Theor Appl Genet (2002) 104:786–796
DOI 10.1007/s00122-001-0842-8

E. Esch · W.E. Weber

Investigation of crossover interference in barley (*Hordeum vulgare* L.) using the coefficient of coincidence

Received: 7 February 2001 / Accepted: 5 May 2001 / Published online: 22 February 2002

Abstract Interference, the interaction between recombination events, was analysed in seven mapping populations of barley (*Hordeum vulgare* L.). The coefficient of coincidence was applied to investigate the type and position of interference within the genome. Interference was analysed with respect to dependence on the recombination fraction, and simulations were used to obtain test statistics which consider the small sample size of 71–150 double haploid lines. In addition to positive interference in intermediate intervals, strong negative interference, i.e. encouraged double recombination, was found in short intervals. The relationship between recombination fraction and interference could not be described with a uniform function, neither for the entire genome nor for individual chromosomes. The analysis of the position of interference within the genome revealed that interference does not act in the same way in the whole genome. Intervals spanning the centromere exhibited significantly higher means for the coefficient of coincidence than intervals within the chromosome arms, especially with regard to small intervals. In general, positive interference was found in the chromosome arms and no or negative interference in the genetically small but physically large centromeric region.

Keywords Crossover interference · *Hordeum vulgare* · Recombination · Negative interference · Coefficient of coincidence

Communicated by H.C. Becker

E. Esch (✉)
Abteilung Angewandte Genetik, Universität Hannover,
Herrenhäuser Str. 2, 30419 Hannover, Germany
e-mail: esch@genetik.uni-hannover.de
Fax: +49-511-7623608

W.E. Weber
Institut für Pflanzenzüchtung und Pflanzenschutz,
Martin-Luther-Universität Halle-Wittenberg,
Ludwig-Wucherer-Str. 2, 06108 Halle, Germany

Introduction

Dense genetic linkage maps and their underlying data sets are an excellent basis for investigating the distribution of recombination events in the genome and their interaction, i.e. interference. This term, introduced in 1916 by Muller, describes the non-random distribution of recombination. Positive interference means that a recombination event inhibits the occurrence of further recombinations in its vicinity, whereas in the case of negative interference, recombination is supported. In eucaryotes, positive interference is generally assumed, and Kosambi's mapping function (Kosambi 1944), which assumes positive interference with dependency on the distance, is widely used for genetic mapping.

Interference can be analysed on a genetical level (i.e. recombination events) or on a cytological level (i.e. chiasmata). It is assumed that both chiasmata and meiotic recombination have the same physical basis. Recently, the 1:1 relationship between recombination and chiasmata has been challenged (Nilsson et al. 1993; Sybenga 1996).

In recent years much progress has been achieved in our understanding of recombination and the processes that occur during meiosis. A comprehensive review, mainly based on results from *Saccharomyces cerevisiae*, has been given by Roeder (1997). In plants also significant progress has been made especially with regard to chromosome pairing and recombination (reviewed by Dawe 1998). The synaptonemal complex appears to play an essential part in the regulation of the frequency and distribution of recombination. It has been suggested that this complex is responsible for transmitting the signal for the positions of recombination. In organisms or mutants without a synaptonemal complex no interference has been observed (Egel 1995; Sym and Roeder 1994). The effect of the synaptonemal complex on recombination seems to be more indirect. It may function as a scaffolding to house and/or stabilize recombination during the final stages (Dawe 1998).

Ott (1997) reviewed the various methods that can be used to investigate interference in genetic maps and de-

Table 1 Characteristics of the datasets used for the analysis of interference in barley

Cross and abbreviation	Markers[a]	Plants	Data points[b] (missing values)	Average marker distance	Singletons[c] (%)[d]	'True' singletons[c] (%)[d]	Reference
Chebec × Harrington C×H	198	120	22,199 (6.6%)	7.1 cM	239 (1.08)	14 (0.06)	Langridge et al. 1996[e]
Clipper × Sahara C×S	141	150	17,921 (15.3%)	8.9 cM	279 (1.56)	7 (0.04)	Langridge et al. 1996[e]
Galleon × Haruna nijo G×H	225	120	22,250 (13.8%)	6.8 cM	344 (1.55)	9 (0.04)	Langridge et al. 1996[e]
Harrington × TR306 H×T	154	150	22,105 (4.3%)	8.3 cM	121 (0.55)	9 (0.04)	Kleinhofs 1994[e]
Igri × Franka I×F	235	71	16,431 (4.2%)	5.2 cM	31 (0.19)	10 (0.06)	Graner et al. 1991, 1994; Graner 1996[e]
Proctor × Nudinka P×N	247	113	25,700 (7.9%)	7.6 cM	302 (1.18)	7 (0.03)	Becker et al. 1995
Steptoe × Morex S×M	326	150	45,879 (6.2%)	3.8 cM	130 (0.28)	–	Kleinhofs et al. 1993

[a] Only one marker of several cosegregating markers was used (see text for further explanation)
[b] Calculated as (no. of markers * no. of plants–no. of missing values)
[c] See text for explanation
[d] Percentage from data points
[e] GrainGenes Database: http://wheat.pw.usda.gov/ggpages/maps.html

fined two different aspects. Interference can refer to the distribution of the numbers of recombination per chromosome or it can refer to the position of recombination events and the distance between them.

In the study presented here, our aim was to examine both the type of interference and its position within the chromosomes of barley (*Hordeum vulgare* L.). For this purpose, coefficients of coincidence were calculated using published data from different mapping projects. Although these datasets represent independent crosses, they partly contain the same markers. This allowed us to study similarities and differences in the degree of interference between populations and, thus, the genetic variability of interference. Since in barley and related cereals numerous studies exist regarding the distribution of recombination and chiasmata at the physical level of the chromosomes, we were able to compare the position of regions with different degrees of interference on the genetic and physical map, respectively. One problem we encountered was the small sample size of the populations, which resulted in some special features of the distribution of the coefficient of coincidence. In addition, we present a suggestion for illustrating the variation of interference along a chromosome.

Materials and methods

Estimation of the coefficient of coincidence in the experimental data

Mapping data from seven different mapping populations of barley available from the GrainGenes database (http://wheat.pw.usda.gov/ggpages/maps.html) were analysed. Table 1 gives an overview of some characteristics of the populations, which all consist of double-haploid (DH) lines. Each DH plant represents an F_1 gamete and, therefore, it is possible to derive the position of recombination events during meiosis directly from the marker genotype.

In the case of cosegregating markers, only the marker with the most data points was used because cosegregating loci lack additional information. The marker order given in the references of the datasets was chosen as fixed throughout the analysis. Datasets and linkage groups were explored separately. All possible combinations of three markers (A, B, C) were formed, and for each triple the gametes were divided into the following classes:

1) gametes without recombination in both intervals with frequency x_0,
2) gametes with recombination between marker A and B with frequency x_1,
3) gametes with recombination between marker B and C with frequency x_2 and
4) gametes with recombination between marker A and B and marker B and C (double recombination) with frequency x_{12}.

The total number of gametes is n. Maximum Likelihood (ML) estimates for the recombination frequencies between markers are given as

$$\hat{r}_{AB} = \frac{x_1 + x_{12}}{n} \quad (1)$$

$$\hat{r}_{BC} = \frac{x_2 + x_{12}}{n} \quad (2)$$

$$\hat{r}_{AC} = \frac{x_1 + x_2}{n} \quad (3)$$

(e.g. Bailey 1961).

Interference was analysed using the coefficient of coincidence, which is defined as the quotient of the number of observed double recombinations and the expected number of double recombinations with independence of recombination in the adjacent intervals AB and BC. It is estimated by

$$\hat{C} = \frac{n x_{12}}{(x_1 + x_{12})(x_2 + x_{12})} \quad (4)$$

Without interference the expectation is one, that is $E(\hat{C}) = 1$. When positive interference occurs, there are fewer double recombinations so that $E(\hat{C}) < 1$. Negative interference results in more double recombinations than expected and $E(\hat{C}) > 1$. In the case of $\hat{r}_{AB} = 0$ or $\hat{r}_{BC} = 0$ the coefficient of coincidence $\hat{C}$ is not defined, because only one interval exists.

Simulation of the distribution of the coefficient of coincidence in small sample sizes

Without interference, the coefficient of coincidence is expected to be 1 irrespective of the interval size, i.e. the recombination frequency between A and C (Haldane 1919). To test $\hat{C}$ in experimental data under the null hypothesis H_0, information about the distribution, especially in small sample sizes as in the barley data, was needed. For this purpose we did computer simulations. Our aim was to show the special features of the distribution under H_0 caused by the small sample size and to obtain a statistic to test our experimental $\hat{C}$.

Without interference, the expected probabilities of the four gamete classes mentioned above can be expressed in terms of r_{AB} and r_{BC}, which are the recombination frequencies between marker A and B and marker B and C, respectively (e.g. Bailey 1961):

1. $1 - r_{AB} - r_{BC} + r_{AB}r_{BC}$
2. $r_{AB}(1 - r_{BC})$
3. $r_{BC}(1 - r_{AB})$
4. $r_{AB}r_{BC}$.

Since the coefficient of coincidence was analysed with dependence on the entire interval AC, the gametic probabilities were expressed in terms of r_{AC}, the recombination frequency between marker A and C. Without interference

$$r_{AC} = r_{AB} + r_{BC} - 2r_{AB}r_{BC} \quad (5)$$

The relative position of marker B in the interval AC can be given with

$$q = \frac{r_{AB}}{r_{AB} + r_{BC}} \quad (6)$$

with $0 < q < 1$. In the case of $q = 0$ or $q = 1$, that is $r_{AB} = 0$ or $r_{BC} = 0$, respectively, only one interval exists. Is B located in the middle of A and C, then $q = 0.5$. The recombination frequencies r_{AB} and r_{BC} can now be expressed in terms of r_{AC} and q:

$$r_{AB} = \frac{1 - \sqrt{1 - 8r_{AC}q + 8r_{AC}q^2}}{4(1 - q)} \quad (7)$$

$$r_{BC} = \frac{1 - \sqrt{1 - 8r_{AC}q + 8r_{AC}q^2}}{4q} \quad (8)$$

Therefore, the expected distribution can be described in terms of r_{AC} and q only. In the computer simulation, certain values for r_{AC} and q were given, and the probabilities of the four gamete classes were calculated. On the basis of these probabilities, a sample of n gametes per each simulation corresponding to n DH lines was created. Gametes were divided into the gametic classes, their frequencies x_0, x_1, x_2 and x_{12} were determined and $\hat{r}_{AC}$ and $\hat{C}$ (according to eqs. (3) and (4)) and

$$\hat{q} = \frac{x_1 + x_2}{x_1 + x_2 + 2x_{12}} \quad (9)$$

were estimated. With $\hat{q} = 0$ or $\hat{q} = 1$, that is $\hat{r}_{AB} = 0$ or $\hat{r}_{BC} = 0$, respectively, $\hat{C}$ is not defined. For each combination of r_{AC} and q, 10^6 simulations were performed. Mean and standard deviation of the estimated parameters were calculated. Subsequently, the bias of the estimated values was analysed for various given r_{AC} and q and sample sizes of 70, 100 and 150 gametes. The mean $\hat{r}_{AC}$ corresponded well with the given value for $r_{AC} \geq 0.075$. With $r_{AC} = 0.075$ and $q = 0.5$, the difference between the estimated mean and the true value was 6.3%, 1.8% and 0.2% with a sample size of 70, 100 and 150, respectively (Table 2). This bias of $\hat{r}_{AC}$, which is a linear function of the observations and therefore should be unbiased, was caused by omitting the cases in which $\hat{C}$ is not defined (see above). If all simulations were used for the estimation of the recombination values, they were in deed unbiased. These cases were not used analogous to the procedure with the experimental data.

The position of marker B only had a limited influence on the estimation of $\hat{r}_{AC}$ as long as q was between 0.35 and 0.65. The standard deviation was hardly influenced. $\hat{C}$ was estimated without bias when $q = 0.5$ and $r_{AC} \geq 0.075$. The estimation was influenced by q. With $q = 0.35$ a positive bias of between 1.9% and 3.5% occurred. The larger the difference from $q = 0.5$ the larger the standard deviation. The increase was between 1.4% and 7.2% with $q = 0.35$ in comparison to $q = 0.5$. The distribution of $\hat{C}$ dependent on $\hat{r}_{AC}$ could be characterized by an increased number of cases with $\hat{C} = 0$ with decreasing recombination frequency and decreasing sample size and $\hat{q}$ near 0 or 1.

The comparison of the experimental distribution of the coefficient of coincidence with the simulated data under H_0 was done by comparing the means of $\hat{C}$ for classes of recombination frequency. Thereby it was possible to take differences in the shape of the distribution (e.g. percentage of $\hat{C} = 0$) within the classes into consideration. For that purpose simulations were done for the estimation of confidence intervals of the means. r_{AC} varied from 0.075 to 0.475 in steps of 0.05 with $q = 0.5$, and 10^6 simulations were done per combination. The experimental data were classified correspondingly for the recombination frequency. For the class $0.05 \leq \hat{r}_{AC} < 0.10$, the results from the simulation with $r_{AC} = 0.475$ were taken and so on until $0.45 \leq \hat{r}_{AC} < 0.50$, for which $r_{AC} = 0.475$ was taken. Only experimental triples with $\hat{r}_{AC} \geq 0.05$ and $0.35 \leq \hat{q} \leq 0.65$ were considered. The number of gametes for the simulation was chosen according to the mean number of gametes from the experimental triples. The mean number of gametes was 97, 102, 91, 132, 68, 91, and 125 for C×H, C×S, G×H, H×T, I×F, P×N and S×M, respectively (see Table 1 for abbrivations). In the experimental data, each class i of the recombination fraction $\hat{r}_{AC}$ contained m_i triples. The 10^6 simulated C values for each class were divided into individual samples of the size m_i, and the mean for each sample was calculated. Confidence intervals were determined from the distribution of the means. The significance level of 5% was corrected for multiple comparisons according to Bonferroni using the number of classes for the recombination fraction, i.e. 9. The size of the confidence intervals increased strongly with decreasing recombination frequency and decreasing number of gametes.

Table 2 Estimated mean and standard deviation for the recombination fraction $\hat{r}_{AC}$ and the coefficient of coincidence $\hat{C}$ from simulation. Parameters for simulation (10^6 for each parameter combination) were $r_{AC} = 0.075$, different numbers of gametes n and different q

n	$\hat{r}_{AC}$		$\hat{C}$	
	$q = 0.35$	$q = 0.5$	$q = 0.35$	$q = 0.5$
Mean				
70	0.0805	0.0797	1.0187	0.9931
100	0.0769	0.0763	1.0203	1.0089
150	0.0754	0.0751	1.0352	1.0124
Standard deviation				
70	0.0297	0.0295	3.9058	3.7833
100	0.0256	0.0256	3.3672	3.3194
150	0.0213	0.0214	2.7450	2.5598

Results

Distribution of the coefficient of coincidence with dependence on the recombination frequency

Very similar distributions of the coefficient of coincidence with dependence on the recombination frequency were obtained for all seven linkage groups and crosses, respectively. A typical example is shown in Fig. 1.

The coefficient of coincidence strongly increased with decreasing interval size – i.e. the smaller the interval the stronger the effect of negative interference. With independence of recombination events over the whole chromosome the expectation is $E(\hat{C}) = 1$, which is shown by the horizontal line in Fig. 1. With the Kosambi function (Kosambi 1944) strong positive interference in small intervals is assumed that linearly decreases with increasing interval size until independence is reached. In Fig. 1 the Kosambi function would result in a straight line from $\hat{C} = 0$ at $\hat{r}_{AC} = 0$ to $\hat{C} = 1$ at $\hat{r}_{AC} = 0.5$. But also shown in Fig. 1 for all values of $\hat{r}_{AC}$ triples with no double recombination at all and therefore $\hat{C} = 0$ were observed. The proportion of these triples is not demonstrat-

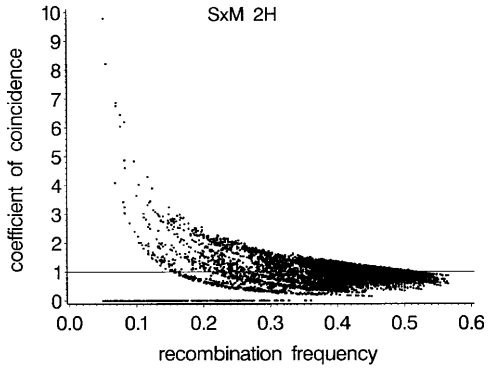

Fig. 1 Relationship between the recombination frequency $\hat{r}_{AC}$ and the coefficient of coincidence $\hat{C}$ for chromosome 2H of the cross S×M. Shown are the values of 12,344 triples with $0.35 \leq \hat{q} \leq 0.65$ and $\hat{r}_{AC} \geq 0.05$. The *horizontal line* shows the expectation without interference [$E(\hat{C}) = 1$]

ed in the figure. With the small sample sizes of between 71 and 150 individuals not all values of the coefficient of coincidence can be obtained for small intervals. When $\hat{r}_{AC}$ is small, $\hat{C}$ will either be zero or will become very large when double recombination occurs. Therefore, coefficients of coincidence of approximately 1, which are expected in the case of independence, could not be obtained at all for small intervals. If the high coefficients of coincidence observed were caused only by the small sample size and statistical variation, then their proportion should have been correspondingly small. To consider the proportion of these triples and the proportion of triples with $\hat{C} = 0$ we divided the recombination fraction into classes and calculated the mean coefficient of coincidence for each class. These means were compared with the confidence intervals from the simulation with independence of recombination. The results for the seven chromosomes in the seven crosses are shown in Fig. 2. When the proportion of triples with $\hat{C} = 0$ was taken into consideration, differences between the linkage groups became obvious. In intermediate intervals with recombination fractions between approximately 0.15 and 0.4 significant positive interference was often found. A few exceptions (S×M 6H, I×F 3H, C×H 2H and 5H) showed negative interference in intermediate intervals with recombination fractions between 0.2 and 0.35. Some chromosomes, S×M 3H and 7H, C×S 2H, showed a pattern similar to the Kosambi function: increasing positive interference with decreasing recombination fraction. Slight negative interference was observed for some linkage groups with recombination fractions between approximately 0.35 and 0.45 (I×F 1H and 7H, G×H 1H and 2H). The groups S×M 4H, P×N 2H 4H 5H 7H, G×H 2H 5H 6H and C×H 7H showed a significant high negative interference in very small intervals. This was particularly the case for the cross P×N. Only in 13 of the 49 linkage groups was independence of recombination in large intervals ($\hat{r}_{AC}$ between 0.45 and 0.5) observed.

A closer look at the negative interference in small intervals revealed that often the triples only showed a single double recombination. This effect can also be due to the presence of so-called singletons (Säll and Nilsson 1994). At these positions a single marker from one parent is flanked on both sides by markers from the other parent – i.e. a single marker is enclosed by a double recombination. Since the occurrence of double recombinations in genetically small intervals in general is assumed to be unlikely, singletons are often regarded as candidates for marker misclassification. The influence of singletons was investigated by repeating the analysis after the singletons were changed into missing values in the datasets. Table 1 gives the number of singletons found in the different populations, which varied between 0.19% and 1.56%. In the genetic maps containing cosegregating markers it was possible to distinguish potential misclassifications from 'true' singletons. A singleton at a map position with cosegregating markers is unlikely to be due to misclassification because the double recombination is confirmed by several markers. These 'true' singletons were not changed into missing values. Their proportion varied between 0.03% and 0.06% (Table 1). Population S×M did not include any cosegregating markers.

In Fig. 3 results from the analysis without singletons are shown for the crosses S×M, P×N and I×F, which contained the highest number of markers. A comparison with Fig. 2 revealed the following differences. As a general effect, a reduction of the means in small intervals with recombination fractions up to 0.2 can be noticed. The cross I×F showed almost no differences. The means of S×M 4H no longer showed significant negative interference, whereas the high negative interference in P×N 2H, 4H, 5H and 7H remained significant. Therefore, it can be concluded that the large coefficients of coincidence were not caused only by singletons.

Position of interference within the genome

With respect to the analysis of the distribution of the coefficient of coincidence with dependence on the interval size, the position of the intervals within the genome was not considered. It was assumed that interference acts in the same way in the whole genome. However, it is of interest to note where the intervals in which there is more double recombination than expected are located in the genetic maps. Intervals with high coefficients of coincidence should be evenly distributed over the genome if interference operates everywhere in the same way. Consequently, the individual coefficients of coincidence were assigned to positions in the genetic maps. The $\hat{C}$ value of each triple was assigned to the mid-marker B because this marker allowed the differentiation between double recombination or no recombination in the entire interval. This resulted in a sample of coefficients of coincidence for each marker, and for each marker the mean for $\hat{C}$ (designated as marker mean) was calculated. For visualization of the location of the marker means in the genetic maps a two-dimensional graph was used. The y-axis indicates the mean for the coefficient of coincidence

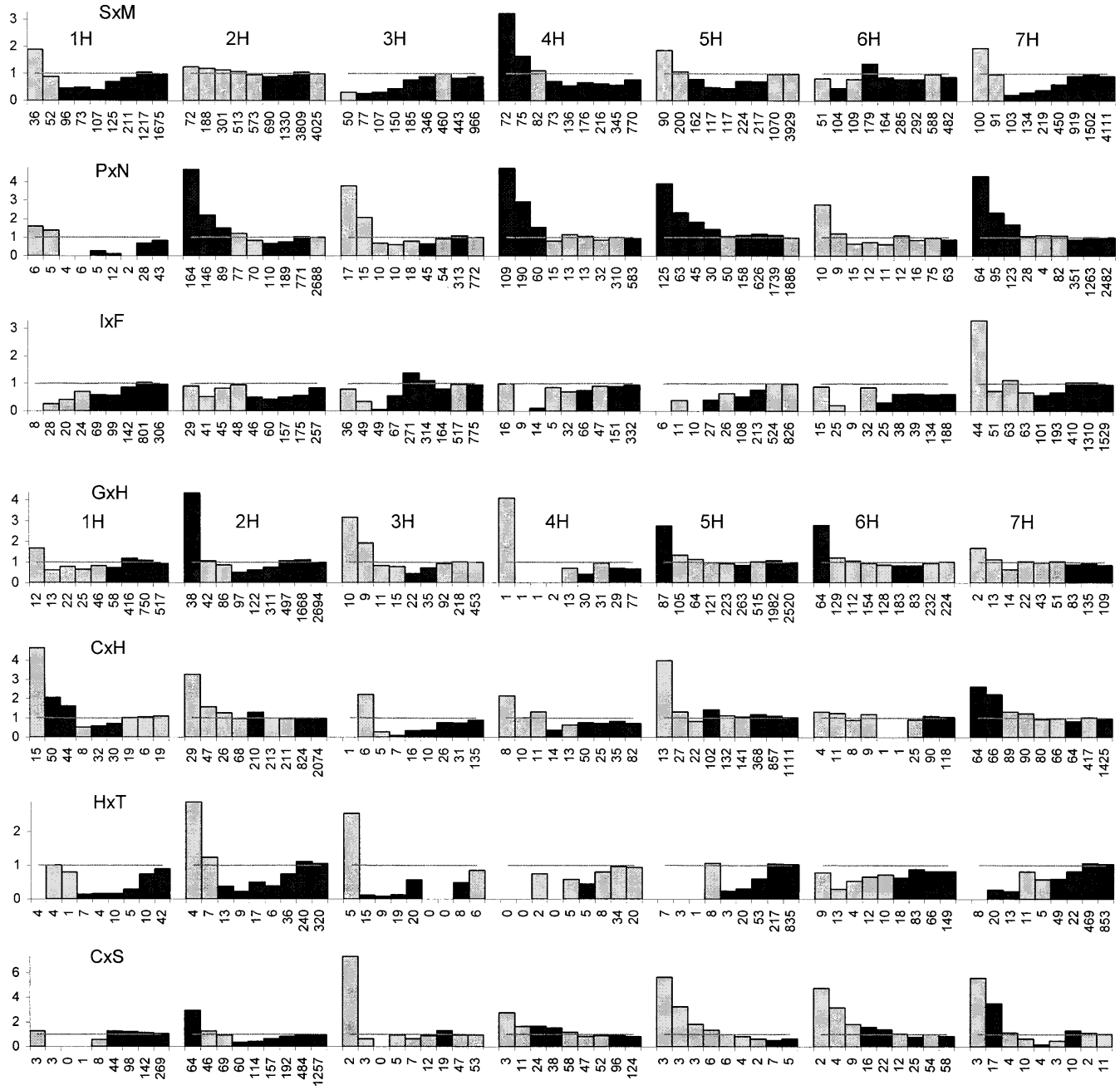

Fig. 2 Mean coefficient of coincidence $\hat{C}$ (*y-axis*) for classes of recombination frequency $\hat{r}_{AC}$ (*x-axis*) for the seven barley chromosomes in the seven crosses analysed. Classes range from 0.05 to 0.50 in steps of 0.05 (from *left to right*). The numbers of triples in the individual classes are given along the *x-axis*. The *lines* indicate the mean values of the simulation with no interference [$E(\hat{C}) = 1$]. Means which significantly deviate are shown in *dark grey*

and the x-axis gives the centiMorgan position of the markers in the genetic map. Figure 4 shows the marker means for three chromosomes from I×F and S×M with many corresponding markers. The markers partially revealed very large differences in their marker means. Markers with high mean values seem to be concentrated in distinct regions rather than evenly distributed. In this context a comparison of the different maps via the same markers could reveal if homologous regions show similar patterns for the marker means. Comparable distributions in homologous regions point against a random distribution of the marker means. Table 3 shows the number of common markers and the comparison of the colinearity of the different maps. The linear order of the markers is in good agreement between the different crosses. The maps of the crosses I×F, P×N and S×M showed no differences at all.

With few exceptions all linkage groups showed the typical pattern for the marker means, shown in Fig. 4. Proximal in the chromosomes existed a region with increased marker means with more or less distinct differences. In many cases the highest mean values were

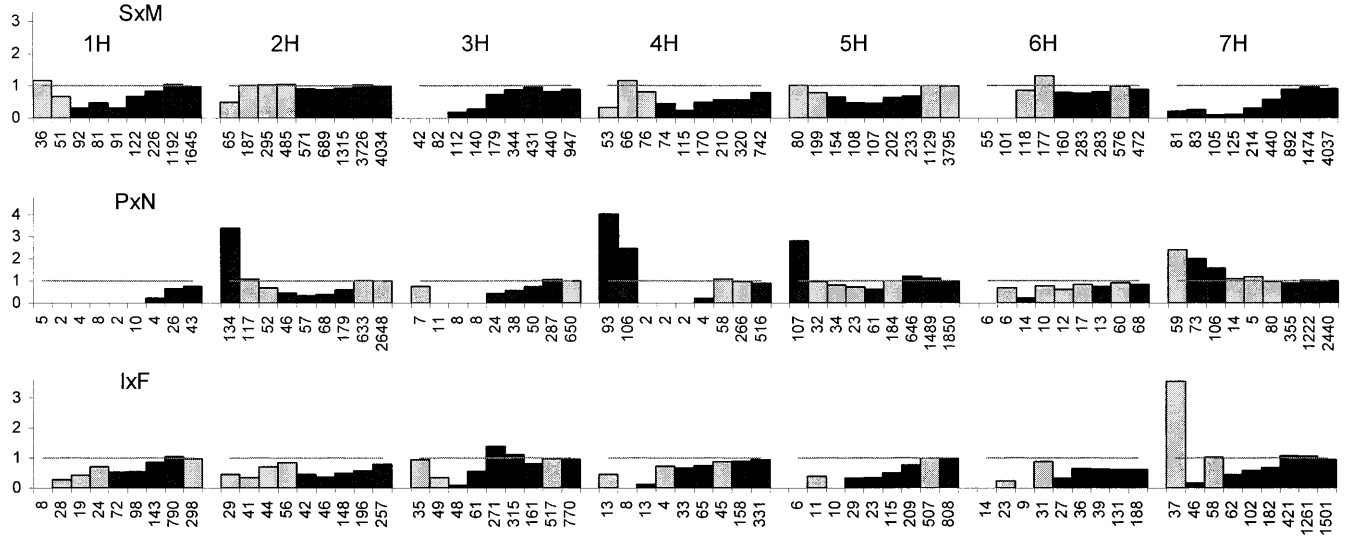

Fig. 3 Mean coefficient of coincidence $\hat{C}$ (*y-axis*) for classes of recombination frequency $\hat{r}_{AC}$ (*x-axis*) for the seven barley chromosomes in three crosses. In the datasets singletons were changed to missing values. Classes range from 0.05 to 0.50 in steps of 0.05 (from *left to right*). The numbers of triples in the individual classes are given along the *x-axis*. The *lines* indicate the mean values of the simulation with no interference [$E(\hat{C}) = 1$]. Means which significantly deviate are shown in *dark grey*

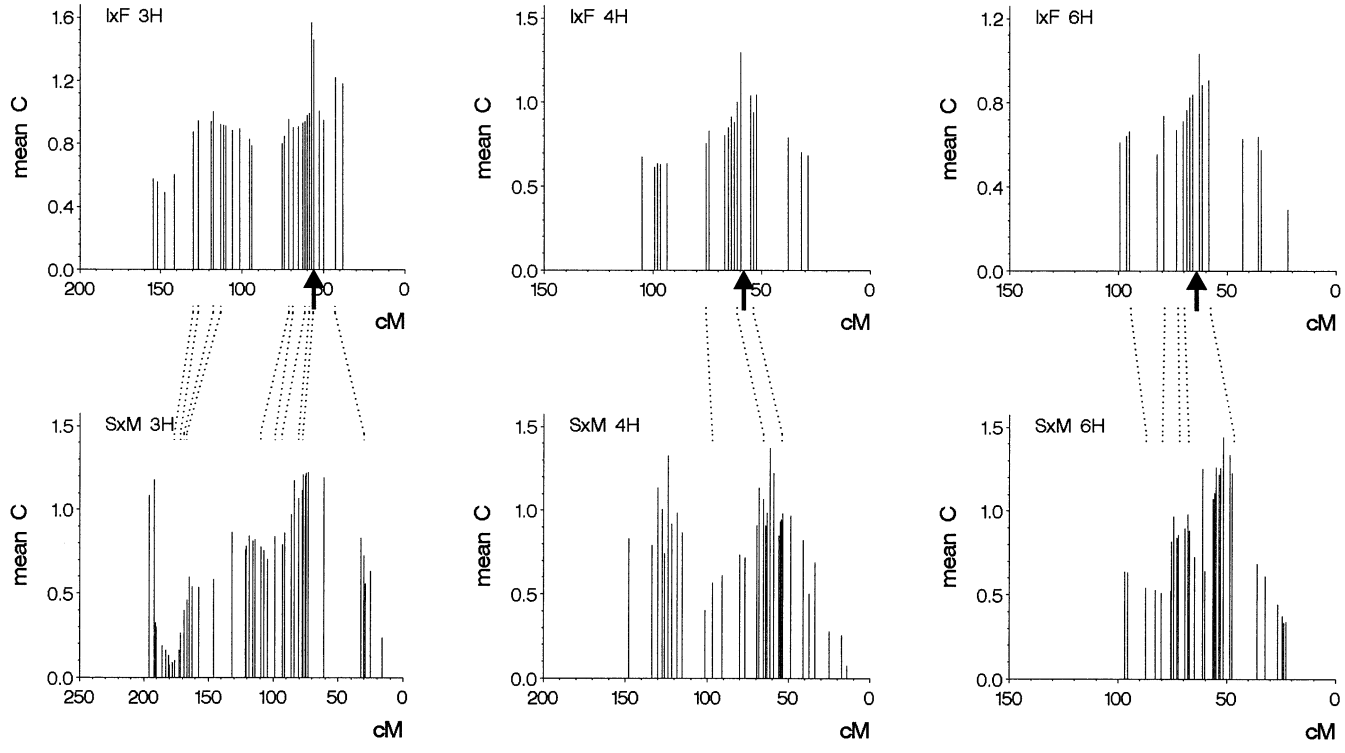

Fig. 4 Mean coefficient of coincidence $\hat{C}$ for triples with the same mid marker. The map position of the marker referred to is given along the *x-axis*. Shown are chromosomes 3H, 4H and 6H in the crosses I×F and S×M. *Dotted lines* connect markers which were mapped in both crosses. The position of the centromere in the I×F map (Kuenzel et al. 2000) is indicated by an *arrow*

above 1. A comparison of the different chromosomes between the populations demonstrated a very good correspondence of the proximal region with increased marker means, as shown in Fig. 4. Some linkage groups also showed distal regions with increased marker means, but no principal correspondence between the crosses was observed. In addition, for the distal markers fewer triples were available for the calculation of the marker means. Therefore, single extreme values which could result from statistical variation had a much larger influence.

A detailed physical map has been published for the cross I×F (Kuenzel et al. 2000). A comparison of the location of the centromere with the location of the proximal region showed that for the chromosomes presented in Fig. 4 the marker with the highest value in the proxi-

Table 3 Comparison of the colinearity of the different maps. Above the diagonal: number of colinear markers/total number of common markers (with a minimum of three markers per chromosome). Below the diagonal: number of colinear chromosomes/total number of chromosomes with at least three common markers

	C×H[a]	C×S	G×H	H×T	I×F	P×N	S×M
C×H		40/51	77/96	14/14	6/6	28/31	22/23
C×S	3/7		41/52	2/3	0/0	20/26	10/13
G×H	1/7	4/7		16/18	9/9	27/28	30/31
H×T	3/3	0/1	2/4		29/31	9/9	49/51
I×F	2/2	0/0	2/2	4/6		0/0	47/47
P×N	4/6	2/5	4/5	3/3	0/0		25/25
S×M	3/4	0/3	4/5	5/7	7/7	5/5	

[a] See Table 1 for abbreviations

Table 4 Influence of the centromere on the coefficient of coincidence $\hat{C}$. Triples from I×F were classified according to whether they spanned the centromere or not. Mean and standard deviation (SD) of $\hat{C}$ were calculated for classes of the recombination fraction $\hat{r}_{AC}$ for the whole genome (1H–7H) and within chromosomes, respectively. The influence of the centromere was tested using the Wilcoxon rank sum test

Chromosome	$\hat{r}_{AC}$	With centromere			Without centromere			*P*-value
		Number of triples	Mean	SD	Number of triples	Mean	SD	
1H–7H	0.05–0.15	138	1.98	2.84	230	0.21	0.91	<0.0001
	0.15–0.25	208	0.96	0.87	268	0.40	0.58	<0.0001
	0.25–0.35	930	1.10	0.49	518	0.41	0.42	<0.0001
	0.35–0.45	2,971	1.01	0.26	1,813	0.85	0.39	<0.0001
	0.45–0.50	3,144	0.98	0.16	1,069	0.83	0.27	<0.0001
1H	0.05–0.15	6	1.24	1.93	30	0	0	–[a]
	0.15–0.25	12	1.53	1.19	32	0.22	0.49	0.0003
	0.25–0.35	78	0.84	0.54	90	0.35	0.46	<0.0001
	0.35–0.45	796	1.06	0.17	147	0.79	0.36	<0.0001
	0.45–0.50	304	0.97	0.11	2	0.74	0	–[a]
2H	0.05–0.15	16	1.69	1.85	54	0.39	1.41	0.0003
	0.15–0.25	48	0.99	0.57	45	0.78	0.60	0.1980
	0.25–0.35	82	0.43	0.32	24	0.53	0.30	0.1468
	0.35–0.45	265	0.51	0.25	67	0.62	0.19	0.892
	0.45–0.50	173	0.94	0.19	84	0.62	0.23	<0.0001
3H	0.05–0.15	38	1.19	2.11	47	0	0	0.0001
	0.15–0.25	68	0.43	0.62	48	0.23	0.42	0.0851
	0.25–0.35	473	1.43	0.22	112	0.43	0.37	<0.0001
	0.35–0.45	328	1.07	0.24	353	0.81	0.39	<0.0001
	0.45–0.50	446	1.00	0.14	329	0.89	0.26	0.0258
4H	0.05–0.15	12	0.97	2.25	13	0.33	1.18	–[a]
	0.15–0.25	13	0.38	0.63	6	0.14	0.34	–[a]
	0.25–0.35	67	0.95	0.35	31	0.30	0.24	<0.0001
	0.35–0.45	150	0.98	0.12	38	0.56	0.15	<0.0001
	0.45–0.50	310	0.98	0.16	22	0.43	0.12	<0.0001
5H	0.05–0.15	–	–	–	17	0.25	0.71	–
	0.15–0.25	–	–	–	37	0.30	0.46	–
	0.25–0.35	11	1.46	0.06	123	0.46	0.48	<0.0001
	0.35–0.45	135	1.16	0.15	602	0.88	0.39	<0.0001
	0.45–0.50	458	1.05	0.06	368	0.91	0.20	<0.0001
6H	0.05–0.15	22	0.60	1.97	18	0.30	0.88	0.9376
	0.15–0.25	19	1.27	0.73	22	0.16	0.36	<0.0001
	0.25–0.35	38	0.65	0.49	25	0.27	0.33	0.0031
	0.35–0.45	124	0.70	0.22	49	0.40	0.15	<0.0001
	0.45–0.50	144	0.66	0.23	44	0.46	0.13	<0.0001
7H	0.05–0.15	44	3.83	3.42	51	0.28	0.91	<0.0001
	0.15–0.25	48	1.58	0.90	78	0.50	0.68	<0.0001
	0.25–0.35	181	0.80	0.42	113	0.42	0.42	<0.0001
	0.35–0.45	1163	1.10	0.17	557	0.95	0.40	0.0116
	0.45–0.50	1309	0.99	0.14	220	0.80	0.29	<0.0001

[a] The test was not performed in cases where the sum of triples with and without centromere was below 40 and the number of triples within a single class was too low, respectively

mal region is located directly at the centromere. Since for the calculation of marker means the coefficient of coincidence of a triple was assigned to the mid marker, all intervals which contributed to the highest marker mean spanned the centromere. For investigating the influence of the centromere in more detail we divided the triples of I×F according to whether they spanned the centromere or not. To consider the size of the interval we also divided the triples into five classes on the basis of their recombination fraction. The comparison was done for the whole genome and for each chromosome separately (Table 4). For the individual chromosomes, in most cases, the mean coefficient of coincidence of the triples spanning the centromere was significantly higher (Wilcoxon test; Wilcoxon 1946) than that of the triples without the centromere. The mean coefficient of coincidence increased with decreasing recombination fraction when the centromere was included, whereas it decreased for triples without centromere. This was particularly obvious for chromosome 7H in which the mean increased up to 3.8 for very short intervals spanning the centromere. Combining the data of the seven chromosomes confirmed this general tendency. Small intervals including the centromere frequently showed means above 1, whereas for triples without the centromere the means generally were below 1. Outside the centromere interference seemed to act similar to the Kosambi function $C = 2r_{AC}$, which means increasing positive interference with decreasing recombination frequency.

Discussion

In general, positive interference (Kosambi's mapping function) or no interference (Haldane's mapping function) is assumed for the construction of genetic maps. Therefore, the question arises whether the high coefficients of coincidence found in the present study, especially for small recombination fractions, are the consequence of negative interference – encouraged double recombination in short intervals – or whether they are caused by other factors.

One possible source of error are double recombinations generated by marker misclassifications. The comparison of results obtained from datasets with and without singletons showed that the significant negative interference in short intervals was not only due to singletons. Therefore, even if all singletons would have been misclassifications, this would have had no effect on the essential results. A whole string of observations indicates that the observed double recombinations in genetically small intervals do in fact occur. In six of the seven datasets, cosegregating markers confirmed the occurrence of singletons. Becker et al. (1995) considered the problem of singletons in detail when adding amplified fragment length polymorphism (AFLP) markers to the P×N map. All singletons were checked and most of them were confirmed. Moreover, the AFLP markers added confirmed 35% of the singletons of the original restriction fragment length polymorphism (RFLP) map (Heun et al. 1991) through the mapping of additional markers between the two flanking recombination events. Becker et al. (1995) came to the conclusion that most of the singletons are caused by double recombination events. The dataset S×M with the highest number of data points showed a distinctly lower proportion of singletons than P×N. Blake (1992; cited in Säll and Nilsson 1994) reported that in an earlier version of the S×M map markers causing too much negative interference had been eliminated.

A comparison of the mean coefficients of coincidence for classes of the recombination fraction (Fig. 2) displayed no common pattern for individual chromosomes of the different datasets. The relationship between recombination fraction and interference can not be described with a uniform function, neither for the entire genome nor for individual chromosomes. The results from our analysis on the location of different degrees of interference in the genome indicate rather that interference does not act in the same way everywhere in the genome. The location of the interval of interest is important as well and the centromere, in particular seems to have an important influence on the degree of interference. Intervals spanning the centromere exhibited higher means for the coefficient of coincidence than did intervals only within the chromosome arms. This difference was particularly obvious for small intervals. More double recombinations took place in the genetically small region around the centromere than within the chromosome arms. However, physically this is a rather large region (Kuenzel et al. 2000).

Reports of negative interference in the literature

Coefficients of coincidence larger than one, indicating negative interference, have also been reported for *Drosophila*. Green (1975) analysed three-point crosses around the centromere of chromosome 3. Combinations of marker triples spanning the centromere showed negative interference with coefficients of coincidence up to 1.5. Interval size varied between 3 cM and 9 cM for a sample size of about 5,000. In addition, three-point crosses in the vicinity of the centromere exhibited negative interference as well. Green (1975) supposed a centromere effect, which facilitates the occurrence of more double recombination events than expected with no interference. Green also cited earlier results from Morgan et al. (1925), who observed the largest coefficient of coincidence for chromosomes 2 and 3 of *Drosophila* in the centromere region. Sinclair (1975) analysed recombination in a 4 cM interval around the centromere of chromosome 3 of *Drosophila melanogaster* in different samples of 4,000–37,000 individuals. The highest values of negative interference were found immediately at the centromere. He already noticed that while the region showing negative interference was genetically small, it represented a large portion of the physical length of the chromosome. Sinclair (1975) mentions possible causes for the high coefficient of coincidence without giving more de-

tails. In the study of Denell and Keppy (1979) these causes were ruled out, and the authors suggest that negative interference may be characteristic of chromosomal regions showing few recombinations in relation to their physical length.

For barley, Søgaard (1974) estimated coefficients of coincidence of 8.7, 2.3 and 1.9 for recombination frequencies of 3.8%, 9.9% and 14%, respectively, on chromosome 7H, and she also noted that the intensity of negative interference appeared to be related to the genetical distance. Since 2,600 plants were analysed, statistical variation was ruled out as an explanation for the negative interference observed. An indication of negative interference was also found at chromosome 5H (Larsson 1985). Säll (1990) used the same markers as Søgaard (1974) for investigating the genetic control of recombination frequencies. For one marker triple the coefficient of coincidence was estimated in four different crosses with 1,200 to 1,600 plants each, and in all crosses negative interference was observed. The experiment was repeated for three crosses, with 2,700–5,500 individuals each. Only one coefficient of coincidence was larger than 1, but the remaining coefficients were still larger than expected values based on the Kosami function. In a theoretical approach, Säll and Bengtsson (1989) showed that heterogeneity of recombination frequencies between plants or within plants within a population can result in coefficients of coincidence larger than one even in the absence of negative interference. This possibility was investigated in the experiments of Säll (1990), but a variation of recombination frequencies within the different crosses was not detected.

The distribution of recombination events in earlier versions of the datasets S×M and P×N including fewer markers has already been analysed by Säll and Nilsson (1994). They observed negative interference for small intervals, but this was attributed to misclassifications of marker data.

In experiments on tagging a resistance gene in *Triticum dicoccoides*, Peng et al. (1999) observed high coefficients of coincidence around the centromere of chromosome 1B, with highest values of negative interference for intervals spanning the centromere. Because of the small sample size of 150 plants, Peng et al. (1999) are cautious with the interpretation of the observed pattern of positive and negative interference. In a more recent publication Peng et al. (2000) report on negative interference in segments spanning or proximal to the centromere in nearly all chromosomes of *Triticum dicoccoides*. These proximal segments comprise about 50–70% of the physical length of the chromosome but only 5–20% of the genetic length (Gill et al. 1996; Lukaszewski and Curtis 1993; Peng et al. 2000). In some chromosomes, additional islands of negative interference were found in median or subterminal regions (Peng et al. 2000).

Models for negative interference

In their paper on *Drosophila melanogaster* Denell and Keppy (1979) propose that negative interference may be a normal characteristic of all regions with few recombinations in relation to physical length. They suggest that when regions with few recombinations do undergo recombination the state of pairing of these regions generates negative interference. This interpretation is based on a modification of the ‘effective pairing’ model of Pritchard (1960), which was devised to account for high coefficients of coincidence in intervals smaller than 0.1 cM. Pritchard (1960) assumes that recombination is restricted to very short ‘effectively paired’ segments of approximately 0.4 cM. These segments are discontinuously distributed such that only a small fraction of the genome is ‘effectively paired’ in any one zygote. The probability of recombination in these segments is very high. Coefficients of coincidence larger than 1 will result when the interval is not considerably larger than the ‘effectively paired’ segment.

Cavalli-Sforza and Jinks (1956) proposed a similar model to explain negative interference extending over larger sections of the genome in *E. coli*. For one marker triple a large variability in recombination frequencies (about 0.04–0.4) and corresponding coefficients of coincidence (about 0.7–16) were observed in 24 different crossings. The relation between recombination frequency and coefficient of coincidence was similar to that shown in figure 1: the smaller the recombination frequency the higher the coefficient of coincidence. The authors explained this observation with incomplete and irregular pairing. Pairing and therefore recombination takes place only in some of the cells, which results in apparent negative interference as well as in problems with the estimation of recombination frequencies and the linear arrangement of markers (Cavalli-Sforza and Jinks 1956). In mapping experiments, problems frequently arise with respect to the ordering of markers in small genetic intervals at centromeric locations (Waugh et al. 1997); this is generally attributed to computational problems in estimating the most likely marker order. If this problem only occurs in the centromeric region and not in other marker dense regions, the explanation from Cavalli-Sforza and Jinks (1956) could be suitable.

The model of Cavalli-Sforza and Jinks (1956) is analogous to the approach published by Säll and Bengtsson (1989). These authors revealed that variation in recombination frequencies within a population may lead to a bias in the estimated degree of interference. The authors concentrated on a bias towards negative interference. A coefficient of coincidence greater than one is expected in a linkage experiment if the offspring arise from heterogeneous meiotic events when there is a positive correlation between recombination frequencies along the chromosome. A large bias can only be produced when the investigated loci are closely linked most of the time, but a very small fraction of gametes is produced with much larger recombination frequencies. Thus, only in specific circumstances will a significant negative interference be observed, and there is an upper limit to the size the bias may reach for every given level of recombination. Applying their model, Säll and Bengtsson (1989) were

able – at least on principle – to explain the high coefficients of coincidence found by Green (1975) and Søgaard (1974). This was not the case with regard to the results obtained by Säll (1990).

Most of the coefficients of coincidence found in the present study were larger than the upper limit of the model and could therefore not be explained with it.

Recent studies on the processes that occur during meiosis (reviewed by Roeder 1997 and Dawe 1998) indicate that recognition of homologous chromosomes precedes and promotes recombination. It has been postulated that the sites of early pairing subsequently serve as sites for the initiation of recombination. The traditional assumption that synapsis is required for recombination seems to be incorrect, and it is now being assumed that the early events of recombination may precede synapsis. Several cytological observations support the view that synapsis initiates at the sites of recombination events. However, in many organisms, the number of sites of synaptic initiation exceeds the number of crossovers.

Sybenga (1999) reported that centromeric regions tend to be late in synapsing or do not synapse at all and often do not show any crossovers. Centromeres actually appear to interfere with synapsis and especially with genetic exchange.

The answer to the question why recombination is suppressed around the centromere still remains unclear. This centromere effect has been documented in a wide range of organisms. (Choo 1998)

Our investigation of barley data revealed positive interference in the chromosome arms and no or mostly negative interference in the centromeric region. Interference in barley seems to depend on the overall frequency of recombination in relation to the physical length. Positive interference appears in physically small regions, where recombination occurs very frequently, but rarely as double or multiple recombination events. In centromeric regions, the frequency of recombination is reduced in relation to the physical length, but when recombination does take place, then it does so mainly in form of double or multiple events.

These results raise the interesting question of whether or not this pattern can be explained by a single model for the regulation of recombination. Does the process of recombination differ between regions with hot spots of recombination and regions with suppressed recombination, respectively?

In addition to our results in barley, we also observed negative interference in rice, rye and sugar beet (Weber and Esch 2000). In view of all the reports on interference it may be postulated that negative interference is a general phenomenon of the genetic small but physical large centromeric region.

Our data show that the use of the Kosambi or Haldane mapping function does not seem to be appropriate in general. Algorithms have to be developed that consider the actual degree of interference during the mapping process. Ordering of markers is usually done under the assumption of no interference, and little is known about the sensitivity to interference. Speed et al. (1992) proved that with a sufficiently large number of individuals the estimation of the marker order assuming no interference remains consistent even when interference occurs. However, the necessary number of individuals is unknown and may be very large in practice. Most likely, the consideration of a reasonable model for interference into the mapping process will allow a more efficient estimation of marker orders (Speed et al. 1992).

Acknowledgements The authors thank Günter Wricke for directing the attention to interference and for encouragement and suggestions. Our thanks are extended to Lore Westphal for critical reading of the manuscript.

References

Bailey NTJ (1961) Introduction to the mathematical theory of genetic linkage. Clarendon Press, Oxford

Becker J, Vos P, Kuiper M, Salamini F, Heun M (1995) Combined mapping of AFLP and RFLP markers in barley. Mol Gen Genet 249:65–73

Cavalli-Sforza LL, Jinks JL (1956) Studies on the genetic system of *Escherichia coli* K-12. J Genet 54:87–112

Choo KHA (1998) Why is the centromere so cold? Genome Res 8:81–82

Dawe RK (1998) Meiotic chromosome organisation and segregation in plants. Annu Rev Plant Physiol Plant Mol Biol 49:371–395

Denell RE, Keppy DO (1979) The nature of genetic recombination near the third chromosome centromere of *Drosophila melanogaster*. Genetics 93:117–130

Egel R (1995) The synaptonemal complex and the distribution of meiotic recombination events. Trends Genet 11:206–208

Gill KS, Gill BS, Endo TR, Taylor T (1996) Identification and high-density mapping of gene-rich regions in chromosome group 1 of wheat. Genetics 144:1883–1891

Graner A, Jahoor A, Schondelmaier J, Siedler H, Pillen K, Fischbeck G, Wenzel G, Herrmann RG (1991) Construction of an RFLP map of barley. Theor Appl Genet 83:250–256

Graner A, Bauer E, Kellermann A, Kirchner S, Muraya JK, Jahoor A, Wenzel G (1994) Progress of RFLP-map construction in winter barley. Barley Genet Newsl 23:53–59

Green MM (1975) Conversion as a possible mechanism of high coincidence values in the centromere region of *Drosophila*. Mol Gen Genet 139:57–66

Haldane JBS (1919) The combination of linkage values, and the calculation of distance between the loci of linked factors. J Genet 8:299–309

Heun M, Kennedy AE, Anderson JA, Lapitan NLV, Sorrells ME, Tanksley SD (1991) Contruction of a restriction fragment length polymorphism map for barley (*Hordeum vulgare*). Genome 34:437–447

Kleinhofs A, Kilian A, Saghai Maroof MA, Biyashev RM, Hayes P, Chen FQ, Lapitan N, Fenwick A, Blake TK, Kanazin V, Ananiev E, Dahleen L, Kudrna D, Bollinger J, Knapp SJ, Liu B, Sorrels M, Heun M, Franckowiak JD, Hoffman D, Skadsen R, Steffenson BJ (1993) A molecular, isozyme and morphological map of the barley (*Hordeum vulgare*) genome. Theor Appl Genet 86:705–712

Kosambi DD (1944) The estimation of map distances from recombination values. Ann Eugenics 12:172–175

Kuenzel G, Korzun L, Meister A (2000) Cytologically integrated physical restriction fragment length polymorphism maps for the barley genome based on translocation breakpoints. Genetics 154:397–412

Larsson HEB (1985) Linkage studies with genetic markers and some *laxatum* barley mutants. Heriditas 103:269–279

Lukaszewski AJ, Curtis CA (1993) Physical distribution of recombination in B-genome chromosomes of tetraploid wheat. Theor Appl Genet 86:121–127

Morgan TH, Sturtevant AH, Bridges CB (1925) The genetics of *Drosophila*. Bibl Genet 2:1–262

Muller HJ (1916) The mechanism of crossing-over. Am Nat 50:193–221, 282–305, 350–366, 421–434

Nilsson NO, Säll T, Bengtsson BO (1993) Chiasma and recombination data in plants: are they compatible? Trends Genet 9:344–348

Ott J (1997) Testing for interference in human genetic maps. J Mol Med 75:414–419

Peng JH, Fahima T, Röder MS, Li YC, Dahan A, Grama A, Ronin YI, Korol AB, Nevo E (1999) Microsatellite tagging of the stripe-rust resistance gene *YrH52* derived from wild emmer wheat, *Triticum dicoccoides*, and suggestive negative crossover interference on chromosome 1B. Theor Appl Genet 98:862–872

Peng JH, Korol AB, Fahima T, Röder MS, Ronin YI, Li YC, Nevo E (2000) Molecular genetic maps in wild emmer wheat, *Triticum dicoccoides*: genome-wide coverage, massive negative interference, and putative quasi-linkage. Genome Res 10:1509–1531

Pritchard RH (1960) Localized negative interference and its bearing on models of gene recombination. Genet Res 1:1–24

Roeder GS (1997) Meiotic chromosomes: it takes two to tango. Genes Dev 11:2600–2621

Säll T (1990) Genetic control of recombination in barley: II. Variation in linkage between marker genes. Hereditas 112:171–178

Säll T, Bengtsson BO (1989) Apparent negative interference due to variation in recombination frequencies. Genetics 122:935–942

Säll T, Nilsson NO (1994) Crossover distribution in barley analysed through RFLP linkage data. Theor Appl Genet 89:211–216

Sinclair DA (1975) Cross over between closely linked markers spanning the centromere of chromosome 3 in *Drosophila melanogaster*. Genet Res 26:173–185

Søgaard B (1974) The localization of eceriferum loci in barley. 3. Three point tests of genes on chromosome 1 in barley. Hereditas 76:41–47

Speed TP, McPeek MS, Evans SN (1992) Robustness of the no-interference model for ordering genetic markers. Proc Natl Acad Sci USA 89:3103–3106

Sybenga J (1996) Recombination and chiasmata: few but intriguing discrepancies. Genome 39:473–484

Sybenga J (1999) What makes homologous chromosomes find each other in meiosis? A review and an hypothesis. Chromosoma 108:209–219

Sym M, Roeder GS (1994) Crossover interference is abolished in the absence of a synaptonemal complex protein. Cell 79:283–292

Waugh R, Bonar N, Baird E, Thomas B, Graner A, Hayes P, Powell W (1997) Homology of AFLP products in three mapping populations of barley. Mol Gen Genet 255:311–321

Weber WE, Esch E (2000) Genetic mapping of Mendelian markers. Vortr Pflanzenzuecht 48:120–131

Wilcoxon F (1946) Individual comparison of grouped data by ranking methods. J Econ Entom 39:269–270