A new index for estimating genetic recombination from chiasma distribution data

PABLO C. COLOMBO*

*Laboratorio de Genética, Departamento de Ciencias Biológicas, Facultad de Ciencias Exactas y Naturales,
Universidad Nacional de Buenos Aires, Int. Güiraldes y C. Norte (1428), Buenos Aires, Argentina*

Chiasma frequency and position are good estimators of genetic recombination in a population. Darlington proposed a recombination index that consisted of the sum of the haploid number (n) and the average chiasma number per nucleus (x); this index $(n + x)$ represents the average number of elements that segregate independently, so that the number of possible gametes equals $2^{(n+x)}$. Here a more accurate approach is intended by proposing an index that takes into account: (a) the proportional length of chromosomes and (b) the accurate position of chiasmata.

(a) Let G be the product of $2(n)(L_i)$, where L_i is the proportional length of each chromosome; if every chromosome has the same length $(L_1 = L_2 = ... = L_n)$, then $L_i = 1/n$, and $G = 2^n$ is the number of possible gametes.

(b) Let G' be the product of $2(n + x)(L_i)$; where L_i here is the length of each one of the $n + x$ chromosome regions that segregate independently. Note that if $L_1 = L_2 = ... = L_n$ and chiasmata are sited in the middle of each chromosome, $L_i = 1/(n + x)$, $G' = 2^{(n+x)}$ is the possible number of gametes, and $lg_2 G' = (n + x)$ is the Darlington index. It is then proposed that $I_r = lg_2 G'$ is a more general index of recombination that represents something like an *effective number* of independently segregating units, weighted by their proportional lengths. A further improvement to this index (I'_r) that renders less dispersed values is also worked out.

The development of simulation programs and the study of real cases (v.gr. *Leptysma argentina*) demonstrated that the index of Darlington (a special case of I_r and I'_r) is *always* an overestimation of the recombination level of a population. Hence indexes I_r and I'_r are proposed for organisms that allow a reasonably accurate measurement of chiasma distribution.

Keywords: chiasma distribution, genetic variability, recombination index.

Introduction

Chiasma frequency and position are usually considered as good estimators of genetic recombination level in a population, especially in organisms where genetic analyses are difficult or impossible to carry out either because good markers are lacking or due to the difficulty in obtaining offspring in captivity, or both (Jones, 1987). However, chiasma frequency may be insufficient when different populations of the same species that show subtle differences in these values are compared. In populational cytogenetic studies, where the calculation of changes in recombination levels is central to the adaptationist arguments (as is the case in *Drosophila* species), it is useful to handle an estimator that can cope with most of the changes that rearrange-

ments cause. Carson (1955, 1958) was the first to report the central-marginal pattern in *D. robusta*, which he compared with that observed by Dobzhansky and co-workers in South American *Drosophila willistoni* (Da Cunha & Dobzhansky, 1954). This pattern was later uncovered in other *Drosophila* species and even in Orthopterans (Hewitt, 1979; Colombo, 1989). As recombination is suppressed in the mutually inverted sequences of inversion heterozygotes, Carson (1955) aimed to express the level of heterozygosity by means of an index of free recombination (IFR), which consisted simply of the proportional length of the genome undisturbed by polymorphic inversions. This index was later used to perform comparisons and correlations with ecological variables to test Carson's hypothesis of *heteroselection*. Irrespective of the final destiny of this hypothesis, the index was shown to be fairly operative, and was still used by Brussard (1984) in a more modern approach to the same problem.

*Member of CONICET (Consejo Nacional de Investigaciones Científicas y Técnicas).

In the case of Orthopterans, chiasma frequency is universally accepted as a recombination level estimator (see Hewitt, 1979 for a review). However, despite the inadequacies mentioned above, this variable shows a flaw in the case of Robertsonian polymorphisms: it does not take into account the variation in the number of linkage groups. Darlington (1932) proposed an index of recombination, which consisted of the sum of the haploid number n and the average chiasma number per population x, and which represents the average number of independently segregating elements, so that the possible number of different gametes is $2^{(n+x)}$. However, this index does not take into account possible variations in chiasma position; in fact, interstitial chiasmata are expected to yield a maximum of recombination and terminal ones a minimum. Tosto & Bidau (1987) developed a modified version of the Darlington index, that consisted of the sum of the haploid number (or the number of segregating groups, in the case of fusion heterozygotes) and the number of *interstitial* chiasmata. This index has two major shortcomings, however: it still ascribes the same weight to *all* interstitial chiasmata (the determination of proximal, interstitial and distal chiasmata is quite subjective) and does not consider proximal chiasmata which, in the case of metacentric chromosomes, also contribute to total recombination. Moreover, the Darlington index (modified or not) gives the same statistical weight to *all* chromosomes, irrespective of their length, which in the largely asymmetrical karyotypes of Orthopterans is an over simplification. A more accurate approach is intended here by working out an index that takes into account (a) the relative size of chromosomes, and (b) the accurate position of chiasmata.

Results

Derivation of recombination index

(a) Let G be the product for all i of $2(n)(L_i)$, where L_i is the proportional length of each chromosome (the sum of all L_i is 1); if all the chromosomes had the same length $(L_1 = L_2 = ... = L_n)$, then $L_i = 1/n$, and $G = 2^n$ is the possible number of different gametes.

(b) Let G' be the product for all i of $2(n+x)(L_i)$, where L_i is now the length of each one of the $n+x$ independently segregating chromosome regions. Note that in the special case in which all the $n+x$ regions have the same length $(L_1 = L_2 = ...L_n)$, i.e. all chromosomes have the same size and all chiasmata are sited in the middle of the chromosome, then $L_i = 1/(n+x)$, the possible number of different gametes would be $G' = 2^{(n+x)}$, and the logarithm base 2 of G' would be $(n+x)$, that is the Darlington index.

Therefore the index I_r is proposed, namely

$$I_r = lg_2 G'$$

where

$$G' = \prod_{i=1}^{n+x} 2(n+x)L_i$$

as a more general index of recombination which represents the *effective number* of independently segregating chromosome segments, weighted by their proportional lengths, per cell. Thus, the mean recombination level in a population sample can be obtained by averaging I_r in a sample of cells among all individuals collected. According to (b), Darlington's index is a special type of index when all independently segregating chromosome segments have the same length. More specifically, the Darlington index (DI) would always be an overestimation of I_r.

Application of recombination index

Of course, I_r would be operative only in those cases in which chiasma positions were determined with accuracy. We can consider that most of the Orthopterans fall into this category, especially as modern image processors simplify and accelerate the measurement and later processing of data. A BASIC program (GW BASIC, Microsoft) is available to calculate this index.

This program was used to obtain the mean I_r in all individuals employed in a chiasma distribution study on the grasshopper *Leptysma argentina* (Colombo, 1991). This species has a basic chromosome number $2n = 21$ ♂/22 ♀ and displays a typical XO/XX sex determination mechanism. Chromosome 1 is metacentric, the remainder being telocentrics. This karyotype is modified by a polymorphic centric fusion between chromosomes 3 and 6 of the standard complement, which determines the presence of three different karyotypes in natural populations, namely: standard homozygotes (BB), heterozygotes (BF) and fusion homozygotes (FF). This fusion occurs in frequencies from 0 to 1 in several Argentine populations of this species (Colombo, 1989). It has been shown in previous papers (Colombo, 1987; 1990) that fusion 3/6 exerts a significant effect on decreasing total chiasma frequency by preventing proximal (with respect to the centromere) chiasma formation. This effect is due to enhanced interference distances and to the operation of trans-centromere interference (Colombo, 1991). In Table 1 the average values of DI and N (the effective number of independently segregating fragments) for each karyotype are given, together with I_r and I_r'. N is much lower than DI, given that distal chiasmata are the most frequent ones, which leads to a rather unequal segment size distribution.

Table 1 Average recombination indexes (I_r and I'_r), over 86 cells of nine fusion homozygotes, 95 cells of 10 heterozygotes and 76 cells of eight basic homozygotes. The Darlington index (DI) and the number of independently segregating fragments (NF) (sum of x_i, see text) are also shown. n = number of cells.

Karyotype	n	I_r	I'_r	DI	N
Basic homozygotes (BB)	76	9.9378	13.7420	24.8026	16.9736
Heterozygotes (BF)	95	8.0948	11.1444	22.5895	14.4632
Fusion homozygotes (FF)	86	6.3457	9.1340	22.1628	13.1628

Three shortcomings of this index may be pointed out. The first one (of theoretical nature) refers to the fact that I_r considers fragments as if they were *actually* independently segregating. Another flaw is that too much statistical weight is attached to too short segments, which diminishes the value of I_r excessively.

The latter problem was solved by employing the following criterion: segments shorter than 1 per cent of the total length of the genome should not be considered by the program.

The former inadequacy was solved by dividing the calculation of the recombination index into two steps: first, by considering the recombination caused by chiasmata *in each chromosome*; second, by taking into account the recombination elicited by independent segregation and by weighting everything by the relative length of each chromosome; this led to a slight modification of the index, as follows

$$I'_r = lg_2\left[\prod_{i=1}^{n} nL_i \prod_{j=1}^{x_i} 2x_i(l_{ij}/L_i)\right]$$

where l_{ij} is the length of each chromosome segment separated by chiasmata, x_i is the number of these fragments, L_i is the proportional length of each chromosome and n is the haploid number of chromosomes (or segregating units, in the case of Robertsonian polymorphisms). Note that if $l_{ij} = l$, $L_i = L$ and $x_i = x$, then $x = 1/(l/L)$ and $2x(l/L) = 1$. Thus I_r would remain

$$I'_r = lg_2 \prod_{i=1}^{n} nLi2^x$$

but all chromosomes with the same length, $n = 1/L$ and $nL = 1$. Then, $I'_r = lg_2(2^x)^n = 2^{(n+x)} = (n+x)$, that is the index of Darlington (DI). Again, DI is a special case (an overestimation) of I'_r. Its application in the case of *Leptysma argentina* produced the results given in Table 1.

In Tables 2 and 3 correlations with both indexes along total, proximal, interstitial and distal chiasma frequencies, and also with the effective number of

Table 2 Correlation of I_r with total, proximal, interstitial and distal chiasm frequency versus the number of independently segregating fragments (N)

	r	Mean
Basic homozygotes		
Total chiasmata	0.2309	14.8026
Proximal chiasmata	0.0016	4.0132
Interstitial chiasmata	0.5655	3.2632
Distal chiasmata	−0.3122	7.5263
NF	0.6109	16.9737
Heterozygotes		
Total chiasmata	0.0408	13.5895
Proximal chiasmata	−0.1248	3.3789
Interstitial chiasmata	0.5844	1.9789
Distal chiasmata	−0.3845	8.2316
NF	0.6473	14.4632
Fusion homozygotes		
Total chiasmata	0.2408	13.1628
Proximal chiasmata	0.2071	3.1279
Interstitial chiasmata	0.4329	1.0349
Distal chiasmata	−0.3352	9.0000
NF	0.6102	12.8256

independently segregating units (sum of x_i), are given for all three karyotypes of fusion 3/6. Both indexes behave in a fairly similar fashion but the second one is notoriously more sensitive to changes in chiasma position and is also more accurate because its variance is lower; this fact is reflected in higher correlation indexes.

A relevant point is the correlations with I chiasmata are good (Table 3); thus Tosto & Bidau's (1987) modification of DI can be considered as an acceptable approximation. But correlation with proximal chiasmata increases in fusion carriers because of the increased number of metacentrics (note that proximal chiasmata do contribute to recombination in metacentrics, but not in telocentrics). This effect is more marked in organisms with an increased proportion of metacentric chromosomes.

Table 3 Correlation of I_r' with total, proximal, interstitial and distal chiasma frequency versus the number of independently segregating fragments (NF)

	r	Mean
Basic homozygotes		
Total chiasmata	0.4698	14.8026
Proximal chiasmata	−0.2393	4.0132
Interstitial chiasmata	0.7476	3.2632
Distal chiasmata	−0.0702	7.5263
N	0.8667	16.9737
Heterozygotes		
Total chiasmata	0.0278	13.5895
Proximal chiasmata	0.0221	3.3789
Interstitial chiasmata	0.6291	1.9789
Distal chiasmata	−0.5691	8.2316
N	0.8574	14.4632
Fusion homozygotes		
Total chiasmata	0.2017	13.1628
Proximal chiasmata	0.2125	3.1279
Interstitial chiasmata	0.7350	1.0349
Distal chiasmata	−0.6381	9.0000
N	0.8658	12.8256

Conclusion

I_r is proposed as an index to measure recombination levels in organisms that are difficult to handle from a genetical standpoint, but which allow the observation and measurement of chiasma distribution. A modification of the Darlington index using interstitial rather than total chiasma frequency is verified as a good approximation to the index proposed here. In the case of *Leptysma argentina* I_r will be used to perform correlations of recombination levels in different populations with morphological as well as ecological parameters.

Acknowledgements

References

BRUSSARD, P. E. 1984. Geographic patterns and environmental gradients: the central-marginal model revisited. *Ann. Rev. Ecol. Syst.*, **15**, 24–64.

CARSON, H. L. 1955. The genetic characteristics of marginal populations of *Drosophila*. *Cold Spring Harb. Symp. Quant. Biol.*, **20**, 276–287.

CARSON, H. L. 1958. Response to selection under different conditions of recombination in *Drosophila*. *Cold Spring Harb. Symp. Quant. Biol.*, **23**, 291–306.

COLOMBO, P. C. 1987. Effects of centric fusions on chiasma frequency and position in *Leptysma argentina* (Acrididae: Orthoptera). I. Spontaneous and stable polymorphic centric fusions. *Genetica*, **71**, 171–179.

COLOMBO, P. C. 1989. Chromosome polymorphisms affecting recombination and exophenotypic traits in *Leptysma argentina* (Orthoptera). *Heredity*, **62**, 289–299.

COLOMBO, P. C. 1990. Effects of centric fusions on chiasma frequency and position in *Leptysma argentina* (Acrididae: Orthoptera). II. Intra- and interchromosome effects. *Caryologia*, **43**, 131–147.

COLOMBO, P. C. 1993. A polymorphic centric fusion enhances chiasma interference in *Leptysoma argentina* (Orthoptera): a chiasma distribution study. *Heredity* (in press).

DA CUNHA, A. B. AND DOBZHANSKY, TH. 1954. A further study of chromosomal polymorphism in *Drosophila willistoni* in relation to environment. *Evolution*, **8**, 119–134.

DARLINGTON, C. D. 1932. *Recent Advances in Cytology*. Churchill Livingstone, London.

JONES, G. H. 1987. Chiasmata. In: Moens, P. (ed.) *Meiosis*. Academic Press, New York.

HEWITT, G. M. 1979. *Animal Cytogenetics. Vol. 3. Insecta 1: Orthoptera*. Berlin, Gebrüder Bornträger.

TOSTO, D. AND BIDAU, C. J. 1987. Análisis de la variabilidad morfológica de razas cromosómicas de *Dichroplus pratensis* en relación a la capacidad potencial de recombinación. II. Zonas Híbridas. *Actas 1 Congr. Entomol.*, **21**, (Tucumán).