GENETICA POLONICA
Vol. 28 1987 No. 3

AN ATTEMPT TO DETERMINE LINKAGES BETWEEN GENES CONTROLLING SOME QUANTITATIVE TRAITS IN RYE (*SECALE CEREALE* L.)

T. RUEBENBAUER, L. KUBARA-SZPUNAR, S. KALETA, K. PAJĄK[2]

Chair of Crop Production, University of Agriculture, Kraków

Summary. The method of determination of linkages between genes controlling quantitative traits using correlation methods was proposed and tested on the results of segregations in five hybrid populations of F_2, originating from crossing homozygous inbred lines of rye. Differences in the correlation coefficients between four quantitative traits (straw length, spike length, spikelet number per spike and 1000-seed weight) for five hybrid populations made it possible to estimate distance of individual genes on the chromosome.

In two previous papers the authors confined themselves to determination of possible linkages between such quantitative traits, as coloration of the auricles, ligule and nodes in rye (Ruebenbauer et al. 1986), as well as pubescence under the spike and ligule shape, and the remaining quantitative traits inherited imperceptibly (Ruebenbauer et al. 1987). In the present paper attention was paid to the possibilities of studying linkages between genes controlling quantitative traits, such as plant height, spike length, spikelet number per spike and 1000-seed weight.

MATERIAL AND METHODS

The detection of the size of linkages between genes controlling quantitative traits, showing the distance of genes from one another on the chromosome, is less accurate, the larger is the modifying variation influencing their size. If not the modifying variation, the size of linked quantitative traits would undergo discriminating segregations and the number of recombinants would show the degree of linkage.

The subject of this work were results of measurements of four quantitative traits: straw length, spike length, spikelet number per spike and 1000-seed weight, performed on single plants of the second generation of five different interlinear hybrids of rye. The lines used for crossings were established homozygotes (25th

[1] Received for publication: September 1986.

[2] First author: Prof. Dr. hab.; second: Dr.; third and fourth: M. Sc. Present address: ul. Św. Marka 37, 31-024 Kraków, Poland.

generation of inbreeding), which was shown by both morphological observations and studies of antigens and isozymes conducted at the Institute of Plant Genetics, Polish Academy of Sciences in Poznań. The initial material of that kind provides a uniform progeny in the F_1 generation and segregations of one type in the F_2 generation, which permitted a precise interpretation of the segregation results with the aim to determine the degree of linkage.

Assuming the existence of the absolute linkage between the genes A and B and genes a and b (cis linkage), one can find the lack of recombinants with the phenotypes Ab and aB. Attributing the increase of the trait on the x-axis to the action of the dominant A genes and the increase of the trait on the y-axis to the action of the dominant B genes, we can find that in the correlation table individual points for the phenotypes ab will be arranged in the left corner of the table, while the values for the phenotypes AB will be in the upper right corner. The correlation coefficient calculated for such a case will take the value of $r = +1$. But if the A and B genes as well as a and b genes are not interlinked, we have to deal with a free segregation of phenotypes in the ratio $9AB : 3aB : 3Ab : 1ab$. That way of segregations presented in the table gives distribution of points in 4 corners of the table, and the calculated correlation coefficient r will be equal to 0.

In the case of the "trans" type linkage (A and b as well as a and B) a positive sign will be replaced by a negative one. However, determination of the gene distance on the chromosomes in comparison with the "cis" type linkage presents larger difficulties which result from significantly smaller differences in the number of individual groups of phenotypes characteristic of limiting states corresponding to

Table 1. Theoretical values of correlation coefficients at various degrees of the "cis" type linkage (in morgans) and numbers of interacting unlinked genes

Linkage in morgans	1	2	3	4	5	6
	r_{xy}	$r_{xy}\ x_1y_0$ $r_{xy}\ x_0y_1$	$r_{xy}\ x_1y_1$	$r_{xy}\ x_2y_1$ $r_{xy}\ x_1y_2$	$r_{xy}\ x_2y_2$	$r_{xy}\ x_2y_0$ $r_{xy}\ x_0y_2$
5	0.9342	0.8792	0.8248	0.6707	0.5454	0.3550
10	0.8700	0.7786	0.6920	0.5026	0.3650	0.2144
15	0.8075	0.6926	0.5878	0.3970	0.2680	0.1502
20	0.7467	0.6179	0.5037	0.3233	0.2074	0.1131
25	0.6875	0.5519	0.4345	0.2685	0.1659	0.0889
30	0.6300	0.4929	0.3764	0.2260	0.1356	0.0718
35	0.5742	0.4398	0.3270	0.1919	0.1125	0.0590
40	0.5200	0.3916	0.2844	0.1639	0.0943	0.0492
45	0.4675	0.3474	0.2474	0.1404	0.0796	0.0413
50	0.4167	0.3069	0.2149	0.1205	0.0675	0.0348
55	0.3675	0.2695	0.1861	0.1033	0.0573	0.0295
60	0.3200	0.2349	0.1605	0.0884	0.0486	0.0249
65	0.2742	0.2027	0.1376	0.0753	0.0411	0.0211
70	0.2300	0.1727	0.1169	0.0637	0.0346	0.0177
75	0.1875	0.1448	0.0983	0.0534	0.0289	0.0148
80	0.1467	0.1187	0.0814	0.0441	0.0239	0.0122
85	0.1075	0.0942	0.0660	0.0358	0.0194	0.0099
90	0.0700	0.0714	0.0520	0.0284	0.0118	0.0078
95	0.0342	0.0501	0.0392	0.0216	0.0085	0.0060

x, y without figures mean linked genes
x, y with figures mean unlinked genes

Table 2. A comparison of empirical correlation coefficients with theoretical values taking into account the numbers of unlinked interacting genes in rye

Correlation coefficient	Hybrids between cultivars: Rogalińskie F_1 × Dańkowskie s231	Uniwersalne 145 × Zeelandzkie E	Kazimierskie H × Zeelandzkie E	Kazimierskie C_3 × Zeelandzkie E	Włoszanowskie C × Zeelandzkie E	Linkage in morgans
			Plant height S — 1000-seed weight M			
Empirical	0.482	0.749	0.091	−0.537	0.407	
Theoretical	0.5037	0.7467	0.1131		0.5037	20
The number of unlinked genes for x	1 or 0	0	2 or 3		1 or 0	
y	0 or 1	0	3 or 2		0 or 1	
			Spike length K — 1000-seed weight M			
Empirical	0.845	−0.026	0.216	0.202	0.501	
Theoretical	0.8700		0.2144	0.2144	0.5026	10
The number of unlinked genes for x	0		2 or 3	2 or 3	2 or 1	
y	0		3 or 2	3 or 2	1 or 2	
			Plant height S — spikelet number L			
Empirical	0.518	0.851	0.558	0.430	0.747	
Theoretical	0.5026	0.8700	0.5026	0.3650	0.7786	10
The number of unlinked genes for x	2 or 1	0	2 or 1	2	1 or 0	
y	1 or 2	0	1 or 2	2	0 or 1	
			1000-seed weight M — spikelet number L			
Empirical	0.288	0.794	0.147	−0.268	0.191	
Theoretical	0.2144	0.8700	0.2144		0.2144	10
The number of unlinked genes for x	2 or 3	0	2 or 3		2 or 3	
y	3 or 2	0	3 or 2		3 or 2	
			Spike length K — spikelet number L			
Empirical	0.747	0.258	0.292	0.933	0.537	
Theoretical	0.7467	0.3233	0.3233	0.7461	0.5037	20
The number of unlinked genes for x	0	1 or 2	1 or 2	0	1 or 0	
y	0	2 or 1	2 or 1	0	0 or 1	
			Plant height S — spike length K			
Empirical	0.613	−0.530	−0.153	0.438	0.258	
Theoretical	0.6300			0.4029	0.2260	30
The number of unlinked genes for x	0			1 or 0	2 or 1	
y	0			0 or 1	1 or 2	

the absolute linkage and free inheritance. For that reason calculations of negative r were omitted in Table 2.

In the case of small distances between genes on the chromosomes the correlation coefficient will be close to 1, and with an increase in the gene distance on the chromosomes that coefficient will take more and more smaller values. The size of that coefficient designated with the symbol r_{xy} is given in the first column of Table 1. On the basis of the data of the quantitative trait segregations concerning 4 mentioned traits the correlation coefficients were calculated between 6 possible interrelations for 5 hybrids of the second generation.

The figures given in Table 2 show the existence of different values of the correlation coefficient with different signs for the same pairs of traits. As mentioned before, the sign of the correlation coefficient at the "cis" linkage is positive, and that one at the "trans" linkage is negative. Rather high negative coefficients between the straw length and spike length as well as between the straw length and 1000-seed weight, beside coefficients with a positive sign indicate linkages in the form of both "cis" and "trans" are possible. In general, a longer straw is accompanied by a longer spike and by a larger 1000-seed weight. However, there is a possibility to combine the trait of short straw with a long spike and a short straw with a large 1000-seed weight.

The existence of different values of the correlation coefficient in five hybrid F_2 populations under study between the analyzed quantitative traits should be explained by the fact that beside 2 pairs of the main linked genes several pairs of independently inherited genes interact with them. These genes may occur in the same homozygous form in the both crossed lines, as a result of which they do not undergo segregation in the F_2 generation and the size of the correlation coefficient is determined only by the degree of linkage between the main genes. That case is designated with the symbol r_{xy} in Table 1, whereas in Table 2 the number of unlinked genes is 0. However, one or more gene pairs may occur in the heterozygous form in the F_1 hybrids, as a result of which in the F_2 generation there will occur segregations, affecting to a larger degree a reduction of the correlation coefficient the larger is the number of free-segregating genes interacting with linked genes. If, for instance, two pairs of genes inherited independently, which is designated in Table 1 with the symbol $r_{xy}x_2y_2$, interact with linked genes, then an increase in the variation of phenotypes in the F_2 generation will cause a change in the correlation coefficient at a significant linkage (5 morgans) from 0.9342 for 0.5454. This corresponds, for instance, to the following formula of the crossed lines:

$$XXYYx_1x_1X_2X_2Y_1Y_1y_2y_2 \times xxyyX_1X_1x_2x_2y_1y_1Y_2Y_2,$$

where x and y without numbers present linked genes, while x and y with numbers — non-linked gene pairs.

These general conditions concerning gene interaction were worked out by a computer and results are given in Table 1, whereas in Table 2 the same values of the coefficients are given in the column "theoretical values". Calculations of the theoretical correlation coefficients were made assuming that: The size of the Z linkage decreases with an increase in the gene distance on the chromosomes, analogically to the value of $1-r$, i. e. $Z=1-r$, providing $0 \leqslant Z \leqslant 1$. The value of the theoretical correlation coefficient was calculated from the formula expressing the covariance to variance ratio, i. e.

$$r=\frac{cov}{var}.$$

In our case:

$cov=(Y-\bar{y})\,[p\,(X,\,Y)\,(X-\bar{x})+p\,(x,\,Y)\,(x-\bar{x})]+(y-\bar{y})\,[p\,(y,\,X)\,(X-\bar{x})+p\,(x,y)\times$
$\times\,(x-\bar{x})]$,

whereas

$(var)^2=\{[p\,(x,\,Y)+p\,(x,\,y)]\,(x-\bar{x})^2+[p\,(X,\,Y)+p\,(X,\,y)]\,(X-\bar{x})^2\}\times$
$\times\,\{[p\,(X,\,y)+p\,(x,\,y)]\,(y-\bar{y})^2+[p\,(X,\,Y)+p\,(x,\,Y)]\,(Y-\bar{y})^2\}$,

where the mean values are:

$$\bar{x}=[p\,(X,\,Y)+p\,(X,\,y)]\,X+[p\,(x,\,Y)+p\,(x,\,y)]\,x\,;$$
$$\bar{y}=[p\,(X,\,Y)+p\,(x,\,Y)]\,Y+[p\,(X,\,y)+p\,(x,\,y)]y\,,$$

providing: $Y>y$; $X>x$,

but the sum of probabilities:

$$p\,(X,\,Y)+p\,(X,\,y)+p\,(x,\,Y)+p\,(x,\,y)=1\,.$$

In the case of linkage with the s value we calculate the probability p, suming up the corresponding elements of the matrix. In that matrix two gametes, paternal and maternal (XY and xy), occur at the frequency of $\frac{1}{2}\,(1-\frac{1}{2}s)$, whereas two recombination gametes Xy and xY — at the frequency of $\frac{1}{4}s$, as a result of which the

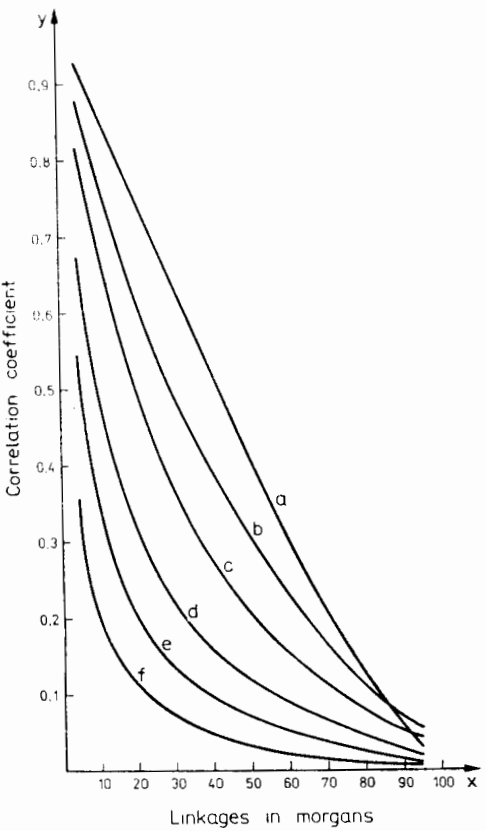

Fig. 1. Curves of dependences between the size of linkages in morgans and that of correlation coefficients for different genes interacting with linked genes

a — *a* curve at the lack of interacting genes, *b* — *a* curve for a single pair of interacting genes, *c* — *a* curve for two pairs of interacting genes, *d* — *a* curve for three pairs of interacting genes, *e* — *a* curve for four pairs of interacting genes, *f* — *a* curve for five pairs of interacting genes

genotypes $XXYY$ and $xxyy$ occur with the probability amounting to $\frac{1}{4}(1-\frac{1}{2}s)^2$. Each next pair of nonlinked genes reduces the number of recessive individuals for a given trait to 25% of the previous value in the way given in Table 1.

On the example of the results given in Table 1 it may be noticed that correlation coefficients decrease with the increase of distance of linked genes on the chromosome as well as with the increase in the number of genes interacting with linked genes. These relationships are presented in Fig. 1.

RESULTS AND DISCUSSION

In our studies we have confined ourselves to the consideration of the effect of interaction of not more than 5 pairs of genes (x_2y_3 or x_3y_2) inherited freely and interacting with coupled genes. This does not exhaust all the possibilities, however, consideration of a larger number of uncoupled genes is of no large importance in relation to a little chance of linkage detectability amounting to 15 - 20 morgans in the cases, when crossed inbred lines give a significant number of multiple heterozygotes in F_1. This finding results from considerations concerning significance of differences between the correlation coefficients, which only slightly decrease with an increase in the population size of F_2. Fig. 2 presents the relationship between the population size (x) and the limiting value (y) proving the significance of differences between the values of the empirical and theoretical correlation coefficient. The

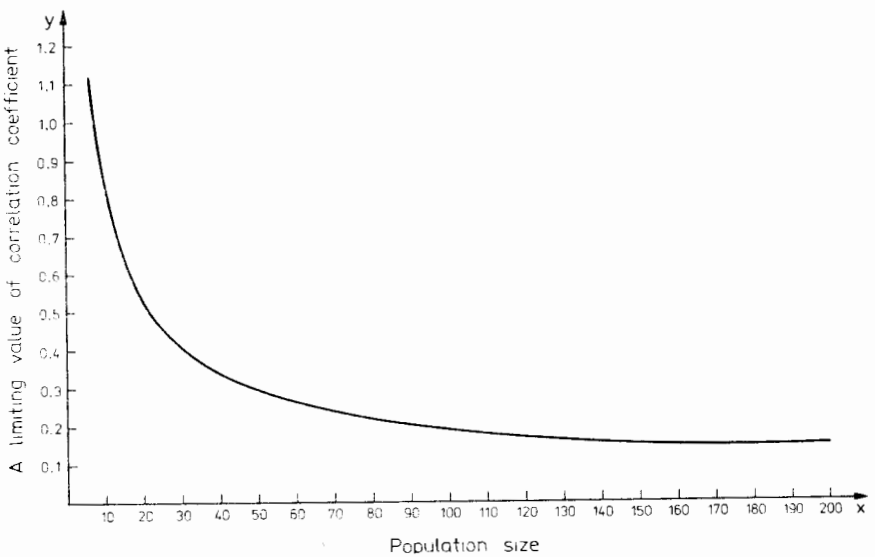

Fig. 2. Dependence between the population size and limiting value of the correlation coefficient

curve presenting that relationship significantly diminishes above $n=40$, attaining the correlation coefficient $r=0.20$ for $n=100$ and $r=0.14$ for $n=200$. It is recommended, however, to manipulate with F_2 populations of over 100 individuals to

determine precisely the empirical correlation coefficient. Table 2 gives the empirical and theoretical correlation coefficients. These last ones were chosen from Table 1 for individual values of morgans in such a way as to make their values to be the most corresponding to the empirical coefficients. The calculated errors of the empirical correlation coefficients permitted to find the lack of significant differences between the theoretical and empirical values, except the value of the correlation coefficient between the spike length and the spikelet number per spike for the hybrid Kazimierskie C×Zeelandzkie E.

It may, therefore, be generally inferred that differences between the empirical correlation coefficients and chosen theoretical coefficients resulted from non-inherited variation. If we assume that there was a more close linkage (for instance, 10 morgans) between the spike length and the spikelet number for the hybrid Kazimierskie C_3×Zeelandzkie E, then there would exist a full agreement between empirical and theoretical coefficients.

In the columns of the attached tables under the values of correlation coefficients are the corresponding numbers of unlinked genes occurring in the F_1 generation in the heterozygous form. This means that individual lines did not differ from one another by gene domination at a corresponding number of loci. For instance, in correlations between the spike length and 1000-seed weight the lines Rogalińskie F_1 and selected Dańkowskie 231 differed only by linked genes, whereas the remaining genes occurred in the F_1 generation in the homozygous form. The lines Kazimierskie C_3 and Zeelandzkie E, beside the way of linked genes domination, differed also by domination of 2 pairs of genes for the spike length and by 3 gene pairs for the 1000-seed weight or vice versa. As result of the so large number of heterozygous loci, not linked in the F_1 generation, the empirical correlation coefficient was very small. The last column of Table 2 gives linkage values in morgans. This value corresponds to that one given in Table 1, for which the correlation coefficients included in Table 2 as theoretical values were chosen. The choice of the respective linkage values cannot be accurate and has the character of a certain freedom consisting mainly in presenting reciprocal linkages in the form given in Fig. 3.

As follows from Fig. 3, between the main genes controlling the spike length and 1000-seed weight there exists linkage amounting to about 10 morgans. The same

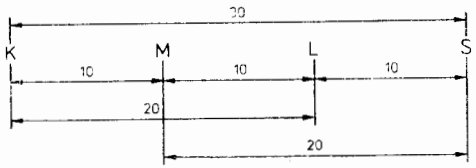

Fig. 3. A suggested position on the chromosome of 4 pairs of genes controlling quantitative traits of rye (distance are given in morgans)

K – gene pair controlling the spike length, *M* – gene pair controlling the 1000-seed weight, *L* – gene pair controlling the spikelet number per spike, *S* – gene pair controlling the plant height

value of linkage exists between the 1000-seed weight and the spikelet number per spike and between the spikelet number per spike and straw length. That order of gene location on the chromosome is supported by the size of linkages in morgans between the spike length and spikelet number per spike, and between the straw length and 1000-seed weight, amounting to 20 morgans each, as well as between the spike length and straw length. The last linkage amounting to 30 morgans is a sum of three distances, each equal to 10 morgans. Our studies carried out on populations consisting of 8 - 41 individuals do not permit to determine precisely the size of linkages; they only give a possibility to verify the suitability of the proposed determination of distance between the genes located on the same chromosomes by correlation. That method requires further testing on larger populations of the F_2 generation. A positive side of the proposed method is a possibility to determine the degree of linkage between genes controlling quantitative traits.

When determining the degree of linkage between genes controlling quantitative traits inherited in a discriminating way, differentiation of recombinants does not cause difficulties. However, in cases with phenotypes having the character of continuous variation their reference to the category of recombinants or parental forms is significantly deliberate. The proposed method consisting in the determination of theoretical correlation coefficients and in the comparison of these values with the empirical values permits not only to determine precisely distances between genes on the chromosomes, but also throws light to the genetic basis concerning the inheritance of quantitative traits.

REFERENCES

1. Ruebenbauer T., Kubara-Szpunar Ł., Kaleta S., Pająk K. (1986). An attempt to determine linkages between genes controlling some qualitative traits of inbred lines of rye (*Secale cereale* L.). I. Linkage of genes controlling anthocyanin coloration of nodes, auricles and ligule. Genetica Polonica, 27: 25 - 43.
2. Ruebenbauer T., Kubara-Szpunar Ł., Kaleta S., Pająk K. (1987). An attempt to determine linkages between genes controlling some qualitative traits of inbred lines of rye (*Secale cereale* L.). II. Linkages of genes controlling pubescence under the spike and the shape of ligule with genes controlling other qualitative traits. Genetica Polonica, 28: 211 - 216.

PRÓBA OKREŚLENIA SPRZĘŻEŃ MIĘDZY GENAMI KONTROLUJĄCYMI NIEKTÓRE CECHY ILOŚCIOWE U ŻYTA (*SECALE CEREALE* L.)

Streszczenie

Przedstawiono metodę wyznaczania sprzężeń między genami kontrolującymi cechy ilościowe, polegającą na określeniu teoretycznych wielkości współczynnika korelacji i porównaniu ich z wielkościami empirycznymi. Metodę tę sprawdzono na wynikach dotyczących rozszczepień

wybranych cech, które obserwowano w pięciu populacjach mieszańców pokolenia F_2, pochodzących z krzyżowania homozygotycznych linii wsobnych żyta. Różnice w wartościach współczynników korelacji między czterema cechami ilościowymi (długość słomy, długość kłosa, liczba kłosków w kłosie, masa 1000 ziarn) dla badanych populacji mieszańcowych pozwoliły na oszacowanie odległości między poszczególnymi genami w chromosomie.

ПОПЫТКА ОПРЕДЕЛЕНИЯ СЦЕПЛЕНИЙ ГЕНОВ, КОНТРОЛИРУЮЩИХ НЕКОТОРЫЕ КОЛИЧЕСТВЕННЫЕ ПРИЗНАКИ У РЖИ (*SECALE CEREALE* L.)

Резюме

В настоящей работе предложен метод определения сцеплений между генами, контролирующими количественные признаки, заключающийся на определении теоретических значений коэффициента корреляции и на сравнению их с эмпиримеческими значениями. Этот метод был проверен на результатах расщеплений выбранных признаков, которые наблюдались в 5 гибридных популяциях поколения F_2, полученных от скрещивания гомозиготных инбредных линий ржи. Выступающие разницы в величине коэффициента корреляции между четырьмя количественными признаками (длиной соломы, длиной колоса, числом колосков в колосе и весом 1000 зёрен) между 5 гибридными популяциями позволили установить расстояние отдельных генов на хромосоме.