ISSN 2959-1864 (Online)
ISSN 2958-0536 (Print)
Volume 2, Number 1
December 2023

Acta Botanica Caucasica

CONTENT

EXPLORING GENETIC DIVERSITY IN AZERBAIJANI BARLEY COLLECTION THROUGH AMPLICON SEQUENCING

Mehraj Abbasov[1],*, Sevda Babayeva[2], Khanbala Rustamov[3], Laman Rasulova[4], Ofelya Rahimova[5], Afat Mammadova[6], Vusala Izzatullayeva[7], Elchin Hajiyev[8], Ramiz Aliyev[9], Zeynal Akparov[10]

[1-10] Ministry of Science and Education of Republic of Azerbaijan Genetic Resources Institute, Azadliq Ave 155, AZ1106, Baku, Azerbaijan;
*Western Caspian University, Department of Natural Sciences, Baku, Azerbaijan;

[1] Correspondence: mehraj_genetic@yahoo.com

DOI: 10.30546/abc.2023.2.1.10.43.
Article info: pp. 38-44
Received 30.05.2023; Received in revised form 23.10.2023; Accepted 28.11.2023

Abstract. *In this reasearch, genotyping was conducted on 86 accessions of wild barley (Hordeum vulgare L. subsp. spontaneum C.Koch) and 85 accessions of cultivated barley (Hordeum vulgare L. subsp. vulgare) collected from 20 different regions in Azerbaijan. Amplicon sequencing technology was employed, revealing 255 specific single-nucleotide polymorphism (SNP) markers of a polymorphic nature. Out of the total options, 66% were transitions, and 34% were transversions. For the studied barley collection, the Genetic Diversity Index (GDI) and Polymorphism Information Content (PIC) were calculated as 0.347 and 0.280, respectively. Notably, the 6H chromosome exhibited the highest number of SNP markers, whereas the 7H chromosome displayed the fewest. Utilizing cluster analysis, the accessions were categorized into three main groups. Remarkably, the study successfully differentiated between genotypes of wild and cultivated barley. Employing the STRUCTURE program, the number of subpopulations (K) ranged from 2 to 5. The model with K = 4 was found to be the most consistent with the collection's structure. The SNP markers acquired through multiplex amplicon sequencing demonstrate robust utility for diverse assessments and analyses within various collections. They prove suitable for exploring genetic variability and relationships, both in barley overall and in H. vulgare and H. spontaneum individually.*

Keywords: *barley, amplicon sequencing, SNP markers, genetic diversity*

INTRODUCTION

Cultivated barley (Hordeum vulgare L. subsp. vulgare) and its wild counterpart (Hordeum vulgare L. subsp. spontaneum (C.Koch)) are both diploid species (2n = 2x = 14) belonging to the genus Hordeum.

History of that cultivated barley, originating around 10,000 years ago, says that it evolved from H. spontaneum, a species still found in the Middle East. H. spontaneum, characterized by small spikelets, predominantly thrives in rocky and gravelly soils, showcasing its adaptability to these harsh environments. Interestingly, it's worth noting that H. spontaneum, a 6-row, 3-spike mutant, does not exhibit the same level of adaptation to these challenging soils. The earliest records of six-row barley, dating back 8000 years, reveal its emergence as a variety or weed in agricultural systems [Komatsuda, 2007; Pourkheirandish, 2007].

Globally, cultivated barley holds the position of the fourth most significant cereal crop and has been embraced as a model plant for genetic and genomic studies within the Triticeae tribe. The barley genome, comprising a substantial 5.3 Gb of genetic material distributed across 7 chromosomes, stands out as one of the largest diploid genomes ever sequenced. Extensive research on the barley genome has been undertaken by the International Barley Genome Sequencing Consortium, leading to the sequencing and assembly of approximately 4.79 GB (~95%) of the total 5.3 GB genome. Within this genome, researchers have identified 39,843 genes, accounting for a mere 1.4% of the entire genetic content. Notably, a significant portion of the genome, approximately 80%, is composed of DNA repeats and transposon elements, underscoring the complexity of barley's genetic makeup [Keller, 2017].

The reduction in funding for sequencing has spurred the development of innovative approaches to harness Next-Generation Sequencing (NGS) technologies in genotyping studies. NGS-based genotyping, a cost-effective and swift method for exploring intricate genomes or populations with limited resources, incorporates diverse techniques like Genotyping by Sequencing (GBS) and targeted amplicon sequencing. Amplicon sequencing, which involves sequencing PCR products (amplicons), enables the examination of genetic variations in specific genome regions [http://schnablelab.plantgenomics.iastate.edu/resources/protocols/.]. This technique relies on PCR amplification using specially designed oligonucleotide primers tailored to target and span the desired genomic regions, followed by sequencing via NGS. Amplicon sequencing proves highly efficient in identifying rare somatic mutations in complex samples.

In the studies on the genetic diversity of barley plants using the high-throughput genotyping method Illumina Golden Gate, researchers identified 1572 single nucleotide polymorphisms (SNPs) [Close, 2009]. Notably, in 2012, Komadran and colleagues developed a SNP panel comprising 9,000 SNPs, followed by Bayer and colleagues' creation of a 50,000 SNP panel for the barley genome in 2017, establishing this panel as the most advanced SNP set employed for studying barley plant diversity [Bayer, 2017].

Azerbaijan is distinguished as an area of barley biodiversity, covering both cultivated and wild varieties. Barley has been discovered in Neolithic sites in the Caucasus region, dating back to the 6th millennium BC, exemplified by locations like Goytepe, Azerbaijan [Kadowaki et al., 2015]. Now in Azerbaijan are cultivated by many sorts of barley, including Pallidium-596, Karabakh-7, Nakhichevan-310, and Cyclone. Exploring the genomic diversity of Azerbaijani barley collections, and utilizing NGS and marker technologies to select forms with diverse alleles, holds significant importance.

The objective of this study was to genotype cultivated and wild barley accessions employing the amplicon sequencing method.

MATERİAL AND METHODS

The study involved 86 accessions of wild barley (Hordeum spontaneum K. Koch) and 85 accessions of cultivated barley (Hordeum vulgare L.) sourced from the Genetic Resources Institute under the Science and Education Ministry of Azerbaijan.

DNA was extracted from seeds that were germinated in the laboratory [Stein, 2001]. The quantity and quality of the extracted DNA were assessed using Nanodrop (Thermo Scientific, 2000).

For the analysis, multiplex amplicon sequencing was conducted following the protocol provided by the Schnable laboratory. This involved a multiplex PCR reaction using SNP-specific primers designed for genomes A, B, and D. The obtained amplicons for each sample were barcoded and pooled. Clonal amplification was achieved by combining the developed library with ion particles using the Ion PGM HI-Q View Template Kit. Sequencing was performed on the Ion PGM platform (Life Technologies) using an Ion 318 chip and a sequencing kit (Ion PGM

Hi-Q Sequencing Kit). SNP calling was done using the UNEAK (Universal Network Enabled Analysis Kit) GBS pipeline, which is part of the TASSEL 3.0 bioinformatics analysis software package [Lu, 2013].

To assess genetic relationships, the Nei genetic distance index was determined, and cluster analysis and the creation of the Neighbor-joining (NJ) dendrogram were conducted using DARwin 6.0 [Perrier and Jacquemoud-Collet, 2006]. Additionally, the genetic structure within the barley collection was examined using the STRUCTURE 2.3.4 software package [Earl, 2012].

RESULTS AND DİSCUSSİON

According to the results of these studies, a collection of 86 wild and 85 cultivated barley accessions, gathered from 20 different regions in Azerbaijan, underwent genotyping through multiplex amplicon sequencing on the Ion PGM platform. This utilized a PCR genotyping panel comprising 365 primers dispersed across the genome. Through amplicon sequencing, 255 SNP markers of polymorphic nature were identified. Interestingly, there was negligible variation in the number of SNP markers along the H chromosome, with the highest count, 45, recorded on chromosome 6H, and the lowest, 29, on chromosome 7H.

These SNP markers in the barley genome, except for A/T substitutions, encompass various substitution types. Transitions accounted for 66% of the total variants, while transversions constituted 34%. Within transitions, the most prevalent replacements were G/A, while in transversions, C/A and G/C were frequent. The Ts/Tv ratio for the 255 single nucleotide markers stood at 1.9 (Table 1).

Table 1.

TYPE AND NUMBER OF SNPS IN BARLEY GENOME

Types	Number	Types	Number
Transition	168	**Transversion**	87
G/A	93	C/A	24
C/T	6	G/C	24
T/C	66	T/G	19
A/G	3	T/A	11
		C/G	4
		G/T	3
		A/C	2
Ts%	65.9	Tv%	34.1
Ts/Tv	1.9		
Total	255		

The observed heterozygosity in the 169 barley accessions ranged from 0 to 0.494, with an average of 0.036 for the entire collection (refer to Table 2). Notably, the main parameters, including the coefficient of genetic diversity (GDI) and the polymorphism information content (PIC), showed variability within the ranges of 0.104-0.500 and 0.099-0.375, respectively. Remarkably, for over half (53%) of the SNP markers, the GDI value exceeded 0.35. Overall, the GDI and PIC for the studied barley collection were calculated as 0.347 and 0.280, respectively, indicating a high level of genetic diversity for biallelic SNP markers.

Table 2.

GENETİC DİVERSİTY PARAMETERS IN BARLEY COLLECTİON

Barley	Number of accessions	Observed heterozygosity (Ho)	Genetic diversity index (GDI)	Polymorphism information content (PIC)
Hordeum L.	169	0.036	0.347	0.280
H. vulgare	85	0.035	0.343	0.272
H. spontaneum	84	0.038	0.247	0.200
H. spontaneum	84	0.038	0.247	0.200

The analysis of genetic diversity across the 7 chromosomes of barley revealed consistently high diversity, with no significant variations observed in the diversity parameters or the number of SNPs. Chromosome 7H exhibited the highest polymorphism despite having the fewest markers (GSR = 0.400; IEP = 0.315), followed by chromosomes 4H and 5H.

The genetic relatedness among the studied barley accessions was explored using various methods, and the results were meticulously analyzed. Within the barley collection, Nei's genetic distance (GD) ranged from 0 to 0.76, with an average of 0.42. Through cluster analysis, the samples were categorized into 3 main groups, leading to the complete separation of wild and cultivated barley genotypes (Fig. 1). Clusters I, III, and IV exclusively comprised samples of the species H. spontaneum, while cluster II grouped genotypes of the species H. vulgare. Although one wild barley sample clustered with cultivated barley genotypes, it formed a distinct subcluster.

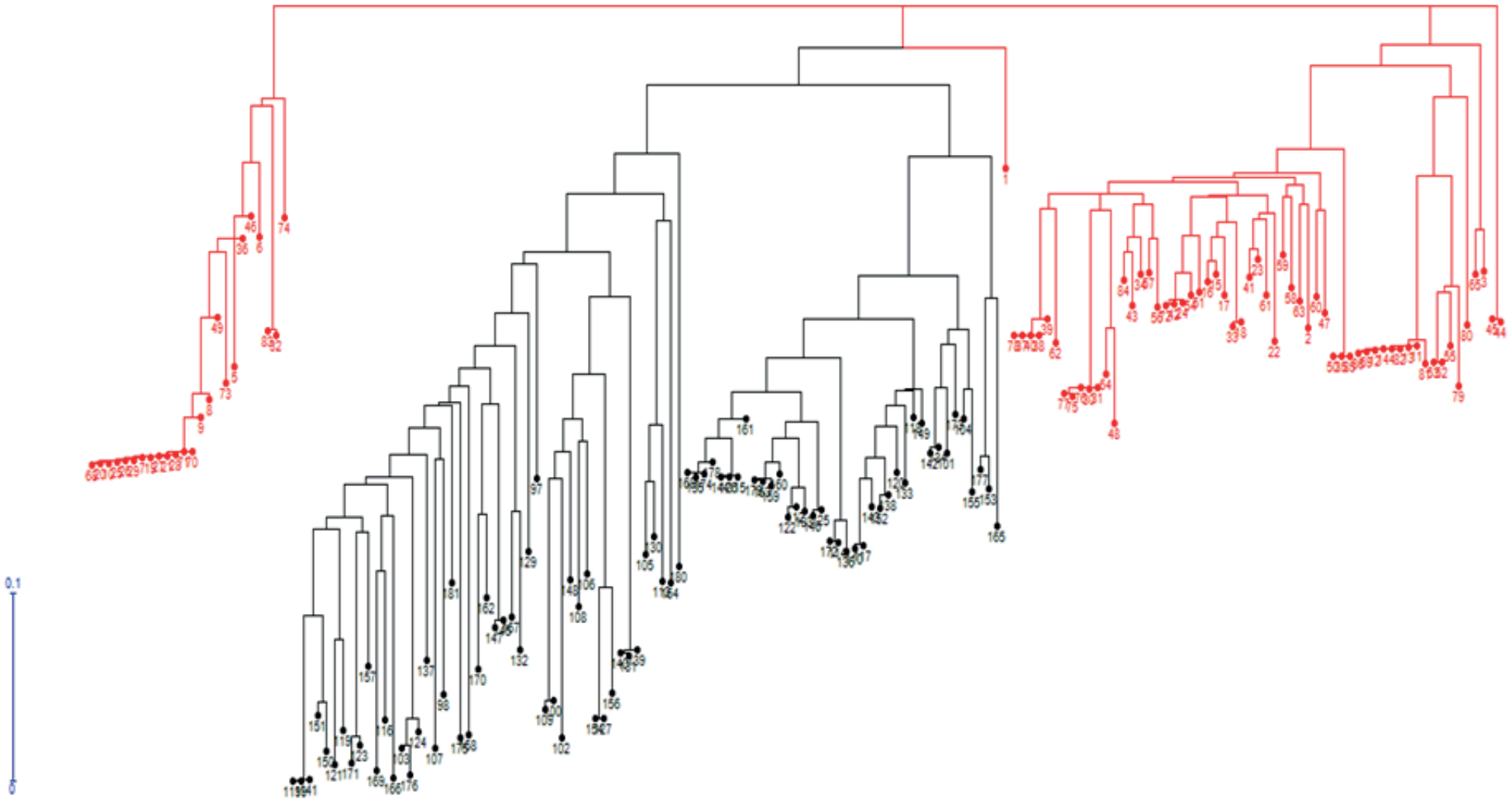

Figure 1. *Dendrogram showing genetic relatedness among 169 barley accessions based on SNP markers. Wild barley accessions of H. spontaneum are shown in red*

Notably, Cluster IV included 2 accessions with an extremely small genetic distance of 0.002. These genotypes, collected around Novkhana's dacha plots, differed by only 1 SNP. Cluster I encompassed 23, and Cluster III included 58 wild barley genotypes, indicating the genetic division of the studied H. spontaneum accessions into two distinct groups or populations. Traditionally, population identification relies on either geographic origin or specimen phenotype. However, it's important to note that the genetic structure of populations doesn't always align with the geographic proximity of samples. Sometimes, populations not geographically separated can diverge genetically due to unspecified barriers to gene flow.

The cultivated genotypes of H. vulgare, clustered in group II, were further subdivided into 2 subclusters comprising 47 and 37 accessions, respectively. These cultivated barley accessions represented various morphobiological groups and different botanical varieties within the species, based on ear color, density, and the number of grain rows. Interestingly, there was no discernible correlation between the clustering of accessions and ear density or the number of rows. However, it's noteworthy that 6 out of 7 genotypes with black ears studied were closely situated within subcluster II. Additionally, all 5 accessions of the species var. nigripallidum were part of this subcluster. Furthermore, distinct groupings were observed among samples belonging to varieties of var. nutans and var. pallidum.

The classification of barley samples was also investigated using the STRUCTURE program. The number of subpopulations (K) ranged from 2 to 5, and the model with K = 4 provided a more coherent representation of the collection's structure (Fig. 2). Higher values of K did not result in substantial changes in the results.

At K = 3, the entire barley collection comprised 67 accessions of H. spontaneum, 35 accessions of H. vulgare, and a group containing both barley species – H. spontaneum and H. vulgare. Upon adding another group (K = 4), wild and cultivated barley distinctly separated, forming two distinct subgroups within each species. Consequently, the STRUCTURE analysis results confirmed the presence of two genetically distinct groups within both cultivated and wild barley species.

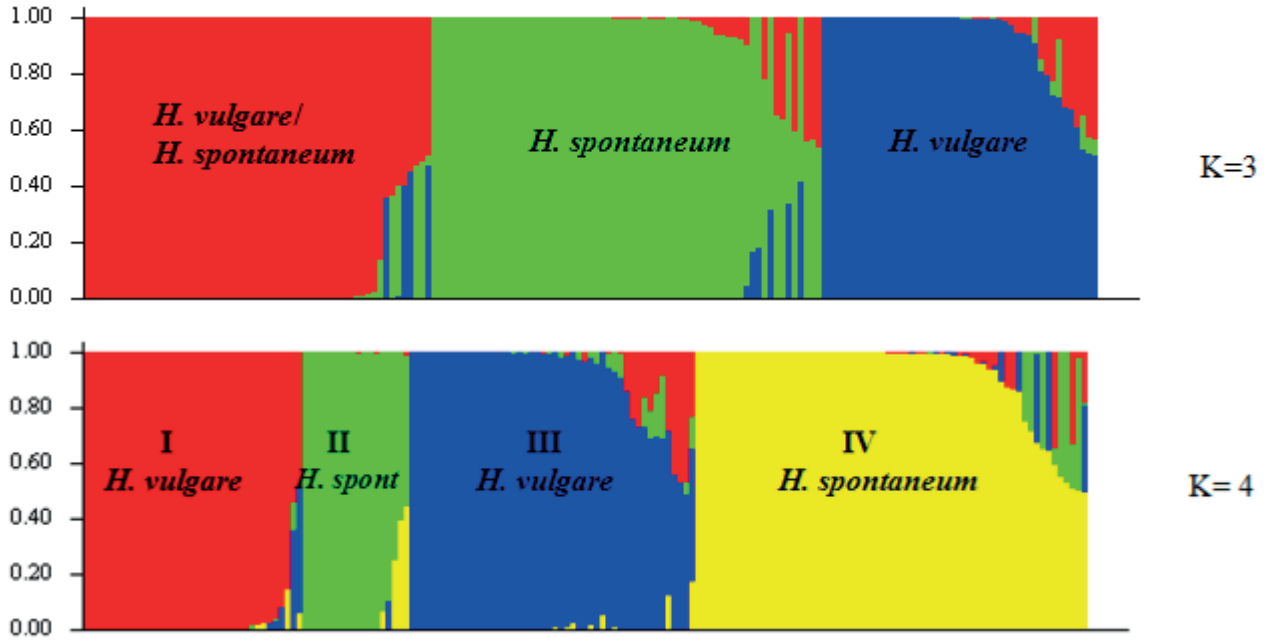

Figure 2. *Analysis using the STRUCTURE program of 169 barley accessions based on 255 SNP markers; K = 3, K = 4.*

To summarize, a collection of H. spontaneum and H. vulgare accessions gathered from 20 different regions in Azerbaijan was genotyped through the multiplex amplicon sequencing method comprising 255 primers distributed across the barley genome. The study revealed high genetic differentiation between the two species, with cultivated barley accessions exhibiting greater genetic diversity than wild barley samples. Additionally, genetically and structurally differentiated groups were identified within both barley species. The genetic diversity index (GDI) and polymorphism information content (PIC) for the studied barley collection were determined to be high. The 6H chromosome displayed the highest number of SNP markers, while the 7H chromosome had the fewest.

CONCLUSİON

The SNP markers obtained from multiplex amplicon sequencing are reliable for diverse assessments and association analyses in various collections. They prove suitable for studying genetic variability and relationships not only in barley overall but also in H. vulgare and H. spontaneum separately. Utilizing the distinct genetic groups identified in cluster and STRUCTURE analyses is recommended for crossbreeding and the creation of trait-specific collections.

REFERENCES

Bayer M. (2017). Development and Evaluation of a Barley 50k iSelect SNP Array / M.M.Bayer, P.Rapazote-Flores, M.Ganal [et al.] // Frontiers in plant science, №8, p.1-10.

Close T. (20090. Development and implementation of high-throughput SNP genotyping in barley / T.J.Close, P.R.Bhat, S.Lonardi [et al.] // BMC Genomics, №10, p.1-13.

Earl D. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method / D.A.Earl, B.M.vonHoldt // Conservation Genetics Resources, №4 (2), p. 359-361.

Genotyping by multiplexing amplicon sequencing: [Electronic resource] / Schnable Lab. (2015). URL: http://schnablelab.plantgenomics.iastate.edu/resources/protocols/.

Gurbanov E., Huseynova H. (2020). New spreading areas of some species in the Botanical-geographical Baku regions of the middle part of the Caspian coast. Acta Botanica Caucasica. Published by State University, Department of Botany and Plant Physiology. Volum 1, № 1.p.4-8.

Gurbanov E. Mammadova Z. (2016) Semishrub-legume-cereal shrub-grass phytocoenoses found in the interzonal vegetation of Azerbaijan. Actual problems of modern chemistry and biology. International Scientific Conference dedicated to the 93rd anniversary of the birth of National leader Heydar Aliyev. Ganja State University. Ganja, 2016. pp. 23-27.

Gurbanov E., Rzayeva A., (2022). Evaluation Of Decorative Properties Of Some Coniferous Plants Introduced To Absheron Peninsula.Acta Botanica Caucasica. Published by State University, Department of Botany and Plant Physiology. Volum 1, № 2. p.3-7.

Kadowaki S, Maher L, Portillo M, Albert RM, Akashi C, Guliyev F, et al. (2015). Geoarchaeological and palaeobotanical evidence for prehistoric cereal storage in the southern Caucasus: the Neolithic settlement of Goytepe (mid 8th millennium BP) // Journal of Archaeological Science, 53:408–25. 10.1016/j.jas.2014.10.021 PubMed PMID: WOS:000348621400036.

Keller B. (2017). Plant science: Genomic compartments in barley / B.Keller, S.G.Krattinger // Nature,. №544 (7651), p.1-5.

Komatsuda T. (2007). Six-rowed barley originated from a mutation in a homeodomain-leucine zipper I-class homeobox gene / T.Komatsuda, M.Pourkheirandish, C.He [et al.] // Proceedings of the National Academy of Sciences, , №104, p. 1424-1429.

Lu F. (2013). Switchgrass genomic diversity, ploidy, and evolution: novel insights from a network-based SNP discovery protocol / F.Lu, A.Lipka, J.Glaubitz [et al.] // PLoS Genetics, №9, p. 1-14.

Perrier X., Jacquemoud-Collet JP, DARwin software: (2006). [Электронный ресурс]/URL: http://darwin.cirad.fr/darwin

Pourkheirandish M. (2007). The importance of barley genetics and domestication in a global perspectives / M.Pourkheirandish, T.Komatsuda // Annals of Bot., №100, p. 999-1008.

Stein N. (2001). A new DNA extraction method for highthroughput marker analysis in a large genome species such as Triticum aestivum / N.Stein, G.Herren, B.Keller // Plant breeding, №120(4), p. 354-356.