



## Genetics and statistical association between lethal alleles and quantitative yield factors in maize (*Zea mays* L.)

Salerno, Juan Carlos<sup>1</sup>; Kandus, Mariana<sup>1</sup>; Boggio Ronceros, R...<sup>2</sup>; Sorarrain, O...<sup>2</sup>; González, Cecilia<sup>2</sup>; Almorza, David<sup>3</sup>

<sup>(1)</sup> Instituto de Genética "Ewald A. Favret"; INTA-Castelar. C.C. 25-1712, Castelar. Argentina.

<sup>(2)</sup> Facultad de Ciencias Agrarias y Forestales. Universidad Nacional de La Plata; Argentina.

<sup>(3)</sup> Departamento de Matemáticas. Universidad de Cádiz, España.

Corresponding author: Ing. Agr. Juan Carlos Salerno  
Instituto de Genética "IGEAF". C.C. 25-1712; Castelar, Pcia. de Buenos Aires, Argentina.  
E-mail: jsalerno@cnia.inta.gov.ar

---

### ABSTRACT

The objective of this work is to study the genetic process related to the hybrid vigour (heterosis) and the identification of quantitative trait loci (QTL) in maize.

A biometric model was built using the mathematical formalism relating to the discrete absorbent Markov chain in canonical form, in order to analyse the evolution of chromosome segments, with recessive lethal genes linked with grain yield factors through generations. The grain yield obtained from an inbred maize line regulated by a balanced lethal system, and other lines without this system as control, were evaluated during five inbreeding generations. The use of the Markov chain theory for the study of evolution during these generations was a different approach of the classical mathematical ones.

Key words: Maize, lethal alleles, heterosis, Markov Chains, QTL.

### RESUMEN

El objetivo de este trabajo es estudiar los procesos genéticos relacionados con el vigor híbrido (heterosis) y la identificación de QTL (Quantitative Trait Locus) en el maíz. Para obtener los resultados, se construyó un modelo biométrico usando los formalismos relacionados a las cadenas absorbentes discretas de Markov en forma canónica, con el fin de analizar la evolución de segmentos cromosómicos, con genes recesivos letales ligados con factores de rendimiento en grano a través de las sucesivas generaciones. El rendimiento en grano obtenido de una línea de maíz regulada por un sistema de letales balanceados y otras líneas sin este sistema como testigo, se evaluó durante 5 generaciones de endocría. El uso de la teoría de las cadenas de Markov con el fin de estudiar la evolución durante estas generaciones fue una aproximación diferente a los métodos matemáticos clásicos.

Palabras claves: maíz, alelos letales, heterosis, cadenas de Markov, QTL.

---

### Introduction

In recent research work, maize has been used as a model for the study of molecular markers for the identification, mapping and manipulation of major

genes affecting quantitative yield factors in plants. Experimental results from numerous mapping studies have shown that quantitative traits are controlled by some factors with major effects (Stuber, 1995). These results are in agreement with previous conclusions

obtained long time ago in our experimental station at Castelar (Argentina) by Favret and Godeck (1959), and Favret and Ryan (1966), in barley. In 1978, we began the present study in maize; some of our findings have been partly reported by Salerno (1981), Salerno and Favret (1984), Salerno and Gardner (1985), and Salerno et al. (1986), with similar conclusions to those reported by Stuber (1995).

The present study focuses on the existence of some linkage groups in which lethal infertile genes associated with yield factors are present. These groups should be considered as special groups of major genes affecting quantitative traits in plants or animals. This hypothesis is supported by the finding of natural balanced lethal systems (Crumpacker, 1967). These systems seem to be useful to maintain the linkage because of the high frequency of lethal genes in the population. They are also useful to show the maximum genetic load (Muller, 1950) that plants or animals can tolerate (Carson, 1967). Concerning genetic load, Wallace (1970, 1981) has published a complete review on this subject. Crumpacker (1967), Gustafsson (1946), and Burnam (1993), among other authors, have revised the genetic load showing that the number of recessive mutations is quite high in different populations, either wild or domesticated, including also human beings.

These considerations moved us to study the maize balanced lethal system as a system itself in order to identify, isolate and keep up through the balanced lethal systems.

It is also useful to predict the number of generations during which a lethal system maintains the linkage, and the distance between genes with two or more lethal alleles at the repulsion phase, so as to get permanent heterozygosis without recombination. Thus, the theory of discrete Markov chains in space and time was used for the analyses (Barucha and Reid, 1960; Bosso et al., 1969; Sorarrain et al., 1980; Boggio et al 1997; Salerno et al., 1999).

## Material and Methods

We have analysed grain yield corresponding to different kinds of selfing cross lines including lethal factors in three maize populations: a) Colossal original population, flint (type of endosperm grain), variety population of open pollination; b) Colossal with genetic load, same population with presence of lethal factors like deficient chlorophyll mutations; c) Colossal without

genetic load, same population without chlorophyll mutations.

From these populations, BLS1 Argentine flint line was obtained for this study. It is regulated by a balanced lethal system located at the pericentric region of chromosome six, and present two varieties: BLS1 1:1, with two lethal alleles; and BLS1 3:1, with only one lethal allele.

The results obtained about the grain yield are compared with that corresponding to different kinds of public lines: N28; B73; MO17 and B14. These are inbred lines used as control during five years at the same location.

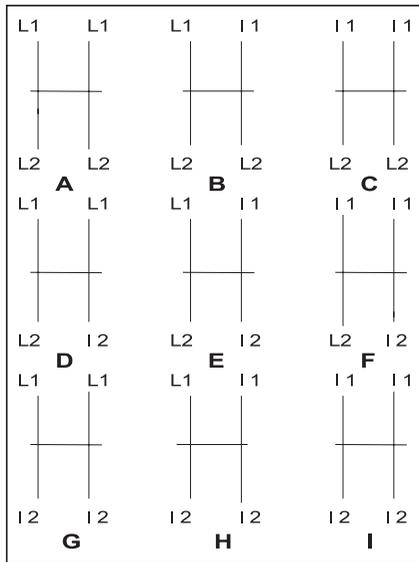
In order to check our findings, we considered another Argentine flint line regulated by a balanced lethal system, BLS14, and it was then compared through a grain yield experiment using the same line without lethal alleles.

In all these lines a randomised block was designed with three replications in a row of 7 x 1.4 meters. Two kernels per gap were planted, leaving one plant per gap at harvest in order to have 60 plants per plot. The experiments were performed at the same station in Castelar.

Finally, we evaluated the combining ability between this balanced lethal lines with public lines with commercial hybrids during two years, at two locations during the first year (Junín and Castelar), and at three locations, the second year (Junín, Castelar and Los Hornos). A randomised block was then designed with two replications in a row of 7 x 1.4 meters. Two kernels per gap were planted, leaving one plant per gap at harvest in order to have 60 plants per plot.

The model for the analysis of the states comprises different kinds of selfing crosses, and it has been performed assuming each state as defined by self fertilisation of any genotype, as shown in figure 1. These genotypes belong to some isolated populations located at a limited closed field. The genotypes entail the action of two genes, and in both loci, lethal alleles  $l_1$  and  $l_2$  have been found in natural form in the population.

**Figure 1.** Means separated chromosomes (actually in the same chromosome with negligible linkage). Cases 1-3. Fully linked chromosomes.



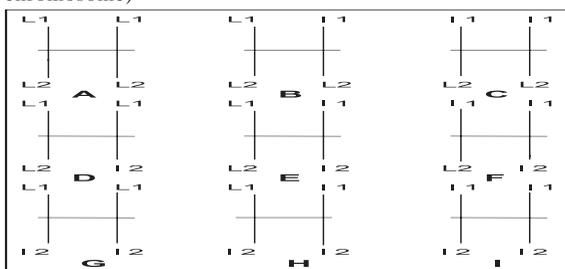
Results

A) About the model

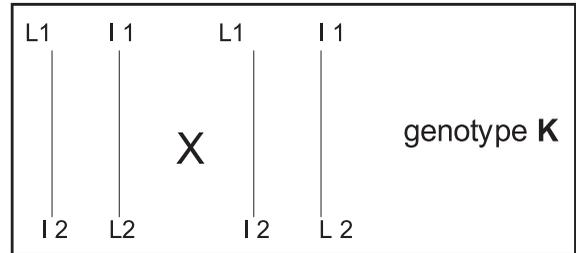
Since the model involves (we have) a self fertilisation, the lethal effect at the homozygous state of alleles  $l_1$  and  $l_2$  will be considered as recessive, since any of them appears in both chromosomes for each locus, in crosses CC, FF and GG. Obviously, the same occurs in II cross, where recessive alleles were found in both loci. These genotypes must be also considered as infertile because they will disappear from the field for the following generation.

According with the absorbent Markov chains theory, genotypes A, C, F, G, H and I act as absorbent states. Also, during self fertilisation of genotype A, the same cross is obtained in the following generation; therefore AA is also an absorbent non-lethal state. Transient states are only BB, DD, EE and also KK (figures 2 and 3).

**Figure 2.** Non balanced system (lethals in the same chromosome)



**Figure 3.** Balanced system (lethal in opposite chromosomes). Fully linked loci



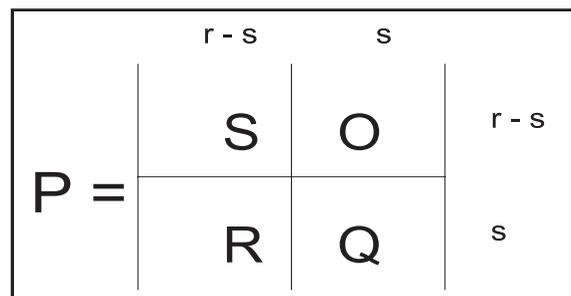
Afterwards, beginning with selfing crosses for the independent segregation case (loci with very weak linkage, practically independent for the segregation), selfing crosses AA, CC, FF, GG, HH, and II can be recognised as absorbent states, while BB, DD, EE are equivalent to KK.

Under these conditions, following Mendel's laws for the calculation of the resulting probabilities for the selfing in the following generation, results are given by the transition matrix P written in canonical form, in the way shown in table 1 and figure 4.

**Table 1.** Matrix P in canonical form

	AA	CC	FF	GG	HH	II	BB	DD	EE
AA	1	0	0	0	0	0	0	0	0
CC	0	1	0	0	0	0	0	0	0
FF	0	0	1	0	0	0	0	0	0
GG	0	0	0	1	0	0	0	0	0
HH	0	0	0	0	1	0	0	0	0
II	0	0	0	0	0	1	0	0	0
BB	1/4	1/4	0	0	0	0	1/2	0	0
DD	1/4	0	0	1/4	0	0	0	1/2	0
EE	1/16	1/16	1/8	1/16	1/8	1/16	1/8	1/8	1/4

**Figure 4.** Matrix P, in canonical form;  $r = 9, s=3$ .



For an absorbing chain, only transient and absorbent states must be considered. The transition matrix P can be written as four sub-matrices: sub-matrix O, a zero matrix of s-(r-s) order; sub-matrix Q, providing the probabilities for the transition between transient states; sub-matrix R, denoting the probabilities between transient and absorbent states; and sub-matrix S, (r-s)·(r-s) identity matrix. From this matrix P it is possible to calculate many characteristic parameters of the evolution process (Boggio et al. 1997). The relative frequencies of the genotypes from the starting to the n-th generation are given by the relation (1) or (2) (Bosso et al. 1969),

$$\Pi_n = \Pi_0 \cdot P^n \tag{1}$$

$$\Pi_{n,k} = \Pi_{0,j} \cdot P_{j,k}^n \tag{2}$$

where  $\Pi_n$  is a row vector with k-th component  $\Pi_{n,k}$ , and  $\Pi_0$  another row vector with j-th component  $\Pi_{0,j}$ .  $\Pi_0$  denotes the relative frequencies of the starting generation. We shall now consider the following situations:

a) Independence between both loci.

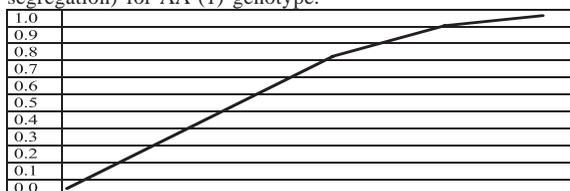
When both loci are independent, starting from vector  $\Pi_0=(0,0,0,0,0,0,1)$ , only genotype E is at the beginning in the field population. The relative frequencies for the non-lethal genotypes are shown in table 2, from the first to the sixth generation.

**Table 2.** Generations for independence between both loci (lethal states have been excluded from the table, and the relative frequencies normalized)

Number of generations						
	1 <sup>st</sup>	2 <sup>nd</sup>	3 <sup>rd</sup>	4 <sup>th</sup>	5 <sup>th</sup>	6 <sup>th</sup>
EE	0.444	0.160	0.050	0.014	0.004	0.0009
DD	0.222	0.240	0.172	0.104	0.055	0.0215
BB	0.222	0.240	0.172	0.104	0.055	0.0215
AA	0.111	0.360	0.605	0.780	0.880	0.9400

In these calculations, non-selectivity between both loci has been assumed; this is, non-selective advantages in frequencies between mono- and double heterozygous genotypes have been considered. In this case, the only absorbent non-lethal genotype increases in the way shown in figure 5.

**Figure 5.** Relative frequencies in the population (independent segregation) for AA (1) genotype.



b) The fully linked genotype E

The fully linked genotype E (figure 2) shows the lethal alleles in the same chromosome. In this case, starting at  $\Pi_0$  with only genotype E in the population, this is, for fully linked loci (negligible crossing-over), and applying Mendel's rules, the elements of the last row of matrix P must be changed in the way shown in table 3.

**Table 3.** Matrix P in canonical form, starting with only genotype E in the population

	AA	CC	FF	GG	HH	II	BB	DD	EE
AA	1	0	0	0	0	0	0	0	0
CC	0	1	0	0	0	0	0	0	0
FF	0	0	1	0	0	0	0	0	0
GG	0	0	0	1	0	0	0	0	0
HH	0	0	0	0	1	0	0	0	0
II	0	0	0	0	0	1	0	0	0
BB	1/4	1/4	0	0	0	0	1/2	0	0
DD	1/4	0	0	1/4	0	0	0	1/2	0
EE	1/4	0	0	0	0	1/4	0	0	1/2

Summing up, starting from the initial vector  $\Pi_0=(0,0,0,0,0,0,1)$  only two absorbent states must be considered, lethal II state and non-lethal AA, already mentioned. In this case, after a few generations, non-lethal genotypes A dominate in the population (table 4), in a similar way as shown in figure 5.

**Table 4.** Generations for the fully linked genotype E (relative frequencies for non-lethal genotypes)

Number of generations						
	1 <sup>st</sup>	2 <sup>nd</sup>	3 <sup>rd</sup>	4 <sup>th</sup>	5 <sup>th</sup>	6 <sup>th</sup>
EE	0.666	0.400	0.223	0.119	0.065	0.037
AA	0.333	0.600	0.770	0.880	0.930	0.960

c) The fully linked genotype K

For the fully linked genotype K (figure 3), with lethal alleles  $l_1$  and  $l_2$  in opposite chromosomes, this is, for the balanced lethal systems, we consider selfing without crossing over. A new matrix P should be defined including the row for lethal alleles  $l_1$  and  $l_2$  in the same chromosome. This new matrix P ( $r = 10$  and  $s = 4$ ) is shown in table 5.

**Table 5.** Matrix P in canonical form, considering fully linked genotype K.

	AA	CC	FF	GG	HH	II	BB	DD	EE	KK
AA	1	0	0	0	0	0	0	0	0	0
CC	0	1	0	0	0	0	0	0	0	0
FF	0	0	1	0	0	0	0	0	0	0
GG	0	0	0	1	0	0	0	0	0	0
HH	0	0	0	0	1	0	0	0	0	0
II	0	0	0	0	0	1	0	0	0	0
BB	1/4	1/4	0	0	0	0	1/2	0	0	0
DD	1/4	0	0	1/4	0	0	0	1/2	0	0
EE	1/4	0	0	0	0	1/4	0	0	1/2	0
KK	0	1/4	0	1/4	0	0	0	0	0	1/2

Then, for the selfing crosses, the absorbent lethal infertile genotypes G and C and the non-lethal balanced genotype K are the ones which grow in the field. In other words, the balanced fully linked K is the only one that survives in the population

It should be interesting to include in the analysis the following theorems related with absorbent Markov chains (Kemeny and Snell, 1960):

**Theorem 1:** In any finite absorbing Markov chain, no matter the initial distribution  $\Pi_0$ , the probabilities that process will be finished in any one of the states belonging to the set of absorbing states tend to 1 as n tends to infinity.

**Theorem 2:** There are numbers  $b > 0$  and  $0 < c < 1$ , so that  $P_{ij}^n \leq b \cdot c^n$  for any transient states  $S_i$  and  $S_j$ , that is,  $P_{ij}^n \rightarrow 0$  as  $n \rightarrow \infty$ .

The former paragraphs and the last theorems allow us to conclude that after a small number of generations (in an interval from six to ten), only genotypes A and K remain in the population. The recombination rate results quite relevant since the crossing leads from genotype K to E in both ways, and from genotype E to the whole dominance in the population; in other words, lethal alleles tend to disappear, and this is the reason for using balanced lethal systems with negligible recombination rate.

The use of balanced lethal systems for the study of short segments is not new, but we have never found previous theoretical studies concerning their persistence in the population and their relation with other genotypes including zero, one or two alleles in a given segment.

Finally, genotypes A and K can be easily detected because genotype K is the only one able to produce

infertile lethal descendents. This property is very significant for future applications.

**B) Grain yield and balanced lethal system**

The BLS1 Argentine line with a balanced lethal system has a significantly higher grain yield when it is compared with the other lines (table 6). This could indicate a selective advantage of the heterozygous segment with lethal alleles in this part of the genome.

**Table 6.** Grain yield (kg/ha) of BLS 1 and public lines.

	GENERATIONS				
	1 <sup>st</sup>	2 <sup>nd</sup>	3 <sup>rd</sup>	4 <sup>th</sup>	5 <sup>th</sup>
BLS1 1:1	3697 A	3747 A	3243 A	3579 A	3538 A
BLS1 3:1	2517 B	2713 A	2116 B	2465 A B	3047 A
N28	2017 C	2240 B	1709 C	2252 B	2675 A
B73	1793 C D	2060 B	853 D	1976 B	2573 A
MO17	1623 D	2007 B	1503 C	1776 B	2566 A
B14	1544 D	1860 B	811 D	1718 B	2506 A
	C.V. = 4,18% msd = 365,92 kg/ha	C.V. = 16,45% msd = 1100,20 kg/ha	C.V. = 5,88% msd = 400 kg/ha	C.V. = 11,24% msd = 1132 kg/ha	C.V. = 22,87% msd = 2565 kg/ha

The BLS14 Argentine line with a balanced lethal system shows significant higher grain yield than the same line without lethal alleles (table 7). This is another proof about the association of grain yield with lethal alleles already mentioned.

**Table 7.** Grain yield (kg/ha) of BLS 14 line, with and without lethal.

	GENERATIONS			
	1 <sup>st</sup>	2 <sup>nd</sup>	3 <sup>rd</sup>	4 <sup>th</sup>
BLS 14 with lethal	3967 A	3313 A	3685 A	4017 A
BLS 14 without lethal	2473 B	1981 B	1672 B	2385 B
	C.V.= 8,47 % msd= 619 kg/ha	C.V.= 4,69 % msd= 282 kg/ha	C.V.= 4,05 % msd= 246 kg/ha	C.V.= 19,23 % Msd= 637 kg/ha

**C) Combining ability**

There is a good combining ability between the lines regulated by a balanced lethal system and public lines (tables 8 and 9). In other words, the grain yield assays of balanced lethal systems (without asterisks) do not show significant differences with the single commercial hybrids obtained using the crossing of pure selected lines in every generation, as usual in commercial lines.

These experiences suggest that the use of lethal alleles could increase yield grain assays. These lethal alleles can be maintained in the population in order to make the handling of hybrids cheaper in grain production, at least in some cases.

**Table 8.** Grain yield assays of maize single hybrids belonging to the Genetics Institute of INTA Castelar, in two locations (Junín and Castelar)

Hybrid	Grain yield (kg/ha)	
	Junín	Castelar
B73CCAYB	6566	5844
<u>PIONEER 3456</u> (*)	6312	7524
CARGILL 270	6149	6023
<u>ACA 923</u> (*)	5826	7026
N28PCAYB	5688	5305
N28BLS16	5525	7407
<u>DEKALB 752</u> (*)	5471	7336
N28PCCB	5395	5999
B73BLS16	5364	8199
BL14MO17	5180	6363
B14PCCB	5133	5521
B14BLS16	5071	5290
B73BLS12	5069	7367
B73BL14	4705	8162
BL14BL16	4635	7367
Labrador 22 INTA	4512	7879
<u>ZENECA 8546</u> (*)	4501	7853
MO17BL16	4168	6507
<u>SURARAMI</u> (*)	4826	6592
N28BLS14	5220	6644
	C.V.= 14,18% msd = 3018 kg/ha	C.V.=14,13% msd = 3887 kg/ha

**Table 9.** Grain yield assays of maize single hybrids belonging to the Genetics Institute of INTA Castelar, in three locations (Junín, Castelar and Los Hornos)

Hybrid	Junín	Castelar	Los Hornos
MO17BLS14 (Labrador 22 INTA)	7916	5575	6330
F2P3069XBLS14	7790	6625	6167
F2664XBLS14	5811	6612	6610
B73 BLS14	7630	5378	6536
F2D762XBLS14	5550	6558	6765
F2P3762XBLS14	6355	6400	6517
BLS14XBLS16	5434	5580	6536
DEKALB 664XBLS14	6698	5841	6167
DEKALB 669XBLS14	6667	5760	6158
NIDERA 888 (*)	6175	5343	5058
DEKALB 664 (*)	6461	5251	5068
PIONEER 3469 (*)	5955	4396	3340
CARGILL 285 (*)	6667	6476	6900
	C.V.=15,89% msd = 4350 kg/ha	C.V.=16,39% msd = 3332 kg/ha	C.V.=12,84% msd=2665 kg/ha

## Conclusions

From the analysis of this work, we have concluded that for a given chromosomal segment containing recessive alleles, in few generations only two genotypes which have been denoted by A and K remain in the population. These genotypes can be easily distinguished, being genotype K an excellent marker for studies on chromosome segments and the association with yield in grain production. There are only two possibilities of getting recessive alleles in two loci defining a chromosome segment: with genotypes E and K, but only genotype K should be useful to maintain the heterozygous condition.

Lethal alleles seem to belong to the genetic load of the population. This genetic load might be useful for genetic studies and grain production in major genes related with recessive alleles since these heterozygotes can be compared with those produced with usual commercial techniques.

The theory of absorbing Markov chains can be used for predictions about the number of generations to maintain the heterozygous state as a function of the recombination rate (crossing-over) in these segments.

## Acknowledgements

This research work was supported by INTA and PICT 10995/02 Grant from ANPCyT (Agencia Nacional de Promoción Científica y Técnica). Mariana Kandus has been granted an ANPCyT scholarship.

## References

- Boggio, R.R., Sorarrain, O., Salerno, J. C. and Favret, E. A. (1997). Theoretical Analysis of Lethal Factors in Plant Populations. *Mathematical Biosciences*, 140, 85-99.
- Bosso, J. A., Sorarrain, O. and Favret, E. A. (1969). Application of finite Markov chains to sib mating populations with selection. *Biometrics*, 22 (1), 17-26.
- Burnam, Ch. R. (1993). Balanced lethal from chromosome 6. *Maize Genetics Cooperation Newsletter*. 67: 101.
- Crumpacker, D. W. (1967). Genetic loads in maize (*Zea mays* L.) and other cross-fertilised plants and animals. *Evolutionary Biology* 1: 396-415.

- Favret, E. A. and Godeck, W. (1959). Índice de mutación espontánea en cebada y otras gramíneas. Revista de Investigaciones Agrícolas T. XIII N° 3.
- Favret, E. A. and Ryan, G. (1966). Possible use of permanent heterozygosity in barley breeding. Barley Newsletter. 10: 122-123.
- Gustafsson, A. (1946). The effect of heterozygosity of variability and vigour. Hereditas. 33: 573.
- Kemeny, J. and Snell, L. (1960). Finite Markov Chains. Princeton, New Jersey: Van Nostrand
- Muller, H. J. (1950). Our load of mutations. Am. J. Hum. Gen. 2: 111-176.
- Salerno J. C. (1981). Utilización de los sistemas letales balanceados en maíz. Acta Jornadas de genética Aplicada del Noroeste Argentino. SAG. 43-51.
- Salerno, J. C. and Favret, E. A. (1984). Introduction among lethal genes in two lines of maize (*Zea mays* L.) Genetics. 107 (1). 93.
- Salerno, J. C., Favret, E. A. and Gardner, C. O. (1986). Heterotic regions in the maize genome. Agron. Abstracts. 80.
- Sorraín, O., Boggio, R.R. and Ocampo, J.C. (1980).
  - Continuous and discontinuous Markov chains for mastitis infection. Mathematics Biosciences (52), 277-287.
- Stuber, Ch. W. (1995) Plants Vol. 11 N° 12, 477.
- Wallace, B. (1970). Genetic load. Its Biological and Conceptual Aspects. Prentice-Hall, Inc. Englewood Cliffs, N.J. Chapter 8, p. 52.