

IDENTIFICATION AND VALIDATION OF INTRA-SPECIES TRANSFERABILITY OF GENOME-WIDE FUNCTIONAL SSR MARKERS IN *GLYCINE MAX*

DR. R. BHOWMICK^{1,2}, K. PAUL³,
S. BHATTACHARJEE^{4*}

¹ Crop Improvement Division, ICAR-VPKAS, Almora-263601

² Crop Improvement Division, ICAR-CRIJAF, Kolkata-700121

³ ICAR-National Institute for Plant Biotechnology, New Delhi-110012

⁴ School of Crop Sciences, ICAR-Indian Agricultural Research Institute, Jharkhand, India-825405

* Corresponding author: Sougata Bhattacharjee,
e-mail: sougata.b@icar.gov.in

*Genic codominant multiallelic markers are essential to identify the genetic variation, population diversity and evolutionary history of a species. Soybean (*Glycine max*) is a major legume crop having importance in both a protein-rich pulse as well as a high recovery oilseed crop. To date, no genome-wide genic SSR markers have been elucidated in this crop of high importance. This article aims to identify and validate regulatory gene-derived SSR markers in soybean. The coding sequences of *Glycine max* were downloaded from PlantTFDB and used for the identification, followed by the localization of SSRs by using a Perl 5 script (MISA, MicroSATellite identification tool). The flanking primers to SSRs were designed and chromosomal distribution and Gene ontology searches were performed using BLAST2GO. Twenty random SSR markers were validated to check cross-species transferability and genetic diversity study was performed. A set of 1138 simple sequence repeat markers from transcription factor coding genes were designed and designated as TF-derived SSR markers. They were anchored on 20 *G. max* chromosomes, and the SSR motif frequency was one per 4.64 kb. Trinucleotide repeats were found abundant and tetra-, as well as pentanucleotide frequency, was least in soybean. Gene Ontology search revealed the diverse role of SSR-containing TFs in soybean. Eight soybean accessions were analyzed for identified twenty candidates for genic SSR diversification, and a principal co-ordinate analysis, a genic dissimilarity-based unweighted neighbour-joining tree, was constructed. Our findings will serve as a potential functional marker resource for marker-assisted selection and genomic characterization of soybean.*

Key words: Transcription Factor (TFs), SSR, Genic Marker, Gene Ontology, Genetic Diversity.

© ІНСТИТУТ КЛІТИННОЇ БІОЛОГІЇ ТА ГЕНЕТИЧНОЇ ІНЖЕНЕРІЇ НАН УКРАЇНИ, 2024

ІДЕНТИФІКАЦІЯ І ВАЛІДАЦІЯ ВНУТРІШНЬОВИДОВОЇ ПЕРЕНΟΣИМОСТІ ЗАГАЛЬНОГЕНОМНИХ ФУНКЦІОНАЛЬНИХ SSR МАРКЕРІВ У *GLYCINE MAX*

Генні кодомінантні мультиалельні маркери надзвичайно важливі для ідентифікації генетичної варіативності, різноманітності популяції та еволюційної історії виду. Соя (*Glycine max*) – це основна рослина з родини бобових, важливість якої полягає в тому, що вона є і зернобобовою рослиною, багатою на білки, і маслянистою культурою з високим виходом олії. Наразі у цій надзвичайно важливій рослині не було виявлено жодних загальногеномних генних SSR маркерів. Мета цієї статті полягає в ідентифікації та валідації регуляторних генних SSR маркерів у сої. Кодуючі послідовності *Glycine max* було завантажено з PlantTFDB і використано для ідентифікації, після чого було проведено локалізацію SSR маркерів за використання скрипту Perl 5 (MISA, інструмент ідентифікації мікросателітів). Було розроблено фланкуючі праймери для SSR і виконано пошук хромосомного розподілу та онтології гену за допомогою BLAST2GO. Двадцять рандомізованих SSR маркерів було валідовано для перевірки переносимості між видами та проведення дослідження генетичної різноманітності. Набір із 1138 маркерів простих повторюваних послідовностей із генів кодування фактору транскрипції було створено і визначено в якості SSR маркерів, отриманих за допомогою ФТ. Вони були зафіксовані на 20 *G. max* хромосомах, і частота мотиву SSR становила 1 на 4,64 кб. Було виявлено велику кількість тринуклеотидних повторів, а частота тетра- і пентануклеотидних повторів у сої була найменшою. Пошук онтології гену продемонстрував різні ролі ФТ, що містили SSR, у сої. Вісім ізолятів сої було проаналізовано щодо двадцяти ідентифікованих кандидатів на генну диверсифікацію SSR, було проведено аналіз основних координат і побудовано дерево незваженим методом приєднання сусідів на основі генної відмінності. Наші результати слугуватимуть потенційним ресурсом функціональних маркерів для маркер-асистованої селекції та геномної характеристики сої.

Ключові слова: фактор транскрипції (ФТ), SSR, генний маркер, онтологія гену, генетична різноманітність.

REFERENCES

Bandelj D, Jakše J, Euphytica BJ, 2004 undefined (2004) Assessment of genetic variability of olive varieties by microsatellite and AFLP markers. Springer 136:

- 93–102. <https://doi.org/10.1023/B:EUPH.0000019552.42066.10>
- Beier S, Thiel T, Münch T et al (2017) MISA-web: A web server for microsatellite prediction. *Bioinformatics* 33:2583–2585. <https://doi.org/10.1093/BIOINFORMATICS/BTX198>
- Bosamia TC, Mishra GP, Thankappan R, Dobaria JR (2015) Novel and Stress Relevant EST Derived SSR Markers Developed and Validated in Peanut. *PLoS One* 10:e0129127. <https://doi.org/10.1371/JOURNAL.PONE.0129127>
- Conesa A, Götz S, García-Gymez J et al Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *academic.oup.com*
- Fredslund J, Madsen LH, Hougaard BK, et al. (2006) A general pipeline for the development of anchor markers for comparative genomics in plants. *BMC Genomics* 7. <https://doi.org/10.1186/1471-2164-7-207>
- Gepts P, Beavis W, Brummer E, Shoemaker R (2005) Legumes as a model plant family. *Genomics for food and feed report of the cross-legume advances through genomics conference*
- Hisano H, Sato S, Isobe S, et al. (2007) Characterization of the soybean genome using EST-derived microsatellite markers. *DNA Research* 14:271–281. <https://doi.org/10.1093/DNARES/DSM025>
- Jin J, Zhang H, Kong L, et al (2014) PlantTFDB 3.0: A portal for the functional and evolutionary study of plant transcription factors. *Nucl Acid Res* 42. <https://doi.org/10.1093/NAR/GKT1016>
- Kujur A, Bajaj D, Saxena M et al Functionally relevant microsatellite markers from chickpea transcription factor genes for efficient genotyping applications and trait association mapping. <http://academic.oup.com>
- Kujur A, Bajaj D, Saxena MS et al (2013) Functionally Relevant Microsatellite Markers From Chickpea Transcription Factor Genes for Efficient Genotyping Applications and Trait Association Mapping. *DNA Res* 20:355–374. <https://doi.org/10.1093/DNARES/DST015>
- Kumar Biswas M, Kumar Nath U, Howlader J et al (2018) Exploration and Exploitation of Novel SSR Markers for Candidate Transcription Factor Genes in Liliium Species. *mdpi.com* 9. <https://doi.org/10.3390/genes9020097>
- Liu N, Cheng F Yun, Guo X, Zhong Y (2021) Development and application of microsatellite markers within transcription factors in flare tree peony (*Paeonia rockii*) based on next-generation and single-molecule long-read RNA-seq. *J Integr Agric* 20:1832–1848. [https://doi.org/10.1016/S2095-3119\(20\)63402-5](https://doi.org/10.1016/S2095-3119(20)63402-5)
- Liu W, Jia X, Liu Z et al (2015) Development and Characterization of Transcription Factor Gene-Derived Microsatellite (TFGM) Markers in *Medicago truncatula* and Their Transferability in Leguminous and Non-Leguminous Species. *Molecules* 20(5): 8759–8771. <https://doi.org/10.3390/molecules20058759>
- Liu YL, Li YH, Zhou GA et al (2010) Development of soybean EST-SSR markers and their use to assess genetic diversity in the subgenus soja. *Agric Sci China* 9:1423–1429. [https://doi.org/10.1016/S1671-2927\(09\)60233-9](https://doi.org/10.1016/S1671-2927(09)60233-9)
- Parmar R, Seth R, Sharma RK (2022) Genome-wide identification and characterization of functionally relevant microsatellite markers from transcription factor genes of Tea (*Camellia sinensis* (L.) O. Kuntze). *Scientific Reports* 12:1–14. <https://doi.org/10.1038/s41598-021-03848-x>
- Saha D, Rana RS, Das S et al (2019) Genome-wide regulatory gene-derived SSRs reveal genetic differentiation and population structure in fiber flax genotypes. *J Appl Genet* 60:13–25. <https://doi.org/10.1007/S13353-018-0476-Z>
- Spitz F, genetics EF-N reviews, 2012 undefined (2012) Transcription factors: from enhancer binding to developmental control. *Nat Rev Genet* 13:613–626. <https://doi.org/10.1038/nrg3207>
- Stefanova P, Taseva M, Georgieva T et al (2013) A modified CTAB method for DNA extraction from soybean and meat products. *Biotechnol Biotechnol* 27(3):3803–3810. <https://doi.org/10.5504/BBEQ.2013.0026>
- Varshney RK, Thiel T, Stein N (2002) In silico analysis on frequency and distribution of microsatellites in ESTs of some cereal species. *Cell Mol Biol Letters* 7(2A):537–546
- Ye J, Fang L, Zheng H, et al (2006) WEGO: A web tool for plotting GO annotations. *Nucl Acid Res* 34(Web Server issue):W293–297. <https://doi.org/10.1093/NAR/GKL031>

Received July 17, 2023

Received October 19, 2023

Accepted September 18, 2024