OPEN ACCESS **Science International**

ISSN 2305-1884
DOI: 10.17311/sciintl.2017.127.132

Research Article
Role of Expressed Sequence Tags in Cotton Improvement

[1]Viralkumar B Mandaliya and [2]Vrinda S Thaker

[1]Gujarat National Law University, Attalika Avenue, Knowledge Corridor, Koba, 382007 Gandhinagar, Gujarat, India
[2]UGC-CAS-Department of Biosciences, Saurashtra University, 360 005 Rajkot, Gujarat, India

Abstract
Cotton is a very economically important crop and cotton scientist, breeders and farmers are in a quest to improve the cotton quality. Genetic marker system will aid a tool to improve cotton crop. Expressed sequence tags (ESTs) are short DNA sequences reverse-transcribed from the cellular mRNA population. The ESTs in cotton had used to study comparative genetic mapping, to study genes involved in cotton fiber development and to develop various PCR-based molecular markers. The highest number of ESTs was reported in *Gossypium hirsutum* (*G. hirsutum*) i.e., 337,811 and lowest in *Gossypium hirsutum* (*G. herbaceum*) i.e., 247. Cotton ESTs in future by locating and detecting agronomic important genes will lead to open up new rise in cotton genomics and transcriptomics research.

Key words: Bio-informatics tools, expressed sequence tags, *Gossypium*, molecular markers, plant breeding

Citation: Viralkumar B Mandaliya and Vrinda S Thaker, 2017. Role of expressed sequence tags in cotton improvement. Sci. Int., 5: 127-132.

Corresponding Author: Gujarat National Law University, Attalika Avenue, Knowledge Corridor, Koba, 382007 Gandhinagar, Gujarat, India
Tel: +91 9510585801

Competing Interest: The authors have declared that no competing interest exists.

Data Availability: All relevant data are within the paper and its supporting information files.

INTRODUCTION

Cotton belongs to the genus *Gossypium* of the family *Malvaceae*[1] and exhibits a very important natural textile fiber source and cotton seed is a significant food source for humans and livestock[2]. A mankind, in his quest to improve cotton quality[3], is looks for variant forms i.e., varieties and hybrids. Plant breeders and farmers have created new varieties through traditional plant breeding. About 150 varieties and hybrids have been released during the last 50 years. Out of these about 30-40 are under large scale cultivation although about 20 varieties and hybrids account for more than 50% of production[4].

In traditional plant breeding crosses between plants are performed. Sexual crossing of such nature are done in an uncontrolled manner and this often leads to a random combination of genes which results in new traits, some of which may be undesirable. Selection and careful evaluation of the offspring is therefore necessary. Traditional plant breeding has gone through many phases, from the era of cross pollination between varieties of the same species to hybridization between different species. However, traditional plant breeding is costly and time consuming and moreover, the selection and evaluation of the new varieties can take several years to achieve[5]. The advancement in DNA technology had leads to a new area of modern plant biotechnology. The application of DNA technology in agricultural research has progressed rapidly over the last 20 years[6].

Nowadays, the cotton scientists have been emphasized on the utilization of public and private sector made cotton hybrids[7]. Numbers of studies have demonstrated that plants exhibiting great phenotypic and genomic variability. Cotton genus (*Gossypium* L.) includes approximately 45 diploid species (2n = 2x = 26) differentiated cytogenetically into 8 genome groups (A-G and K) and 5 allotetraploid species (2n = 4x = 52)[8]. In order to exploit this diversity an efficient molecular marker system is required[9]. In plant breeding the development of molecular marker systems facilitated the selection and evaluation process greatly. These molecular tools have increased the speed and precision for achieving desired agronomic traits.

Restriction fragment length polymorphism (RFLPs) belongs to the first generation of hybridization-based markers that developed in humans in the 1980s[10] and thereafter used in plant biotechnology[11]. The RFLP where size based variation in DNA fragments produced by a digestion of DNA[12]. The RFLPs have been used extensively to compare genomes in the cotton plants[1]. The advantages of RFLPs include detecting unlimited number of loci, codominant and the use of probes from other species. However, RFLPs are expensive, time consuming and labour intensive. The PCR based marker systems are more rapid and requires less plant material. The first of PCR based marker was known as rapid amplified polymorphic DNAs (RAPDs) and are produced by PCR using genomic DNA and arbitrary primers[13-15]. However, the results from RAPDs were not reproduced in different laboratories. The RFLPs and RAPDs have been used to map or tag agronomically important genes including resistance genes against viruses, bacteria, fungi, nematodes and insects[16,17].

Amplified fragment length polymorphism (AFLPs) combines both PCR and RFLP and it is generated by digestion of PCR amplified fragments using restriction enzymes[18]. For example, AFLPs have been used to assess the levels of genetic diversity within and between cotton hybrids[1]. The AFLPs are highly reproducible and this enables rapid generation and high frequency of identifiable AFLPs, making it an attractive technique for identifying polymorphisms and for determining linkages by analyzing individuals from a segregating population[17].

Another class of molecular markers which depends on the availability of short oligonucleotide repeat sequences in the genome of plants is the simple sequence repeat (SSR) polymorphism or microsatellites[19,20]. The SSR markers are fairly cheap and no sequence information is required for their detection. The SSR gives good polymorphism as well as requiring only a small quantity of DNA to start with. However, similar to RAPDs the major problems encountered with SSR that its reproducibility in different laboratories were low.

Currently plant biologists are exploiting the use of expressed sequence tags (ESTs) as markers in gene discovery research[5]. For example, a recently described set of ESTs from cotton fiber[21] provides a valuable new resource for developing PCR-based DNA markers for fiber genes.

WHAT ARE ESTs?

The ESTs are short DNA sequences corresponding to a fragment of a complimentary DNA (cDNA) molecule and which may be expressed in a cell at a particular given time. The ESTs are currently used as a fast and efficient method of profiling genes expressed in various tissues, cell types or developmental stages[22].

The concept of using cDNAs as a route to expedited gene discovery was first demonstrated in the early 1980s[23]. In 1990, Sydney Brenner proposed that an obvious method for characterizing the 'important' part of the human genome would involve looking at messengers from the expressed genes-thus advocating the application of high-throughput

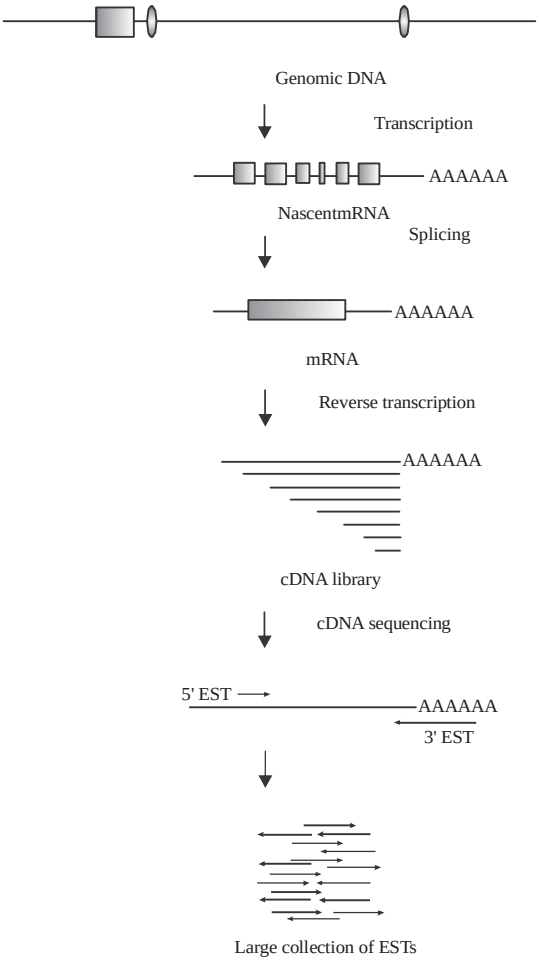

Fig. 1: EST sequencing

methods for transcriptome sampling. Mark Adams first used the term EST in relation to gene discovery and the human genome project in 1991[23].

ESTs are typically unedited, automatically processed, single-read sequences produced from cDNAs (small DNA molecules reverse-transcribed from the cellular mRNA population). Gene discovery via ESTs is comprised of four steps which include (i) The construction of cDNA libraries and single-pass sequencing of (randomly) selected clones, (ii) EST quality check the removal of vector and low quality sequences, (iii) The alignment of ESTs to identify the number of represented genes and (iv) The annotation of these genes or the partial sequences which are available thereof Fig. 1 summarizes EST sequencing.

EST sequencing initially favored the 5′ end of directionally cloned cDNAs because the 5′ sequences are likely to contain more protein coding sequence than the 3′ ends, which often contain significant untranslated regions (UTRs). Improvements in the techniques for cDNA preparation and the arrival of capillary-based sequencing have driven the evolution of high-throughput sequencing for ESTs and especially plant ESTs, while the 3′ end of the cDNA clone is often preferred because it is likely to offer more unique sequence (in many cases, the UTR) and can be used to distinguish between gene paralogues. The EST sequencing strategies in which both ends of the cDNA are sequenced are also becoming widespread. Thus, EST is a sequence representative of the corresponding cDNA clone and can be used for its characterizing with various bioinformatics tools and softwares[24].

Needles in the EST haystack exploited using bioinformatics tools and software packages: Some specific plant EST databases with their websites were enlisted in Table 1. The majority of databases are generally contains ESTs from various plants. Cotton Gen is specific for cotton crop. There are several bioinformatics tools and software packages that useful for EST annotation and protein sequence prediction, database creation, to facilitates the analysis of large volume of ESTs or to predict the transcript from ESTs, etc. are enlisted into Table 2.

Udall *et al*.[31] reported approximately 185,000 *Gossypium* EST sequences comprising >94,800,000 nucleotides were amassed from 30 cDNA libraries constructed from a variety of tissues and organs under a range of conditions, including drought stress and pathogen challenges.

The number of cotton ESTs available in NCBI dbEST database were comparatively complied in Table 3 (accessed on primariliy on 20th January, 2014 and then on 17th February, 2017). The highest number of ESTs was reported in *G. hirsutum* and lowest in *G. herbaceum*. There were no change in records observed in G. *raimondii* and *G. herbaceum* i.e., 63,577 and 247, respectively in 2014 and 2017. The number of ESTs were computationally aligned and trimmed to remove unessential sequences from ESTs because vector and low-quality sequences need to be removed from the raw sequence data as well as bacterial sequences or other contamination should be removed[32].

APPLICATIONS OF COTTON ESTS

Development of molecular markers: ESTs allow the efficient development of highly valuable molecular markers, because genes often represent single or low-copy sequences. Often EST-based RFLP markers allow comparative mapping across different species. The ESTs also allow a computational approach to the development of SSR (simple sequence repeat) and SNP (single nucleotide polymorphism) markers[24]. The

Table 1: Some specific plant EST databases

Plant EST database	Website
Plant GDB	http://www.plantgdb.org
dbEST	http://www.ncbi.nlm.nih.gov/dbEST/index.html
NCBI unigenes	http://www.ncbi.nlm.nih.gov/UniGene
CottonGen	https://www.cottongen.org
B-EST barley database	http://:pgrc.ipk-gatersleben.de/est/est/login
REDB (rice EST database)	http://www.ncpgr.cn
ESTree db	http://www.itb.cnr.it/estree
The Crop EST database (CR-EST)	http://apex.ipk-gatersleben.de/apex/f?p=116:1

Table 2: Bioinformatics tools and software packages to exploit ESTs

Names	Usage
EST-PAC[25]	EST annotation and protein sequence prediction
GO-Diff[26]	Mining functional differentiation between EST-based transcriptomes
EST2uni[27]	Automated EST analysis and database creation
JUICE[28]	Facilitates the analysis of large volumes of information in an EST project
Pairagon+N-SCAN_EST[29]	To predict complete transcripts by extending or merging EST alignments
ESTpass[30]	A web-based server for processing and annotating

Table 3: Number of ESTs in cotton (*Gossypium* sp.)

	Hirsutum	*Raimondii*	*Arboreum*	*Barbadense*	*Herbaceum*
Taxa id no. in NCBI	txid3635	txid29730	txid29729	txid3634	txid76056
Number of ESTs available noted on 20th January, 2014	270,519	63,577	40,910	1,023	247
Number of ESTs available noted on 17th February, 2017	337,811	63,577	64,798	39,115	247

available sequence information allows the design of primer pairs, which can be used to screen cultivars of interest for length polymorphisms.

Chee *et al.*[33] had developed PCR-based markers from known function EST sequences for the cultivated tetraploid cotton species *Gossypium barbadense* and *Gossypium hirsutum*. Here, outcomes suggested that digestion of PCR-amplified sequences offers one means by which cotton genes can be mapped to their chromosomal locations more quickly and economically than by RFLP analysis. Zhang *et al.*[34] and Qureshi *et al.*[35] reported that cotton EST-SSR markers were derived from *Gossypium arboreum* and *Gossypium hirsutum* and *Gossypium barbadense* and their study effectively proved that EST-SSRs are valuable for genetic diversity analysis and genetic mapping.

High-throughput transcript profiling: ESTs also provide the main resource for the construction of cDNA arrays in plants. The construction and use of such EST arrays for high-throughput transcript profiling can be divided in four general steps: (i) Identification of a non-redundant set of cDNA clones, (ii) Synthesis and deposition of hybridization targets on an appropriate surface, (iii) Preparation of mRNA from the tissue of interest, labelling of the hybridization probe and the hybridization to the array and finally (iv) Data acquisition and evaluation. Arpat *et al.*[21] reported that genetic characterization of rapid cell elongation in cotton fibers and approximately 14,000 unique genes were assembled from 46,603 expressed sequence tags (ESTs) from developmentally staged fiber cDNAs of a cultivated diploid species (*Gossypium arboreum* L.).

Biological interpretation of expression data: Biological experimentation are carried out *in vitro*, *ex vivo*, in lab or in field conditions[36-49]. Expression data are expected to yield insights into regulatory processes during plant development and stimulus response. To reach that goal, it is necessary to compare the pre-processed array data with known models of metabolic and regulatory networks as depicted in databases or the general literature and to confirm or reject specific hypotheses. Many successful examples have been provided already, for example, *Gossypium hirsutum* derived EST-SSRs can be used in identification of quantitative trait loci (QTLs) and comparative genomics studies of diploid and allotetraploid cotton[50].

CONCLUSION

The application of molecular markers in cotton plants has tremendous utility to cotton scientists for improving the plant. EST based gene discovery will leads to advance our understanding of the complexity of biological and cellular processes that are required for cotton fiber development and important genes. The EST based markers like SNP and EST-SSR that able to locate and detect agronomic traits and obtain a transcript map, which can be directly compared with earlier

detected quantitative trait loci. From this, it could be foresee that cotton ESTs in future will provide the new horizons in cotton genomics and transcriptomics research.

SIGNIFICANCE STATEMENTS

This study will help cotton scientist, breeders and farmers to improve the cotton quality. In this study, genetic marker system analysed to improve cotton crop. The ESTs in cotton had used to study comparative genetic mapping which helps in future.

ACKNOWLEDGMENT

The authors would like to thank the Government of Gujarat for financial support in form of Centre for Advanced Studies in Plant Biotechnology and Genetic Engineering (CPBGE) programme and funding research project GG-R1/RP/VM-1.

REFERENCES

1. Zhang, H.B., Y. Li, B. Wang and P.W. Chee, 2008. Recent advances in cotton genomics. Int. J. Plant Genom. 10.1155/2008/742304.
2. Sunilkumar, G., L.M. Campbell, L. Puckhaber, R.D. Stipanovic and K.S. Rathore, 2006. Engineering cottonseed for use in human nutrition by tissue-specific reduction of toxic gossypol. Proc. Natl. Acad. Sci. USA., 103: 18054-18059.
3. Bhatt, K.R. and V.S. Thaker, 2008. Relationship between gibberellic acid and water amount in the cotton seed. Russian J. Plant Physiol., 55: 808-813.
4. Murugkar, M., B. Ramaswami and M. Shelar, 2006. Liberalization, biotechnology and the private seed sector: The Case of India's cotton seed market (No. 06-05). Indian Statistical Institute, New Delhi, India.
5. Ayeh, K.O., 2008. Expressed Sequence Tags (ESTs) and Single Nucleotide Polymorphisms (SNPs): Emerging molecular marker tools for improving agronomic traits in plant biotechnology. Afr. J. Biotechnol., 7: 331-341.
6. Zidani, S., A. Ferchichi and M. Chaieb, 2005. Genomic DNA extraction method from pearl millet (*Pennisetum glaucum*) leaves. Afr. J. Biotechnol., 4: 862-866.
7. Hussein, E.H.A., M.H.A. Osman, M.A. Hussein and S.S. Adawy, 2007. Molecular characterization of cotton genotypes using PCR-based markers. J. Applied Sci. Res., 3: 1156-1169.
8. Guo, W., C. Cai, C. Wang, L. Zhao, L. Wang and T. Zhang, 2008. A preliminary analysis of genome structure and composition in *Gossypium hirsutum*. BMC Genomics, Vol. 9. 10.1186/1471-2164-9-314.
9. Mace, E.S., L. Xia, D.R., Jordan, K. Halloran and D.K. Parh *et al*., 2008. DArT markers: Diversity analyses and mapping in *Sorghum bicolor*. BMC Genomics, Vol. 9. 10.1186/1471-2164-9-26.
10. Botstein, D., R.L. White, M. Skolnick and R.W. Davis, 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am. J. Hum. Genet., 32: 314-331.
11. Weber, D. and T. Helentjaris, 1989. Mapping RFLP loci in maize using BA translocations. Genetics, 121: 583-590.
12. De Martinville, B., A.R. Wyman, R. White and U. Francke, 1982. Assignment of first random Restriction Fragment Length Polymorphism (RFLP) locus ((D14S1) to a region of human chromosome 14. Am. J. Hum. Genet., 34: 216-226.
13. Jacobson, A. and M. Hedren, 2007. Phylogenetic relationships in *Alisma* (*Alismataceae*) based on RAPDs and sequence data from ITS and *trn*L. Plant Systemat. Evol., 265: 27-44.
14. Welsh, J. and M. McClelland, 1990. Fingerprinting genomes using PCR with arbitrary primers. Nucleic Acids Res., 18: 7213-7218.
15. Williams, J.G.K., A.R. Kubelik, K.J. Livak, J.A. Rafalski and S.V. Tingey, 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res., 18: 6531-6535.
16. Mohan, M., S. Nair, J.S. Bentur, U.P. Rao and J. Bennett, 1994. RFLP and RAPD mapping of the rice *Gm*2 gene that confers resistance to biotype 1 of gall midge (*Orseolia oryzae*). Theor. Applied Genet., 87: 782-788.
17. Mohan, M., S. Nair, A. Bhagwat, T.G. Krishna, Y. Masohiro, C.R. Bhatia and T. Sasaki, 1997. Genome mapping, molecular markers and marker-assisted selection in crop plants. Mol. Breed., 3: 87-103.
18. Vos, P., R. Hogers, M. Bleeker, M. Reijans and T. van de Lee *et al*., 1995. AFLP: A new technique for DNA fingerprinting. Nucleic Acids Res., 23: 4407-4414.
19. Hearne, C.M., S. Ghosh and J.A. Todd, 1992. Microsatellites for linkage analysis of genetic traits. Trends Genet., 8: 288-294.
20. Tautz, D. and M. Renz, 1984. Simple sequences are ubiquitous repetitive components of eukaryotic genomes. Nucleic Acids Res., 12: 4127-4138.
21. Arpat, A.B., M. Waugh, J.P. Sullivan, M. Gonzales, D. Frisch and D. Main *et al*., 2004. Functional genomics of cell elongation in developing cotton fibers. Plant Mol. Biol., 54: 911-929.
22. Adams, M.D., J.M. Kelley, J.D. Gocayne, M. Dubnick and M.H. Polymeropoulos *et al*., 1991. Complementary DNA sequencing: Expressed sequence tags and human genome project. Science, 252: 1651-1656.
23. Rudd, S., 2003. Expressed sequence tags: Alternative or complement to whole genome sequences? Trends Plant Sci., 8: 321-329.
24. Kisiel, A. and J. Podkowiski, 2005. Expressed sequence tags and their application for plant research. Acta Physiol. Planta., 27: 157-161.

25. Strahm, Y., D. Powell and C. Lefevre, 2006. EST-PAC a web package for EST annotation and protein sequence prediction. Sour. Code Biol. Med., Vol. 1. 10.1186/1751-0473-1-2.
26. Chen, Z., W. Wang, X.B. Ling, J.J. Liu and L. Chen, 2006. GO-Diff: Mining functional differentiation between EST-based transcriptomes. BMC Bioinfo., Vol. 7. 10.1186/1471-2105-7-72.
27. Forment, J., F. Gilabert, A. Robles, V. Conejero, F. Nuez and J.M. Blanca, 2008. EST2uni: An open, parallel tool for automated EST analysis and database creation, with a data mining web interface and microarray expression data integration. BMC Bioinfo., Vol. 9. 10.1186/1471-2105-9-5.
28. Latorre, M., H. Silva, J. Saba, C. Guziolowski and P. Vizoso *et al*., 2006. JUICE: A data management system that facilitates the analysis of large volumes of information in an EST project workflow. BMC Bioinfo., Vol. 7. 10.1186/1471-2105-7-513.
29. Arumugam, M., C. Wei, R.H. Brown and M.R. Brent, 2006. Pairagon+N-SCAN_EST: A model-based gene annotation pipeline. Genome Biol., Vol. 7. 10.1186/gb-2006-7-s1-s5.
30. Lee, B., T. Hong, S.J. Byun, T. Woo and Y.J. Choi, 2007. ESTpass: A web-based server for processing and annotating Expressed Sequence Tag (EST) sequences. Nucleic Acids Res., 35: W159-W162.
31. Udall, J.A., J.M. Swanson, K. Haller, R.A. Rapp and M.E. Sparks *et al*., 2006. A Global Assembly of Cotton ESTs. Cold Spring Harbor Laboratory Press, USA.
32. Sreenivasulu, N., P.K. Kishor, R.K. Varshney and L. Altschmied, 2002. Mining functional information from cereal genomes: The utility of expressed sequence tags. Curr. Sci., 83: 965-973.
33. Chee, P. W., J. Rong, D. Williams-Coplin, S.R. Schulze and A.H. Paterson, 2004. EST derived PCR-based markers for functional gene homologues in cotton. Genome, 47: 449-462.
34. Zhang, Y., Z. Lin, W. Li, L. Tu, Y. Nie and X. Zhang, 2007. Studies of new EST-SSRs derived from *Gossypium barbadense*. Chinese Sci. Bull., 52: 2522-2531.
35. Qureshi, S.N., S. Saha, R.V. Kantety and J.N. Jenkins, 2004. EST-SSR: A new class of genetic markers in cotton. J. Cotton Sci., 8: 1475-1479.
36. Patel, K.G., V.B. Mandaliya, G.P. Mishra, J.R. Dobaria and T. Radhakrishnan, 2016. Transgenic peanut overexpressing *mtlD* gene confers enhanced salinity stress tolerance via mannitol accumulation and differential antioxidative responses. Acta Physiol. Plant., Vol. 38.
37. Mandaliya, V.B. and V.S. Thaker, 2016. Molecular markers in male sterility: Step towards crop improvement. Int. J. Mol. Biol. Biochem., 4: 1-12.
38. Jhala, V.M., V.B. Mandaliya and V.S. Thaker, 2015. Simple and efficient protocol for RNA and DNA extraction from rice (*Oryza sativa* L.) for downstream applications. Int. Res. J. Biol. Sci., 4: 62-67.
39. Mandaliya, V.B., V.M. Jhala and V.S. Thaker, 2015. Molecular characterization of genetic male sterile lines using RAPD, ISSR and SRAP markers. J. Cotton Res. Dev., 29: 1-6.
40. Chariya, L.D., V.B. Mandaliya and V.S. Thaker, 2013. Conversion of monomorphic band into polymorphic pattern using nucleotide sequencing data in *Musa* varieties. Genet. Plant Physiol., 3: 77-89.
41. Mandaliya, V.B., R.V. Pandya and V.S. Thaker, 2011. CSNP: A tool for harnessing the genetic potential of cotton. Cotton Res. J., 2: 1-14.
42. Bhatt, K.D., S.K. Girnari, V.B. Mandaliya, L.D. Chariya and V.S. Thaker, 2011. Use of RAPD marker to confirm mutation in morphological variants on Neem tree. Electron. J. Plant Breed., 2: 473-478.
43. Nakum, N.M., V.B. Mandaliya, R.V. Pandya and V.S. Thaker, 2011. A simple and efficient protocol for high quality of DNA from *Vitis quandrangularis* L. Electron. J. Plant Breed., 2: 87-95.
44. Girnari, S.K., K.D. Bhatt, V.B. Mandaliya, L.D. Chariya and V.S. Thaker, 2011. RAPD markers studies in some selected medicinally important plants. J. Scient. Agric. Res., 72: 5-14.
45. Mandaliya, V.B., R.V. Pandya and V.S. Thaker, 2010. SNP: A trend in genetics and genome analyses of plants. Gen. Applied Plant Physiol., 36: 159-166.
46. Mandaliya, V.B., R.V. Pandya and V.S. Thaker, 2010. Comparison of cotton DNA extraction method for high yield and quality from various cotton tissue. J. Cotton Res. Dev., 24: 9-12.
47. Mandaliya, V.B., R.V. Pandya and V.S. Thaker, 2010. Genetic diversity analysis of cotton (*Gossypium*) hybrids using RAPD markers. J. Cotton Res. Dev., 24: 127-132.
48. Ramavat, J.M., R.K. Ramanuj, V.B. Mandaliya, L.D. Chariya, Y.M. Bapodariya and V.S. Thaker, 2010. A rapid and realible protocol for DNA extraction from *Catharanthus roseus*. J. Scient. Agric. Res., 71: 47-52.
49. Ramanuj, R.K., J.M. Ramavat, Y.M. Bapodariya, L.D. Chariya, V.B. Mandaliya and V.S. Thaker, 2010. Rapid *in vitro* propagation of *Ocimum sanctum* L. through multiple shoot induction. J. Scient. Agric. Res., 71: 5-11.
50. Han, Z., C. Wang, X. Song, W. Guo and J. Gou *et al*., 2006. Characteristics, development and mapping of *Gossypium hirsutum* derived EST-SSRs in allotetraploid cotton. Theor. Applied Genet., 112: 430-439.