***Short Research Article***

**Genetic diversity study in tomato (s*olanum lycopersicum* l.) Discovered by simple sequence repeats (ssr) markers**

**Abstract**

Tomato (*Solanum lycopersicum* L.) is a globally important vegetable crop and enhancing its genetic diversity is critical for developing improved cultivars. This study evaluated the genetic diversity among eight tomato genotypes DVRT 2, IIHR 335, ATL 17-06, GAT 5, GAT 8, ATL 18-04, NTL 12-02 and GP 11 using 15 polymorphic Simple Sequence Repeat (SSR) markers out of 47 screened. A total of 87 alleles were detected, with an average of 6.00 alleles per locus, indicating substantial polymorphism. Major allele frequency ranged from 0.062 to 0.167, averaging 0.135, suggesting balanced allele distribution. The PIC values ranging from 0.501 to 0.845 demonstrated high genetic variability. Particularly, SSR287, SSR598 and TES856 were identified as highly informative markers with PIC values above 0.8. Cluster analysis grouped the genotypes into three main clusters, with genetic distances ranging from 0.67 to 1.00. DVRT 2 exhibited the highest divergence (1.00). The findings underscore a wide genetic base among the genotypes, offering valuable insights for future breeding programs aimed at improving yield potential, stress resilience and nutritional quality in tomato.

**Key Words**

Tomato, Markers, Diversity, Alleles, PIC, Clusters

**Introduction**

Vegetables are progressively acknowledged as a vital component for nutritional security. Because of its economic and nutritional worth, fruit and vegetable production currently has a higher worldwide value than all food grains. Among all, tomato [*Solanum lycopersicum* (L.), 2n = 2x = 24] is one of the most widely grown vegetable crop in both tropics and sub tropics of the world.

In plant breeding, the development of molecular marker systems significantly accelerated the selection and evaluation processes. In contrast to conventional breeding techniques, which mostly depend on phenotypic selection also it is time consuming, resource intensive and environmental influences might have an impact. Molecular markers have proven to be valuable tools in the evaluation of genetic variation both within and between species. These molecular tools have improved the speed and accuracy of achieving desired agronomic traits. A molecular marker refers to a specific DNA sequence located at a specific position on a chromosome, linked to a particular gene or trait. It represents genetic variation, which may arise from mutations, insertions, deletions, or alterations in genomic loci and can be detected. Genetic markers can vary in length from short sequences like single nucleotide polymorphisms (SNPs) to the most abundant and widely used form of genetic variation which is more complex and variable regions, such as minisatellites. These markers may or may not directly relate to the observable traits of an organism. They offer several advantages over traditional phenotype-based approaches, as they are consistent and can be detected in all tissues, regardless of the cell's state of growth, differentiation, development or defense.

Molecular markers have numerous uses in plant breeding, such as analyzing germplasm, defining genetic purity, confirming hybrids and concussing the evolutionary connections between genotypes. Breeding operations have been much more effective because to molecular marker technology, which permits breeders to rapidly and consistently screen vast populations for precise traits. This has also helped to create superior cultivars with increased yield, stress tolerance and nutritional quality.

**Methodology**

The present investigation scrutinizes the inheritance pattern of morphological characters and analyze the molecular diversity of parental genotypes during the *kharif*-*rabi* season of 2024-25 at the Main Vegetable Research Station, Anand Agricultural University, Anand. Eight diverse parents of tomato *viz*., DVRT 2, IIHR 335, ATL 17-06, GAT 5, GAT 8, ATL 18-04, NTL 12-02 and GP 11 were used for present investigation was undertaken with objective to study the parental diversity using molecular markers. The details of the parents are given in Table 1.

**Parental Genetic Diversity Based on Molecular Markers**

As part of the formation of a molecular tool kit for considering diversity within the tomato parental genotype, SSR (Simple Sequence Repeat) technology was employed on eight tomato parental genotypes that were hired to produce hybrids. To study genetic diversity at the molecular level, tender leaves from three-week-old seedlings were unruffled for DNA isolation.

**Glassware, plasticware and reagents**

The glassware exploited was Borosilicate from Schott Duran, Germany, although the plasticware was obtained from Eppendorf, Germany. All chemicals and reagents employed were of extra pure or molecular biology grade quality, assimilated from multiple companies in India and Abroad. Before use, glassware, plasticware and reagents were pasteurized through double autoclaving and stored conferring to their specific requirements.

**Genomic DNA extraction**

Genomic DNA extraction from leaf samples of all genotypes under examination was directed at the Department of Agril. Biotechnology, Anand Agricultural University, Anand. Leaves were composed at the four-leaf stage and DNA isolation was completed using the Cetyl Trimethyl Ammonium Bromide (CTAB) protocol manufacturing by Doyle and Doyle (1990).

**Protocol for genomic DNA extraction**

Genomic DNA extraction from composed leaves was achieved using the Cetyl Trimethyl Ammonium Bromide (CTAB) technique (Doyle and Doyle, 1990), adapted as follows.

1. Each leaf sample weighing 1.0 g was finely severed into small pieces using sterilized scissors and ground with liquid nitrogen (N2) using a mortar and pestle until a fine powder was found
2. The powder was elated into a 2 ml sterile tube (Eppendorf), followed by the addition of 800 μl pre-warmed CTAB solution laterally with 2 μl Proteinase-K
3. The tubes were originally nurtured at 45°C for 15 minutes followed by incubation at 65°C for 60 minutes in a water bath
4. An equivalent volume of Chloroform: Isoamyl alcohol (24:1) was supplementary to each tube and mixed gently by capsizing for 5-10 minutes, followed by centrifugation at 12,000 rpm for 15 minutes at 4°C
5. The supernatant was prudently transferred and relocated to new tubes by means of a micro pipette [Eppendorf (100-1000) μl]
6. Steps 4 and 5 were frequent twice
7. The supernatant was composed into new tubes and an equal volume of ice-cold Isopropanol was supplementary to each tube to precipitate DNA
8. The solution was hatched overnight at -20°C
9. The occasioned solution was centrifuged at 10,000 rpm for 10 minutes at 10°C
10. The supernatant was rejected and 100 μl ethanol (70%) was supplementary to wash the pellet.
11. The pellets were air-dried and resuspended in 100 μl of 1X TE buffer
12. To acquire RNA-free DNA samples, 1 μl RNase was added and reserved at 37°C for 60 minutes followed by incubation at 65°C for 10 minutes in a water bath
13. Finally, the DNA stock was deposited at -20°C and used to make a working stock (50 ng/µl) for PCR amplification

SSR (Simple Sequence Repeat) marker technology was employed on eight tomato parental genotypes. Tender leaves from three-week-old seedlings were unruffled for DNA isolation. Genomic DNA extraction from leaf samples of all genotypes under examination was directed at the Department of Agril. Biotechnology, Anand Agricultural University, Anand. Leaves were composed at the four-leaf stage and DNA isolation was completed using the Cetyl Trimethyl Ammonium Bromide (CTAB) protocol manufacturing by Doyle and Doyle (1990). After the extraction process, the quality calculation and quantification of DNA were directed using agarose gel electrophoresis and a spectrophotometer, correspondingly. A working DNA solution of 50 ng/µl was equipped for polymerase chain reaction (PCR). The extracted DNA from particular tomato parental genotypes was amplified using simple sequence repeat markers (SSR). The SSR primers exploited to examine the polymorphism between the different tomato genotypes are detailed in Table 2.

The molecular genetic analysis involved evaluating DNA polymorphism by comparing amplified fragment lengths against a 100 bp DNA ladder, with allele sizes determined based on migration patterns. Genetic parameters including major allele frequency, expected heterozygosity (Jaccard, 1908)., observed heterozygosity and polymorphism information content (PIC) were calculated using PopGene software (Liu and Muse, 2005), with PIC assessed following Botstein *et al*. (1980). A dissimilarity matrix was generated using DARWIN 6.0 with 1000 bootstrap iterations and a Neighbor-Joining dendrogram was constructed to analyze phylogenetic relationships. This comprehensive approach provided insights into genetic diversity, polymorphism patterns and phylogenetic associations among the studied genotypes.

**Results and Discussion**

**Marker data analysis using SSR markers**

The assessment of parental divergence is an important and primary goal in any plant breeding programme. Molecular markers initiated valuable foil to morphological and physiological description of cultivars because they permit cultivar identification in plant growth. The resolution of this study was to assess the genetic diversity across eight genotypes (Agong *et al*. 2001). Initially, a set of 47 SSRs were screened on eight DNA samples to identify reproducible and scorable primer. During primer screening, 36 primers amplified successfully of which 21 primers were found monomorphic; while, remaining 15 were polymorphic. All the polymorphic markers were eventually amplified through PCR to analyse genetic diversity among tomato genotypes. Various bio statistical parameters *viz*., effective number of alleles, allele frequency, observed heterozygosity, expected heterozygosity, Nei’s gene diversity and PIC values were calculated (Song *et al*. 2006).

Molecular weight of the amplified PCR products fluctuated from 164bp (SSR111 and SSR598) to 526 bp (TES872), which imitated notable difference in the number of repeats amongst alleles. The result existed in Table 3 revealed a total 87 alleles were perceived across the 15 polymorphic SSR loci found in the genotypes under examination. The number of alleles per locus reached from 3 (TES332) to 10 (SSR287) with an average of 6.00 alleles per locus. The actual number of alleles, which accounts for the relative frequencies of alleles was vacillated from 2.333 (TES332) to 7.111 (SSR287) with an average of 4.97, signifying considerable allelic diversity. The major allele frequency was ranged from 0.062 (SSR287) to 0.167 (Lega005, TES1276, Leta007 and SSR111) with an average of 0.135. This comparatively low frequency proposed a stable allele distribution and specified a high level of genetic diversity. The observed heterozygosity was mostly low with most markers unveiling values of 0.00 except for SSR287 (1.00), SSR598 (0.429) and TGS602 (0.167) shimmering a limited incidence of heterozygous individuals between the genotypes studied. The expected heterozygosity, oscillated from 0.571 (TES332) to 0.859 (SSR287) with an average value of 0.777 (Table 3). Polymorphism Information Content (PIC) restrained the clarity of markers, ranged from 0.501 (TES332) to 0.845 (SSR287) with an average of 0.744 (Reddy *et al*. 2014, Parmar *et al*. 2010 & Rai *et al*. 2016). Maximum number of markers displayed PIC values above 0.7, suggesting they are highly edifying for genetic studies. SSR287, SSR598 and TES856 presented PIC values ≥0.8, making them predominantly valuable for association studies and selection in breeding programs (Raveendar *et al*., 2016, Osie *et al*. 2014). The overall high average PIC value also specified the occurrence of diverse alleles within the population (Table 3)..

**Construction of dendrogram**

The dissimilarity matrix and cluster analysis for SSR data of eight tomato genotypes were performed using DARWIN 6 software. The resulting matrix illustrated the genetic relationships among the genotypes based on SSR marker profiles with dissimilarity values ranging from 0.67 to 1.00 as mentioned in Table 4. Lower values denoted closer genetic relationships, while higher values reflected greater divergence. Among all pairs, DVRT 2 showed the highest genetic divergence (1.00) when compared with several genotypes. Genotypes such as NTL 12-02, GP 11 and IIHR 335 displayed high dissimilarity with most genotypes, suggesting a distinct genetic background. These findings revealed substantial genetic diversity among the evaluated genotypes, offering valuable potential for broadening the genetic base in tomato breeding programs (Fig. 1). Based on the dendrogram analysis, the evaluated genotypes were classified into three distinct main clusters, A, B and C reflecting varying levels of genetic relatedness. Cluster A included the genotypes GP 11 and NTL 12-02 and Cluster B comprised GAT 8 and GAT 5, suggesting minimal genetic divergence. Cluster C exhibited greater diversity and was further subdivided into two sub-clusters: C1 and C2. Sub-cluster C1 included IIHR 335 and ATL 18-04, while sub-cluster C2 comprised ATL 17-06 and DVRT 2. The formation of sub-clusters within Cluster C highlighted a broader spectrum of genetic variability among these genotypes compared to those in Clusters A and B (Table 5).

**Conclusions**

Molecular characterization revealed high level of genetic diversity with major allele frequencies ranging from 0.062 to 0.167. The Polymorphism Information Content (PIC) values ranged from 0.501 to 0.845 with an average of 0.744, indicating that the chosen markers were highly informative for selection of genotypes in breeding programmes. Markers SSR287, SSR598 and TES856 were the most informative (PIC>0.8), making them valuable for marker assisted selection in tomato improvement. Cluster analysis using DARWIN 6 grouped the eight genotypes into three main clusters, confirming wide genetic divergence. DVRT 2 was found to be the most genetically divergent and grouped in a separate sub-cluster, making it a valuable donor parent for future breeding programmes.

**Disclaimer (Artificial intelligence)**

Author(s) hereby declare that NO generative AI technologies such as Large Language Models (ChatGPT, manuscript.

**References**

Agong S. G., Schittenhelm S. & Friedt. W. (2001). Genotypic variation of Kenyan tomato (*Lycopersicon esculentum* L.) germplasm. *Journal of Food Technolology*,6, 13–17.

Botstein, D., White, R. L., Skolnick, M., & Davis, R. W. (1980). Construction of a linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics*, 32 (3), 314-331.

Doyle, J. J., & Doyle, J. L. (1990). Isolation of plant DNA from fresh tissue. *Focus*, 12, 13-15.

Liu, K., & Muse, S. V. (2005). Power Marker: an integrated analysis environment for genetic marker analysis. *Bioinformatics*, 21 (9), 2128-29.

Osei, M. K., Bonsu, K. O., Agyeman, A. & Choi H. S. (2014). Genetic diversity of tomato germplasm in Ghana using morphological characters. *International Journal of Plant Soil Science*,3: 220–231.

Parmar, P., Oza, P., Vishal., Chauhan, V. & Patel, A. D., Kathiria, K. B., Subramanian R. B. (2010). Genetic diversity and DNA fingerprint study of tomato discerned by SSR markers. *International Journal of Biotechnology and Biochemistry*,6: 657-666.

Rai, G. K., Jamwal, D., Singh, S., Parveen, A., Kumar, R. R., Singh, M., Rai, P. K. & Salgotra, R. K. (2016). Assessment of genetic variation in tomato (*Solanum lycopersicum* L.) based on quality traits and molecular markers. *Journal of Breeding and Genetics*, 48, 80- 89.

Raveendar, S., Jong-Wook Chung., Gi-An Lee1., Jung-Ro Lee1., Kyung-Jun Lee1., Myoung-Jae Shin., Yang-Hee Cho., & Kyung-Ho M. (2016). Genetic diversity and population structure of Asian tomato accessions based on Simple-Sequence Repeats. *Plant Breeding and Biotechnology*, 4 (3), 306-314.

Reddy, B. R., Reddy, M. P., Begum, H. & Sunil, N. (2013). Genetic diversity studies in tomato (*Solanum lycopersicum* L.). *Journal of Agriculture and Veterinary Science*, 4, 53-55.

Song J., Chen J., Chen H. Y., Liu Y. & Zhuang T. M. (2006). Research of genetic diversity of tomato using SSR markers. *Journal of Shanghai Jiao tong University*. 24, 524−528.

**Table 1. Details of the parents used in the present study**

|  |  |  |
| --- | --- | --- |
| **Sr. no.** | **Genotypes** | **Source** |
| **1** | DVRT 2 | ICAR, Varanasi |
| **2** | IIHR 335 | ICAR, Bengaluru |
| **3** | ATL 17-06 | MVRS, Anand |
| **4** | GAT 5 | MVRS, Anand |
| **5** | GAT 8 | MVRS, Anand |
| **6** | ATL 18-04 | MVRS, Anand |
| **7** | NTL 12-02 | VRS, Navsari |
| **8** | GP 11 | MVRS, Anand |

**Table 2.** **List of SSR primers along with their sequence used for the present study**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Sr. No.** | MarkerName | Tm | **Sr. No.** | MarkerName | Tm | **Sr. No.** | MarkerName | Tm | **Sr. No.** | MarkerName | Tm |
| **1.** | SSR19 | 59.4 | **13.** | SSR47 | 56.5 | **25.** | A1895126 | 63.2 | **37.** | TGS467 | 52.4 |
| 57.3 | 55.9 | 59.7 | 57.2 |
| **2.** | SSR22 | 61.4 | **14.** | TES332 | 49.7 | **26.** | SSR111 | 57.3 | **38.** | SSR331 | 53.8 |
| 57.3 | 51.8 | 56.5 | 57.3 |
| **3.** | SSR92 | 56.5 | **15.** | TES856 | 47.7 | **27.** | TES1192 | 53.8 | **39.** | SSR28 | 58.3 |
| 59.7 | 53.8 | 54.4 | 53.2 |
| **4.** | LEgata002 | 57.9 | **16.** | TES1276 | 52.4 | **28.** | SSR94 | 55.3 | **40.** | SSR20 | 58.4 |
| 55.6 | 51.8 | 57.3 | 57.3 |
| **5.** | TGS993 | 53.6 | **17.** | Lega006 | 51.8 | **29.** | SSR9 | 55.9 | **41.** | LEaat007 | 58.4 |
| 55.3 | 49.7 | 57.9 | 53.8 |
| **6.** | SSR304 | 57.1 | **18.** | TGS500 | 57.3 | **30.** | SSR598 | 55.9 | **42.** | TGS602 | 53.5 |
| 55.8 | 57.8 | 57.3 | 53.8 |
| **7.** | Lega005 | 51.8 | **19.** | Leta007 | 56.6 | **31.** | SSR300 | 57.3 | **43.** | TES734 | 51.1 |
| 52.8 | 57.3 | 55.9 | 54.4 |
| **8.** | TES872 | 53.8 | **20.** | A1773078 | 59.3 | **32.** | SSR76 | 55.3 | **44.** | TGS504 | 51.8 |
| 49.7 | 62.1 | 57.3 | 53.5 |
| **9.** | SSR287 | 52.4 | **21.** | AW037347 | 59.3 | **33.** | SSR63 | 57.3 | **45.** | LEta016 | 57.3 |
| 52.4 | 53.7 | 55.9 | 48.5 |
| **10.** | SSR86 | 55.3 | **22.** | SSR115 | 56.5 | **34.** | LEaat003 | 50.5 | **46.** | SSR104 | 53.8 |
| 61.4 | 57.5 | 46.6 | 47.7 |
| **11.** | AW034362 | 56.7 | **23.** | TGS862 | 56.2 | **35.** | TGS0070 | 47.7 | **47.** | TGS740 | 53.8 |
| 57.9 | 54.8 | 49.7 | 59.8 |
| **12.** | SSR99 | 55.9 | **24.** | SSR310 | 51.8 | **36.** | TGS2259 | 55.9 |
| 56.5 | 51.8 | 52.4 |

**Table 3. Analysis of polymorphic SSR markers**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Sr.no. | Marker | MW | Observed number of alleles | Effective number of alleles | Majorallele frequency | Observed heterozygosity | Expected hetero-zygosity | PIC |
| 1 | Lega005 | 205-213 | 5 | 4.500 | 0.167 | 0 | 0.778 | 0.744 |
| 2 | TES872 | 454-526 | 6 | 5.333 | 0.125 | 0 | 0.812 | 0.786 |
| 3 | SSR287 | 240-278 | 10 | 7.111 | 0.062 | 1.000 | 0.859 | 0.845 |
| 4 | TES332 | 183-186 | 3 | 2.333 | 0.143 | 0 | 0.571 | 0.501 |
| 5 | TES856 | 238-245 | 7 | 6.400 | 0.125 | 0 | 0.844 | 0.825 |
| 6 | TES1276 | 248-256 | 5 | 4.500 | 0.167 | 0 | 0.778 | 0.744 |
| 7 | Lega006 | 212-218 | 4 | 2.579 | 0.143 | 0 | 0.612 | 0.570 |
| 8 | TGS500 | 235-247 | 7 | 6.400 | 0.125 | 0 | 0.844 | 0.725 |
| 9 | Leta007 | 246-251 | 4 | 3.600 | 0.167 | 0 | 0.722 | 0.671 |
| 10 | SSR111 | 164-187 | 6 | 6.000 | 0.167 | 0 | 0.833 | 0.710 |
| 11 | SSR598 | 164-192 | 7 | 6.125 | 0.071 | 0.429 | 0.837 | 0.816 |
| 12 | SSR300 | 247-253 | 6 | 4.571 | 0.125 | 0 | 0.781 | 0.754 |
| 13 | SSR331 | 171-178 | 6 | 4.500 | 0.143 | 0 | 0.816 | 0.791 |
| 14 | TGS602 | 275-280 | 6 | 5.333 | 0.083 | 0.167 | 0.792 | 0.763 |
| 15 | LEaat007 | 261-275 | 5 | 7.111 | 0.143 | 0 | 0.776 | 0.740 |

**Table 4. Nei’s dissimilarity coefficient matrix of eight tomato genotypes based on SSR analysis**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Genotypes** | **DVRT 2** | **IIHR 335** | **ATL 17-06** | **GAT 5** | **GAT 8** | **ATL** **18-04** | **NTL** **12-02** | **GP 11** |
| **DVRT 2** | 0.00 |  |  |  |  |  |  |  |
| **IIHR 335** | 0.93 | 0.00 |  |  |  |  |  |  |
| **ATL 17-06** | 1.00 | 0.97 | 0.00 |  |  |  |  |  |
| **GAT 5** | 0.97 | 0.93 | 0.87 | 0.00 |  |  |  |  |
| **GAT 8** | 0.93 | 1.00 | 1.00 | 0.67 | 0.00 |  |  |  |
| **ATL 18-04** | 1.00 | 0.80 | 1.00 | 0.87 | 0.73 | 0.00 |  |  |
| **NTL 12-02** | 1.00 | 1.00 | 0.93 | 0.87 | 0.83 | 0.77 | 0.00 |  |
| **GP 11** | 0.93 | 1.00 | 1.00 | 0.93 | 0.80 | 0.87 | 0.77 | 0.00 |

**Table 5.** **Distribution of parental genotypes to different clusters based on Nei’s dissimilarities**

|  |  |  |  |
| --- | --- | --- | --- |
| **Sr. no** | **Main cluster** | **Sub cluster** | **Genotypes** |
| **1** | A | GP 11, NTL 12-02 |
| **2** | B | GAT 8, GAT 5 |
| **3** | C | C1 | IIHR 335, ATL 18-04 |
| C2 | ATL 17-06, DVRT 2 |

**Figure 1. Dendrogram representing the association between the eight parental genotypes of tomato**

**DVRT 2**

**NTL 12-02**



**GAT 8**

**GAT 5**

**IIHR 335**

**ATL 18-04**

**ATL 17-06**

**DVRT 2**