Evaluation of linkage in F_2 progenies by the chi-square criterion with codominant genes involved

Evaluation of the significance of genetic linkage using the chi-square criterion is a long-established practice. Nevertheless, I think that its application for F_2 progenies in cases where one or both of the genes analysed are codominant is problematic.

The common practice consists in calculation of the joint segregation chi-square (JSCS), when the theoretically expected numbers are determined on the presumption that the two genes assort independently, using the experimentally obtained marginal (row and column) sums. In a classical case when each gene has a recessive and a dominant allele, we have to analyse a 2x2 table of 4 phenotypical classes, the JSCS having 1 degree of freedom. If one allele of the first gene exhibits dominance, while alleles of the second gene are codominant, we obtain a 2x3 table with 6 classes and a JSCS with 2 degrees of freedom. If both genes are codominant, we have a 3x3 table with 9 classes and a JSCS with 4 degrees of freedom. (If we use not JSCS but the partitioned chi-square method, we obtain the same number of degrees of freedom for the part of the total chi-square left after subtraction of the chi-squares reflecting deviation of segregation of individual genes from the relevant classical models assumed, 3:1 or 1:2:1).

The problem consists in the fact that in all the cases only one degree of freedom is used by variation of genetic linkage, which depends on one parameter, the remaining degrees of freedom reflecting a stochastic variation of the experimental numbers not corresponding to linkage. I can illustrate this by the following example. Imagine we have obtained the following joint segregation numbers for two codominant genes:

	B1/B1	B1/B2	B2/B2
A1/A1	8	24	8
A1/A2	24	32	24
A2/A2	8	24	8

or, if we reduce this case *ad absurdum,* the following table:

	B1/B1	B1/B2	B2/B2
A1/A1	0	40	0
A1/A2	40	0	40
A2/A2	0	40	0

Both examples, the former not being too unnatural, exhibit no deviation from the 1:2:1 ratio for both genes and no genetic linkage, and are symmetrical on both diagonals. However, three values that could be calculated, namely, the total chi-square for the null hypothesis of classical segregation for both genes and the absence of linkage (8 degrees of freedom), its part after subtracting the 1:2:1 chi-squares for both genes (being equal to zero in our cases) (4 degrees of freedom), and JSCS (4 degrees of freedom), are identical and equal to 6.4 for the former table and 160 for the latter.

Of course, these deviations from the expected numbers

10	20	10
20	40	20
10	20	10

lack a biological sense (although a sophisticated mode of genotype viability can be imagined when monoheterozygotes are of advantage, but such a mode does not occur in nature). Nevertheless, it is evident that in the case of 2x3 and 3x3 tables, JSCS, as well as other mentioned chi-squares, can be contributed to by pure stochastic fluctuations of the phenotype numbers which do or do not imitate linkage. It should be noted, however, that this overestimation might be substantial only if the linkage is weak or absent. A strong linkage would lead to a large value of the JSCS much exceeding the contribution of the stochastic fluctuations considered. Hence, the significance of linkage is overestimated by the use of the JSCS for 2x3 or 3x3 tables. I and my colleagues failed to find in the literature available to us a method of obtaining in such cases the very component of the total chi-square which reflects solely linkage. It is possible that it has been proposed but has not penetrated into textbooks, as codominant loci became familiar to classic geneticists only after biochemical markers came into widespread use. I would appreciate hearing from anyone who knows a solution to the problem or would care to discuss the issues involved.

It seems that in the cases considered, the evaluation of the statistical significance of linkage by a maximum likelihood estimation of the recombination fraction and its standard error can be useful rather than JSCS. The standard error is obtained from the second derivative of the likelihood function by the very parameter evaluated (recombination fraction in our case), i.e. only information relevant to genetic linkage is taken into account Measurement of the deviation of the recombination fraction from 0.50 divided by its standard error is in fact the use of Student's t-test, that is adequate because the binomial distribution with the expectation not far from 0.5 is nearly symmetrical and close to the normal one.

O.E. Kosterin
Institute of Cytology & Genetics, Lavrentiev ave. 10,
Novosibirsk 630090, Russia
E-mail: kosterin@gene.nsk.su