second law were encountered. Parental combinations of non-allelic genes were, in some cases, found to be commoner among the gametes than were non-parental combinations. The latter ranged in frequency, according to the particular genes concerned, from nearly zero to nearly a half of the gametic output. All genes which show linkage in segregation belong to a linkage group within which they can be arranged unambiguously in a strict linear order. Genes which belong to different linkage groups commonly show independent segregation. Those in the same linkage group may show apparent independent segregation if they are far apart in the array, but will each show linkage to intermediate genes.

The number of linkage groups is characteristic of the species, being one in *Escherichia coli* and its viruses (such as T4 phage), four in *Drosophila melanogaster*, seven in *Pisum sativum* and ten in *Zea mays*. Except in special systems (see Rees and Jones, 1977), the number of linkage groups corresponds to the number of different chromosomes, the haploid number in each species. The genes of each linkage group are specified by segments of the DNA constituting the chromosomes. Recombination between

linked genes occurs when homologous chromosomes come together and pair intimately at one stage of the sexual cycle. The process involved in the exchange of information between the paired chromosomes is the subject of this book. While it has many derivative functions in fertility, physiology and evolution, its basic function lies in successfully closing the cycle of sexual reproduction. It may be assumed that because of its central role in sexual reproduction and its ubiquity, the mechanism of recombination is evolutionarily conservative. Hence the mechanism may be analysed by considering the manifestations of recombination wherever and however they can be observed (review: Emerson, 1969).

At first the event could be described in simple terms as one reciprocal exchange between two homologous chromosomes occurring with similar probability at the junction between any two gene loci. Measurements of frequency could therefore be expressed as genetic maps in which the successive loci were ordered at distances approximately proportional to the recombination between them. It was shown fairly early, making use of attached-X stocks in *Drosophila melanogaster* and of trisomics in *Zea mays*, that the event of crossing over occurred at the four strand stage of meiosis, between a chromatid of each of two homologues each consisting of two chromatids.

Advance in understanding the mechanism of recombination has come from two main sources and a few subsidiary ones. One main source was the close study of events when all products of each meiosis can be recovered and the relationship between the products analysed in detail.

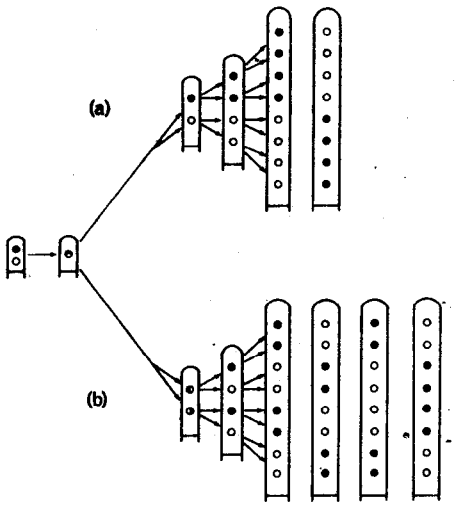

Fig. 1.1 Diagram of segregation in ordered asci: (a) at division I; (b) at division II.

A few species of fungi are the chief contributors. The other main source has been the study, especially in bacteria and their phages, of mutants which affect recombination. Behind both, in interpretation, was the realization that each chromatid is a double stranded structure, a molecule of DNA. Therefore purely mechanical break and rejoin, or even copy choice, theories were defective and had to give way to biochemical systems in which molecular events were catalysed by enzymes.

In some species of fungi, such as *Neurospora crassa* and *Sordaria fimicola*, the products of each meiosis are retained within a linear ascus in such a way that, subject to accidents, products each from one of the DNA strands of the homologues are arranged in a linear order. Counting the spores as 1 to 8 from apex to base of the ascus, the plane of the first division of meiosis lies between 4 and 5, while the planes of the second division lie between 2 and 3 and between 6 and 7 (Figs 1.1 and 1.2).

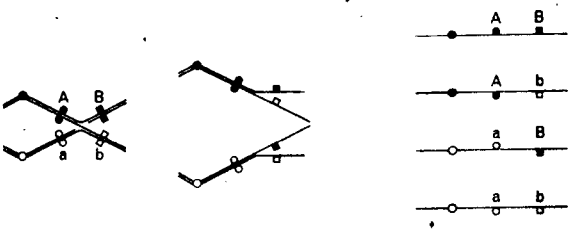

Fig. 1.2 Diagram to show relation of first and second division segregation of loci to a chiasma (or cross over). The centromeres and loci proximal to the chiasma show first division segregation, while distal loci show second division segregation. A and a are alleles at a proximal locus; B and b are alleles at a distal locus.

Each pair of spores (1 and 2, 3 and 4, 5 and 6, 7 and 8), which are sisters by a post-meiotic mitosis, represents at each homologous point the information carried by the two chains of DNA of a given chromatid as it goes through the later stages of meiosis. Segregation for one factor difference (say $+ \times m$) normally shows six types of ascus (Fig. 1.3), two with first and four with second division segregation of the factor followed (Table 1.1), all asci showing a 4:4 ratio. The frequency of second division segregation is characteristic of the particular locus, ranging from zero for those adjacent to the centromere to about 67% for those remote from the centromere. Moreover, the two first division classes are equal in frequency and the four second division classes are equal.. These equalities hold unless there is some genetic cause which biases the orientation of bivalents at meiosis and so makes the direction of segregation non-random, as in *Bombardia lunata* (Catcheside, 1944) where the segregation of a factor (*rubiginosa*) affecting ascospore colour can be followed (Table 1.1).

When regular segregation at more than one locus is followed in ordered asci, precise information is obtainable with respect to linkage and possible

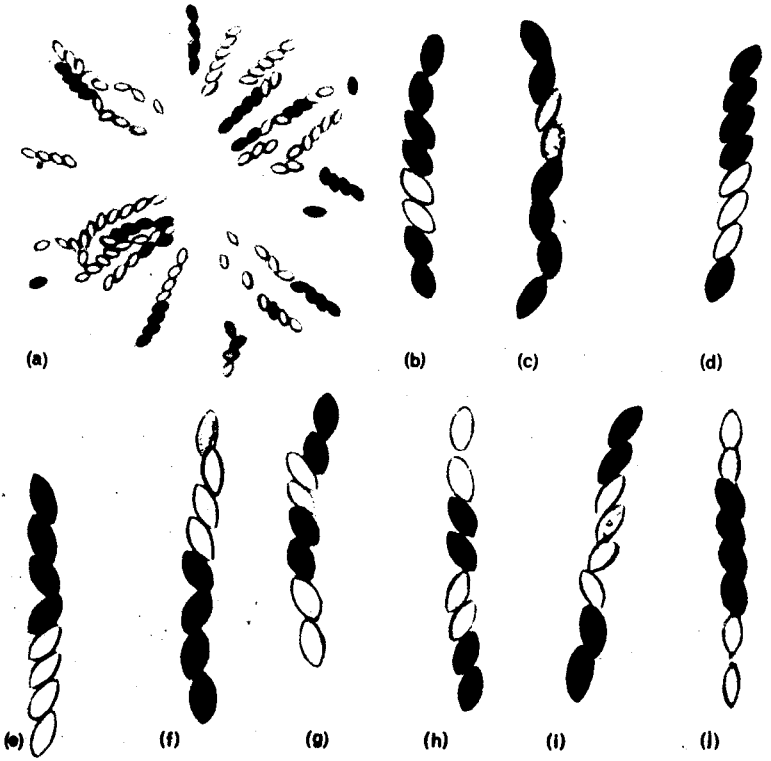

Fig. 1.3 *Neurospora crassa* asci from the cross *lysine-5* × +. The mutant *lysine-5* interferes with maturation of ascospores as well as causing a requirement for lysine; ascospores carrying *lysine-5* appear pale in asci which contain dark ripe normal ascospores. (a) is a cluster of asci. (e)–(j) are the six types shown by normal 4+ : 4*lys-5* asci. (b)–(d) are examples of abnormal asci, respectively two that are 6+ : 2*lys-5* and one that is 5+ : 3*lys-5*, presumably due to conversion of *lys-5* to normal. Apex of each ascus is above.

interference in crossing over. For two loci, there are 36 possible arrangements which may be grouped into seven modes (Table 1.2), according to whether segregation at each locus is first or second division and whether genes are recombined. The numbers of arrangements and modes increase rapidly with larger numbers of loci, to 32 modes for three loci and 172 modes for four loci. Considering the case of two loci, linkage is shown by the marked inequality of modes 1 and 2 in which both loci show first division segregation (Table 1.2). With two linked loci interference

Table 1.1 The six normal classes of segregation for a one factor difference in an ordered ascus, with illustrative data for the *mating type* locus of *Neurospora crassa* (Lindegren, 1932) and the *rubiginosa* locus ($r \equiv a$) of *Bombardia lunata* (Zickler, 1934).

Spore	Segregation					
	Division I		Division II			
1	A	a	A	a	A	a
2	A	a	A	a	A	a
3	A	a	a	A	a	A
4	A	a	a	A	a	A
5	a	A	A	a	a	A
6	a	A	A	a	a	A
7	a	A	a	A	A	a
8	a	A	a	A	A	a
N. crassa mt	105	129	9	5	10	16
B. lunata r	1687	1271	1518	1246	1174	1308

Table 1.2 The seven modes of normal segregation observed for two loci in ordered asci. The cross is AB × ab. Change of order between planes of first (I) and or second (II) division give rise to equivalent permutations.

Mode	1	2	3	4	5	6	7
II →	AB	Ab	AB	AB	AB	Ab	AB
	AB	Ab	AB	AB	AB	Ab	AB
	AB	Ab	Ab	aB	ab	aB	ab
I →	AB	Ab	Ab	aB	ab	aB	ab
	ab	aB	aB	Ab	AB	Ab*	Ab
II →	ab	aB	aB	Ab	AB	Ab	Ab
	ab	aB	ab	ab	ab	aB	aB
	ab	aB	ab	ab	ab	aB	aB
Permutations	2	2	8	8	4	4	8
Unlinked	$(1-x)(1-y)$	$(1-x)(1-y)$	$(1-x)y$	$x(1-y)$	$\frac{xy}{4}$	$\frac{xy}{4}$	$\frac{xy}{2}$
Linked opposite sides of centromere	2 $(1-x)(1-y)$	2 0	$(1-x)y$	$x(1-y)$	$\frac{xy}{4}$	$\frac{xy}{4}$	$\frac{xy}{2}$
Linked same side of centromere	$(1-x)(1-y+x)$	0	$(1-x)(y-x)$	$\frac{x(y-x)}{2}$	$x(1-y+x)$	0	$\frac{x(y-x)}{2}$

Note: The frequencies of the seven modes are given to a first approximation, neglecting double or more complex cross overs in all segments, in terms of the frequency of second division segregation of each locus, being x for Aa and y for Bb.

between two adjacent segments may be studied since the centromeres, which always segregate at division I, provide a third marked point.

Other fungi, such as yeast and *Ascobolus immersus*, keep the products of meiosis together in an unordered way. Thus with one locus segregating only one arrangement, 2A:2a or 4A:4a, is detected. With two loci three arrangements are detected. These are: 2AB:2ab, the parental ditype, comprising modes 1 and 5; 2Ab:2aB, the non-parental ditype, comprising modes 2 and 6; and 1AB:1ab:1Ab:1aB, the tetratype, comprising modes 3, 4 and 7. However, order can be imposed on such unordered tetrads or octads by means of genetic loci that show no recombination with the centromeres of the chromosomes in which they lie.

Among the first fruits of close study of *Neurospora crassa* was the confirmation of reciprocity in crossing over involving two out of four chromatids at any one place, with two, three or four strand relations showing in multiple cross overs. These occurred approximately in the 1:2:1 ratio predicted by chance coincidence of different events equally likely to involve any pair of non-sister chromatids. Extensive results from *Neurospora crassa*, summarized by Bole-Gowda, Perkins and Strickland (1962), show 423 two strand, 759 three strand and 329 four strand double cross overs as the pooled totals of a large number of studies. This represents a significant excess of two strand doubles over four stranded ones. However, other data obtained by Knapp and Möller (1955) with *Sphaerocarpus donnellii*, Strickland (1958) with *Aspergillus nidulans* and Ebersold and Levine (1959) with *Chlamydomonas reinhardi* do not show chromatid interference, though in some cases there is an excess of two stranded doubles.

In a small proportion of asci abnormal segregations are encountered in respect of a single factor difference, most being departures from the normal 4:4 segregation. These are usually 6+:2*m*, 2+:6*m*, 5+:3*m*, 3+:5*m* or abnormal 4+:4*m* (Table 1.3; some examples in Fig. 1.3). The last shows differences in respect of the locus concerned in two pairs of spores which by other criteria are sisters. The abnormalities are interpreted to arise by conversion of one gene to its allele. Lindegren (1953) followed Winkler (1932) in calling the phenomenon, expressed as departure from 4:4 or 2:2 segregation, gene conversion. Zickler (1934) was the first to prove instances of conversion. He observed 6+:2*m* in 2.46% of asci segregating at the *lactea* locus and in 0.45% at the *rubiginosa* locus, in *Bombardia lunata*. Although he also saw 2+:6*m* asci, no estimates of frequencies were reported because of possible confusion with asci containing unripe spores. However, the notion of conversion did not become accepted at all generally until later, beginning with Mitchell's (1955) work on *pyridoxin* mutants in *Neurospora crassa*. A notable feature of conversion at a locus is the correlation with crossing over in the immediate neighbourhood, making it possible to consider that the two were different consequences of one and the same event.

When recombination was found to occur at sites within a locus

Table 1.3 The normal (I) and five modes of abnormal (II–VI) segregation for a one-factor difference observed in an ordered ascus, with data observed for the *gray* locus (a = mutant, A = normal) of *Sordaria fimicola* by Kitani *et al.* (1962). The numbers of permutations of each mode take into account the various possible arrangements in an ascus; bracketed spores are sisters.

Mode	I	II	III	IV	V	VI
	A	A	A	A	A	A
	A	A	A	A	A	A
	A	A	a	A	A	A
	A	A	a	A	a	a
	a	A	a	A	a	A
	a	A	a	a	a	a
	a	a	a	a	a	a
	a	a	a	a	a	a
Permutations	6	4	4	24	24	48
Observed (i)	—	98	13	118	20	—
(ii)	—	←		141		→ 9
Frequencies inferred per 10^4	—	5	0.8	6	1	0.7

(Mitchell, 1955; Roman, 1956), the events were usually not reciprocal and also often required two or three classical exchanges over a very short region to account for some of the products. Examples of this sort gave rise to explanations of negative interference, one event increasing the chance of another close by. The non-reciprocal events showing, in a tetrad, as $1m^1:2m^2:1+$ for example, are more readily interpreted as gene conversion. They are the result of conversion of one mutant gene to wild type due to interaction with the allelic gene, rather than purely the result of a physical exchange. The occurrence of recombination between different sites within a locus allowed the construction, using frequencies of recombinants, of fine structure maps, an undertaking most fully pursued for the rII loci in T4 bacteriophage (Benzer, 1961). Moreover, at the molecular level, recombination may occur between adjacent nucleotides. This was demonstrated first by Yanofsky (1963) in *Escherichia coli*, using mutants causing different substitutions of one normal amino acid in the A polypeptide of tryptophan synthetase.

Theories of recombination began with ideas of breakage of non-sister chromatids at precisely corresponding sites followed by reunion in a new way. Apart from the problem of explaining how the exchanges occur with absolute precision so that unequal recombination does not occur, this mechanism does not account for non-reciprocal recombination. It was made to account for the occurrence of apparent multiple recombination

in a short segment (negative interference) by supposing that pairing was intermittent, as indeed it must be, and that in paired regions exchanges could be clustered.

Lederberg (1955) suggested that recombination in bacteria might be related to the process of replication of the hereditary material, which might be copied first from one parent and then, further along, from the other parent. This copy choice hypothesis was a revival of Belling's (1928) hypothesis of crossing over. Freese (1957) applied the idea to allelic recombination in *Neurospora crassa*. The hypothesis assumed that replication was conservative so that, by copying the hereditary material of a parental chromosome so as to leave the latter intact, a process of switching between two homologues might operate when these were close together. Some lack of synchrony between the synthesis of two daughters might allow both to copy from the same parent over a short interval, so generating 3:1 ratios. Repeated switching would account for multiple exchanges. This hypothesis suffers from serious faults. For example it takes no account of the evidence that both DNA (Meselson and Stahl, 1958) and chromosomes (Taylor, *et al.*, 1957) replicate semiconservatively. It does not explain post-meiotic segregation, nor the inequality of 6:2 and 2:6 ratios. It does not explain why recombination occurs with molecular precision and that actual exchanges of material occur. It does not predict that different pairs of chromatids are involved, apparently at random at different places along a bivalent. Modifications to the copy choice theory can be made to accommodate these and other objections, but it becomes cumbersome, inelegant and hardly plausible.

All subsequent theories propose the formation of a hybrid overlap (as a joint molecule or a heteroduplex) between a segment of a chain from one DNA molecule and its complement from another DNA molecule. The joint molecule is held together by hydrogen bonding. The manner by which such a joint molecule is established may not be the same in all organisms, but it seems to be a common ingredient in all recombination processes. The idea that recombination involved the formation of a hybrid segment occurred independently to several people at about the same time, notably Holliday (1962, 1964), Meselson (1964), Whitehouse (1963) and Taylor *et al.* (1962). All made more or less detailed proposals. The notion has the explicit merit of assuming that the process is precise at the molecular level, so that the breaks, later inferred, appear to be between exactly corresponding nucleotides of the two parents. No other processes offer the appearance of such exactness. The hypothesis also has the result that the joint molecule may have mispaired bases, due to genetic differences between the parents. Correction of these by removal and replacement might occur either by excision of just one of the mispaired bases or by excision of a length of one strand including a mispaired base. Repair of the gaps would complete the conversion. Failure of correction of a mispaired base during the meiotic stages would lead to post-meiotic segregation, following replication of the chromosome or chromosomes that

carried mispaired bases through the interphase.

Whitehouse (1965) pointed out that the five kinds of aberrant ascus observed by Kitani *et al.* (1962) in $+ \times g$ crosses of *Sordaria fimicola* are just what may be expected if a section of hybrid DNA is formed in two chromatids, provided that correction of mispairing does not always occur and that when it does it may occur in either direction, to normal or to mutant. If there were no correction of the mispairs, asci with $4+:4g$ and with half of them not in pairs would result (class VI in Table 1.3). The $5+:3g$ and $3+:5g$ ratios would result from correction at meiosis in one chromatid but not in the other, while the $6+:2g$ and $2+:6g$ would result from correction at meiosis in both chromatids in the same direction. Correction in both chromatids, but in opposite directions, would yield normal $4+:4g$ asci. The relative frequencies of the aberrant asci agree with this explanation. If the frequencies with which the hybrid DNA at the mutant site remained hybrid, became pure normal or pure mutant were x, y and z respectively the expected frequencies of the classes II to VI would be y^2, z^2, $2xy$, $2xz$ and x^2 respectively, assuming that the correction mechanism acts independently in the two chromatids. The observed frequencies are fitted, for the $+ \times g$ crosses in *Sordaria*, by $x = 0.27$, $y = 0.56$ and $z = 0.17$. The difference between y and z, the relative frequency of correction of the mispairing in opposite directions, might be related to the particular nature of the mutational change, whether a transition or transversion, or a deletion or insertion of bases, or to an extraneous difference. The post-meiotic segregation, implying the persistence of mispairing throughout meiosis until the synthetic (S) phase of the ensuing mitosis, could be due to failure of the postulated endonuclease to reach its substrate during a critical period of the meiotic prophase.

Whitehouse's analysis assumes that the mispaired bases in the two chromatids are alike, a consequence of his particular theory of how the segments of hybrid DNA are formed. Emerson (1966) showed that the analysis could be made more general, for the case in which the mispair in one chromatid is not like that in the other. This involves four unknown quantities, respectively p the probability of repair, with r the probability of repair to wild type, at one hybrid site, and q the probability of repair, with s the repair to wild type, at the other hybrid site. Applying the analysis to the *Sordaria fimicola* data of Kitani *et al.* (1962) and some data of Yu-Sun and himself on *Ascobolus immersus*, Emerson (1966) showed that good agreement to the data is given if $p \neq q$ and $r \neq s$ (Table 1.4). On the other hand, the *Ascobolus* data are not in agreement with $p = q$ and $r = s$, though this is fairly satisfactory for *Sordaria*.

For some time it was thought that conversion and crossing over could be independent processes. However, although the two types of events appear to be separately influenced by some conditions, the weight of evidence is that they are both the expression of a unitary process. Conversion is polarized, declining from more frequent conversion at one end of a locus to a lower frequency at the other (Lissouba *et al.*, 1962; Murray, 1963). So