Study of Molecular Diversity in High Zinc Rice (*Oryza sativa* L.) Genotypes

Yashraaj Sharma, S. K. Singh, Prudhvi Raj Vennela, Riddha Dey, D. K. Singh and Sonali Habde[*]

Department of Genetics and Plant Breeding, Institute of Agricultural Sciences,
Banaras Hindu University, Vanarasi, India
*Corresponding Author E-mail: sonali.habde@gmail.com

ABSTRACT

Assessment of genetic diversity is a key factor in germplasm characterization and conservation. An investigation was carried out with 20 rice genotypes constituting checks collected from different sources using 13 microsatellite markers to study the nature and magnitude of genetic divergence. The results showed that all the SSRs were polymorphic with different alleles among the genotypes studied showing the robust nature of microsatellites in revealing polymorphism. The average PIC value was 0.349 and it ranged from a low of 0.090 (RM OSR) to a high of 0.572 (RM 495). The Jaccard's similarity coefficient ranged from 0.231 to 1. The UPGMA dendrogram revealed two major groups with five clusters and the wide range of dissimilarity values (0.01-0.88) which represents a high degree of diversity among the genotypes. The results of the genetic diversity can be exploited for the selection of the parents for developing high zinc rice variety through molecular breeding program.

Key words: Microsatellites, Polymorphism, PIC value, UPGMA dendrogram.

INTRODUCTION

Rice (*Oryza Sativa* L.) is major staple food grain crops as it feeds more than half of world's population. According to the 3rd Advance Estimates of production of rice for 2015-16 released by the Department of Agriculture, Cooperation and Farmers Welfare, the estimated production of rice during 2015-16 is 103.36 million tones. Zinc deficiency is a well-documented problem in food crops, leading to decreased crop yields and nutritional quality. The International Zinc Nutrition Consultative Group estimates that 26% of the Indian population is at risk of inadequate zinc intakes. Zinc deficiency may cause large reductions in crop quality and yield.

Molecular Marker based Genetic Diversity Analysis (MMGDA) also has potential for assessing changes in genetic diversity over time and space[4]. Molecular markers have proven to be powerful tools in the assessment of genetic variation and in the elucidation of genetic relationships within and among species.

Information regarding genetic variability at molecular level can be exploited to identify and develop genetically unique germplasm that compliments existing cultivars. The use of molecular markers has been suggested for precise and reliable characterization and discrimination of rice genotypes[7]. For genetic variability assessment, molecular markers are extensively used because they are not affected by environmental factors.

Microsatellites (SSRs) are the marker of choice because of their advantages over other markers. The SSRs are most suitable for rice because of their reproducibility, multi-allelic nature, hypervariablility, co dominant inheritance, relative abundance, and genome-wide coverage[10]. Due to their co-dominant inheritance and amenability to high throughput, SSRs have become a tool of choice for investigations of critical importance to crop germplasm managers, such as the establishment of unique genetic identities or fingerprints, determination of genetic relatedness between accessions, and the assessment of genetic diversity contained within a collection. The high levels of variability and reproducibility associated with SSR markers will permit them to serve as anchor markers between different genetic maps within a specific crop[2].

The present investigation was carried out to identify the suitable SSR primers for genetic analysis of high zinc rice and to measure the genetic diversity and relatedness among twenty high zinc rice genotypes using SSR markers.

MATERIAL AND METHODS

The experimental material for this investigation comprised of twenty genotypes were raised in Randomized complete Block Design with three replications during kharif season 2016 at Agricultural Research Farm, BHU, Varanasi, (U.P.), India. The molecular analysis was carried out at the Molecular Biology Laboratory (Niche Area Lab) of the Department of Genetics and Plant Breeding, Institute of Agricultural Sciences, BHU, Varanasi (U.P.), India. A total of thirteen simple sequence repeat (SSR) markers were used for studying molecular diversity. The details of SSR primers used are presented in Table 1.

Table 1: Details of the microsatellite primers used in present studies

Microsatellite locus	Forward/ Reverse	Sequence 5'--------------> 3'	Annealing Temperature (°C)
RM495	Forward	AATCCAAGGTGCAGAGATGG	55
	Reverse	CAACGATGACGAACACAACC	
RM1	Forward	GCGAAAACACAATGCAAAAA	55
	Reverse	GCGTTGGTTGGACCTGAC	
RM283	Forward	GTCTACATGTACCCTTGTTGGG	61
	Reverse	CGGCATGAGAGTCTGTGATG	
RM259	Forward	TGGAGTTTGAGAGGAGGG	55
	Reverse	CTTGTTGCATGGTGCCATGT	
RM312	Forward	GTATGCATATTTGATAAGAG	55
	Reverse	AAGTCACCGAGTTTACCTTC	
RM5	Forward	TGCAATTCTAGCTGCTCGA	57
	Reverse	GCATCCGATCTTGATGGG	
RM237	Forward	CAAATCCCGACTGCTGTCC	55
	Reverse	TGGGAAGAGAGCACTACAGC	
RM431	Forward	TCCTGCGAACTGAAGAGTTG	55

	Reverse	AGAGCAAAACCCTGGTTCAC	
RM154	Forward	ACCCTCTCCGCCTCGCCTCCTC	61
	Reverse	CTCCTCCTCCTGCGACCGCTCC	
RM452	Forward	CTGATCGAGAGCGTTAAGGG	61
	Reverse	GGGATCAAACCACGTTTCTG	
RM489	Forward	ACTTGAGACGATCGGACACC	55
	Reverse	TCACCCATGGATGTTGTCAG	
OSR13	Forward	CATTTGTGCGTCACGGAGTA	53
	Reverse	AGCCACAGCGCCCATCTCTC	
RM338	Forward	CACAGGAGCAGGAGAAGAGC	55
	Reverse	GGCAAACCGATCACTCAGTC	
RM55	Forward	CCGTCGCCGTAGTAGAGAAG	55
	Reverse	TCCCGGTTATTTTAAGGCG	
RM514	Forward	AGATTGATCTCCCATTCCCC	55
	Reverse	CACGAGCATATTACTAGTGG	

Isolation of DNA

Young leaves were collected from 12 days old rice seedlings and immediately stored at -20 ^{0}C till further processing. The DNA was extracted following CTAB extraction method with few modifications in composition of DNA extraction buffer[3]. 1 gm fresh leaves were taken and ground in pre-heat 500 µl of CTAB (cetyl tri-methyl ammonium bromide) buffer in mortar pestle. Transfer the homogenate to 65°C water bath for 30 minutes incubation after mixing and thoroughly vortex. Equal volume of Phenol: chloroform: Isoamyl alcohol mixture (25:24:1 v/v) was added and mixed. It was centrifuged and supernatant was separated to new labeled tube. 800 µl of Chloroform: Isoamylalcohol mixture was added and centrifuged. Supernatant was taken in 1.5 ml autoclaved centrifuge tube. 2/3 volume of chilled isopropanol was added and kept in the deep freezer until DNA was precipitated out. It was centrifuged, supernatant was discarded and DNA pellet settled in the bottom of the tube. DNA pellet was washed by using 200 µl 70% of ethanol and kept it overnight at room temperature for drying. Finally DNA pellet was dissolved in 50 µl of TE buffer and stored at -20 °C. The DNA quality estimation was done using Biophotometer plus.

PCR amplification

The PCR assays were conducted with rice microsatellite markers to selectively amplify in vitro a specific segment of the total genomic DNA to a billion fold. Polymerase chain reaction (PCR) was performed in 15 µl reaction volume with 1X buffer, MgCl2, dNTP, 1.25 µM of primer pairs, HPLC water, Taq DNA polymerase and DNA template (Table 2). The most essential requirement of PCR is an presence of a pair of short (about 20-25 nucleotides) primers having sequence complementary to either end of the target DNA segment known as template DNA supposed be synthesized in large amount.

All the steps were kept as such except the annealing temperature for PCR programming. Annealing temperature was determined based on the GC content of the primer as:

$$Tm = [2 \times (A+T) + 4 \times (G+C)] - 4$$

The above formula provides the preliminary information but not the exact annealing temperature. Therefore, the correct annealing temperature was determined based on best PCR amplification. All the amplifications were performed in the Eppendorf Thermo-cycler (USA) using the thermal cycler program. After the completion of the PCR, the products were stored at -20ºC until the gel electrophoresis was done.

Table 2: Lists of the various components of PCR master mixture (15 μl)

S.No.	Reagents	Stock Concentration	Required concentration	Quantity(μl)
1.	Assay buffer	10x	1x	1.50
2.	Mgcl$_2$	25Mm	.3mM	0.20
3.	DNTPs	10Mm	133μM	0.20
4.	Taq polymerase	5 U	1 U	0.20
5.	Primer(Forward)	10μM	1.25μM	1.0
6.	Primer(Reverse)	10μM	1.25μM	1.0
7.	Genomic DNA	15 ng/μl	15 ng/μl	1.0
8.	HPLC water	-	-	9.90
	Total			15.0

SSR marker Analysis

The amplified DNA products generated through SSR primers were resolved through electrophoresis in 2.5 per cent agarose gel prepared in TAE buffer. Ethidium bromide solution at final concentration of 0.03 ng/μl was added to the agarose solution. For electrophoresis, 15μl of the PCR product was mixed with 2μl of 6X loading dye (bromophenol blue) and loaded in the slot of agarose gel. In order to determine the molecular size of amplified products, the gels were loaded with 1μg of a 50 bp DNA marker. Gel electrophoresis was carried out at a constant voltage of 65 V for about 3.5 hours and the gels were visualized under a UV light source in a gel documentation system (Gel Doc^{TM}XR$^+$, BIO-RAD, USA). Finally, the images of amplification products were captured for evaluation of diversity among the genotypes.

Band position in comparative SSR profile for each genotype and primer combination was scored from the respective gel images. The amplified fragments were scored as '1' for the presence and '0' for the absence of a band generating the 0 and 1 matrix. These binary data matrix was then utilized to generate genetic similarity data among the 20 lines of rice genotypes.

Genetic dissimilarity

Polymorphic SSR markers generated the binary data matrixes which were subjected to further analysis using NTSYS-pc version 2.11W. To calculate the jackard's coefficient SIMQUAL programe was used.

Cluster analysis based on UPGMA

UPGMA-based clustering was done using SAHN module of NTSYSpc for dendogram construction. In Unweightedpair-group average (UPGMA) clusters are joined based on the average distance between all members in the two groups.

Polymorphic information content (PIC)

Polymorphic information content (PIC) values were calculated for each SSR marker by using following formula[1]: PIC $=1 – \Sigma Pij2$, where pij is the frequency of jth allele at the locus I and summation extends over n alleles.

RESULTS AND DISCUSSION

Molecular diversity study has been carried on by using chemically designed molecular markers. Molecular markers are the complimentary sequences of DNA which lie close to the particular gene or QTL. So by annealing the primer we can amplify our target gene. In the present investigation rice microsatellites or SSR markers were used to assess genetic diversity among 20 rice genotypes from various sources. A total of 15 primers were utilized to assess genetic diversity among 20 rice genotypes. Thirteen primers showed polymorphism in these genotypes and two primers (RM 312 and RM 55) were not amplified and cluster analysis was done to construct dendogram using Jaccard's similarity coefficient.

Diversity analysis

A dendrogram (figure ii) based on Jackard's dissimilarity coefficient was prepared using UPGMA. Twenty rice genotypes were

grouped into two main clusters i.e. cluster I and cluster II with dissimilarity coefficient (0.01). Cluster I was further sub-divided into two minor sub-groups IA and IB with dissimilarity coefficient (0.28). Cluster IA was further sub-divided into two subgroups i.e. IA-1 and IA-2 (0.34). Cluster IA-1 contains four genotypes and cluster IA-2 contains five genotypes. Cluster IB contains three genotypes. Cluster II was also further sub-divided into two minor sub-groups i.e. II-A and II-B (0.43) with 4 genotypes in each. This indicated presence of considerable diversity in the genotypes studied. Therefore, it is essential in order to select desirable genotypes from the most diverse genotypes for utilization in breeding programmes.

On the basis of dendrogram (figure ii) most diverse cultivar was BRRI Dhan-64 and R-RHZ-7 followed by BRRI Dhan-64 and Swarna. The highest similarity was observed between cultivar NDR-97 and Dudhkandar followed by NDR-97and CGZR-1. Similar result was found by Matin *et al.*2012, by using 18 SSR markers among 12 rice genotypes with 4 clusters and the wide range of dissimilarity values (0.14-0.89).

Dissimilarity coefficient

The dissimilarity coefficient was used to determine the level of relatedness among the genotypes studied. The dissimilarity coefficient varies from one to zero, close to zero shows high similarity while close to one shows high dissimilarity. The average of dissimilarity coefficient varies from 0.61 to 0.83. The total average of dissimilarity coefficient of all 20 genotypes is 0.72. The dissimilarity coefficient varied from the largest value 1.0 between the cultivar BRRI Dhan-64 and R-RHZ-7 followed by value 0.95 between the cultivar BRRI Dhan-64 & Swarna, and Gora (white) & Sambha Mahsuri which shows high dissimilarity between them showing that they are highly dissimilar from each other. The lowest value 0.23 was found between NDR-97 & Dudhkandar followed by 0.25 and 0.26 between the cultivar NDR-97 and CGZR-1, and between Swarna & Kalamkati IRGC 45975-1 respectively. Cultivar NDR-97 shows highest similarity with cultivar Dudhkandar (0.23). Similar results were found for aromatic landraces of rice[7].
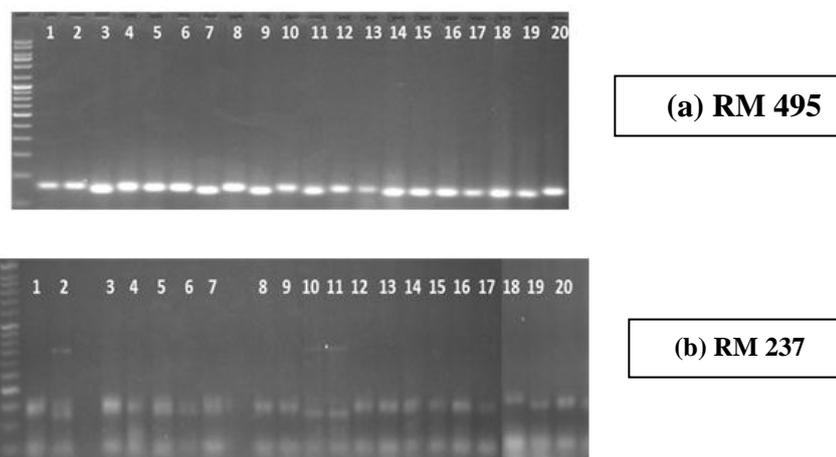
Fig: i) Some SSR Markers (a) RM 495, (b) RM 237 obtained in 20 rice genotypes. Lane 1-20 represents rice genotypes

1=CGZR-1, 2=NDR-97, 3=IR-96248-16-3-3-1-B, 4=Dudhkandar, 5=R-RHZ-7, 6=IR-92978-192-1-2(R-306), 7=IR-82475-110-2-2-1-2, 8=Swarna, 9=Kalamkati IRGC 45975-1, 10=IR-96248-16-3-3-2-B, 11=IR-92960-75-1-3, 12=Gora(white), 13=Sathi, 14=Black Gora, 15=BRRI Dhan-72, 16=BRRI Dhan-62, 17=MTU 1010, 18=SambhaMahsuri, 19=BRRI Dhan-64, 20=Bansphul.
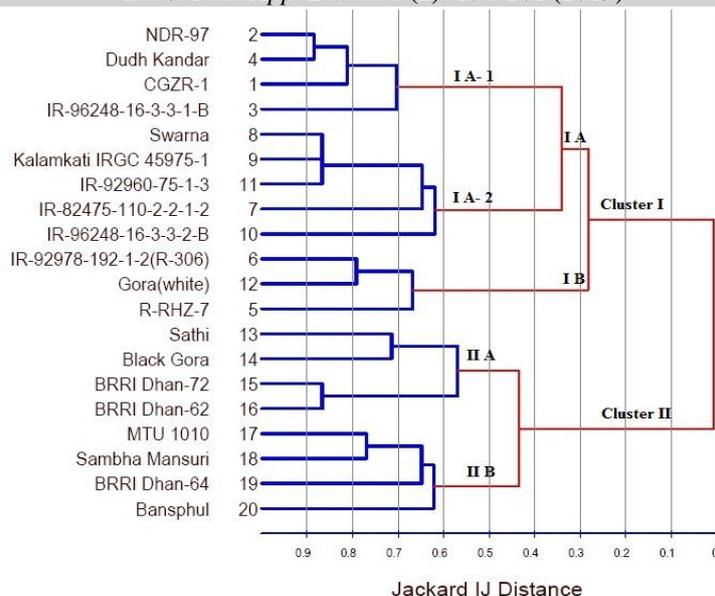
Fig: ii) Dendrogram based on Jackard dissimilarity coefficient

Table 3: Jackard IJ Distance

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Average D²
1		0.250	0.50	0.429	0.824	0.733	0.833	0.647	0.722	0.667	0.789	0.824	0.750	0.667	0.789	0.722	0.80	0.778	0.813	0.813	0.703
2			0.429	**0.231**	0.765	0.571	0.706	0.50	0.667	0.688	0.667	0.765	0.833	0.765	0.737	0.737	0.813	0.789	0.824	0.824	0.661
3				0.563	0.765	0.824	0.625	0.737	0.588	0.688	0.667	0.895	0.950	0.833	0.857	0.857	0.882	0.647	0.750	0.750	0.727
4					0.706	0.60	0.647	0.333	0.529	0.706	0.529	0.706	0.778	0.706	0.684	0.684	0.824	0.857	0.895	0.895	0.647
5						0.538	0.765	0.722	0.789	0.889	0.789	0.667	0.750	0.824	0.905	0.850	0.941	0.90	**1.0**	0.944	0.807
6							0.750	0.533	0.706	0.813	0.706	0.417	0.643	0.733	0.778	0.706	0.786	0.895	0.941	0.80	0.709
7								0.588	0.588	0.833	0.50	0.765	0.833	0.895	0.857	0.80	0.944	0.905	0.947	0.824	0.769
8									0.267	0.563	0.267	0.647	0.722	0.722	0.556	0.471	0.765	0.864	0.952	0.842	0.616
9										0.563	0.267	0.722	0.789	0.722	0.632	0.556	0.765	0.810	0.90	0.706	0.647
10											0.563	0.889	0.889	0.824	0.647	0.722	0.714	0.706	0.813	0.813	0.736
11												0.789	0.850	0.789	0.632	0.556	0.833	0.864	0.90	0.842	0.674
12													0.571	0.824	0.905	0.850	0.875	0.952	0.882	0.813	0.777
13														0.571	0.789	0.647	0.875	0.90	0.882	0.813	0.781
14															0.563	0.563	0.615	0.706	0.733	0.813	0.730
15																0.267	0.385	0.529	0.778	0.778	0.688
16																	0.60	0.684	0.842	0.842	0.682
17																		0.462	0.786	0.692	0.756
18																			0.50	0.688	0.760
19																				0.714	0.834
20																					0.800

1=CGZR-1, 2=NDR-97, 3=IR-96248-16-3-3-1-B, 4=Dudhkandar, 5=R-RHZ-7, 6=IR-92978-192-1-2(R-306), 7=IR-82475-110-2-2-1-2, 8=Swarna, 9=Kalamkati IRGC 45975-1, 10=IR-96248-16-3-3-2-B, 11=IR-92960-75-1-3, 12=Gora(white), 13=Sathi, 14=Black Gora, 15=BRRI Dhan-72, 16=BRRI Dhan-62, 17=MTU 1010, 18=SambhaMahsuri, 19=BRRI Dhan-64, 20=Bansphul

PIC Value

The polymorphic information content (PIC) was employed for each locus to evaluate the information of each marker and its discriminatory ability. The PIC for the *i*th marker is calculated by using formula[11].

PIC = 1 – Σ*Pij*

Where *j*= 1, 2...n and *Pij* is the frequency of the *j*th pattern for the *i*th marker and the summation extends over (n) patterns[8].

The PIC value is an indication of diversity and frequency among the varieties[9]. The PIC value of each marker can also be evaluated on the basis of its alleles and diverse for all SSR loci. Since, a marker with fewer alleles has less power to distinguish several samples and alleles present at low frequency have less power to distinguish, a higher PIC is assigned to a marker with many alleles and with alleles

present roughly equal proportions in the population[5].

The largest PIC value was observed for locus RM 495 (0.572) followed by RM 259 and RM 55 (0.527) and lowest by RM OSR (0.090) followed by RM 1 (0.346). So PIC value ranged from 0.572 to 0.090 with a mean value of 0.349. Thirteen primers showed polymorphism and the number of alleles ranged from 2 to 3 with an average of 2.13 (table 4). A total of 32 alleles were amplified from 20 genotypes and this demonstrates considerable variability among genotype.

Table 4: Allele and Polymorphism information content (PIC) of SSR primers used in the present study

Primer	Allele	PIC
RM 495	3	0.572
RM 1	2	0.346
RM 283	3	0.368
RM 259	3	0.527
RM 5	3	0.527
RM 237	3	0.460
RM 431	2	0.365
RM 154	2	0.375
RM 452	3	0.500
RM 489	2	0.365
OSR 13	2	0.090
RM 338	2	0.374
RM 514	2	0.370

CONCLUSION

Twenty rice genotypes were grouped into two main clusters i.e. cluster I and cluster II with dissimilarity coefficient of 0.01. The main cluster i.e., Cluster I was further sub-divided into two minor sub-groups IA and IB with dissimilarity coefficient of 0.28. Cluster IA was further sub-divided into two subgroups i.e. IA-1 and IA-2 (0.34). The second main cluster was also further sub-divided into two minor sub-groups i.e. II-1 and II-2 (0.43).

This indicated presence of considerable diversity in the genotypes studied. On the basis of dendogram the highest similarity was observed between NDR-97 and Dudhkandar followed by NDR-97 and CGZR-1. The most diverse genotypes were BRRI Dhan-64 and R-RHZ-7. The largest PIC value was observed for locus RM 495 (0.572) followed by RM 259 and RM 55 (0.527) and lowest by RM OSR (0.090) followed by RM 1 (0.346).

Genetic diversity among 20 rice genotypes was determined using 13 SSR markers. The results reveal that genotypes have a great genetic diversity. SSR markers provided very useful data to exploit genotypes for future research on rice breeding. BRRI Dhan-64 and R-RHZ-7 show higher level of dissimilarity and hence can be used as a donor lines for the particular traits like high zinc.

REFERENCES

1. Anderson, J. A., Churchill, G. A., Autrique, J. E., Tanksley, S. D. And Sorrells, M. E., Optimizing parental selection for genetic linkage maps. *Genome*, **36(1):** 181-186 (1993).
2. Beckmann, J. S., & Soller, M., Toward a unified approach to genetic mapping of eukaryotes based on sequence

tagged microsatellite sites. *Nature Biotechnology*. **8(10):** 930-932 (1990).

3. Doyle, J.J., and Doyle, J.L., A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bulletin*. **19:** 11-15 (1987).

4. Duwick, D. N., Genetic diversity in major farm crops on the farm and reserve. *Econ Bot*. **32:** 161–178 (1984).

5. Jiang, S. K., Cheng, H., Wang, Y. J., Chen, W. F., Zeng, J. X., Development of a Highly Informative Microsatelllites (SSR) Marker Framework for Rice (*Oryza sativa* L.) Genotyping. *Agric. Sci. China*. **9(12):** 1697-1704 (2010).

6. Karkousis, A., Barr, A. R., Chalmers, K. J., Ablett, G. A., Holton, T. A., Henry, R. J., Lim, P. and Langridge, P., Potential of SSR markers for plant breeding and variety identification in Australian Barley germplasm. *Australian J. of Agriculture Research*. **54:** 1197-1210 (2003).

7. Patel, N. B., Agrawal, N., and Shrivastava, R. Molecular Characterization of Aromatic Landraces of Rice (*Oryza sativa* L.) Using Microsatellite Markers. *Journal of Rice Research*. **8:** 2 (2015).

8. Peng, J.H., Lapitan, N.L.V., Characterization of ESTderived microsatellites in the wheat genome development of eSSR markers. *Funct Inter Genom* **5:** 80-96 (2005).

9. Pervaiz, Z.H., Rabbani, M.A., Pearce, S.R., and Malik, S.A., Determination of genetic variability of Asian rice (*Oryza sativa* L.) varieties using microsatellite markers. *African Journal of Biotechnology*. **8(21):** 5641-5651 (2009).

10. Powell, W., Machray, G. C., and Provan, J., Polymorphism revealed by simple sequence repeats. *Trends Plant Sci*. **1:** 215-222 (1996).

11. Weir, B. S., Genetic Data Analysis II: Materials for Discrete Population Genetic Data. *Sinauer Associates, Sunderland, MA*. (1996).