Qabul qilindi: 2-oktabr 2026-yil • Ma’qullandi: 7-oktabr 2026-yil • Nashr qilindi: 15-oktabr 2026-yil

SELECTION OF SSR MARKERS FOR ASSESSING GENETIC DIVERSITY OF RADISH (*RAPHANUS SATIVUS* L.) LANDRACES: AN ANALYSIS OF PUBLISHED DATA

Tajiboyeva Dilafruz Baxromjon qizi

Master’s graduate of Andijan Institute of Agriculture and Agrotechnologies, Uzbekistan.

Kamilov Murodjon Muqimjonovich

Deputy Director of the Andijan Scientific Experimental Station of the Research Institute of Vegetable, Melon and Potato Crops; Doctor of Philosophy (PhD) in Agricultural Sciences, Uzbekistan.



Abstract. The reliability of SSR-based diversity studies depends largely on the choice of markers. Five SSR marker sets published for radish in 2011–2021 were compared by marker type, number of markers developed and tested, size of the test panel, number of alleles, polymorphism information content (PIC) and map position. Genomic SSRs showed more alleles and higher PIC values, while EST-SSRs had the advantage of known positions on the genetic map. Seven selection criteria were formulated and a stepwise workflow was proposed for building a working set of 20–30 markers to assess local radish forms of the Fergana Valley.

Keywords: radish, Raphanus sativus, SSR markers, marker selection, polymorphism information content, genetic diversity, landraces.

Introduction

Assessment of genetic diversity is a necessary step before local germplasm can be conserved and used in breeding. For radish, several marker systems have been applied to this task. Amplified fragment length polymorphism (AFLP) was used to compare accessions from different parts of the world [1]. Insertion/deletion (InDel) markers made it possible to build a core collection of 125 accessions that retained all 70 alleles of an 886-accession

collection [2]. Sequencing-based methods now provide tens of thousands of SNPs: 52 559 SNPs were identified among 520 accessions by ddRAD-Seq [3], and genotyping-by-sequencing was applied to a geographical core set of 152 accessions [4]. SSR sequencing combines the two approaches and was used to analyse 939 accessions with 38 SSR loci [5].

Despite the progress of sequencing methods, simple sequence repeat (SSR) markers remain a practical choice for many laboratories. They are codominant and multiallelic, give reproducible results, require only PCR and electrophoresis equipment, and are inexpensive per sample. SSR markers developed in radish are often transferable to related Brassicaceae species: in one study, 64.2% of EST-SSR primer pairs amplified products in *Raphanus*, *Brassica* and *Arabidopsis* accessions [6]. The availability of radish genome sequences [7] further simplifies the search for new SSR loci and the location of existing ones.

Several independent SSR sets have been published for radish, but they differ in origin (genomic DNA or expressed sequences), number of markers, level of polymorphism and availability of map positions. The results of a diversity study depend strongly on which markers are used, so the choice of markers should be justified before laboratory work begins.

The **aim** of this work was to compare published radish SSR sets and to justify selection criteria and a working set of markers for assessing the genetic diversity of local radish forms of the Fergana Valley.

Materials and Methods

Publications on the development and use of SSR markers in radish were searched in PubMed Central, J-STAGE, SpringerLink, MDPI and AGRIS. Works were included if they reported the development of a radish SSR set or its use on a panel of radish accessions with data on polymorphism or map position. Five SSR sets published in 2011–2021 met these criteria [5, 6, 8–10].

The sets were compared by marker type, number of markers developed and tested, number and type of accessions in the test panel, number of alleles per locus (Na), polymorphism information content and map position. PIC was considered as reported by the authors; in general it is calculated as $PIC = 1 - \Sigma\, p_i^2 - \Sigma\,\Sigma\, 2p_i^2p_j^2$, where p_i and p_j are the frequencies of the i-th and j-th alleles. Expected heterozygosity was considered as $He = 1 -$

Σp_i^2. Markers with PIC > 0.5 are usually regarded as highly informative.

Results

Characteristics of published SSR sets. The main characteristics of the five sets are summarized in Table 1.

Table 1.

Comparison of SSR marker sets published for radish

Source	Marker type	Number of markers	Test panel	Polymorphism	Map position
Shirasawa et al., 2011	EST-SSR	3800 developed, 630 polymorphic loci mapped	mapping population	630 polymorphic loci	9 linkage groups, 1129.2 cM
Jiang et al., 2012	EST-SSR (cDNA library)	125 developed, 28 tested	32 genotypes	2–7 alleles, mean PIC 0.500	not reported
Bae et al., 2015	genomic SSR	27 selected	144 cultivars	2–18 alleles (mean 7.77), PIC 0.491–0.906	not reported
Wang et al., 2015	EST-SSR	[to be specified]	93 accessions	cultivated and wild species separated	[to be specified]
Li et al., 2021	SSR sequencing	38	939 accessions	European wild radish most diverse	[to be specified]

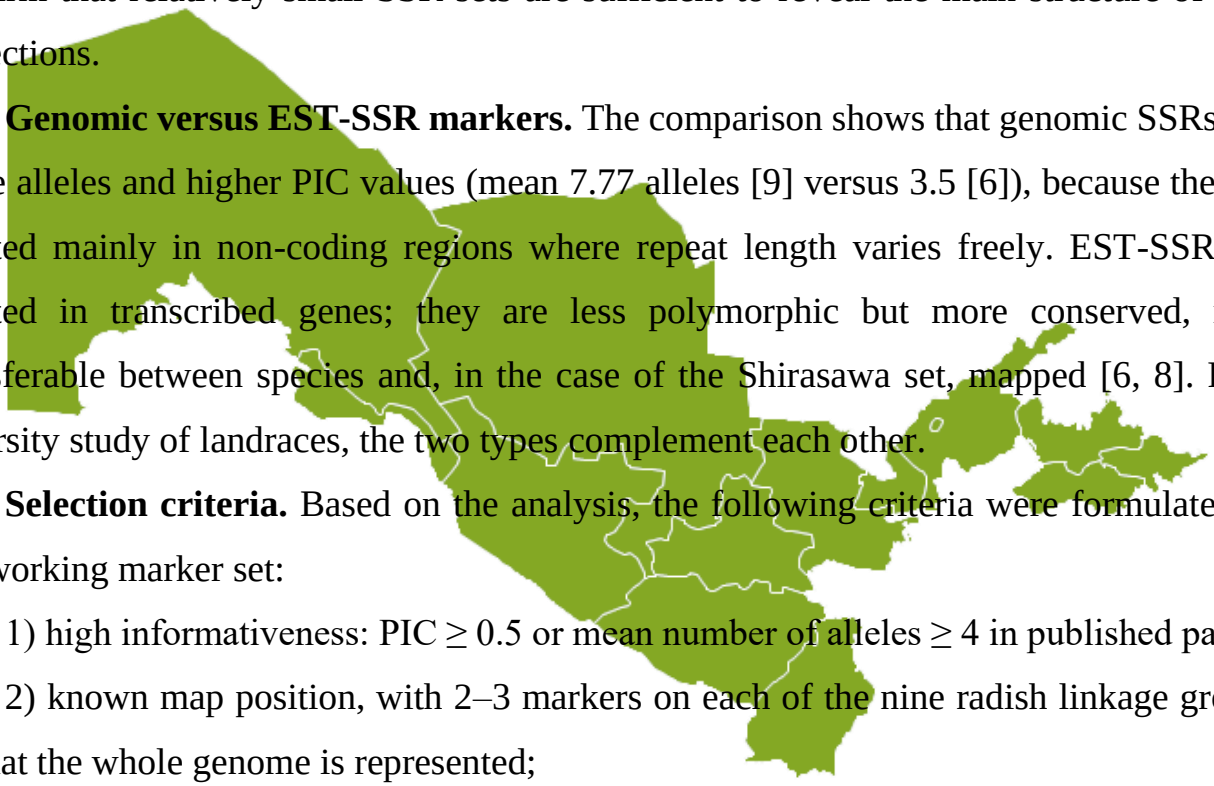
The largest resource was developed by Shirasawa et al., who collected 26 606 ESTs, designed 3800 SSR markers and mapped 630 polymorphic SSR loci together with 213 previously reported loci on nine linkage groups covering 1129.2 cM, with an average distance of 1.3 cM between loci [8]. This set is valuable mainly because each marker has a known map position, which allows markers to be distributed evenly over the genome.

Jiang et al. built a taproot cDNA library, sequenced 10 052 clones and found 179 SSRs; 110 of 125 primer pairs (88%) gave clear products. Of 28 markers tested on 32 genotypes, the number of alleles ranged from 2 to 7 (mean 3.5) and PIC from 0 to 0.825 (mean 0.500) [6]. Thus, about half of these markers would meet the usual threshold of high informativeness.

Bae et al. developed genomic SSR markers from an SSR-enriched library, selected 27 markers by polymorphism on 11 cultivars and applied them to 144 cultivars. The number of alleles per marker ranged from 2 to 18 (mean 7.77), PIC ranged from 0.491 to 0.906, and all

64 big radish cultivars were distinguished [9]. This set has the highest informativeness among those analysed, but map positions of the markers were not reported.

Wang et al. used EST-SSR markers on 93 accessions and showed clear separation of cultivated radish from wild *Raphanus* species [10]. Li et al. applied SSR sequencing with 38 loci to 939 wild, semi-wild and cultivated accessions and found that European wild radish is separated from cultivated radish and has higher genetic diversity [5]. These studies confirm that relatively small SSR sets are sufficient to reveal the main structure of large collections.



Genomic versus EST-SSR markers. The comparison shows that genomic SSRs give more alleles and higher PIC values (mean 7.77 alleles [9] versus 3.5 [6]), because they are located mainly in non-coding regions where repeat length varies freely. EST-SSRs are located in transcribed genes; they are less polymorphic but more conserved, more transferable between species and, in the case of the Shirasawa set, mapped [6, 8]. For a diversity study of landraces, the two types complement each other.

Selection criteria. Based on the analysis, the following criteria were formulated for the working marker set:

1) high informativeness: PIC ≥ 0.5 or mean number of alleles ≥ 4 in published panels;

2) known map position, with 2–3 markers on each of the nine radish linkage groups, so that the whole genome is represented;

3) single-locus amplification giving one clear product, which simplifies scoring;

4) mainly di- and trinucleotide repeats; trinucleotide repeats usually give fewer stutter bands and more reliable allele calling;

5) differences in product size that allow several markers to be combined in one multiplex PCR;

6) stable amplification under one or two annealing temperatures, to reduce the number of PCR programmes;

7) no evidence of null alleles or frequent failed amplification in published data.

Proposed workflow. A stepwise procedure was proposed for building and applying the working set (Table 2). The 27 genomic SSR markers of Bae et al. [9] form the core

because of their high PIC; linkage groups not covered by them are completed with mapped EST-SSR markers of Shirasawa et al. [8]. To check map positions of the genomic markers, their primer sequences can be aligned to the radish genome [7].

Table 2.

Workflow for selecting and applying SSR markers

Step	Content	Expected output
1. Compilation	Primer sequences of the 27 genomic SSRs [9] and mapped EST-SSRs [8] are compiled; positions are checked against the genome [7]	list of 40–50 candidate markers covering all 9 linkage groups
2. Pilot screening	Candidates are tested on 8 morphologically different local samples and 2 check cultivars	markers with clear, polymorphic products
3. Final set	Markers are selected by criteria 1–7 and grouped into multiplex sets	working set of 20–30 markers
4. Genotyping	Whole collection; at least 10 plants per sample genotyped individually	allele data matrix
5. Analysis	Na, Ne, Ho, He, PIC (GenAlEx); population structure (STRUCTURE); PCoA, UPGMA, AMOVA	genetic diversity indices and grouping of samples

Sampling within accessions. Radish is a cross-pollinated crop, and most of its genetic variation is found within accessions: in the genotyping-by-sequencing study of 152 accessions, 77.49% of variation was within individuals and only 6.51% among geographic regions [4]. Therefore, genotyping of a single plant or a small bulk per accession would underestimate diversity. Individual genotyping of at least 10 plants per sample is recommended; bulked DNA may be used only for the pilot screening of markers.

Sources of error. The main limitations of SSR data are null alleles, size homoplasy (identical fragment length with different sequence), stutter bands and scoring errors. To control them, each PCR plate will include the same reference genotypes, about 10% of samples will be genotyped twice, and alleles will be binned by size using the reference genotypes.

Discussion

Sequencing-based SNP panels provide far more loci than SSRs and give a detailed picture of genome-wide diversity [3, 4]. However, they require sequencing services and bioinformatic support, which increases cost and time. For a first assessment of a regional

collection of 40–50 samples, a well-chosen set of 20–30 polymorphic SSRs is sufficient to estimate diversity, reveal population structure and select genetically distant parents, as shown by studies of much larger collections with 27–38 SSR loci [5, 9].

The combination of highly polymorphic genomic SSRs with mapped EST-SSRs uses the strengths of both marker types. In the future, the same loci can be analysed by SSR sequencing [5], which increases resolution by revealing sequence variation within fragments of equal length, and the data can be linked with genes controlling important traits, such as bolting time and disease resistance.

Conclusions

1. Five published radish SSR sets differ in informativeness and map information: genomic SSRs give up to 18 alleles per locus and PIC up to 0.906, while EST-SSRs provide known positions on nine linkage groups. This comparison shows that the two types should be used together.

2. Seven criteria were formulated for selecting markers, including PIC ≥ 0.5, even genome coverage and single-locus amplification. Their use ensures that the working set is both informative and technically reliable.

3. A stepwise workflow was proposed, leading to a working set of 20–30 SSR markers and individual genotyping of at least 10 plants per sample. With this approach, reliable assessment of genetic diversity of Fergana Valley radish landraces across the whole genome is achieved.

References

1. Wang N., Kitamoto N., Ohsawa R., Fujimura T. Genetic diversity of radish (Raphanus sativus) germplasms and relationships among worldwide accessions analyzed with AFLP markers // Breeding Science. – 2008. – Vol. 58. – P. 107–112.

2. Lee Y.J., Mun J.H., Jeong Y.M., Joo S.H., Yu H.J. Assembly of a radish core collection for evaluation and preservation of genetic diversity // Horticulture, Environment, and Biotechnology. – 2018. – Vol. 59, No. 5. – P. 711–721.

3. Kobayashi H., Shirasawa K., Fukino N., Hirakawa H., Akanuma T., Kitashiba H. Identification of genome-wide single-nucleotide polymorphisms among geographically

diverse radish accessions // DNA Research. – 2020. – Vol. 27, No. 1. – dsaa001.

4. [Authors to be specified]. Radish geographical core set genotyped by sequencing // Genetic Resources and Crop Evolution. – 2021. – DOI: 10.1007/s10722-021-01212-6.

5. Li X., Wang J. et al. SSR-sequencing reveals the inter- and intraspecific genetic variation and phylogenetic relationships among an extensive collection of radish (Raphanus) germplasm resources // Biology. – 2021. – Vol. 10, No. 12. – 1250.

6. Jiang L., Wang L., Liu L., Zhu X., Zhai L., Gong Y. Development and characterization of cDNA library based novel EST-SSR marker in radish (Raphanus sativus L.) // [journal to be specified]. – 2012.

7. Kitashiba H., Li F., Hirakawa H. et al. Draft sequences of the radish (Raphanus sativus L.) genome // DNA Research. – 2014. – Vol. 21. – P. 481–490.

8. Shirasawa K. et al. An EST-SSR linkage map of Raphanus sativus and comparative genomics of the Brassicaceae // DNA Research. – 2011. – Vol. 18, No. 4. – P. 221–232.

9. Bae K.-M., Sim S.-C., Hong J.-H., Choi K.-J., Kim D.-H., Kwon Y.-S. Development of genomic SSR markers and genetic diversity analysis in cultivated radish (Raphanus sativus L.) // Horticulture, Environment, and Biotechnology. – 2015. – DOI: 10.1007/s13580-015-0089-y.

10. Wang Q., Zhang L., Zheng P. Genetic diversity and evolutionary relationship analyses within and among Raphanus species using EST-SSR markers // Molecular Breeding. – 2015. – Vol. 35. – 62.