RESEARCH ARTICLE

SABRAO Journal
of Breeding and Genetics
46 (1) 56-66, 2014

DEVELOPMENT OF INTRON-FLANKING EST-SPECIFIC PRIMERS FROM DROUGHT RESPONSIVE EXPRESSED SEQUENCE TAGS (ESTs) IN SAFFLOWER

M.Y. DUDHE[1], O. SUDHAKARBABU[2] and C. SUDHAKAR[2]*

[1]Directorate of Oilseeds Research, Rajendranagar, Hyderabad -500030, India
[2]Sri Krishnadevaraya University, Anantapur-515 003, India
*Corresponding author's email: chintasudhakar@yahoo.com

SUMMARY

The available drought responsive ESTs (673) sequences of *Carthamus tinctorius* were aligned with the genomic sequences of two model plants *Arabidopsis thaliana, Medicago truncatula*, two cereal crop plants rice, sorghum and one oil yielding crop soybean for the development of intron- flanking EST-primers. We have developed total of 50 orthologous conserved primers in our study. We have identified highest number of conserved primers from soybean after *Arabidopsis* which reflects conserved evolution of drought responsive conserved sequences or primers in safflower and soybean. There may be similar patterns of synteny across the safflower, soybean genomes at least for drought responsive genes. We have successfully obtained 52% of e-PCR products (> 300 bp) in *Arabidopsis*, 85% in rice, 80% in sorghum, 92% in soybean and 85% in *Medicago.* Based on our results, we recommend that for *in silico* ILP markers identification in safflower, soybean genomic resources can be used rather than model crop plant. Out of 50 annotated sequences containing ILP markers, only 23 sequences were identified for expressed proteins and can be considered as safflower stress-responsive proteins/genes. The present study highlights the importance of investigating the genes which interact with ILP markers and suggests following the behavior of marker-related genes at protein level. The ILP markers developed from our study will be of particular use for molecular breeding, screening and development of drought tolerant genotype in safflower.

Keywords: *Carthamus tinctorius*, drought stress, ESTs, ILP marker, *in silico*, synteny

Manuscript received: May 29, 2013; Decision on manuscript: February 2, 2014; Manuscript accepted: April 1, 2014.

Communicating Editor: Akshaya Biswal

INTRODUCTION

Safflower (*Carthamus tinctorius* L.) in India is cultivated mainly for the high quality edible oil extracted from the seeds. In some part of the country it has been grown for centuries for the orange-red dye (carthamin) extracted from its brilliantly colored flowers (Mukta *et al.*, 2012) and for its oil rich in polyunsaturated fatty acids (linoleic acid, 78%). Many types of molecular markers have been developed since 1980, such as restriction length polymorphism (RFLP), (Botstein *et al.*, 1980), random amplified polymorphic DNA (RAPD), (Williams *et al.*, 1990), amplified fragment length polymorphism (AFLP), (Vos *et al.*, 1995), simple sequence repeat polymorphism (SSR), (Becker and Heun,1995), intron length polymorphism (ILP), **(**Choi *et al*., 2004) etc. Molecular markers obtained from the non-functional region of the genome are mostly preferred for physical phylogenetic analysis and inferring taxon

relatedness. Particularly the introns (originally thought of as junk DNA) that are widespread and abundant in eukaryotic genomes (Hawkins *et al.*, 1988) are now-a-days commonly used as sources of DNA polymorphism. Since introns have no functional significance (although may influence the level of the gene expressions), they are common variable than coding sequences (Wang *et al*., 1988). Introns between genotypes show polymorphism due to differences in sequences, number of repetitions of simple sequences in length. But it is easy to detect intron length polymorphism (ILP) between the populations of genotypes using conventional PCR. To amplify introns by PCR, primers are generally designed based on the flanking exons-called as exon-primed intron-crossing PCR (EPIC-PCR) (Palumbi *et al.*, 1995). The advantage of this technique is that the primers designed based on the genome sequences of model plant like *Arabidopsis* could also be used to amplify introns from the non-model plants.

Among the major oilseed crops safflower is relatively more adaptable to drought and saline conditions (Weiss *et al.*, 2000; Mohammad *et al*., 2011) however, with lower crop yield. Moisture deficit stress is one of the major abiotic factors for severe yield loss in safflower. There has been only limited information with respect to abiotic stresses and genetic control of tolerance in safflower (Singh *et al.*, 2007). As for the plants with unknown genomic sequences, developing ILP markers is limited by lack of genomic sequences (Li and Brutnell, 2011). For the development of ILPs we have used our laboratory generated drought responsive 673 Expressed Sequence Tags (ESTs) (Thippeswamy *et al.*, 2012). To our knowledge, there is no published information on the development of drought responsive ILPs marker in safflower. Also, the number and kind (either gene specific or sequence specific) markers developed are less in safflower compared to other oilseeds crop. By considering the above facts, the primary goal of the present study was to develop a set of ILPs markers from our already published drought responsive EST library generated from drought tolerant safflower cultivar A1. Second, to study distribution on chromosomes for the ESTs containing ILPs by using the whole genome sequences of rice, sorghum, soybean and *Medicago*. Third, to functionally annotate EST derived ILPs for their putative functions in drought ESTs library.

MATERIALS AND METHODS

Development of intron-flanking EST-primers

We have used two model plants *Arabidopsis thaliana*, *Medicago truncatula*, two cereal crop plants Rice, Sorghum and one oil yielding crop soybean for the development of intron- flanking EST-primers. The genome sequences (with intron and UTRs) of *Arabidopsis*, rice, sorghum, soybean and *Medicago* were downloaded from reference databases to find out the homolog sequences corresponding to the drought ESTs sequences of *Carthamaus tinctorius* (Safflower) using BLASTN search. We used a high e-value (10^{-10}) for the BLASTN to remove prologues. Intron–flanking EST-primers were designed based on the "exon/exon" junction site information, inferred from the pair-wise alignments between the ESTs of the safflower and their homolog exonic sequences from all the crops. As per this pair-wise alignment strategy the predicted "exon/exon" junctions from safflower ESTs are essentially corresponded to "exon/intron" junctions (splice sites) in all genomic sequences. The forward and reverse primers were designed using Primer3 (Rozen *et al.*, 2000) with default setting based on the EST sequences that flanks at least two exon/exon junctions. These primers were expected to span the predicted safflower intron regions in a PCR. Percentage of the identified primers was calculated as total number of primers identified to total number of sequence used for the identification of conserved primers expressed in percentage.

In silico PCR and mapping of ILPs markers

We tested the designed primers by electronic PCR (e-PCR). Further, we have validated our results by using online PCR (http://biocompute.bmi.ac.cn/lab/MFEprimer-2.0/) genomic sequences for *Arabidopsis* and rice, respectively. PCR conditions used as,

optimum annealing temperature 55 – 61°C with 60°C as the optimum, and a minimum GC content of 30%, with 50% being the optimum. The ILP marker containing EST sequences were BLAST-searched against the genome sequences at Phytozome (http://www.phytozome.net) (9.1 version) to find each ILP marker's chromosome position on rice, sorghum, soybean and *Medicago*. The default setting parameters were used for the entire BLAST search for each of the ILP marker containing EST sequences.

Assessment of functional relevance of EST sequences having ILP markers

Sequences containing ILP were physically searched and sorted out. Each sequence from drought responsive EST library containing ILP were used for similarity search using BLAST2GO (Conesa *et al.,* 2005) to identify their putative function and were run against the non-redundant (nr) protein database of the NCBI (http://www.ncbi.nlm.nih.gov/blast). The obtained hits were compiled (Conesa *et al.,* 2008). Sequences not showing any match were considered as unique to the safflower species.

RESULTS

In current study, EST-PCR primers were designed based on the predicted exon/intron junction sites in drought responsive ESTs of safflower. Using BLASTN, out of 673 droughts responsive EST sequences of safflower, we detected less than 1% homologs with *Arabidopsis*, rice, sorghum, soybean and *Medicago*, it may be due to fact that the drought responsive ESTs were limited. We have identified 16 conserved primers from *Arabidopsis*, 13 from soybean, 7 each from rice, sorghum and *Medicago* are presented in Table1. Percentage of the identified primers was highest for *Arabidopsis* followed by soybean. Base pair length of the sequences containing conserved primers was highest for *Arabidopsis* followed by soybean. Details of conserved primers identified from reference genome of *Arabidopsis*, rice, sorghum, soybean, *Medicago* and in *silico* PCR results are presented in Table 2. The product size ranged from 228-787 bp in case of *Arabidopsis*, 214-926 bp for rice, 267-1664bp for sorghum, 238-1329 bp for soybean and for *Medicago* 316-1276 bp. Another area of research to study distribution on chromosomes for the ESTs containing ILPs by using the whole genome sequences of rice, sorghum, soybean and *Medicago*. The ILP containing EST sequences were BLAST-searched against the genome sequences at Phytozome to find each ILPs marker's chromosome position on rice, sorghum, soybean and *Medicago* (Figures 1, 2, 3 and 4). For rice seven ILPs were localized on chromosome, 2, 3, 6, for sorghum, 1, 4 and 10, for soybean one to eight and for *Medicago* 1,5 and 8. It is interesting to note that drought responsive safflower EST containing conserved primers are same for rice and sorghum (Table 2) but the distributions on chromosomes of these identified ILPs are different in Figures 1 and 2.

A total of 50 ILP markers were identified from drought responsive ESTs in our study which were used for similarity search to identify their putative function. The ILP marker containing sequences were matched to proteins having distinct molecular functions such as binding, catalytic, different biological processes, and cellular and sub-cellular organization. Out of 50 identified sequences, only 23 sequences were identified for annotated proteins and can be considered as safflower stress-responsive proteins/gene are presented in Table 3. For rest of the sequences no information is available in non-redundant (nr) protein database and can be considered unique to safflower.

Table 1. Number of conserved primers identified and other attributes studied.

	Conserved primers identified	Percentage of the identified primers	Base pair length of the sequences containing conserved primers	Average length of the sequence in base pair
Arabidopsis	16	2.75	9175	573.43
Rice	7	1.04	4210	601.42
Sorghum	7	1.04	4210	601.42
Soybean	13	1.93	8478	652.15
Medicago	7	1.04	4809	687.0

Table 2. Conserved primers identified from reference genome of *Arabidopsis*, rice, sorghum, soybean, *Medicago* and *in silico* PCR results.

Sequence ID	Conserved primer identified from reference genome	Sequence	Length	tm	Product Size (bp)	Sequence Length (bp)
Arabidopsis						
GW584142	1(2969228 - 2969551)	F-CAGAAAGTGTGCCAGAGTTCA	21	59.07	236	329
		R-GAATATCTCCCACGGCATGA	20	60.83		
GW584142_1	2(4820384 - 4820676)	F-CAGAAAGTGTGCCAGAGTTCA	21	59.07	230	323
		R-GAATATCTCCCACGGCATGA	20	60.83		
GW584199	3 (10069929 - 10070266)	F-ATTCTTGGATGGTGGTGTCC	20	59.64	228	292
		R-TCGCGTAACATCTGAAGCAC	20	60.02		
GW584053	4(1361565 - 1361948)	F-TTGGTGGCGGTAAATCAACT	20	60.37	318	337
		R-CACCACGGATCTTCTTCGAT	20	60.07		
GW584053_1	5 (16871751 - 16872648)	F-TTGGTGGCGGTAAATCAACT	20	60.37	364	383
		R-CACCACGGATCTTCTTCGAT	20	60.07		
GW584256	6(2818418 - 2819058)	F-ACCAGCAAGATCCTCTCGAA	20	59.95	787	897
		R-TTCCGGAGTCTTTCCTTTGA	20	59.78		
GW584250	7(2824001 - 2824662)	F-CTTCCACGTTGCTGAGGTTT	20	60.29	392	640
		R-GCAATCACCTCGTCAATTCC	20	60.47		
GW584250_1	8 (6754088 - 6754919)	F-CTTCCACGTTGCTGAGGTTT	20	60.29	413	661
		R-GCAATCACCTCGTCAATTCC	20	60.47		
GT155311	9(6758937 - 6759333)	F-ATCCTGGATCAAGCCCTTCT	20	60.04		
		R-TTCCCTCAAAGCAAACATCC	20	60.05		
GT155311_1	10(2937736 - 2938205)	F-CTGGATCAAGCCCTTCTCAG	20	59.94	363	396
GT155311_1		R-CTTAAGCTCGGCACCAGAAG	20	60.15		
EL611934	11 (1604490 - 1605194)	F-CTGAAGCCTGGTGATTTGGT	20	60.11	324	469
		R-GGTGGACGGATCCCTAATTT	20	60.02		
EL611934_1	12(15394753 -	F-GAGTCAAGGCCATGGAGGT	19	60.06	575	704

	15395206)					
		R-TCTTTGGCCAGTTGAAATGC	20	61.15		
EL611906	13(15398455 - 15398878)	F-TGACTACCTTCTCCGCAACA	20	59.44	383	453
		R-AATCCGAATAAACGCACTCG	20	60.1		
EL611906_1	14(23920205 - 23920766)	F-TGACTACCTTCTCCGCAACA	20	59.44	353	423
		R-AATCCGAATAAACGCACTCG	20	60.1		
EL611884	15(23246704 - 23247071)	F-ATGCACAAGAGACGTGCAAC	20	59.91	296	561
		R-CGAACAGTCTTGTTGCTGGA	20	60.02		
EL611875	16(23246704 - 23247071)	F-GCTCATCAACTGCAGGAACA	20	59.99	284	367
		R-TGATGACTGAATCACCTCGAA	21	59.23		
Rice						
GW584250	1(5920104 - 5920576)	F-CTTCCACGTTGCTGAGGTTT	20	60.29	825	1082
		R-GCAATCACCTCGTCAATTCC	20	60.47		
GT155311	2(24564077 - 24564472)	F-GCTCGATGGATTTGAAGCAT	20	60.19	214	472
		R-CTTAAGCTCGGCACCAGAAG	20	60.15		
GT155311_1	3(32288239 - 32289012)	F-CTGGATCAAGCCCTTCTCAG	20	59.94	285	395
		R-CTTAAGCTCGGCACCAGAAG	20	60.15		
GT155488	4(19617563 - 19618428)	F-CCCTCGTCCCATAACTGTGA	20	60.91	655	773
		R-CACTCACCACCAGGTCCTTT	20	60		
GT155488_1	5 (3690655 - 3691593)	F-CCCTCATGTCAACCGTACTGA	21	60.96	811	865
		R-CACTCACCACCAGGTCCTTT	20	60		
EL611934	6(35123478 - 35124473)	F-CTGAAGCCTGGTGATTTGGT	20	60.11	410	937
		R-GGTGGACGGATCCCTAATTT	20	60.02		
EL611934_1	7 (35123478 - 35124473)	F-CTGAAGCCTGGTGATTTGGT	20	60.11	926	995
		R-TCTTTGGCCAGTTGAAATGC	20	61.15		
Sorghum						
GW584250	1 (7324804 - 7325295)	F-CTTCCACGTTGCTGAGGTTT	20	60.29	1364	1597
		R-GCATTGTCCTTGGCAATCTT	20	60.08		
GT155311	2(51844248 - 51844643)	F-CTGGATCAAGCCCTTCTCAG	20	59.94	267	490
		R-CTTAAGCTCGGCACCAGAAG	20	60.15		
GT155311_1	3(6455264 - 6455836)	F-CTGGATCAAGCCCTTCTCAG	20	59.94	276	395
		R-CTTAAGCTCGGCACCAGAAG	20	60.15		
GT155488	4(4880272 - 4881207)	F-CCCTCGTCCCATAACTGTGA	20	60.91	453	571
		R-CACTCACCACCAGGTCCTTT	20	60		
GT155428	5 (4509057 - 4510563)	F-CACAGTGCAAGGGTGAGAAA	20	59.87	762	935
		R-GGGCATTCCTTCATCGTTTA	20	59.9		
EL611934	6(66073094 - 66074104)	F-CTGAAGCCTGGTGATTTGGT	20	60.11	413	1505

		R-GGTGGACGGATCCCTAATTT	20	60.02		
EL611934_1	7(66073094 - 66074104)	F-CTGAAGCCTGGTGATTTGGT	20	60.11	947	1010
		R-TCTTTGGCCAGTTGAAATGC	20	61.15		
Soybean						
GW584142	1(5898854 - 5899257)	F-CAGAAAGTGTGCCAGAGTTCA	21	59.07	238	331
		R-GAATATCTCCCACGGCATGA	20	60.83		
GW584112	2(525242 - 525942)	F-TCCAGTCGCCAAACTCTTTC	20	60.38	333	403
		R-ACAAGACGGTCTGGCAGTTT	20	59.77		
GW584014	3(18271149 - 18272876)	F-GTGGCCAAAGCTTCATTGAG	20	60.78	654	700
		R-AAACAGCTGAACCAGCAACC	20	60.3		
GT155311	4 (186109 - 187526)	F-ATCCTGGATCAAGCCCTTCT	20	60.04	1329	1725
		R-TTCCCTCAAAGCAAACATCC	20	60.05		
GT155311_1	5(9190334 - 9191026)	F-ATCCTGGATCAAGCCCTTCT	20	60.04	1287	1416
		R-TTCCCTCAAAGCAAACATCC	20	60.05		
GT155503	6(48496994 - 48497724)	F-AGCCGAGCCTCTTCTTCTCT	20	59.86	486	692
		R-GAGCTCCCTTTCCTGTGATG	20	59.8		
GT155503_1	7(10813685 - 10814778)	F-AGCCGAGCCTCTTCTTCTCT	20	59.86	526	730
		R-GAGCTCCCTTTCCTGTGATG	20	59.8		
GT155488	8 (6946166 - 6947281)	F-CCCTCGTCCCATAACTGTGA	20	60.91	978	1093
		R-CACTCACCACCAGGTCCTTT	20	60		
GT155488_1	9(433555 - 433983)	F-CCCTCGTCCCATAACTGTGA	20	60.91	1000	1115
		R-CACTCACCACCAGGTCCTTT	20	60		
GT155488_2	10(45558189 - 45558664)	F-TTTCTGAGGATGGATTTGTGAG	22	59.17	345	428
		R-CACTCACCACCAGGTCCTTT	20	60		
EL611934	11(43470311 - 43470929)	F-CTGAAGCCTGGTGATTTGGT	20	60.11	337	475
		R-GGTGGACGGATCCCTAATTT	20	60.02		
EL611884	12(37092636 - 37094450)	F-ATGCACAAGAGACGTGCAAC	20	59.91	354	618
		R-CGAACAGTCTTGTTGCTGGA	20	60.02		
EL611909	13(37092636 - 37094450)	F-GATGGCCTCTTTGACTGAGC	20	59.96	1045	1813
		R-CACAAACACCCGTTTCAACA	20	60.44		
Medicago						
GW584165	1(16217942 - 16219240)	F-TGTACAATCGCCGAAACAAG	20	59.73	316	411
		R-TTACGCATGCATTTCTCCAA	20	60.22		
GW584053	2 (8208607 - 8209844)	F-TTGGTGGCGGTAAATCAACT	20	60.37	1276	1298

		R-CACCACGGATCTTCTTCGAT	20	60.07		
GW584256	3 (12391891 - 12392282)	F-ACCAGCAAGATCCTCTCGAA	20	59.95	1218	1237
		R-CCCTACGAACCTCCTCTTCC	20	60.07		
GT155399	4 (12399698 - 12400092)	F-GTGGAGCGATGGGTATTCTG	20	60.48	371	391
		R-ATCCATGTCAGCCTTGTTCC	20	59.93		
GT155399_1	5 (17554171 - 17554634)	F-GTGGAGCGATGGGTATTCTG	20	60.48	374	394
		R-ATCCATGTCAGCCTTGTTCC	20	59.93		
EL611934	6 (2189918 - 2191108)	F-CTGAAGCCTGGTGATTTGGT	20	60.11	328	463
		R-GGTGGACGGATCCCTAATTT	20	60.02		
EL611884	7(2189918 - 2191108)	F-ATGCACAAGAGACGTGCAAC	20	59.91	972	1190
		R-GCCACTTCACATTTCCACCT	20	59.97		

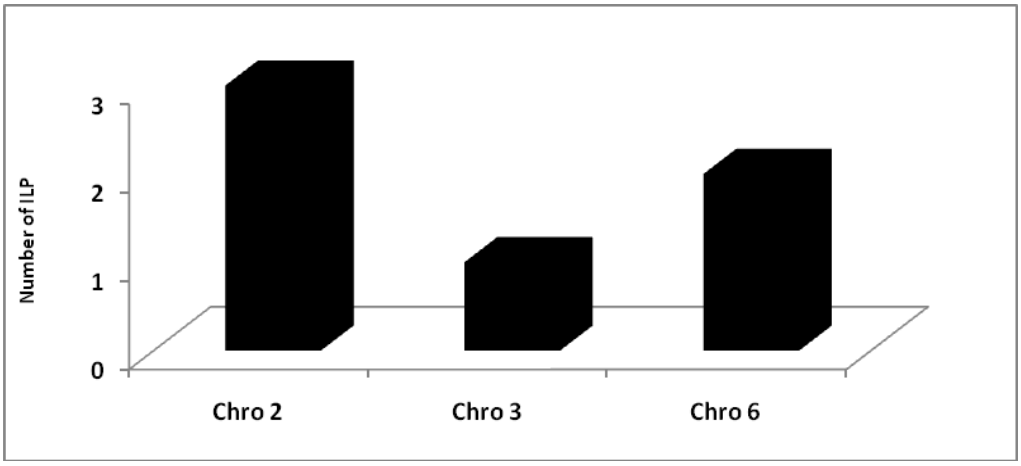

Figure 1. Distribution pattern of seven identified ILP markers on rice chromosomes.

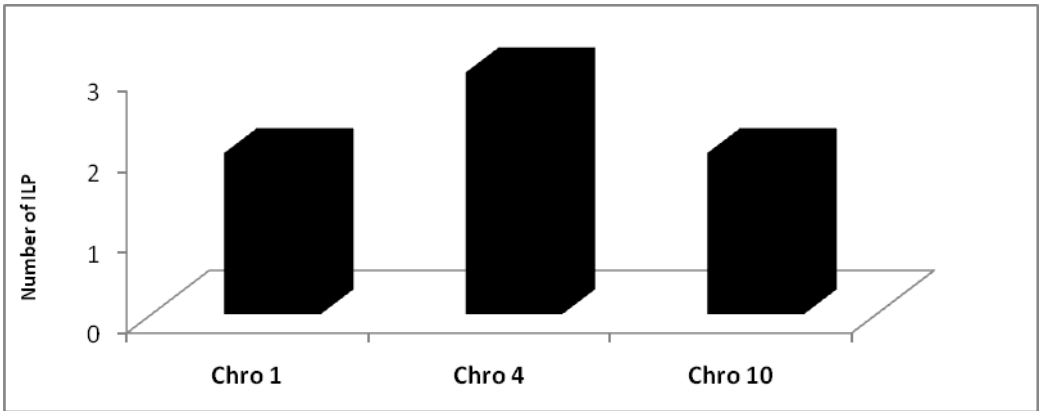

Figure 2. Distribution pattern of seven identified ILP markers on sorghum chromosomes.

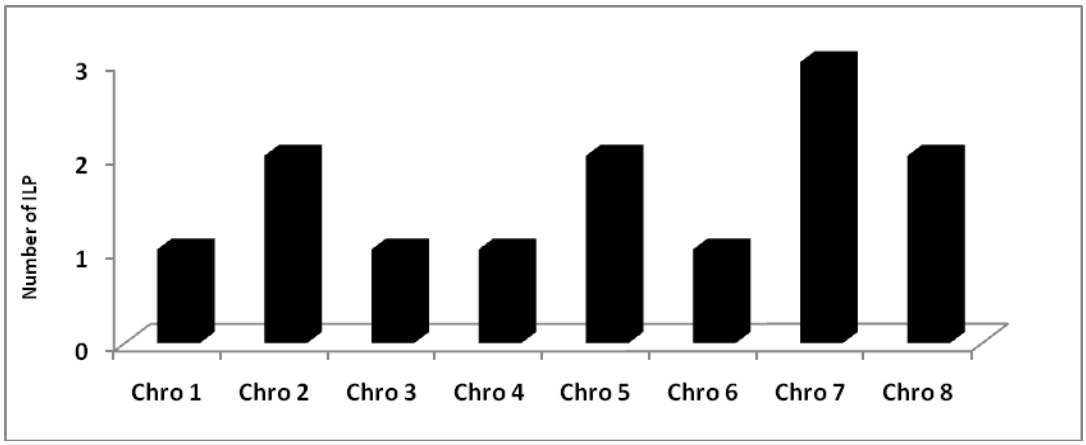

Figure 3. Distribution pattern of thirteen identified ILP markers on soybean chromosomes.

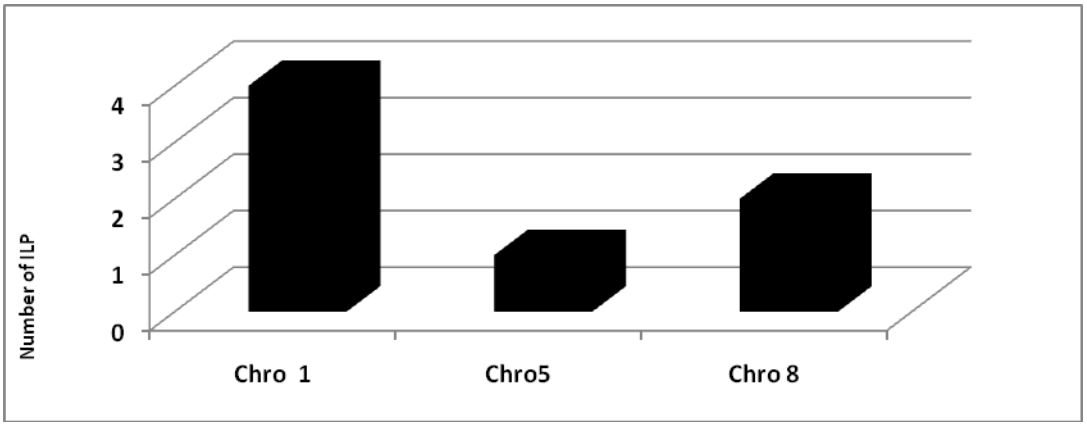

Figure 4. Distribution pattern of seven identified ILP markers on *Medicago* chromosomes.

Table 3. Functional annotations of ESTs containing ILP markers in drought responsive ESTs.

S. No	Seq. Name/Acc.#	Seq. Description	Seq. Length(bp)	Min. e Value
1	GW584142	protein	330	2.89E-66
2	GW584142_1	protein	330	2.89E-66
3	GW584199	glycolate oxidase	521	1.70E-59
4	GW584053	40s ribosomal protein s24-2-like	569	4.20E-33
5	GW584053_1	40s ribosomal protein s24-2-like	569	4.20E-33
6	GW584256	skp1-like protein 3	656	3.95E-69
7	GW584250	atp synthase subunit mitochondrial-like	668	3.55E-94
8	GW584250_1	atp synthase subunit mitochondrial-like	668	3.55E-94
9	GT155311	26s protease regulatory subunit 8	655	5.37E-99
10	GT155311_1	26s protease regulatory subunit 8	655	5.37E-99
11	EL611934	26s protease regulatory subunit 6a homolog a-like	490	1.65E-113
12	EL611934_1	26s protease regulatory subunit 6a homolog a-like	490	1.65E-113
13	EL611906	ribulose- -bisphosphate carboxylase oxygenase small subunit protein	646	5.26E-81
14	EL611906_1	ribulose- -bisphosphate carboxylase oxygenase small subunit protein	646	5.26E-81
15	EL611884	40s ribosomal protein s8	734	7.13E-99
16	EL611875	small nuclear ribonucleoprotein d2	548	1.87E-59
17	GT155488	eukaryotic translation initiation factor 5a-like	626	1.63E-68
18	GT155488_1	eukaryotic translation initiation factor 5a-like	626	1.63E-68
19	GW584014	ribulose-phosphate 3- chloroplastic-like	785	1.89E-88
20	GT155503	photosystem ii 22 kda protein	751	5.08E-99
21	EL611909	glucose-6-phosphate phosphate translocator-like protein	777	5.32E-149
22	GW584165	proteasome subunit beta type-4-like	690	5.40E-100
23	GT155399	ras-related protein ara-3	745	2.24E-64

DISCUSSION

Moisture deficit stress is one of the major abiotic factors for severe yield loss in safflower. There has been only limited information with respect to abiotic stresses and genetic control of tolerance in safflower (Singh *et al.,* 2007). Due to generation and availability of huge genomic information for the crop plants, the *in silico* studies i.e. performed on computer or via computer simulation are important areas of interest for genomic researchers for comparative genomics study (Dudhe *et al.,* 2012) and for molecular markers development. ESTs sequences generated from plants are valuable sources of information for PCR-based gene-specific marker studies and gene mapping (Kumar *et al.,* 2012). Expressed sequence tags (ESTs) based approach could give substantial information not only about stress responsible gene expression but also about novel genes. However, as result of unavailability of genomic DNA sequences in many non-model plants, the potential utility of PCR-based EST markers (in terms of their ability to generate high frequency of band polymorphism) has not been fully

exploited. It is generally believed that intron regions are more divergent than exons (Choi *et al*., 2004). Therefore, the EST-PCR primers that are designed anneal to exons, amplify across introns regions and resulted in relatively higher band polymorphism.

We have developed total of 50 orthologous conserved primers in our study. We have successfully obtained 52% of e-PCR products (> 300 bp) in *Arabidopsis,* 85% in rice, 80% in sorghum, 92% in soybean and 85% in *Medicago*. Further, we have validated our results by using online PCR genomic sequences for *Arabidopsis* and rice, respectively in order to test the working nature of the primer and found to be yielding suitable amplicon size. Recently, (Kumar *et al*., 2012) successfully exploited ILP markers for the characterization of *Artemisia annua* genotypes. In our study, all the identified ILP markers from rice, sorghum and *Medicago* were distributed on three chromosomes but the chromosome number varies as per the species. The most varied distribution pattern was identified for soybean, where all thirteen ILPs which were localized on eight chromosomes or at least one marker per chromosome. We have identified highest number of conserved primers/markers from soybean after *Arabidopsis* which also reflects conserved evolution of drought responsive conserved sequences or primers in safflower and soybean. Another reason could be safflower and soybean both are oil yielding crop and there may be the similar patterns of synteny across the safflower, soybean genomes at least for drought responsive genes. Based on our results we recommend that in future for *in silico* ILP markers identification in safflower, soybean genomic resources can be used rather than model crop plants. In our earlier studied carried out for identification of conserved domains in safflower we have shown that the conserved domains or regions identified from safflower are common in maize, rice and *Arabidopsis* (Dudhe *et al.*, 2012) which supports our present findings. When we searched information on the identified proteins through literature search, four proteins have proven role under drought stress or for other stresses in other crop plants. The plant protein showing the strongest response to drought stress was identified as ribulose- bisphosphate carboxylase oxygenase small subunit protein in rice under drought stress (Kuixian Ji *et al.,* 2012). In *Arabidopsis* 26S proteasome subunit, RPN1a is required for optimal plant growth under water and for salinity stress response (Songhu *et al.,* 2009). The ribulose-3-bis-phosphate carboxylase/oxygenase small subunit (Rbcs), is acute sensitive to combined light and drought stress in *Arabidopsis* (Michel *et al*., 2007). Eukaryotic translation initiation factor 5a-like shows, plant growth regulation and stress responses in plants (Takahashi *et al.,* 1999). Isoenzyme replacement of glucose-6-phosphate dehydrogenase in the cytosol improves stress tolerance in plants (Judith *et al*., 2009).

Annotated ESTs containing ILP markers in drought responsive EST linked to above candidate protein/genes can be considered as potential candidate genes for drought tolerance. To validate the results obtained in this study, the wet lab studies are recommended for future. This study highlights the importance of investigating the genes which interact with ILP markers and suggests following the behavior of marker-related genes at protein level. Identification of genes underlying drought stress will facilitate understanding of molecular mechanisms and will lead to genetic improvement of safflower through marker-assisted selection (MAS). The present study gains importance by adding the ILP markers to the genomic resources of safflower for future use in breeding and safflower improvement. The ILP markers developed from our study will be of particular use for molecular breeding, screening and development of drought tolerant genotype in safflower.

REFERENCES

Becker J, Heun M (1995). Barely microstatellite: Allele variation and Mapping. *Plant Mol. Biol.* 27:835-845.

Botstein DRL, White M, Stoni-lick, Davis RW (1980) Construction of genetic linkages map in man using restriction fragment length polymorphisms. *Ameri. J. Human Gene.* 32: 314-331.

Choi HKD, Kim T, Uhm T, Limpens H, Lim JH, Mun P, Kalo RV, Penmesta A, Seres O, Kulikova (2004). A sequence–based Genetic

map of *Medicago truncatuta* and comparison of markers coloinearity with *M. sativa. Genetics* 166:1463-1502.

Conesa A, Gotz S, Garca-Gomez JM, Terol J, Talon M, Robles M (2005). Blast2go: A universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*, 21:3674-3676

Conesa A, Götz S (2008). Blast2GO: A Comprehensive suite for functional analysis in plant genomics. *Int. J. Plant Genomics*: 1-13.

Dudhe MY, Meena HP, Ranganatha ARG, Mukta N, Lavanya C (2012). *In silico* -identification of conserved domains from EST database in safflower. *J. Oilseeds Res.* 29 (Special issue): 178-181.

Hawkins JD (1988). A survey on intron and exon length. *Nucleic Acids* Res. 9893-9908.

Kumar J, Bajaj P, Singh H, Mishra G P, Srivastava R B, Naik PK (2012). Utilization of intron - flanking EST-specific markers in the genetic characterization of *Artemisia annuna* genotypes from the trans- Himalayan region of Ladakh, India. *J. Enivorn. Biol*. 33:991-997.

Li P, Brutnell, TP (2011). Setaria viridis and Setaria italica, model genetic systems for Panicoid grasses, *J. Exp. Bot.*,62, 3031–7.

Judith S, Hardy S, Zeina T, Weis E, Schaewe AV (2009). Isoenzyme replacement of glucose-6- phosphate dehydrogenase in the cytosol improves stress tolerance in plants. *Proc. Nat. Aca. sci*. 19: 8061-8066.

Kuirian Ji, Yangyang Wanga, Sun Quiaojun Louc, Hanwei Meic, Shihua Shen, Hui Chen (2012). Drought–responsive mechanism in rice genotypes with constrating drought tolerance during reproductive stage. *J. Plant Physio*.169: 336-344.

Mohammad MM, Vahid T, Aghafakhr M, Mohammad RS (2011). Safflower species (*Carthamus oxyacanthus*). A possible source of drought tolerance for arid environments .*Aust. J. Crop Sci*. 81055-1063.

Mukta, N, Reddy, PA, Lavanya, C, Dudhe, MY (2012). Variability for DUS characteristics in released varieties of safflower (*Carthamus tinctorius* L.). *J. Oilseeds Res.* 29 (Special issue): 133-135.

Palumbi SR (1995). Nucleic acids II: The polymerase chain reaction, In molecular systematic Eds: D Hillis and Mortiz C) Sunderland MA 205-247 *Physiology.* 50:9: 1721–1725.

Rozen S, Skaletsky H, (2000). Primer on the WWW for general users and for biologist programmers.*Methods Mol.Biol*. 132: 365-386.

Ruckle ME, De Marco SM, Larkin RM (2007). Plastid signals Remodel light signaling and are essential for efficient chloroplast Biogenesis in Arabidopsis. *The Plant Cell* 19: 3944-3960.

Singh V, Nimbkar N. (2007). Safflower (*Carthamus tinctorius* L.) In:Singh R. J. (Ed.). *Genetic Res., Chro. Eng. and Crop Imp. Oilseed Crops* 4, CRC, Boca Raton, FL 33487-7742, USA, pp. 185.

Songhu W, Kurepa J, Smalle JA (2009). The Arabidopsis 26S proteasome Sub unit required for optimal plant growth and stress response. *Plant Cell Physiol.* 50: 9: 1721–1725.

Takahashi T, Kakehi J (2009). Polyamines: ubiquitous polycations with unique roles in growth and stress responses. *Annals of Bot.* 105: 1-6.

Thippeswamy M, Sivakumar M, Sudhakarbabu O, Reddy CP, Veeranagamallaiah G, Pandurangaiah M, Ramya M, Nareshkumar A, Kirankumar T, Sudhakar C (2012). Generation and analysis of drought stressed subtracted expressed sequence tags from safflower (*Carthamus tinctorius* L.).*Plant Grow. Reg.* 69: 29-41.

Vos PR, Hogers M, Bleeker, Reijians T, Van dee Lee M, Hornes A, Frijters J, Pot J, Peleman M, Kuiper M, Zabeace (1995). AFLP: A new technique for DNA finger printing. *Nucleic Acids Re.* 23: 4407-4414.

Wang DG, Fan JB, Siao CJ, Berno H, Young P, Sapolsky G, Ghandour N, Perkins E, Winchester E, Spencer J, Kruglyak L,Stein L, Hsie T, Topaloglou E, Hubbell E, Robinson M, Mittmann M, Moris MS,n Shen N, Kil bum D, Rioux J, Nusbaum C, Rozen S, Hudson J, Lipshutz R, Chee M, Lander ES (1998). Large scale identification, Mapping, and genotyping of single – nucleotide polymorphism in the human genome. *Science* 280: 1077-1082.

Weiss, EA (2000). Oilseed crops, 2nd Ed. Blackwell, *Oxford*, pp 364.

Williams JGK, Kub Selik AR, Livak KJ, Rafalski JA, Tingey SV (1990). DNA polymorphism amplified by arbitrary primers useful as genetic markers. *Nucleic Acids Res.* 6531-6535.