

# Genetic differentiation in populations polymorphic for Robertsonian translocations

Hans-Rolf Gregorius and  
Sven Herzog

Abteilung für Forstgenetik und Forstpflanzenzüchtung  
Georg-August-Universität Göttingen Büsgenweg 2,  
3400 Göttingen, Federal Republic of Germany.

**Robertsonian translocations are characterized by centric fusion of two non-homologous acro- or telocentric chromosomes. Such translocations can be maintained in populations if in the chromosomally heterozygous individuals meiosis takes place regularly. In this case the question arises as to the existence of secondary isolating mechanisms that inhibit genetic exchange between the translocation chromosomes and their non-fused counterparts. The present paper treats this question by analyzing simultaneous changes at two gene loci, one located on each of the two constituent chromosomes. The analytic results are discussed in terms of opportunities for genetic differentiation to occur within a population polymorphic for Robertsonian translocations and are opposed to current theories on this subject.**

## INTRODUCTION

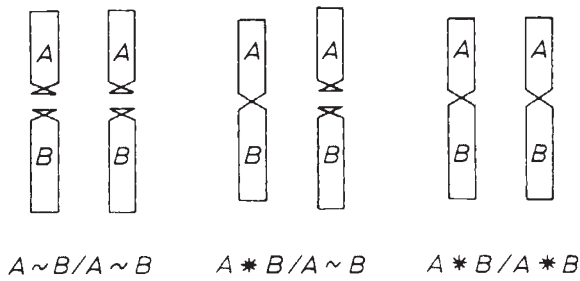
The centric fusion of two non-homologous acro- or telocentric chromosomes forming one meta- to subtelocentric chromosome results in the reduction of chromosome number in the set, while the "nombre fondamental" (Matthey 1945), *i.e.*, the total number of chromosome arms, remains constant. Chromosomal rearrangements of this type were described first by Robertson (1916) in grasshoppers. Since then, such "Robertsonian translocations" have been reported for many animal and plant taxa (see *e.g.*, White, 1973; John, 1976) and are assumed to be one of the fundamental mechanisms in animal karyotype evolution.

There is evidence that heterozygous Robertsonian translocations have the capacity to mal-segregate at meiosis but also to form balanced meiotic systems with regular segregation in the meiotic products. This means that during meiosis the homologous chromosome arms of the fused and the non-fused chromosomes are able to form stable symmetric trivalents if some basic requirements are fulfilled: the arm lengths of the fused chromosomes must be nearly equal, the frequency of chiasmata should not greatly exceed 1.0, and chiasmata should be localized preferably distally (White, 1973; Kaina and Rieger, 1979). By this mechanism, balanced systems of "Robertsonian

polymorphisms" with regular meiotic and mitotic segregation as well as homologous recombination can be generated and maintained in a population.

Chromosomal rearrangements may, of course, have different population genetic or evolutionary consequences, depending on whether meiotic segregation for the "chromosomally heterozygous" individuals occurs in a regular or an irregular fashion. As King (1987) emphasized, reproductive isolation mediated by chromosome rearrangements and thus the initiation of speciation processes may occur only for irregular segregation, whereas with regular segregation balanced (persistent) chromosomal polymorphisms are expected to evolve. However, it remains an open question as to whether such balanced polymorphisms entail unrestricted genetic exchange among the chromosomally heterozygous and the two chromosomally homozygous types.

The present paper will address this problem by considering the simultaneous changes at two gene loci assumed to be located at the two non-homologous chromosomes involved in the Robertsonian translocation system. It will be shown that the special recombination events between the two loci introduced by the chromosomally heterozygous and the two chromosomally homozygous types govern the amount and speed of genetic



**Figure 1** Schematic representation of Robertsonian translocations and their associations in the corresponding three chromosomal genotypes.

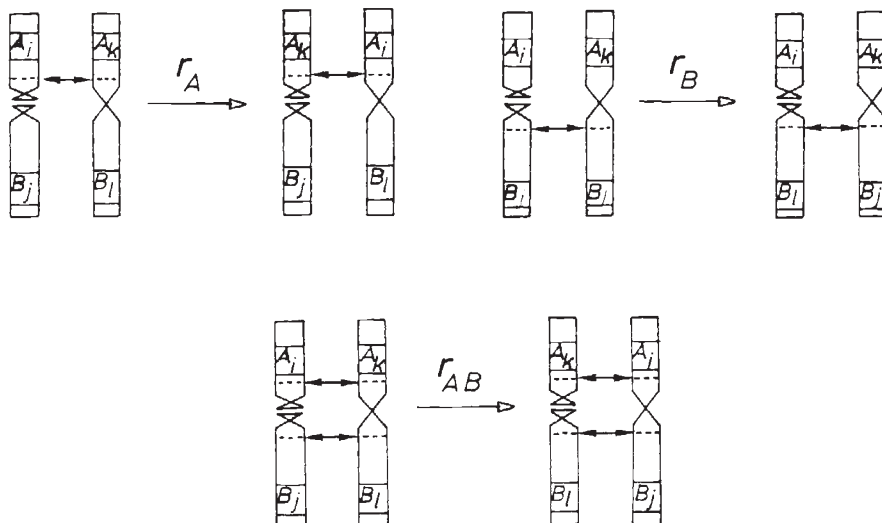
exchange at each of the loci in a population polymorphic for Robertsonian translocations.

**FORMAL REPRESENTATION**

Let *A* and *B* be two non-homologous chromosomes which, as a result of (centric) chromosomal fusion, may also occur in the population as a single chromosome denoted by *A\*B*. Hence, the gametes formed in this population may contain either the fused type *A\*B* or a type *A~B* in which *A* and *B* are not fused. As a result the population consists of two kinds of "chromosomal homozygotes", namely *A\*B/A\*B* and *A~B/A~B*, and one "chromosomal heterozygote" *A\*B/A~B* (fig. 1). Two gene loci, one on each of the two chromosomes, will be considered, and

for simplicity the alleles at these loci will again be identified by the symbols *A* and *B*, so that *A<sub>i</sub>* and *B<sub>j</sub>* denote an allele at the *A*- and *B*-locus, respectively. The two-locus gametic types *A<sub>i</sub>B<sub>j</sub>* will be distinguished with respect to the two chromosomal types in which they may occur, and this is reflected by the corresponding notation *A<sub>i</sub>\*B<sub>j</sub>*, *A<sub>i</sub>~B<sub>j</sub>*.

Because the gametes formed may be either of type *A\*B* or of type *A~B*, five recombination frequencies have to be distinguished depending on whether the diploid genotype is chromosomally homozygous or heterozygous. Thus the recombination frequency between the *A* and *B* locus is  $\frac{1}{2}$  for the chromosomal homozygote *A~B/A~B*, but it may differ from  $\frac{1}{2}$  in the chromosomal homozygote *A\*B/A\*B*. Concerning the chromosomal heterozygote *A\*B/A~B*, three recombination events have to be taken into account. One leads to a genetic exchange between the gametic types *A\*B* and *A~B* only at the *A*-locus by cross-overs located between the *A*-locus and the centromere. Analogously, genetic exchanges between the two gametic types are restricted to the *B*-locus by cross-overs located between this locus and the centromere. As a third recombination event between the two gametic types, a genetic exchange may occur at both loci simultaneously when cross-overs occur between the *A*-locus and the centromere as well as between the *B*-locus and the centromere. In a chromosomal homozygote this event is usually not considered as a recombination since all gametes formed carry the same chromosomal types



**Figure 2** Schematic representation of recombination events among the gametes of a chromosomal heterozygote and associated recombination frequencies. The symbol  $\leftrightarrow$  indicates cross-over positions.

and since, at the loci under observation, a double exchange is tantamount to no change. Hence, the following recombination frequencies (also see fig. 2) are taken into consideration:

$r$  = recombination fraction among the gametes of the chromosomal homozygote  $A * B / A * B$ .

$r_A$  = relative frequency of genetic exchanges realized at the  $A$ -locus only among the gametes of a chromosomal heterozygote  $A * B / A \sim B$ .

$r_B$  = relative frequency of genetic exchanges realized at the  $B$ -locus only among the gametes of a chromosomal heterozygote  $A * B / A \sim B$ .

$r_{AB}$  = relative frequency of genetic exchanges realized at both loci simultaneously among the gametes of a chromosomal heterozygote  $A * B / A \sim B$ .

$$r' = r_A + r_B + r_{AB}$$

Each diploid two-locus genotype will thus produce gametic types according to the frequencies given in table 1, and the frequencies of the gametic types as they are represented in the zygotes produced by the population will be denoted by

$g_{i*j}$  = relative frequency of  $A_i * B_j$ -gametes among the zygotes;

$g^*$  =  $\sum_{ij} g_{i*j}$  or relative frequency of gametes containing chromosomal type  $A * B$  among the zygotes;

$g_{i\sim j}$  = relative frequency of  $A_i \sim B_j$ -gametes among the zygotes;

$g^-$  =  $\sum_{ij} g_{i\sim j} = 1 - g^*$  or relative frequency of gametes containing the  $A \sim B$  chromosomal combination;

$g_{ij} = g_{i*j} + g_{i\sim j}$  or relative frequency of  $A_i A_j$ -gametes among the zygotes.

It follows that the corresponding allelic frequencies at the two loci are

$$a_i^* = \sum_l g_{i+l}, \quad b_j^* = \sum_k g_{k+j},$$

so that

$$\sum_i a_i^* = \sum_j b_j^* = g^*,$$

$$a_i^- = \sum_l g_{i-l}, \quad b_j^- = \sum_k g_{k-j},$$

so that

$$\sum_i a_i^- = \sum_j b_j^- = g^-$$

$$a_i = a_i^* + a_i^-, \quad b_j = b_j^* + b_j^-.$$

The dynamics of the gametic frequencies will be considered under the absence of selection at both loci, regular segregation, and random mating in a population reproducing in separated generations. The transition equations ((A1a) and (A1b)) and their analysis are presented in the Appendix.

As was to be expected from the absence of selection, the overall allelic frequencies  $a_i$  and  $b_j$  at the  $A$ - and  $B$ -locus do not change over the generations, and the same holds true for the frequencies  $g^*$  and  $g^-$  of the gametic types  $A * B$  and  $A \sim B$ , respectively. On the other hand, the allelic frequencies specific to the gametic types  $A * B$  and  $A \sim B$  (i.e.  $a_i^*$  etc.) may change if they differed initially (i.e.  $a_i^* \neq a_i^-$  and/or  $b_j^* \neq b_j^-$ ). This depends solely on the recombination fractions  $r_A$ ,  $r_B$ ,  $r_{AB}$  characteristic of the chromosomal heterozygote: the recombination fraction  $r$  of the chromosomal homozygote  $A * B / A * B$  has no effect on the dynamics of the allelic frequencies (see equation (A2) in the Appendix).

In particular, if  $r_{AB} > 0$  or both  $r_A > 0$  and  $r_B > 0$ , then the allelic frequencies within each of the gametic types  $A * B$  and  $A \sim B$  converge at both loci to their respective overall frequencies. Thus  $a_i^* / g^*$  and  $a_i^- / g^-$  converge to  $a_i$ , and  $b_j^* / g^*$  and  $b_j^- / g^-$  converge to  $b_j$ . Hence,  $r_{AB} = 0$  is a necessary condition for constancy of the above allelic frequencies at any one of the two loci. If in addition  $r_A = 0$  and  $r_B > 0$ , then the alleles at the  $A$ -locus maintain their initial frequencies, while at the  $B$ -locus they converge to their respective overall

**Table 1** Frequencies of gametic types in the gametic output of the various two-locus genotypes

Genotype	Gametes from	
	Coupling	Recombination
$A_i * B_l / A_k * B_l$	$\frac{1}{2}(1-r)A_i * B_l, \frac{1}{2}(1-r)A_k * B_l$	$\frac{1}{2}rA_i * B_l, \frac{1}{2}rA_k * B_l$
$A_i \sim B_l / A_k \sim B_l$	$\frac{1}{4}A_i \sim B_l, \frac{1}{4}A_k \sim B_l$	$\frac{1}{4}A_i \sim B_l, \frac{1}{4}A_k \sim B_l$
$A_i * B_l / A_k \sim B_l$	$\frac{1}{2}(1-r')A_i * B_l, \frac{1}{2}(1-r')A_k * B_l$	$\frac{1}{2}r_A A_i \sim B_l, \frac{1}{2}r_B A_k * B_l$
		$\frac{1}{2}r_{AB} A_i * B_l, \frac{1}{2}r_{AB} A_k \sim B_l$
		$\frac{1}{2}r_{AB} A_i \sim B_l, \frac{1}{2}r_{AB} A_k * B_l$

frequencies  $b_j$ . This role played by the two loci is reversed if  $r_B = 0$  and  $r_A > 0$ . Only if all three recombination fractions are zero (and thus  $r' = 0$ ) can all alleles at both loci maintain their initial frequencies. This is tantamount to the absence of recombination between the two loci among the gametes produced by the chromosomal heterozygote.

Again these results conform with intuition when viewed in the light of the particular recombination fractions utilized here.

SIMULTANEOUS CHANGE AT THE TWO LOCI

The main concern in studies of simultaneous change at loci lies in the specification of stochastic associations of genes on gametes. The present model differs from conventional models in that two types of gametes,  $A * B$  and  $A \sim B$ , can be distinguished for which stochastic associations of genes can be considered separately. Stochastic independence of the two loci among  $A * B$ -gametes would thus be characterized by  $g_{i+j}/g^* = (a_i^*/g^*) \cdot (b_j^*/g^*)$ . Similarly, stochastic independence among  $A \sim B$ -gametes means  $g_{i\sim j}/g^- = (a_i^-/g^-) \cdot (b_j^-/g^-)$ . A third possibility arises when the two types of gametes are considered as indistinguishable, in which case stochastic independence is indicated by  $g_{ij} = a_i \cdot b_j$ . This relates to the conventional situation and will be referred to as “ $AB$ -independence”, while the other two forms of stochastic independence will be addressed by the prefix  $A * B$  and  $A \sim B$ , respectively.

As is shown in the Appendix, the gametic frequencies always converge to limiting values, so that considerations of stochastic independence of the loci can be restricted to the limiting gametic

frequencies approached over the generations. These equilibrium frequencies will be denoted by  $\hat{g}$ , while frequency symbols without a circumflex refer to the respective initial frequencies. The results derived in the Appendix are summarized in table 2.

As is suggested by the free recombination among gametes from the chromosomal homozygote  $A \sim B/A \sim B$ , the association of genes at  $A \sim B$ -gametes always approaches a state of stochastic independence. This does not carry over to  $A * B$ -gametes unless recombination is realized among the gametes of the chromosomal homozygote  $A * B/A * B$  ( $r > 0$ ) or the chromosomal heterozygote ( $r' > 0$ ). In fact, if recombination is absent for these two chromosomal genotypes ( $r = r' = 0$ ), the frequencies of  $A * B$ -gametes remain constant over the generations so that an initially present non-random association of genes on  $A * B$ -gametes is maintained.

However, even if  $A * B$ -independence is approached, the gametic frequencies among  $A * B$ - and  $A \sim B$ -gametes need not be identical. This is a consequence of initial differences in the allelic frequencies between the two gametic types in combination with certain recombination patterns of the chromosomal heterozygote that maintain these differences at one but not at the other locus (as was addressed explicitly in the last section).

On the other hand, despite the fact that  $A * B$ - and  $A \sim B$ -gametes may differ in frequency, overall stochastic independence (i.e.  $AB$ -independence) is reached asymptotically as long as recombination takes place for the chromosomal heterozygote ( $r' > 0$ ). This holds irrespective of the recombination pattern realized among the constituents  $r_A$ ,  $r_B$ , and  $r_{AB}$  of  $r'$ . While the previous results may be amenable to intuition after some thought, the last state-

Table 2 Asymptotic association of the two gene loci

Type of association	Conditions	Equilibrium frequencies
$A \sim B$ -independence	always	$\hat{g}_{i\sim j}/g^- = a_i \cdot b_j$ if $r_{AB} > 0$ or $r_A \cdot r_B > 0$ $\hat{g}_{i\sim j}/g^- = a_i \cdot (b_j^-/g^-)$ if $r_{AB} = r_B = 0 < r_A$ $\hat{g}_{i\sim j}/g^- = (a_i^-/g^-) \cdot b_j$ if $r_{AB} = r_A = 0 < r_B$ $\hat{g}_{i\sim j}/g^- = (a_i^-/g^-) \cdot (b_j^-/g^-)$ if $r' = 0$
$A * B$ -independence	$r + r' > 0$	$\hat{g}_{i+j}/g^* = a_i \cdot b_j$ if $r_{AB} > 0$ or $r_A \cdot r_B > 0$ $\hat{g}_{i+j}/g^* = a_i \cdot (b_j^*/g^*)$ if $r_{AB} = r_B = 0 < r_A$ $\hat{g}_{i+j}/g^* = (a_i^*/g^*) \cdot b_j$ if $r_{AB} = r_A = 0 < r_B$ $\hat{g}_{i+j}/g^* = (a_i^*/g^*) \cdot (b_j^*/g^*)$ if $r' = 0$
$AB$ -independence	$r' > 0$	$\hat{g}_{ij} = a_i \cdot b_j$
$A \sim B$ -dependence	never	
$A * B$ -dependence	$r = r' = 0$ , $g^* \cdot g_{i+j} \neq a_i^* \cdot b_j^*$	$\hat{g}_{i+j} = g_{i+j}$
$AB$ -dependence	$r = 0$ , $a_i^* \neq a_i$ , $b_j^* \neq b_j$	$\hat{g}_{ij} = g_{i+j} + (a_i^- \cdot b_j^-/g^-)$ if $r = 0$ $\hat{g}_{ij} = (a_i^* \cdot b_j^*/g^*) + (a_i^- \cdot b_j^-/g^-)$ if $r > 0$

ment is probably non-trivial since it includes the case  $r = 0$  in which recombination is prohibited for the gametes of the chromosomal homozygote  $A * B / A * B$ .

Finally, the asymptotic frequency distributions among  $A * B$ - and  $A \sim B$ -gametes are identical for all initial conditions only if either  $r_{AB} > 0$  or  $r_A, r_B > 0$ , in which case  $AB$ -independence prevails, of course. Thus, this is the only case for which the gametic frequency distributions of all three classes  $A * B$ ,  $A \sim B$ , and  $AB$  of gametes become asymptotically identical.

## DISCUSSION

At first sight, Robertsonian polymorphisms in the form studied here may be considered to resemble common models of recombination modification, where the three chromosomal diplotypes correspond to a diploid modifier locus with two alleles represented by the gametic types  $A * B$  and  $A \sim B$  (see Nei 1967 and cf. Liberman and Feldman, 1986, for a more recent treatment of the subject). However, neither the structure nor the conceptual basis of these models appears to allow for a re-interpretation that would make them compatible with the cytogenetic mechanisms characteristic of Robertsonian translocations. For the present we shall therefore refrain from comparing the present results with those obtained for the common models of recombination modification.

In order to evaluate the evolutionary significance of the present findings, it is useful to envisage two populations belonging to the same biological species, where one population is fixed for the gametic type  $A * B$  and the other is fixed for the gametic type  $A \sim B$ . In two such populations, loci located on the  $A$ - or  $B$ -sections may differ in allelic frequencies up to the extreme case where the populations do not share certain alleles. Now, if the isolating barriers are removed and the two populations hybridize to form a single population, the question arises as to the degree of genetic mixing that will be realized.

For example, in red deer (*Cervus elaphus* L.) and Sika deer (*Cervus nippon* Temminck) chromosome numbers of  $2n = 68$  and  $2n = 64$ , respectively, were found in non-hybrid populations. In hybrid populations, however, chromosome numbers vary between 64 and 68, depending on the presence or absence of two submetacentric Robertsonian translocation pairs or single chromosomes (Harrington and Power 1985). On the basis of morphological criteria, both red and Sika deer are

considered as different species. Yet, from cytogenetic and biochemically genetic studies it can be concluded that both "species" are capable of free interbreeding, and no evidence for pre- or post-zygotic incompatibility was observed, so that they ought to be considered as a single biological species (Herzog, unpublished data).

Robertsonian translocations provide an efficient tool for studying this question, since the two gametic types  $A * B$  and  $A \sim B$  representing the original populations can be traced in the hybrid population. The degree of genetic mixing can thus be measured by the amount of allelic exchanges among the two gametic types, which, in turn, is determined by the amount of homologous recombination detectable in the gametic output of the chromosomal heterozygotes  $A * B / A \sim B$ . This became evident in the transition equations for the allelic frequencies, since the change in these frequencies turned out to depend solely on the recombination fractions  $r_A + r_{AB}$  for the  $A$ - and  $r_B + r_{AB}$  for the  $B$ -locus. Note that  $r_A + r_{AB}$ , for example, measures exactly the amount of allelic exchange between the gametic types  $A * B$  and  $A \sim B$  at the  $A$ -locus. Hence, on the basis of the two gametic types  $A * B$  and  $A \sim B$ ,  $r_A + r_{AB} = 0$  would conserve the allelic frequencies of the two initial populations and would thus prevent genetic mixing in the hybrid population at this particular locus. Though essentially trivial, this aspect deserves emphasis, since it is sometimes not recognized as a major evolutionary consequence of the absence of recombination.

The last statements focus on genetic mixing at a single locus and may thus be viewed as "intra-locus" phenomena of genetic exchange triggered by (homologous) recombination. On the other hand, recombination refers by definition to "inter-locus" associations, i.e., associations *between* rather than *within* gene loci. It is thus tempting to conclude that each type of association implies the other. Yet, the present analyses reveal that this has to be viewed with caution. If  $r' > 0$ , complete inter-locus mixing in the sense of stochastic independence or linkage equilibrium between loci is approached asymptotically at all three gametic levels  $A * B$ ,  $A \sim B$ , and  $AB$  (see Table 2). However,  $r' > 0$  includes the case  $r_{AB} = r_A = 0 < r_B$  for which genetic exchange at the  $A$ -locus and thus intra-locus exchange is prohibited.

Therefore, complete inter-locus mixing between two gene loci need not imply complete intra-locus mixing at each of the two loci (note that the term "mixing" as used here should not be confused with "recombination"). In the above

example this holds true despite the fact that, in the gametic output of the chromosomal heterozygote, the recombination fraction between the two loci (which is given by  $r_A + r_B$ ) is positive. The reason for this is to be found in the fact that  $r_{AB} = r_A = 0 < r_B$  excludes recombination between loci on the chromosome *A* but allows for it on the chromosome *B*. The same applies of course to the *A*- and *B*-section, respectively, of the chromosome  $A * B$ . It follows that Robertsonian translocations may aid the maintenance of the genetic identity of previously isolated populations at some gene loci after removal of the isolation, while at other loci complete mixing may evolve over the generations. Clearly, here the notion of genetic identity refers to the two gametic types only, and, when applied to the diploid individuals, characterizes properties of the two chromosomal homozygotes.

In this way, two subpopulations corresponding to the two chromosomal homozygotes can be distinguished which may differ genetically at certain loci and are furnished with the capacity to even amplify this initial difference by successive differential accumulation of mutations on the gametic types  $A * B$  and  $A \sim B$ . Notwithstanding random mating and regular segregation in the total population, genetic differentiation between cytologically defined subpopulations can thus evolve which, at later stages, may lead to the initiation of sympatric speciation processes. This result should be viewed in connection with King's (1987) statement that "if any of these (chromosomal) rearrangements produce a balanced meiotic system, in which normal segregation of the meiotic products occurs, they will no longer be able to form a postmating isolating mechanism". As the above considerations demonstrate, this is not of general validity.

*Acknowledgement* This work was supported by a grant from the Deutsche Forschungsgemeinschaft. The authors gratefully recognize the helpful comments of an anonymous reviewer.

## REFERENCES

- HARRINGTON, R. AND POWER, M. 1985. Karyotyping within the genus *Cervus*. *1st Symposium on Genetics of Wild Animals (Cytogenetics and Biochemical Genetics)*. Gießen, pp. 66-75.
- JOHN, B. 1976. *Population Cytogenetics*. Edward Arnold Limited, London.
- KAINA, B. AND RIEGER, R. 1979. Chromosomenmutation, Karyotyp-Polymorphismus und Speciation. *Biol. Zbl.*, 98, 661-697.

- KING, M. 1987. Chromosomal rearrangements, speciation and the theoretical approach. *Heredity*, 59, 1-6.
- LIBERMAN, U. AND FELDMAN, M. W. 1986. A general reduction principle for genetic modifiers of recombination. *Theoret. Popul. Biol.*, 30, 341-371.
- MATTHEY, R. 1945. L'évolution de la formule chromosomiale chez les vertébrés. *Experientia* 1, 50-56 and 78-86.
- NEI, M. 1967. Modification of linkage intensity by natural selection. *Genetics*, 57, 625-641.
- ROBERTSON, W. 1916. Chromosome studies. I. Taxonomic relationships shown in the chromosomes of *Tettigidae* and *Acrididae*: V-shaped chromosomes and their significance in *Acrididae*, *Locustidae*, and the *Gryllidae*. *J. Morph.*, 27, 179-331.
- WHITE, M. J. D. 1973. *Animal Cytology and Evolution*. Cambridge University Press, London.

## APPENDIX

The frequency  $g'_{i*j}$  of the gametic type  $A_i * B_j$  in the next generation of zygotes resulting from random mating and regular segregation is obtained as follows:

$$\begin{aligned}
 g'_{i*j} = & (g_{i*j})^2 + (1-r)g_{i*j} \sum_{k \neq i, l \neq j} g_{k*l} \\
 & + g_{i*j} \sum_{l \neq j} g_{i*l} + g_{i*j} \sum_{k \neq i} g_{k*j} + r \sum_{l \neq j, k \neq i} g_{i*l}g_{k*j} \\
 & + g_{i*j}g_{i\sim j} + (1-r')g_{i*j} \sum_{k \neq i, l \neq j} g_{k\sim l} \\
 & + (1-r' + r_A)g_{i*j} \sum_{l \neq j} g_{i\sim l} \\
 & + (1-r' + r_B)g_{i*j} \sum_{k \neq i} g_{k\sim j} + r_B \sum_{l \neq j, k \neq i} g_{i*l}g_{k\sim j} \\
 & + r_{AB}g_{i\sim j} \sum_{k \neq i, l \neq j} g_{k*l} + (r_B + r_{AB})g_{i\sim j} \sum_{l \neq j} g_{i*l} \\
 & + (r_A + r_{AB})g_{i\sim j} \sum_{k \neq i} g_{k*j} + r_A \sum_{l \neq j, k \neq i} g_{i\sim l}g_{k*j}.
 \end{aligned}$$

Carrying out the sums, one obtains

$$\begin{aligned}
 g'_{i*j} = & (g_{i*j})^2 + (1-r)g_{i*j}(g^* - a_i^* - b_j^* + g_{i*j}) \\
 & + g_{i*j}(a_i^* - g_{i*j}) + g_{i*j}(b_j^* - g_{i*j}) \\
 & + r(a_i^* - g_{i*j})(b_j^* - g_{i*j}) + g_{i*j}g_{i\sim j} \\
 & + (1-r')g_{i*j}(g^{\sim} - a_i^{\sim} - b_j^{\sim} + g_{i\sim j}) \\
 & + (1-r' + r_A)g_{i*j}(a_i^{\sim} - g_{i\sim j}) \\
 & + (1-r' + r_B)g_{i*j}(b_j^{\sim} - g_{i\sim j}) \\
 & + r_B(a_i^* - g_{i*j})(b_j^{\sim} - g_{i\sim j}) \\
 & + r_{AB}g_{i\sim j}(g^* - a_i^* - b_j^* + g_{i*j}) \\
 & + (r_B + r_{AB})g_{i\sim j}(a_i^* - g_{i*j}) \\
 & + (r_A + r_{AB})g_{i\sim j}(b_j^* - g_{i*j}) \\
 & + r_A(a_i^{\sim} - g_{i\sim j})(b_j^* - g_{i*j})
 \end{aligned}$$

which, after collection of terms yields the final form

$$g'_{i*j} = g_{i*j} \cdot (1 - r \cdot g^* - r' \cdot g^-) + g_{i\sim j} \cdot r_{AB} \cdot g^* + r \cdot a_i^* \cdot b_j^* + r_A \cdot a_i^- \cdot b_j^* + r_B \cdot a_i^* \cdot b_j^- \quad (\text{A1a})$$

The transition equation for  $g_{i\sim j}$  can be directly obtained from the above equation by replacement of \* by  $\sim$ ,  $\sim$  by \*, and  $r$  by  $\frac{1}{2}$ . Thus

$$g'_{i\sim j} = g_{i\sim j} \cdot (1 - \frac{1}{2} \cdot g^- - r' \cdot g^*) + g_{i*j} \cdot r_{AB} \cdot g^- + \frac{1}{2} \cdot a_i^- \cdot b_j^- + r_A \cdot a_i^* \cdot b_j^- + r_B \cdot a_i^- \cdot b_j^* \quad (\text{A1b})$$

Forming the respective sums, one arrives at the transition equations for the allelic frequencies:

$$\begin{aligned} a_i^{*'} &= a_i^* + (r_A + r_{AB}) \cdot (a_i^- \cdot g^* - a_i^* \cdot g^-) \\ b_j^{*'} &= b_j^* + (r_B + r_{AB}) \cdot (b_j^- \cdot g^* - b_j^* \cdot g^-) \\ a_i^{-'} &= a_i^- + (r_A + r_{AB}) \cdot (a_i^* \cdot g^- - a_i^- \cdot g^*) \\ b_j^{-'} &= b_j^- + (r_B + r_{AB}) \cdot (b_j^* \cdot g^- - b_j^- \cdot g^*) \end{aligned}$$

so that

$$g^{*'} = g^*, \quad g^{-'} = g^-, \quad a_i^{\prime} = a_i, \quad b_j^{\prime} = b_j.$$

### Convergence

Recalling that  $a_i^* + a_i^- = a_i$ ,  $b_j^* + b_j^- = b_j$ , and  $g^* + g^- = 1$ , the above transition equations for the allelic frequencies can be rewritten in the form

$$\begin{aligned} a_i^{*'} - a_i \cdot g^* &= (a_i^* - a_i \cdot g^*)(1 - r_A - r_{AB}) \\ b_j^{*'} - b_j \cdot g^* &= (b_j^* - b_j \cdot g^*)(1 - r_B - r_{AB}) \\ a_i^{-'} - a_i \cdot g^- &= (a_i^- - a_i \cdot g^-)(1 - r_A - r_{AB}) \\ b_j^{-'} - b_j \cdot g^- &= (b_j^- - b_j \cdot g^-)(1 - r_B - r_{AB}) \end{aligned} \quad (\text{A2})$$

which shows that the dynamics of the allelic frequencies are not affected by the recombination fraction  $r$ , and that convergence to  $a_i g^*$ , etc., occurs whenever at least one of the recombination fractions involved is positive. Otherwise, if e.g.  $r_A = r_{AB} = 0$ , the initial frequencies of the alleles at the  $A$ -locus remain constant over the course of the generations. If convergence to  $a_i g^*$  occurs, this has as a consequence that the frequency of  $A_i$  both among the  $A * B$  and among the  $A \sim B$  gametic types approaches the initial frequency of  $A_i$ , i.e.  $a_i$ , among all zygotes in the population.

Since all allelic frequencies converge to a unique equilibrium or remain constant, these limiting frequencies can be inserted into the transition equations (A1) for the gametic frequencies in order to simplify the analysis of convergence. The cases  $r_{AB} > 0$  and  $r_{AB} = 0$  will be considered separately.

**$r_{AB} > 0$ :** In this case  $a_i^* b_j^- = a_i^- b_j^* = a_i b_j g^* g^-$ ,  $a_i^* b_j^* = a_i b_j \cdot (g^*)^2$ , and  $a_i^- b_j^- = a_i b_j \cdot (g^-)^2$ , so that the transition equations (A1) can be rewritten in the form

$$\begin{aligned} g'_{i*j} - a_i b_j g^* &= (g_{i*j} - a_i b_j g^*)(1 - r g^* - r' g^-) \\ &\quad + (g_{i\sim j} - a_i b_j g^-) r_{AB} g^* \\ g'_{i\sim j} - a_i b_j g^- &= (g_{i\sim j} - a_i b_j g^-)(1 - \frac{1}{2} g^- - r' g^*) \\ &\quad + (g_{i*j} - a_i b_j g^*) r_{AB} g^- \end{aligned}$$

Thus

$$\begin{aligned} |g'_{i*j} - a_i b_j g^*| + |g'_{i\sim j} - a_i b_j g^-| \\ \leq |g_{i*j} - a_i b_j g^*| \cdot (1 - r g^* - (r_A + r_B) g^-) \\ + |g_{i\sim j} - a_i b_j g^-| \cdot (1 - \frac{1}{2} g^- - (r_A + r_B) g^*). \end{aligned}$$

For  $r > 0$  or  $r_A + r_B > 0$  both factors on the right side of the inequality are constant and  $< 1$ . Hence,  $g_{i*j}$  and  $g_{i\sim j}$  converge to  $a_i b_j g^*$  and  $a_i b_j g^-$ , respectively, and  $g_{ij}$  converges to  $a_i b_j$ . On the other hand, if  $r = r_A + r_B = 0$  then  $g_{i\sim j}$  converges to  $a_i b_j g^-$ , and, since  $r' = r_{AB} > 0$ , this entails that  $g_{i*j}$  must also converge to  $a_i b_j g^*$ .

**$r_{AB} = 0$ :** The transition equations for  $g_{i*j}$  and  $g_{i\sim j}$  are now mutually independent and linear. Thus, convergence is again guaranteed, and the equilibrium frequencies  $\hat{g}$  approached are as follows:

- (i) If  $r = r' = 0$ , the gametic frequencies  $g_{i*j}$  as well as all allelic frequencies remain constant.  $g_{i\sim j}$  converges to  $a_i^- b_j^- / g^-$  and, in general,  $\hat{g}_{ij} \neq a_i b_j$ .
- (ii) If  $r > 0 = r'$ , all allelic frequencies remain constant, and  $g_{i*j}$  and  $g_{i\sim j}$  converge to  $a_i^* b_j^* / g^*$  and  $a_i^- b_j^- / g^-$ , respectively. Again, in general,  $\hat{g}_{ij} \neq a_i b_j$ .
- (iii) If  $r_A > 0 = r_B$ , the allelic frequencies at the  $B$ -locus remain constant, so that  $\hat{g}_{i*j} = a_i b_j^*$ ,  $\hat{g}_{i\sim j} = a_i b_j^-$ , and thus  $\hat{g}_{ij} = a_i b_j$ .
- (iv) If  $r_B > 0 = r_A$ , the allelic frequencies at the  $A$ -locus remain constant, so that  $\hat{g}_{i*j} = a_i^* b_j$ ,  $\hat{g}_{i\sim j} = a_i^- b_j$ , and thus  $\hat{g}_{ij} = a_i b_j$ .
- (v) If  $r_A \cdot r_B > 0$  then  $\hat{g}_{i*j} = a_i b_j g^*$ ,  $\hat{g}_{i\sim j} = a_i b_j g^-$ , and thus  $\hat{g}_{ij} = a_i b_j$ .