Detection of Segregation distortion through molecular markers

Chitti Bharat Kumar¹ and Shiva Kumar K.²

Abstract: Segregation distortion refers to deviation of observed marker genotypic frequencies from the expected Mendelian ratios in a segregating population. This phenomenon has attracted considerable attention as an evolutionary force that can cause an increase in the population frequency of the heterozygous allele or heteromorphic chromosome in both animals and plants. Errors in marker genotyping, statistical analysis, residual variability in parental lines and mutation in the binding site of DNA marker. In addition to these, some of other reasons like environmental factors, transgenic silencing, and chromosome instability resulting from aneuploid, unstable translocation will result into distorted segregation. It is one of most important factors influencing the precision of genetic mapping, by affecting the genetic distance between markers and the order of markers on linkage groups. Thus, it is advisable to exclude markers that exhibit segregation distortion during mapping.

An eclectic way is adding up of segregation-distorted markers the map to be constructed with normally segregated markers and determine the fate of the distorted markers, based on how much the map was affected. Hence, there is no proper standard method for deleting the distorted markers from a genetic map. SD can be detected using statistical tools such as chi-square test of goodness of fit, if the observed segregation ratio significantly deviates from the expected ratio, it can be conclude that molecular markers under study are showing segregation distortion. The main scope of this article is to provide basic idea about Segregation distortion, their causes and the ways to overcome this problem.

key words: chi-square, DNA marker, Segregation distortion

INTRODUCTION

Genetic segregation experiments in plant species commonly used for understanding the inheritance of traits. A basic assumption in these experiments is that each gamete developed from megasporogenesis has an equal chance of fusing with a gamete developed from microsporogenesis, and every zygote formed has an equal chance of survival. If gametic and or zygotic selection occurs whereby certain gametes or zygotic combinations have a reduced chance of survival, progeny distributions are skewed and are said to exhibit segregation distortion. The law of segregation, which is the most fundamental law in Mendelian genetics, relies on: (1) a predictable transmission of alleles from a parent to its offspring, and (2) a predictable formation of genotypes from the transmitted alleles. Segregation distortion, a deviation of the observed genotypic frequencies from their expected values, violates the law of segregation that tends to render the conventional genetic theory and analysis to be invalid. Segregation distortion genes are widely distributed in plant as well as animal genome and functions by building competition among gametes for preferential fertilization or abortion of female or male gametes or zygotes. Competition may occur due to gametophyte genes expressed in the haploid gametes resulting in distorted segregation ratios.

Many genes expressed during post meiosis of the micro-spore and pollen development in angiosperms. Genetic difference among pollen may lead to gametophyte competition and selection, which result in non-random fertilization. Alternatively, hybrid sterility genes that cause the abortion of specific gamete or zygote genotypes can give rise to segregation distortion. Generalization of underlying mechanisms in segregation distortion is do not distort meiosis per se but rather alter the products of meiosis by causing chromosome breakage and or aborting gametes that do not carry the driving allele. Segregation distortion have been observed in populations of a wide variety of organisms including
animals, plants, fungi, insects, and mammals and function by their effect on competition among gametes for preferential fertilization.

Segregation distortion may have been important for the evolution of many fundamental aspects of sexual reproduction, including the evolution of sex and recombination, the evolution of heteromorphic sex chromosomes, sex ratio evolution, mate choice and reproductive isolation (Taylor and Ingvarsson, 2003).

SEGREGATION DISTORTION IN CROPS

First report of segregation distortion in crop plants on the basis of linkage between the gametophyte factor Ga1 and the Su allele for starchy endosperm in Maize, in this system pollination occurs with either Ga1 pollen only, or with ga1 pollen only, led to normal genotypic ratios. Nevertheless, pollen-tube growth is faster in pollen with Ga1 than ga1, a mixture of Ga1 and ga1 pollen led to an excess of the genotypes with the linked Su allele. In recent years, segregation distortion has been discovered in a widening variety of plant species. Rice (Harushima et al., 1996), maize (Matsushita et al., 2003), Barely Hordeum vulgare L.) in allohexaploid oat (Avena sativa), Sibov et al., 2003, (Graner et al. 1991), (Pawlowshi et al., 1998), Brassica crops (Tonguc et al., 2003; Zhang et al., 2003)

SEGREGATION DISTORTION AND LINKAGE MAPPING

Linkage map of a species indicates the position of its known genes or genetic markers relative to each other in terms of recombination frequency, rather than as specific physical distance along length of each chromosome, which is useful for identifying the location of genes responsible for phenotypic expression of traits. Greater the frequency of recombination (segregation) between two genetic markers, the farther apart they are assumed to be. Conversely, the lower the frequency of recombination between the markers, the smaller the physical distance between them. Historically, the markers originally used were detectable phenotypes (derived from coding DNA sequences; eventually, confirmed or assumed noncoding DNA sequences such as microsatellites or those generating restriction fragment length polymorphisms (RFLPs) have been used.

The construction of a linkage map requires a segregating plant population. The parents selected for generating the mapping population will differ for one or more traits of interest or contrasting or diverse or both diverse an contrasting trait of interest. Population sizes used in preliminary genetic mapping studies generally range from 50 to 250 individuals, however larger populations about 1000 plants are required for high-resolution mapping. If the map will be used for QTL studies (which is usually the case), then an important point to note is that the mapping populations must be phenotypically evaluated (i.e. trait data must be collected) before subsequent QTL mapping. Once polymorphic markers have been identified, they must be screened across the entire mapping population. This is known as marker ‘genotyping’ of the population.

The expected segregation ratios for codominant and dominant markers

<table>
<thead>
<tr>
<th>Population Type</th>
<th>Codominant markers</th>
<th>Dominant markers</th>
</tr>
</thead>
<tbody>
<tr>
<td>F2</td>
<td>1:2:1 (AA:As:aa)</td>
<td>3:1 ( (_:_b)</td>
</tr>
<tr>
<td>Backcross</td>
<td>1:1(Cc:cc)</td>
<td>1:1 (Dd:dd)</td>
</tr>
<tr>
<td>Recombinant inbred</td>
<td>1:1 (EE:ee)</td>
<td>1:1 (FF:ff)</td>
</tr>
<tr>
<td>or double haploid</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Significant deviations from expected ratios can be analyzed using chi-square tests. Generally, markers will segregate in a Mendelian fashion although distorted segregation ratios may be encountered frequently; those markers are not taken for the construction of linkage map. Segregation distortion has frequently been found during the construction of genetic linkage maps. If a gene that causes segregation distortion is segregating population, then markers close to it would tend to exhibit distorted ratios because of tight linkage between the marker and the gene that cause segregation distortion. And if several populations are segregating for the same gametophyte factors or other unknown genes that cause segregation distortion, then these populations will exhibit segregation distortion at the same chromosomal regions. Molecular-marker analysis in several populations is useful for finding common regions with segregation distortion known as segregation distortion regions (SDRs) and for future identification of yet unknown genes that cause segregation distortion in these regions.

DISTRIBUTION AND DIRECTION OF DISTORTED MARKERS

The distribution of distorted markers along linkage groups may divide into two main types: one is that deviations from Mendelian ratios at individual loci are numerous, directed towards either parental class and widespread between chromosomes. This reason is not universal other type is markers distorted at the
same direction are clustered in a small chromosomal region (Harushima et al., 1996; Matsushita et al., 2003; Sibov et al., 2003), called a segregation distortion region (SDR). This type is usually regarded to be related to the existence of segregation distortion loci (SDLs). In SDRs, selective effect of the SDL determines the direction of distorted markers. Extreme segregation deviations in parents and in the heterozygote have been widely reported in different crop species.

SEGREGATION DISTORTION AND MARKER TYPE

A genetic marker is a biological feature which can be determined by their allelic forms, used as experimental probe or tag to keep track of an individual. Described as a variation, which may arise due to mutation or alteration in the genomic loci. Marker generally refers to molecular markers (DNA markers) a short DNA sequence, such as a sequence surrounding a single base-pair change (single nucleotide polymorphism, SNP), or a long one, like minisatellites. In view of molecular marker, the scale and extent of segregation distortion varied significantly among different data. A mapping strategy combining methylation sensitive and insensitive enzymes with the AFLP technique would provide complete map coverage and allow for the targeting of specific regions based on genomic methylation patterns.

SEGREGATION DISTORTION AND POPULATION TYPE

The proportion of loci showing significant segregation varies greatly with the species, population type and specific cross. Several different populations may utilize for mapping within a given plant species, with each population type possessing advantages and disadvantages. F2 populations derived from F1 hybrids, and backcross (BC) populations, derived by backcrossing the F1 hybrid to one of the parents are the simplest types of mapping populations developed for self-pollinating species.

Their main advantages are that they are easy to construct and require only a short time to produce. Inbreeding from individual F2 plants up to F5 or F6 generations allows the construction of recombinant inbred (RI) lines, which consist of a series of homozygous lines, each containing a unique combination of chromosomal segments from the original parents. The length of time needed for producing RI populations is the major disadvantage, because usually six to eight generations are required and in cross pollinated crops difficult to develop due to severe inbreeding depression. Doubled haploid (DH) populations are regenerated by induction of chromosome doubling from pollen grains, though, the production of DH populations is only possible in species that are amenable to tissue culture e.g. cereal species such as rice, barley and wheat.

The major advantages of RILs and DHs populations are homozygous or ‘true-breeding’ lines that can multiply and reproduce without genetic change. This allows for the conduct of replicated trials across different locations and years. Thus, both RILs and DH populations represent ‘permanent sources’ for QTL mapping. Furthermore, seed from individual RILs or DHs lines may transfer between different laboratories for further linkage analysis and the addition of markers to existing maps, ensuring that all collaborators examine identical material. The disadvantage of double haploid (DH) populations and populations of recombinant inbred lines (RIL) are they usually have extreme segregation distortion.

In DH populations, many recessive lethal genes become homozygous and express, leading to a high percentage of markers showing segregation distortion that there are many recessive lethal genes, which are related to the characteristic of cross-pollination (self-incompatibility). The percentage and composition of distorted markers are quite different between a DH (44%) and test cross (TC) population from the same clone, which extensively supports the existence of strong zygotic selection, but does not exclude gametic selection. Comparison between genetic maps of backcross (BC1) and DH populations from the same Gossypium hirsutum X Gossypium barbadense interspecific cross in cotton got similar results. Backcross populations usually have relatively fewer segregation distortions.

In order to obtain larger phenotypic differences coupled with higher polymorphisms, crosses should be made between wild species with the cultivars, between subspecies, and even between species for linkage analysis, which resulted in genomic structure mutants, such as heterogeneous recombination and chromosome rearrangement, in favor of extreme segregation distortion. an good example in Brassica species is construction of genetic map with three intraspecific (Brassica oleracea) F2 populations and one interspecific (B. oleracea; B. insulari) F2 population. Results indicated that 59% of the markers of the interspecific F2 population showed significant segregation distortion, whereas only 7%, on average,
of markers found in the three intraspecific populations.

MECHANISMS/REASONS OF SEGREGATION DISTORTION

Segregation distortion may arise as a result of variety of physiological and genetic factors. Mechanisms for preferential segregation include pollen tube competition, Pollen lethal, preferential fertilization and selective elimination of zygote. In maize, most commonly reported genetic factors associated with the distorted segregation ratio are gametophytic factors (ga). It can affect either the male or female germ line, or it is a result of post zygotic selection. SD is often attributed to pollen-pistil incompatibilities, gametic competition, negative epistatic interactions or gamete abortion.

There are other chromosomal causes of nonrandom segregation preferential selection in maize due to the presence of a large heterochromatic knobs as well as another piece of chromatin of undetermined size on the long arm on chromosome 10. In plants heteromorphic of this “abnormal 10” chromosome, markers in coupling with the knob have an increased chance of being included in the functional basal megaspore; homozygotes for “abnormal 10” apparently behave normally. A gene associated with the knob, or the knob itself, seems to induce the formation of spindle fiber attached points at locations other than the centromere called Neocentromeres. These centromeres increase the chances for the knobbed chromosome to be moved into the basal megaspore. The selection advantage of certain pollen tube (in growing down style) over other of different genetic constitution may also alter the normal segregation ratio (Pollen tube competition). Pollen killer genes in wheat, tomato and tobacco and spore killer gene in Neurospora also cause distorted segregation ratios.

unequal reproduction: In Drosophila unequal reproduction of two kinds of gametes in some heterozygotes frequently contain a structural rearrangement of chromosomes. As consequences, certain meiotic products show differential survival (Meiotic drive). The frequency of affected genotypes in the population may be altered if meiotic drive operates. During genotypic analysis of a population, errors in marker genotyping and statistical analysis may cause systematic deviations from the expected segregation ratio (Sibov et al., 2003). Besides experimental techniques, residual heterozygosity in parental lines has resulted in complicated allelic assignment and has yielded subsequent difficulties in linkage determination. A mutation within the binding site for a DNA marker may also cause segregation distortion.

Clusters of segregation distortion markers usually result from the selective effect of Segregation distortion loci (SDL). Depending on the time of action of the SDL, this selective effect behaves as: (1) pollen abortion; (2) pollen tube competition (3) competitive fertilization; or (4) zygotic selection. Gametic and zygotic selection may be controlled by SDLs that work before and after fertilization, respectively. SDLs expressed before fertilization can change only the genotypic ratio of zygotes indirectly, by altering the gamete ratio; SDLs expressed after fertilization affect the genotypic ratio of zygotes directly. Both gametic and zygotic selections give rise to most of the distorted segregations. In maize 14 SDRs have been detected of which four were located in relevant regions where they are gamete genes suggesting the existence of segregation distortion loci in these regions. In rice, 13 gamete genes have been identified and mapped. It is now believed that distorted segregation at a specific marker locus is mainly due to the linkage of this marker to a gametophyte gene located nearby on the same chromosome, although other factors, such as complementary genes, duplicate genes, chromosomal abnormality and competitive fertilization between marker genotypes, have been suggested as the cause of distortion. In the case of cotton, reports indicated the existence of loci conditioning lethality, partial male and (or) female sterility, or pollen spine development on either chromosome of the homologous pair. Chromosomal regions of distorted segregation on chromosomes 2, 3 and 7, and linkage group A1 in a cotton genetic map constructed from a Gossypium hirsutum X G. barbadense interspecific BC1 population. Further studies are being performed to validate whether these regions are SDRs and the existence of gamete genes in these regions.

Viability selection after fertilization may be more important than gametic selection popularly in outcrossed crops. Inbreeding depression decreases viability of the homozygote and results in distorted segregation and often expressed in early life stages. In Scots pine, only about 15% of self-pollinated embryos develop into mature seeds, compared to about 75% of wind-pollinated seeds. Strong heterosis possessed by the heterozygote could explain partly the segregation distortion towards a heterozygote genotype in a segregation population. Besides the reasons mentioned, environmental factors,
heterogeneous recombination, gene diversification, transgenic silencing (Pawlowshi et al., 1998), selection pressure during in vitro androgenesis (Zhang et al., 2003) and chromosome instability resulting from aneuploidy and unstable translocation (Tonguc et al., 2003) could also give rise to distorted segregation.

**GENETIC MAPPING OF SEGREGATION DISTORTION LOCI**

Chromosomal regions that cause distorted segregation ratios in early life stages may be preferred as distortion loci. If an SDL segregates in a population, markers linked to this the Loci will show distorted segregation. The closer the linkage between a marker and an SDL, the more conspicuous is the expected distortion of segregation, and the larger is the chi-square value. So the location of a SDL could be primarily determined when the $\chi^2$ values are plotted against the location of the marker. If an SDL segregates in several populations, corresponding chromosomal regions of distorted segregation ratio will be identified in these populations. So this strategy can be used to determine the existence and location of an SDL. In maize detected several regions of distorted segregation ratio by comparing the distribution of distorted markers in four populations (F2Syn3, RIL, F6:7, F2), and inferred that these chromosomal regions were correlated with corresponding gamete genes. By comparative mapping, chromosomal regions linked to microspore embryogenic ability were identified in *Brassica napus* L. (Zhang et al., 2003).

**HOW TO LOCATE AN SDL ON A LINKAGE GROUP?**

If you want estimate recombination values and the differential fertilization ability of the male gametes using $F_2$ and $F_3$ segregation populations derived from a cross between a mutant and a tester line. three types (low, normal and high) populations will produced with respect to the segregation ratio of the recessive allele of a marker among F3 lines derived from F2 heterozygous plants. Then the recombination value between the lethal factor and the marker could be estimated from the relative proportion of these types.

Another more convenient method to estimate the recombination between the lethal factor locus and neighboring markers. Only F2 segregation data were needed to estimate the relative viability or fertilization ability of gametes or zygotes affected by the lethal factor in population by using the maximum likelihood method and the expectation conditional maximization (ECM) algorithm. The selection models (gametic selection, gametic and zygotic selection,) of a lethal factor locus could also be determined. Using this method lethal factor locus successfully located a on chromosome III in rice that caused partial gametic selection in both the male and female sides, and indicated that the fertilization chance of a male or female gamete possessing the lethal factors was, on average, 41.5% of that of the normal gamete.

in order to estimate locations and genetic effects of multi-SDLs in a backcross population. The multipoint method which allows efficient use of a map of partially or fully informative marker loci by using this method, it is possible to analyze efficiently the number, positions and effects of SDLs in organisms for which a high-resolution marker map has been developed and where inbred line crosses can be performed easily. With some changes, analysis can be easily extended to a general full sib family or to the sifting of an outcrossing individual.

**SEGREGATION DISTORTION AN EVOLUTIONARY FORCE**

Genetic elements that distort Mendelian segregation to enhance their own transmission (so-called ‘selfish’ genetic elements) are thought to be a potent evolutionary force (Sandler & Novitski, 1957). Theory suggests they may be important for the evolution of many fundamental aspects of sexual reproduction including the evolution of sex and recombination, evolution of heteromorphic sex chromosomes, sex ratio evolution, mate choice, and the reproductive isolation (Taylor and Ingvarsson, 2003). However the fact that Mendelian inheritance is nearly universal points to the opposite interpretation, that segregation distortion is a rare curiosity with little or no evolutionary importance. There are two explanations for importance of SD in evolution, but would appear to be rare. First, non-Mendelian elements might spread rapidly to fixation without being observed. Second, non-Mendelian elements might be deleterious, which creates selection of alleles at other loci that suppress their effects so-called genetic conflict. Taylor and Ingvarsson (2003) cite two explanations for why segregation distortion factors may be important in evolution, but go unobserved. First, meiotic drive elements without deleterious effects might rapidly become fixed. Second, genetic conflict could arise due to meiotic drive elements having deleterious effects, which would then cause inadvertent selection for alleles at other loci that suppress their effects.
IMPACT ON GENETIC MAP AND IMPROVEMENT

Segregation distortion is one of the factors influencing the precision of genetic mapping, reported in many crops. It has been shown that the genetic distance between markers and the order of markers on linkage groups may be affected by segregation distortion in coffee inferred that segregation distortion selection was in favor of recombinant genotypes and/or against parental genotypes, resulting in a high estimation of recombination fraction values. In *Brassica napus* L. indicated that small but significant segregation distortions resulted in reduced estimation of the recombination fraction, and extreme segregation distortions towards the same parental allele contributed to an additional source of spurious linkage. Data from diploid alfalfa (*Medicago sativa*) showed that segregation distortion extremely in favour of heterozygous individuals could artificially link genetic regions that turned out to be unlinked. Quantitative trait loci (QTL) mapping is an important application of genetic mapping. For a correctly inferred (map distance and marker order) map, influence of segregation distortion on QTL analysis could be negligible. But if the recombination fractions or, worse, the order of marker loci are inferred incorrectly, basic assumptions of QTL analysis do not hold and results will be imprecise at best. So, much attention should be paid to selecting.

PROPER METHODS FOR GENETIC MAPPING ANALYSIS

As the simplest remedy, markers that show obvious segregation distortion are often excluded from the map, but this treatment usually reduces coverage of the genome and qualitative or quantitative trait loci might be missed. An eclectic way is adding segregation distorted markers to a map constructed with normally segregated markers and determine the fate of the distorted markers, based on how much the map was affected. But there is no a proper standard for deleting the distorted markers from a genetic map. Effects of segregation distortion on linkage were estimated with independent \( \chi^2 \) test and LOD values, and the use of a more stringent linkage test was recommended for markers showing extreme segregation distortion, such as higher LOD values or lower maximum linkage distance. Comparative mapping is a useful tool to identify spurious linkage. Loose were identified in maize. This kind of comparative mapping is usually performed among different types of populations from the same cross or among different populations of the same type. Using a selective population is another way to reduce the influence of segregation distortion on a genetic map.

Based on the soybean genetic map constructed with SSR, AFLP and RFLP markers. evaluated the RIL population of soybean ‘NJRIKY’ with its segregation data of RFLP markers, and adjusted this population from 201 to 184 family lines by deleting distorted family lines, and constructed the genetic map of soybean more properly using the adjusted RIL population in 2003. several workers developed the maximum likelihood models for mapping genetic markers showing segregation distortion under gametic selection and zygotic selection in backcross and F2 populations, respectively. They also showed that the estimation of recombination fractions between co-dominant markers is less affected by selection than that of dominant markers. If possible, use only co-dominant markers to construct genetic maps in species. Based on these models, developed computer software, Mapdisto, for genetic map construction with segregation data including distorted markers. After several improvements, this software can now solve map construction and map drawing and deal with F2, BC1, DH, and single seed descent (SSD) populations. Using the maximum-likelihood formulas an independent, non-mathematical method called color mapping excluded the misleading linkages, caused by extreme segregation distortion, on three linkage groups in diploid alfalfa (*Medicago sativa*) and estimated the genetic distances more precisely.

CONCLUSION

Compared to animals, research on segregation distortion in plants still remains at its initial stage. Several questions remain unclear, such as: (1) Although the phenomenon of segregation distortion has been reported in many crops, the number and genetic effects of SDL are not clear; (2) only a few SDLs have been located on linkage groups and results are lacking at DNA sequence level; (3) only a few types of mapping populations can be analyzed with the available methods for mapping SDLs, and the effects of interactions between SDLs located on different chromosomes cannot be estimated yet.

Further studies on segregation distortion in plants will enhance our understanding of hybrid lethality and deviant gene segregation during introgression of target genes from wild species to cultivars; facilitate the location and identification of new SDLs, the correct evaluation and utilization of selective effects
Detection of Segregation distortion through molecular markers

of SDLs, and production of precise genetic maps and QTL results. All these are important for genetics and improvement of plants. The identification, location and genetic effect evaluation are all based on dense molecular genetic maps. By now, genetic maps have been constructed for almost all important plant species, but most of the research is at framework stage. With the construction of high-resolution genetic maps in more species, rapid progress in segregation distortion will be achieved.

Knowledge of the chromosomal locations and activation of segregation distortion loci will aid breeders in designing appropriate crossing schemes and help them predict the frequency at which a given allele will be transmitted to the progeny. Once the performance of a genotype can be predicted, a breeder may be able to maximize transmission of desired allele or to preferential exclude deleterious alleles.

REFERENCE


Masumi Yamagishi, Yoshinobu Takeuchi, Isao Tanaka and Masahiro Yano, (2003), Segregation F2 and doubled haploid population of temperate japonica rice. 71: 37-41.


Taylor DR and Ingvarsson PK. (2003), Common features of segregation distortion in plants and animals. Genetica 117: 27–35.


