SciMedCentral

International Journal of Plant Biology & Research

Research Article

# Computational in Sight into Identification and Analysis of SSR-FDM in *Citrus limon*

**Manisha Ray, Smruti P. Pradhan, and Souvagyalaxmi Sahoo***

*Tectona Biotech Resource Center, Bhubaneswar, India*

***Corresponding author**
Souvagyalaxmi Sahoo, Tectona Biotech Resource Center, Bhubaneswar, India

**Submitted:** 16 January 2019

**Accepted:** 31 January 2019

**Published:** 02 February 2019

**ISSN:** 2333-6668



**Keywords**
• SSRs; ESTs; Genomics; MISA; Plant breeding; Primer

**Abstract**

SSRs or microsatellites identification and its functional analysis has a key role in different sectors of genomics such as genome organization, gene regulation, quantitative genetic variation, evolution of genes and plant breeding sectors. Therefore, computational approach was undergone for identification and analysis of SSR within functional domain of *Citrus limon* (C. *limon*) of family Rutaceae, is one of the vital and effective medicinal plant. Total 1644 numbers of extracted ESTs of C. *limon* were validated through Tandem repeat finder and VecScreen which have been assembled in CAP3 that resulted 55 contigs and 1183 singletons. Afterwards, total 420 SSRs were identified as SSR-ESTs using MISA tool and also detected 75.23% of mononucleotide SSR motifs with most ample sort of repeats such as di- (9.52%), tri- (7.61 %) and tetra- (0.95 %) nucleotide. Ultimately, 128 SSR sequences have been selected with appropriate primer properties which would be used as markers to look at transferability to related species. Further, the useful functional annotation was performed using Blast2GO. These findings would assist to understand the significance of SSR markers and also to facilitate the evaluation of genetic range in medicinal plant flora.

## ABBREVIATIONS

EST: Expressed Sequence Tag; SSR: Simple Sequence Repeats; NCBI: National Centre for Biotechnology Information; DbEST: Database of EST; MISA: Micro Satellite Identification Tool; KEGG: Kyoto Encyclopaedia of Genes and Genomes; BP: Biological Process; MF: Molecular Function; CC: Cellular Components

## INTRODUCTION

Plant oriented natural resources are vital for human life. Especially in the last century, the irresponsible use of natural resources has become one of the alarming problems which are a threat to nature and the environment. The wide variety of plant derived medicaments has expanded slowly to come upon needs [1]. Thus an expertise of the patterns of genetic variant within and among populations of medicinal plant life is essential for devising most effective genetic resource control strategies for his or her conservation, sustainable usage and genetic improvement [2].

Citrus fruits are one of the international's most essential fruit crops, and are regarded for their nutritive values and unique aroma. Citrus is specially consumed as clean fruit or juice. Many in vivo and in vitro researches have proved that citrus fruit is effective against many chronic diseases, like cancers and vascular illnesses. Lemon could be very rich in important natural compounds, which include citric acid, ascorbic acid, minerals, flavonoids, and crucial oils. Therefore, the new Citrus cultivars have been mainly developed for fresh consumption i.e. to screen these plants in order to validate their use in food and medication and to and to show the active ingredients by the way of characterizing their constituents. The unique tendencies which include their phenolic compound and specially the flavonoids contents led to their use in new fields inclusive of pharmacology and food era [3].

Although if, have a look at on taxonomic type of *Citrus limon,* it represents the complicated, debatable and ambiguous taxonomy as it consists of a number of the most commercially crucial fruits [4]. This purpose prompted to work on molecular marker evaluation on this present characteristic because taxonomic category offers the records for future breeding, genetic improvement etc, so to enhance this observe the following analyses have been taken in to attention.

Expressed collection tags (ESTs) are sub sequence of cDNA instructions that offer direct facts of gene expression and additionally function resources of microsatellites or the simple sequence repeats (SSRs), are the short DNA sequences with 1-6 base pairs of length. Several studies advise that the plenty of SSRs were found in non-coding regions of the genome sequences and have a wide application in the area of plant genetic studies which includes genetic variant, linkage mapping, gene tagging, evolution and breeding as they have multi-allelic, reproducible and co-dominant inheritance properties [5].

EST–SSR markers are anticipated to own excessive interspecific transferability as they belong to conserved genic areas of the genome [6], thus the objectives of this work focused on the *in silico* identification of EST- SSR markers of *Citrus limon*.

Also the primer designing from EST-SSRs turned into one of the prospective elements of this study because in expressed DNA areas the present primer sequences are anticipated to be quite well conserved, hence it improving the threat of marker transferability across taxonomic boundaries [7]. The final element is the functionality annotation of SSR-FDM, which gives the facts approximately the involvement of EST-SSRs in distinct metabolic features and throws a course to research the genetic capability of *C. limon*.

## MATERIALS AND METHOD

### Retrieval of EST sequences

The Expressed sequence tag (EST) sequences of *Citrus limon* were retrieved from EST database (dbEST) (https://www.ncbi.nlm.nih.gov/nucest/?term=) of National Centre for Biotechnology Information (NCBI) web server (https://www.ncbi.nlm.nih.gov/).

### Detection of repeat locations

The accumulated EST sequences of *Citrus limon* were subjected to for the elimination or deletion of repeat regions within the nucleotide sequences through the usage of Tandem Repeats Finder (TRDB) (https://tandem.bu.edu/trf/trf.html) that's a application to find or show the repeated sample of one or greater nucleotides in DNA sequences.

### Screening of vector regions

After the deletion of tandem repeats containing sequences the EST sequences were again analyzed to screen the vector regions through VecScreen (https://www.ncbi.nlm.nih.gov/tools/vecscreen/) which is a system to find the section of nucleotide, which may be a vector contaminated vicinity or the infection rate is more at that precise segment.

### Sequence assembly analysis

The remaining EST sequences of *C. limon* were taken for assembly analysis by using CAP3 (http://doua.prabi.fr/software/cap3) sequence assembly program, which permits to assemble a set of contiguous or contigs sequence as well as the singleton sequences.

### Detection of SSR containing EST sequences

The resulted contigs and singleton sequences were subjected to further analysis to find out those sequences which contained the single sequence repeats (SSR) sequences via Microsatellite identification tool (MISA) (http://pgrc.ipk-gatersleben.de/misa/). It allows the identification and localization of perfect microsatellites as well as compound microsatellites which are interrupted via a certain wide variety of bases.

### Retrieval of primer sequences

The amassed SSR containing EST sequences were again computed in Primer 3 (http://bioinfo.ut.ee/primer3-0.4.0/) to collect the appropriate primer sequences or the forward and reverse primer from the given nucleotide sequences.

### SSR-FDM analysis

The functional annotation of considered Primer sequences turned into performed through Blast2GO (https://www.blast2go.com/) analysis. Blast2GO is a bioinformatics platform for high-quality functional annotation and evaluation of genomic datasets. So, this could offer all of the useful facts for selected sequences.

## RESULTS AND DISCUSSION

### Sequence retrieval and validation

There were total 1644 number of EST sequences of *Citrus limon* were retrieved from EST database of NCBI and were analyzed through Tandem Repeat Finder to find out the sequence in which one or more nucleotides were repeated at a phase, because these tandem repeats can be found not only in intergenic regions however also in each of the non coding and coding regions of an expansion of different genes and these repeat expansion sicknesses are a set of human genetic problems caused by long and highly polymorphic tandem repeats, such as if the repeat is present in an exon or coding part, then Huntington Disease (HD) or spinobulbar muscular atrophy (SBMA) is happened and if repeat is outside of the open reading frame myotonic dystrophy (DM) or Fragile X syndrome (FXS) can caused [8]. Thus, to overcome all above the complications, there were total 210 numbers of tandem repeats contained EST sequences were removed manually. After removal of tandem repeat sequences, remaining 1434 numbers of sequences were once more analyzed through VecScreen [9] to discover and eliminate the vector infected sequences i.e. the segment of nucleotide which may be a strong vector origin and might have more chance to contaminate by vector, thus out of 1434 number of sequences 63 numbers of sequences were deleted as vector contaminated sequences and remaining 1371 number of EST sequences of *Citrus limon* were went for further analysis (Figure 1).

### Sequence assembly analysis and identification of SSR

The completion of retrieval and validation of EST sequences of *Citrus limon,* initiated the analysis of remaining 1371 numbers of sequences for meeting evaluation through CAP3 program which collect the reads of EST sequences and predicts the contigs and singleton sequences, because the software has a functionality to clip 5′ and 3′ low-quality regions of reads, uses base quality values in computation of overlaps between reads, construction of more than one sequence alignments of reads, and generation of consensus sequences. This program also uses forward–opposite constraints to correct assembly errors and hyperlink contigs [10]. So, on the premise of all of the functions of CAP3 application, there were total 55 numbers of contigs and 1183 number of singleton sequences were predicted from 1371 numbers of EST sequences and these predicted contigs and singleton sequences were subjected for identification of SSR sequences i.e. those contigs and singleton sequences contains single sequence repeats (SSR) or microsatellites, are extensively-used marker device in plant genetics and forensics and beneficial for primer design [11], were identified through MISA, for the reason that it could identify the SSR containing sequences from both contigs and singletons [12] and right here it resulted 420 numbers of SSR containing sequences, which is termed as SSR-ESTs (Figure1).

### Primer designing

The fundamental parameters for primer pair design have been as follows: a minimum range of SSR pattern repeats of 10 for di-

nucleotides, seven for tri-nucleotides, four for tetra-nucleotides, minimum and most product sizes of 103–250 bp (optimal: 150 bp); primer length of 18–25 bases (optimal: 21 bases); GC content of 57.45% –61.76% (optimal: 50%); annealing temperatures of 31.82°C - 60°C (optimal: 56°C); and default values for the other parameters [13]. Thus, by following these above criteria, out of 420 EST-SSR 128 numbers of sequences were considered, which were gave appropriate forward and reverse primers through Primer3, because Primer3 software has been broadly used for primer layout, often in high-throughput genomics programs [14] (Table 1) (Figure 1).

## Frequency distribution of SSRs

The diagnosed SSRs (Microsatellites) were analyzed by MISA tool, which were gave mononucleotides, dinucleotides, trinucleotides, tetranucleotides and compounds. Out of 420 SSR containing sequences, the highest proportion were presented 316 numbers of mononucleotide repeats (MNR) (75.23%), 40 numbers of dinucleotide repeats (DNR) (9.52%), 32 numbers of trinucleotide repeats (TNR) (7.61%), and 4 numbers of tetranucleotide repeats (TNR) (0.95%) were observed.

## Functional domain analysis of SSR markers

128 numbers SSR-ESTs sequences were assigned with gene ontology terms for the functional domain annotation through BLAST2GO, has the capability to produce high throughput useful annotation statistics [15] but among 128 only 115 numbers of EST-SSRs were analyzed in BLAST2GO (Figure 1). The evaluation

**Table 1:** List of Primers designed from SSR-EST sequences though Primer3 Program.

| Sl. No. | ID | SSR | F.P (5′-3′) | length | Tm (%) | GC (%) | R.P (5′-3′) | length | Tm (°C) | GC (%) | P.S (BP) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | GO25 6710.1 | P3 | AGGGTCTGCTGC TATTCACA | 20 | 58.47 | 50.00 | TGAACATTACAGG CCTCTCTCA | 22 | 59.88 | 45.45 | 169 |
| 2 | DC89 1737.1 | P2 | AAGCAAGCTGAT GTGCCTCT | 20 | 60.16 | 50.00 | ATCCGTGCATTTA GCAATCC | 20 | 59.93 | 45.00 | 230 |
| 3 | DC89 1733.1 | P1 | TGTGCGCGTAGA AAATAAGG | 20 | 58.98 | 45.00 | CAGTTCCCTCCAA TCGTGTT | 20 | 59.97 | 50.00 | 159 |
| 4 | DC89 1731.1 | P1 | ACGTCGACGGTG AATAATCC | 20 | 59.82 | 50.00 | TTCATTCACAAGC AGGCAAG | 20 | 59.99 | 45.00 | 216 |
| 5 | DC89 1730.1 | P2 | TGAAGGAGACGC ACTTGATG | 20 | 59.98 | 50.00 | AGATGCAACAGG CAACATGA | 20 | 60.27 | 45.00 | 181 |
| 6 | DC89 1729.1 | P1 | CGCTCCTGGCTTT CTTCTTA | 20 | 59.72 | 50.00 | CATTTTACAACCC CGGAATG | 20 | 60.05 | 45.00 | 167 |
| 7 | DC89 1725.1 | P1 | ATGCTTGTACCTA CCCCCAAC | 21 | 60.12 | 52.38 | TCAAATAGAGCA GGAATTTTTCTCA | 25 | 60.56 | 32.00 | 150 |
| 8 | DC89 1724.1 | P1 | CCGGTGATGAAG AAGATGGT | 20 | 59.93 | 50.00 | CATCTGCCTGTTG AGGTCAC | 20 | 59.26 | 55.00 | 168 |
| 9 | DC89 1710.1 | P1 | GTGATGGTGGTT CCCGATAC | 20 | 60.06 | 55.00 | AACAGGAACCGA TCAAAACG | 20 | 59.97 | 45.00 | 155 |
| 10 | DC89 1703.1 | P2 | GTATCGGGGTCA TGGTGAAC | 20 | 60.06 | 55.00 | AAGGGTTTTACCC TGCGTTT | 20 | 59.87 | 45.00 | 170 |
| 11 | DC89 1695.1 | P3 | CTGTGGTCGATG GAGAAGGT | 20 | 60.11 | 55.00 | GCAGCCATCATCA ATGTGAA | 20 | 60.64 | 45.00 | 161 |
| 12 | DC89 | P1 | AGCTGCTGCTAA | 20 | 60.0 | 55.0 | CCTGCCAGGATGA | 20 | 60. | 50.0 | 186 |
| | 1690. | | CTCGCTTC | | 7 | 0 | AGAACAT | | 07 | 0 | |
| | 1 | | | | | | | | | | |
| 13 | DC89 | P1 | GTGATACCGGTC | 20 | 60.0 | 55.0 | TCACTTTGCCAGA | 20 | 59. | 50.0 | 208 |
| | 1673. | | GGAGAGAA | | 7 | 0 | GTCGATG | | 98 | 0 | |
| | 1 | | | | | | | | | | |
| 14 | DC89 | P3 | GGAGCTCAAGGA | 20 | 59.9 | 55.0 | AACCCATCCCCGT | 20 | 59. | 45.0 | 159 |
| | 1665. | | GACGTTTG | | 9 | 0 | TTTTATC | | 89 | 0 | |
| | 1 | | | | | | | | | | |
| 15 | DC89 | P1 | TACCCGGGGGAG | 20 | 60.0 | 50.0 | GCCTCGTCCATGT | 20 | 60. | 50.0 | 199 |
| | 1644. | | AAAAATAC | | 1 | 0 | TCAGTTT | | 12 | 0 | |
| | 1 | | | | | | | | | | |
| 16 | DC89 | P1 | ACCAGTCGTGAT | 20 | 59.9 | 50.0 | GAACCCTCAAGCA | 20 | 59. | 55.0 | 152 |
| | 1840. | | GGAAAAGG | | 7 | 0 | CCTCAAC | | 70 | 0 | |
| | 1 | | | | | | | | | | |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 17 | DC89 | P1 | GAAGGGAATGGG | 20 | 60.1 | 50.0 | TTCCCAGCAATAC | 20 | 59. | 50.0 | 194 |
| | 1826. | | GACAACTT | | 7 | 0 | CCTGTTC | | 93 | 0 | |
| | 1 | | | | | | | | | | |
| 18 | DC89 | P2 | TCTCTGATACCGC | 20 | 59.9 | 50.0 | GCCGGAGGATCTT | 20 | 60. | 50.0 | 174 |
| | 1817. | | CTGCTTT | | 8 | 0 | GTTGTAA | | 07 | 0 | |
| | 1 | | | | | | | | | | |
| 19 | DC89 | P3 | TGTTTGGTTCTGG | 20 | 60.2 | 45.0 | CTTCCAATGGAGC | 20 | 60. | 50.0 | 159 |
| | 1815. | | TCAAGCA | | 8 | 0 | AGGAAGA | | 33 | 0 | |
| | 1 | | | | | | | | | | |
| 20 | DC89 | P1 | TGGGAAGGTATG | 20 | 60.0 | 55.0 | CACCATGTGCCGT | 20 | 59. | 50.0 | 176 |
| | 1805. | | CCTCAGTC | | 7 | 0 | GTTATTC | | 85 | 0 | |
| | 1 | | | | | | | | | | |
| 21 | DC89 | P1 | AGGCCTCTGCAT | 20 | 59.9 | 50.0 | AACATTCAAAGCC | 20 | 59. | 45.0 | 183 |
| | 1803. | | ATCTCGAA | | 4 | 0 | CACCAAG | | 97 | 0 | |
| | 1 | | | | | | | | | | |
| 22 | DC89 | P1 | CTTGCGGAAAGC | 20 | 60.2 | 50.0 | CCTTTGTGGCAGG | 20 | 60. | 50.0 | 188 |
| | 1787. | | TGAGAAAG | | 6 | 0 | CCTAATA | | 09 | 0 | |
| | 1 | | | | | | | | | | |
| 23 | DC89 | P1 | TCCGAAATTCTTG | 20 | 60.0 | 45.0 | TTAAAGTGCACGG | 20 | 59. | 50.0 | 206 |
| | 1786. | | GAACTGG | | 4 | 0 | CACAGTC | | 91 | 0 | |
| | 1 | | | | | | | | | | |
| 24 | DC89 | P1 | ACAAACACAACG | 20 | 60.0 | 40.0 | ACAGACCCAAGT | 20 | 59. | 55.0 | 199 |
| | 1785.1 | | CCAATTCA | | 1 | 0 | GAGGATGG | | 96 | 0 | |
| 25 | DC89 | P3 | TGGATCGTGTGA | 20 | 59.9 | 40.0 | ACCTCTGCTTTTT | 20 | 60. | 45.0 | 221 |
| | 1783. | | TTGGAAAA | | 0 | 0 | GCAGCAT | | 02 | 0 | |
| | 1 | | | | | | | | | | |
| 26 | DC89 | P1 | GAATCCCAGGGT | 20 | 60.0 | 55.0 | AAGCAAGGCCAA | 20 | 60. | 45.0 | 167 |
| | 1776. | | AGGGTGTT | | 5 | 0 | ATGATGAC | | 08 | 0 | |
| | 1 | | | | | | | | | | |
| | DC89 | P1 | TATTGCCGTTCAG | 20 | 59.5 | 45.0 | CCACTTCACCAGC | 20 | 61. | 60.0 | 222 |
| 27 | 1774. | | TTGGTTG | | 8 | 0 | CAGTCAC | | 76 | 0 | |
| | 1 | | | | | | | | | | |
| 28 | DC89 | P1 | GATATGATGCCG | 20 | 59.9 | 45.0 | GGTTCTCCAGTTG | 20 | 59. | 50.0 | 188 |
| | 1766. | | GTTTTGCT | | 3 | 0 | CACCAAT | | 97 | 0 | |
| | 1 | | | | | | | | | | |
| 29 | DC89 | P1 | AGATGGAGACAA | 20 | 59.9 | 55.0 | CGTTTTTCACGCA | 20 | 60. | 45.0 | 244 |
| | 1563. | | CCCGTGAC | | 7 | 0 | GCACTAA | | 05 | 0 | |
| | 1 | | | | | | | | | | |
| | DC89 | P1 | GGTATTCCTGCTG | 20 | 59.7 | 55.0 | CTCCAAAATCTCC | 20 | 59. | 50.0 | 150 |
| 30 | 2725. | | GCTTGTC | | 0 | 0 | GTCTTGC | | 81 | 0 | |
| | 1 | | | | | | | | | | |
| 31 | DC89 | P3 | TTTCGGATCAGG | 20 | 60.0 | 45.0 | AGGGAACGGTGA | 20 | 59. | 50.0 | 238 |
| | 2345. | | GAGAAATG | | 1 | 0 | ACATGAAG | | 97 | 0 | |
| | 1 | | | | | | | | | | |
| 32 | DC89 | P1 | GCCGTTGGCAAT | 20 | 59.9 | 45.0 | GGCAATAGGAAG | 20 | 60. | 50.0 | 160 |
| | 2325. | | AAGAATGT | | 7 | 0 | AGCACGAA | | 35 | 0 | |
| | 1 | | | | | | | | | | |
| 33 | DC89 | P1 | TGTTCCGATTTGG | 20 | 60.3 | 45.0 | GGACAACTTTCCC | 20 | 60. | 50.0 | 239 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 2322. | | GTCAACT | | 5 | 0 | CCTCAAT | | 17 | 0 | |
| | 1 | | | | | | | | | | |
| 34 | DC89 | P1 | CTCCGGCAACAT | 20 | 60.0 | 45.0 | CACTTTGACCTTC | 20 | 59. | 50.0 | 213 |
| | 2319. | | TCATTTCT | | 7 | 0 | GCCTTTC | | 85 | 0 | |
| | 1 | | | | | | | | | | |
| | DC89 | P1 | GCCCATCTCATG | 20 | 59.8 | 50.0 | TAAAATACGGGCC | 20 | 59. | 45.0 | 189 |
| 35 | 2317. | | ATGAACCT | | 9 | 0 | AAACTGG | | 83 | 0 | |
| | 1 | | | | | | | | | | |
| | DC89 | P1 | CAAATGGAGCGG | 20 | 59.9 | 50.0 | TCAGCCCATTTGA | 20 | 60. | 45.0 | 235 |
| 36 | 2465. | | AAGCTAAG | | 7 | 0 | AAAGTCC | | 05 | 0 | |
| | 1 | | | | | | | | | | |
| 37 | DC89 2463.1 | P1 | CGCAAGGGCAAA TCACTATT | 20 | 60.10 | 45.00 | GCAGACTCTAGGG CAGGTTG | 20 | 60.01 | 60.00 | 218 |
| 38 | DC89 2455.1 | P1 | AATTCCGTCGTC GTTTTCAG | 20 | 60.11 | 45.00 | CTTCCCTGTGCAG TGCAGTA | 20 | 60.05 | 55.00 | 232 |
| 39 | DC89 2450.1 | P1 | GCGTTTGAGCCT GGTTTTAG | 20 | 59.88 | 50.00 | AACAAAGCCAAA ACCTCGTG | 20 | 60.15 | 45.00 | 156 |
| 40 | DC89 2432.1 | P1 | GCAGTGGATCAT CGCCTAAT | 20 | 60.07 | 50.00 | TGACTGCCAGAAG AGCTCAA | 20 | 59.86 | 50.00 | 203 |
| 41 | DC89 2423.1 | P1 | TTAGCACCACAG CATCCAAG | 20 | 59.86 | 50.00 | GCTGACACTGCTT CAACCAA | 20 | 60.03 | 50.00 | 227 |
| 42 | DC89 2411.1 | P1 | GATCAAGAAGCT TCGCAAGG | 20 | 60.10 | 50.00 | CGGCTTGATTTCA ACCTGAT | 20 | 60.07 | 45.00 | 160 |
| 43 | DC89 2400.1 | P1 | CATCAGCTCCAG TGACTCCA | 20 | 59.98 | 55.00 | CTCCTGGCTGGAT AGGACTG | 20 | 59.82 | 60.00 | 171 |
| 44 | DC89 2399.1 | P1 | CTTGCAGATCAG CTTCTGGA | 20 | 59.27 | 50.00 | CCAAGGACATATC CCCACTG | 21 | 60.19 | 55.00 | 240 |
| 45 | DC89 2376.1 | P1 | GTGAAGAGGGTG GTGGTCAT | 20 | 59.82 | 55.00 | CCTGCCCTGACCT AAAAACA | 20 | 60.10 | 50.00 | 246 |
| 46 | DC89 2373.1 | P1 | GACTCCCGGTGT TTTTCTGA | 20 | 60.09 | 50.00 | TCCAAATTGATCA GCAGCAG | 20 | 59.95 | 45.00 | 164 |
| 47 | DC89 2369.1 | P1 | GCTTCAGCTTGA AACCCAAT | 20 | 59.32 | 45.00 | CCAAAAAGTCGAT TCCGTCA | 20 | 61.00 | 45.00 | 240 |
| 48 | DC89 2716.1 | P1 | CCCTATACCTGTG CCATGCT | 20 | 59.98 | 55.00 | CCGGACCTTAGAG CAGTCAG | 20 | 60.01 | 60.00 | 211 |
| 49 | DC89 2712.1 | P1 | GCCTGCTTCTCTG TTTTCCTT | 21 | 60.01 | 47.62 | ACATTGCACAAAT CCACAGC | 20 | 59.58 | 45.00 | 153 |
| 50 | DC89 2705.1 | P1 | CTTTAGTGTGCCG CCGTTAT | 24 | 60.15 | 50.00 | CACGTCTCACTCG CTGGTTA | 25 | 60.05 | 55.00 | 177 |
| 51 | DC89 2697.1 | P1 | TGGGGGTAGATA GGGGTAGG | 20 | 60.03 | 60.00 | TTGTGGCAACAGG TATCCAA | 20 | 59.96 | 45.00 | 168 |
| 52 | DC89 2692 | P1 | AACAATTTGACG CCGATCTC | 20 | 60.08 | 45.00 | TTGTCAGGCTTCG ACCTCTT | 20 | 59.99 | 50.00 | 161 |
| 53 | DC89 2688 | P1 | GGGTTATTTGCTT TGGTTCG | 20 | 59.45 | 45.00 | TGATTCCAGAAGG CCAAAAC | 20 | 60.05 | 45.00 | 190 |
| 54 | DC89 2682.1 | P2 | GTCGTGGTCCTG GTTTCTGT | 20 | 60.01 | 55.00 | CCCCCAAAGTCAG CAAATAC | 20 | 59.43 | 50.00 | 204 |
| 55 | DC89 2672.1 | P1 | ACCCTTGGGGAC CATATCTC | 20 | 60.01 | 55.00 | CTCTCGAAGCTCT GCTGGAT | 20 | 59.85 | 55.00 | 166 |
| 56 | DC89 2652.1 | P1 | AAGTTTGTGGGG ATGCAAAG | 20 | 59.97 | 45.00 | TGGGCCAGACTGG ATAAAAG | 20 | 60.07 | 50.00 | 194 |
| 57 | DC89 2650.1 | P1 | TAGCCAATGAGG CACATCAG | 20 | 59.82 | 50.00 | TCTCCCCCACATG ATCCTAA | 20 | 60.27 | 50.00 | 170 |
| 58 | DC89 2647.1 | P1 | GCCTCATCTTGGT TTGAGGA | 20 | 60.20 | 50.00 | GGGTCTCAAGCGT CAACATAA | 21 | 60.13 | 47.62 | 208 |
| 59 | DC89 2644.1 | P1 | CCTTTGATGGCTG TTGTCCT | 20 | 60.11 | 50.00 | AAAGCTAGAGGG GCTGGAAG | 20 | 59.98 | 55.00 | 246 |
| 60 | DC89 2636.1 | P3 | GCTGTCAAGGCT CTTGGTTC | 20 | 60.00 | 55.00 | TTTTGCCTACCCA CTCCTTG | 20 | 60.10 | 50.00 | 203 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 61 | DC89 2635.1 | P3 | CCCAACCCTTTCCTTCTTTC | 20 | 59.91 | 50.00 | TGCTGTCCTTTCAACACCAA | 20 | 60.28 | 45.00 | 188 |
| 62 | DC89 2621.1 | P1 | GGACCCAGAGAGTGAAGCTG | 20 | 59.99 | 60.00 | CGAGAAGCCATGCTCCTTAC | 20 | 59.98 | 55.00 | 188 |
| **63** | DC892614.1 | P1 | AGCTTCGGTGGTTCTGTTTG | 20 | 60.29 | 50.00 | ATTGTCACAGCCCAGGAAAG | 20 | 60.11 | 50.00 | 156 |
| 64 | DC89 2282.1 | P1 | GCCTTCTGCTTGTGTCCTTC | 20 | 60.00 | 55.00 | CAATGGGATTATGGGAATCG | 20 | 59.97 | 45.00 | 177 |
| 65 | DC89 2281.1 | P3 | AGTCCAAACAACGGCTATGG | 20 | 59.99 | 50.00 | CCGTCACATTCAAACCAGTG | 20 | 60.00 | 50.00 | 181 |
| 66 | DC89 2278.1 | P1 | GCTCAGACATGGATGAAGATGA | 22 | 60.23 | 45.45 | AATTTCTTGAGCACCGCATC | 20 | 60.22 | 45.00 | 150 |
| 67 | DC89 2266.1 | P1 | CTTCAATCCCATCTGCCCTA | 20 | 60.03 | 50.00 | GGCGTCCCAAATCTTATGAA | 20 | 59.90 | 45.00 | 141 |
| 68 | DC89 2250.1 | P4 | GGGCTCTAAGCATTGTCTGC | 20 | 59.98 | 55.00 | AGTTTCTGCCTCGTGCTGAT | 20 | 60.02 | 55.00 | 193 |
| 69 | DC89 2247.1 | P2 | TTTAGAGACGGCGGCTAATG | 20 | 60.36 | 50.00 | ACTGGGATTTCTGCACTGCT | 20 | 59.87 | 50.00 | 210 |
| 70 | DC89 2239.1 | P1 | ATCGAGCGTCTTGAAGCAAT | 20 | 59.98 | 45.00 | TATGCACCCAACATGGAAAA | 20 | 59.79 | 40.00 | 192 |
| 71 | DC89 2230.1 | P2 | TTCGGTATCAGAGCCGAACT | 20 | 59.84 | 50.00 | TGGAGTTGTGGAACTGCTTG | 20 | 59.87 | 50.00 | 201 |
| 72 | DC89 2213.1 | P1 | TCTGGTTGTCTCTGCCAATG | 20 | 59.83 | 50.00 | CATTTCGGATTCGACTTTCC | 20 | 59.50 | 45.00 | 214 |
| 73 | DC89 2208.1 | P3 | CACTATGGCGTTGACATTGG | 20 | 59.99 | 50.00 | AATGCAGGCCAACAACCTAC | 20 | 60.00 | 50.00 | 204 |
| 74 | DC89 2195.1 | P1 | CAGCATTGGTGTTGGTGTTC | 20 | 60.01 | 50.00 | CCTGCAGCAAAGACAAGACA | 20 | 60.18 | 50.00 | 199 |
| 75 | DC89 2183.1 | P1 | TGCTCATACGCAGCAATCTC | 20 | 60.13 | 50.00 | GTCGCACACCTTTCTCCATT | 20 | 60.12 | 50.00 | 192 |
| 76 | DC89 | P1 | TTTTTGGGCTGGA | | 60.0 | 45.0 | GACGGTTCGTGAA | | 59. | 50.0 | 244 |
| | 2179. | | AACTGAC | 20 | 9 | 0 | TCATCTG | 20 | 09 | 0 | |
| | 1 | | | | | | | | | | |
| 77 | DC89 | P1 | GGACGCAGGGTT | 20 | 60.0 | 50.0 | CACCAAAACCTCA | 21 | 59. | 47.6 | 249 |
| | 1643. | | AGATTTCA | | 7 | 0 | GCCTTATG | | 62 | 2 | |
| | 1 | | | | | | | | | | |
| 78 | DC89 | P1 | GCCCAGGAGCAG | 20 | 60.1 | 50.0 | TGTTGGTTCCTTC | 20 | 59. | 45.0 | 190 |
| | 1641. | | ATTCATAA | | 8 | 0 | ACCACAA | | 98 | 0 | |
| | 1 | | | | | | | | | | |
| 79 | DC89 | P1 | GAGACGCAATGG | 20 | 59.8 | 50.0 | GCAAGTGACCCTG | 20 | 59. | 50.0 | 202 |
| | 1638. | | AGAAAAGG | | 1 | 0 | GCATAAT | | 96 | 0 | |
| | 1 | | | | | | | | | | |
| 80 | DC89 | P1 | TACTGTTTTTGCC | 20 | 60.1 | 50.0 | TTACAAGCACGCA | 20 | 60. | 50.0 | 223 |
| | 1625. | | GGTCTCC | | 1 | 0 | GGTGGTA | | 32 | 0 | |
| | 1 | | | | | | | | | | |
| 81 | DC89 | P1 | GCTATCGATTGC | 20 | 60.1 | 50.0 | TCAAAGATGCCTG | 20 | 59. | 45.0 | 155 |
| | 1624. | | CCAAGTGT | | 0 | 0 | ATGCTTG | | 95 | 0 | |
| | 1 | | | | | | | | | | |
| 82 | DC89 | P1 | CGATCCTGGACC | 20 | 60.0 | 55.0 | ATAAAGGCAAAG | 20 | 59. | 45.0 | 244 |
| | 1617. | | CAGAAGTA | | 7 | 0 | GCAGCTCA | | 98 | 0 | |
| | 1 | | | | | | | | | | |
| 83 | DC89 | P1 | TTGAGGAGGCTT | 20 | 59.9 | 45.0 | CAGGCTCGTCCGA | 20 | 60. | 55.0 | 178 |
| | 1613. | | CTTTTCCA | | 3 | 0 | CTTTTAG | | 01 | 0 | |
| | 1 | | | | | | | | | | |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | DC89 | P1 | CCAGAAACCTCC | 20 | 59.8 | 50.0 | ATCCATGTGCACC | 20 | 59. | 45.0 | 194 |
| 84 | 1587. | | AAGCAAAG | | 5 | 0 | AGAAACA | | 97 | 0 | |
| | 1 | | | | | | | | | | |
| 85 | DC89 | P1 | GCCACGAGTGAT | 20 | 59.6 | 45.0 | CAGCCATAAGCCG | 20 | 60. | 55.0 | 156 |
| | 1566. | | GAAATTGA | | 5 | 0 | GTAGGTA | | 11 | 0 | |
| | 1 | | | | | | | | | | |
| 86 | DC89 | P1 | AAGTGGACGAAA | 20 | 60.1 | 50.0 | TCAAACCCCTCGT | 20 | 60. | 50.0 | 153 |
| | 1557. | | TGGAGACG | | 1 | 0 | ACGTCAT | | 38 | 0 | |
| | 1 | | | | | | | | | | |
| 87 | DC89 1552.1 | P1 | GATTGGATCCGC AGATGTTT | 20 | 59.90 | 45.00 | CATACGGCTCCTC GTTCAAT | 20 | 60.10 | 50.00 | 201 |
| 88 | DC89 1538.1 | P1 | AGTGCTCCTTCG AGTCCTCA | 20 | 60.14 | 50.00 | TCATCCTCACCGG AGTTTTC | 20 | 60.05 | 50.00 | 191 |
| 89 | DC89 2368.1 | P1 | GGACCAAGATGT GGGAAAGA | 20 | 59.90 | 50.00 | TGAAACGCAGGA TGTCAGTC | 20 | 59.84 | 50.00 | 225 |
| 90 | DC89 2361.1 | P1 | CGCACTGAACAA TGGAGAGA | 20 | 59.98 | 50.00 | CAACGGGTAAGA CCTGCATT | 20 | 59.99 | 50.00 | 181 |
| 91 | DC89 2350.1 | P1 | GTGTCGGCTTCTT ACGTGCT | 20 | 60.46 | 55.00 | GCCACTTCTCGGA ACTTGAG | 20 | 59.99 | 55.00 | 241 |
| 92 | DC89 2346.1 | P1 | AACCCACAGAGG CGTAACAC | 20 | 60.04 | 55.00 | CTAGCACTGGGGA AGGTCAG | 20 | 59.86 | 60.00 | 207 |
| 93 | DC89 2591.1 | P1 | AAAGATGGCCAA GCCAAGAT | 22 | 60.96 | 45.00 | CACCAACGCTTCA GTAGTGC | 24 | 59.51 | 55.00 | 245 |
| 94 | DC89 2585.1 | P1 | CCAAAACCAAAT CGCTCTTC | 20 | 59.69 | 45.00 | GGACTTGTGGGTT GCTCATT | 20 | 59.97 | 50.00 | 150 |
| 95 | DC89 2580.1 | P1 | GATGTCATGGCT CAGCTCAA | 20 | 59.95 | 50.00 | TGACGTGTCTCTT CGACAGG | 20 | 60.02 | 55.00 | 232 |
| 96 | DC89 2150.1 | P1 | CAGAAGTTCAGG AGGGCAAG | 20 | 59.98 | 55.00 | CGGAAAACTGAG ACCCTTGA | 20 | 60.22 | 50.00 | 227 |
| 97 | DC89 2144.1 | P1 | AGACTCTCGTTGT CGCCATT | 20 | 59.87 | 50.00 | GTCTTGATGTGGC TTGCTGA | 20 | 59.99 | 50.00 | 211 |
| 98 | DC89 2572.1 | P1 | TTGCTCAGGCTG CTGTAGAA | 20 | 59.89 | 50.00 | CTGCTTCAGCTTC TGGCTCT | 20 | 60.04 | 55.00 | 225 |
| 99 | DC89 2568.1 | P2 | CCTCAACTGTGA ATCGAGCA | 20 | 59.98 | 50.00 | TCCCGTTGTCTTTT GGTTTC | 20 | 59.95 | 45.00 | 170 |
| 100 | DC89 2133.1 | P1 | CCGGAGTCTGAT CTCAAAGC | 20 | 59.95 | 55.00 | CTGGCTAGGGTCT CCAAGTG | 20 | 59.86 | 60.00 | 187 |
| 101 | DC89 2126.1 | P1 | TCTCTCTTCTCGC CTCTTCG | 20 | 59.97 | 55.00 | ACGGAGCCTCCTA GCTTCTC | 20 | 60.12 | 60.00 | 218 |
| 102 | DC89 2122.1 | P1 | CCAGATCGTTTTG ACCACCT | 20 | 59.97 | 50.00 | TTCACAGGGTGCT TCTGATG | 20 | 59.83 | 50.00 | 159 |
| 103 | DC89 2118.1 | P1 | TGTGAAGCTGCC TTTGTGTC | 20 | 60.03 | 50.00 | CATCTAACCCCCG TAGCTCA | 20 | 60.09 | 55.00 | 154 |
| 104 | DC89 2115.1 | P3 | AAGCAGGGAGTT TGGATGTG | 20 | 60.11 | 50.00 | CAGGAGTGGGAA GTTGGAGA | 20 | 60.23 | 50.00 | 204 |
| 105 | DC89 2841.1 | P1 | GGTTCCAAGACA CGGAAAAA | 20 | 59.95 | 45.00 | TATCCTGGAAAAC GGGATGA | 20 | 60.27 | 45.00 | 249 |
| 106 | DC89 2108.1 | P1 | AGTCCCCAAGGG AAAAAGAA | 20 | 59.91 | 45.00 | TCTGTAGCATTGC AGCGAGT | 20 | 59.77 | 50.00 | 202 |
| 107 | DC89 2824.1 | P2 | TAAATTGGGTCC GTGAGGTT | 20 | 59.29 | 45.00 | TCCGCTCAAATTA GGACCTG | 20 | 60.21 | 50.00 | 173 |
| 108 | DC89 2823.1 | P1 | TGATGGGATTGA GAGGTGGT | 20 | 60.33 | 50.00 | CTTGACTGGGATC CAATGCT | 20 | 60.07 | 50.00 | 195 |
| 109 | DC89 2094.1 | P1 | GAGAAGCAGCAA CAAGCACA | 20 | 60.33 | 50.00 | CTGTTGCATCAAT GGTGGAG | 20 | 60.11 | 50.00 | 183 |
| 110 | DC89 2087.1 | P2 | TGTGAAGAAGCC CACTGTTG | 20 | 59.87 | 50.00 | TTTCCTCCATTGC TTTTTGG | 20 | 60.05 | 40.00 | 236 |
| 111 | DC89 | P1 | TCTTTGGGAGTG | 20 | 60.0 | 50.0 | CTGCAGAAAGAA | 20 | 59. | 45.0 | 165 |
| | 2798. | | GCAATAGG | | 7 | 0 | TGCCACAA | | 99 | 0 | |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | | | | | | | | | | |
| 112 | DC89 | P3 | ATGGCTCGAGAA | 20 | 59.9 | 45.0 | TAATGTCCCCAGG | 20 | 59. | 55.0 | 161 |
| | 1526. | | ATCGAGAA | | 2 | 0 | CACTACC | | 81 | 0 | |
| | 1 | | | | | | | | | | |
| 113 | DC89 | P1 | ACATGCACTTAC | 20 | 60.0 | 45.0 | ATCGAGCACAGG | 20 | 59. | 50.0 | 174 |
| | 1502. | | CCATGCAA | | 0 | 0 | CCAAGTAT | | 72 | 0 | |
| | 1 | | | | | | | | | | |
| 114 | DC89 | P1 | GTACAGGCCGTG | 20 | 60.0 | 55.0 | GCCTTAGCATCTG | 20 | 60. | 55.0 | 184 |
| | 1468. | | TCTGGAAT | | 0 | 0 | CCTTGAG | | 12 | 0 | |
| | 1 | | | | | | | | | | |
| 115 | DC89 | P1 | CCAATGGCTTCCT | 20 | 59.8 | 45.0 | CACACCTTCATGC | 20 | 59. | 50.0 | 191 |
| | 1456. | | CAATGAT | | 9 | 0 | ATTGGAC | | 97 | 0 | |
| | 1 | | | | | | | | | | |
| 116 | DC89 | P1 | CAAGTGCAGCCA | 20 | 59.4 | 50.0 | CGTATTGCATCGG | 20 | 59. | 45.0 | 242 |
| | 1455. | | ATGCTATC | | 5 | 0 | AAGTTGA | | 69 | 0 | |
| | 1 | | | | | | | | | | |
| 117 | DC89 | P1 | GGTTGTTGGCTA | 20 | 59.7 | 50.0 | CCTGAATTTGGTT | 20 | 60. | 50.0 | 164 |
| | 2047. | | AGGCAAAG | | 5 | 0 | GGTGGAG | | 34 | 0 | |
| | 1 | | | | | | | | | | |
| 118 | DC89 2032. 1_ | P1 | CGGTACTCCTGC CATGACTT | 20 | 60.1 3 | 55.0 0 | GCCTCGATGTCCT TGTTGTT | 20 | 60. 12 | 50.0 0 | 165 |
| 119 | DC89 1987. 1 | P3 | ATCAATGGTTTG GCTTCTCG | 20 | 60.0 7 | 45.0 0 | CAACCATGGCAGC TACAGAA | 20 | 59. 86 | 500. 00 | 200 |
| 120 | DC89 1986. 1 | P3 | ACGGAGACGATG GTGAAAAC | 20 | 59.9 7 | 50.0 0 | CTCTCACCGTCGG ATTGATT | 20 | 60. 07 | 50.0 0 | 175 |
| 121 | DC89 1973. 1 | P1 | GATGCAACAGCT GATGAGGA | 20 | 59.9 5 | 50.0 0 | CAATTAGGCTTCT CCGCTTCT | 21 | 59. 99 | 47.6 2 | 170 |
| 122 | DC89 1957. 1 | P1 | TCTCTTCTGGCCC TCTTCAA | 20 | 60.0 7 | 50.0 0 | GAACAACGGCAG TCAACAAA | 20 | 59. 74 | 45.0 0 | 241 |
| 123 | DC89 1956. 1 | P1 | GGTGAGCCTGTC ATATTTCG | 20 | 58.1 7 | 50.0 0 | TTTGATTCCGGCA CTTTACC | 20 | 59. 94 | 45.0 0 | 178 |
| 124 | DC89 1986. 1 | P1 | TAGTGTTCGGAT GCTTGTGG | 20 | 59.7 2 | 50.0 0 | CTGAAGACAGCTT CCTGAGC | 20 | 57. 90 | 55.0 0 | 159 |
| 125 | DC89 1945. 1 | P1 | TAACCGATGGAA GGTTTTCG | 20 | 59.9 3 | 45.0 0 | CGATTGCACGTTC TTGATGT | 20 | 59. 72 | 45.0 0 | 248 |
| 126 | DC89 1944. 1 | P1 | GCGTTTTGGTAAT TGGGAGA | 20 | 59.9 4 | 45.0 0 | GAGTCAAACGAG GCAGGCTA | 20 | 60. 54 | 55.0 0 | 165 |
| 127 | DC89 2768. 1 | P1 | GCACTTCCTGAT GGGAAGAA | 20 | 60.2 0 | 50.0 0 | AAGCCAAACCAA CATCAACC | 20 | 59. 84 | 45.0 0 | 164 |
| 128 | DC89 2758. 1 | P2 | GAAAACCCTAAC GGGGAAGA | 20 | 60.2 9 | 50.0 0 | GAAGCCTCAGAA CCAGATGC | 20 | 59. 96 | 55.0 0 | 206 |

(Footnotes: [a] F.P – Forward Primer, [b] R.P – Reverse Primer, [c] TM - Melting Temperature, [d] P.S - Product Size)

process consists of alignment, mapping, annotation and so on of given sequences with the aid of using unique packages like BLAST, InterProScan etc (Figure 2). Under BLAST2GO the functional analysis of considered sequences were done through InterProScan program, as it uses the databases like pattern scan, Signal PHMM, TMHMM, HMM Panther, and FPrintScan for functional domain analysis [7]. Here, the associated metabolic pathways and the enzyme codes for the EST-SSRs were additionally studied in BLAST2GO via KEGG database as KEGG database is a collection of organic pathways, chemical materials, diseases, drugs and many others [16].

There were total 913 numbers of mapped and annotated GO terms were analyzed out of which 392 numbers of biological process, 389 numbers of molecular functions and 132 numbers of GO terms for cellular components were analyzed for 146 numbers of EST-SSR sequences.

**Biological processes**

A biological process (larger processes) is a series of events accomplished by one or more (multiple) ordered assemblies of molecular function. In biological process, the most frequently observed functions were, Translation (13 SSR-ESTs), Transmembrane transport (6 SSR-EST), Fatty acid biosynthetic

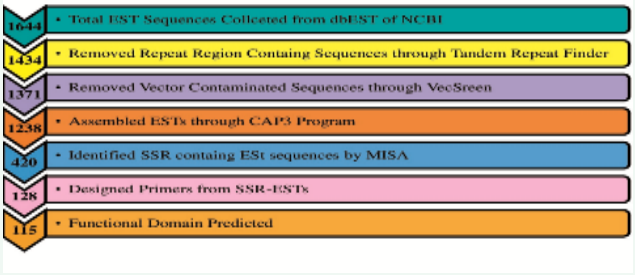


**Figure 1** The results of whole process, from sequence retrieval to Blast2Go analysis, carried out in the present work.

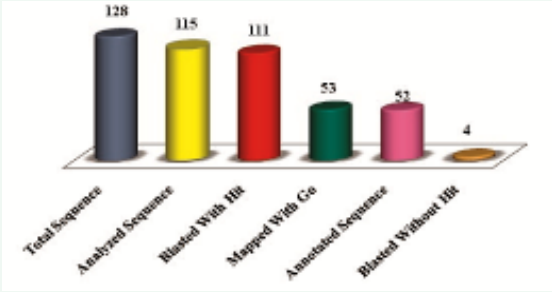


**Figure 2** Representation of Functional annotation of SSR-EST through Blast2Go.

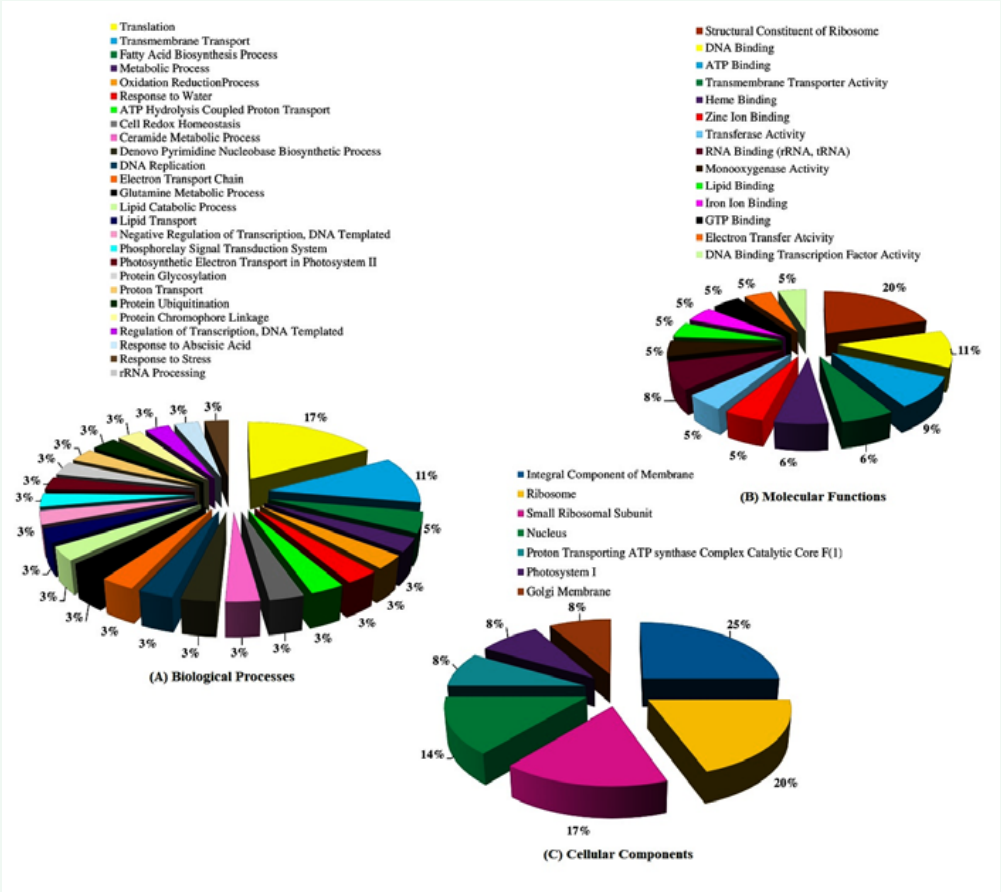


**Figure 3** Graphical interpretation of SSR-EST involved metabolic functions (a) the more efficiently occurred Biological Process (b) Majorly resulted Molecular Functions (c) Highly presented Cellular Components.

**Table 2:** Detected metabolic pathways and enzyme codes through Blast2Go KEGG pathway analysis.

| Sl. No. | Metabolic Pathway | Enzyme Code (EC Number) | Sequence ID |
|---|---|---|---|
| 1 | Pyrimidine metabolism | ec:2.1.1.45- synthase | DC891710.1 |
| | | ec:3.1.3.5-uridine5'-nucleotidase | DC891710.1 |
| | | ec:6.3.4.2- synthase (glutamine hydrolysing) | DC891710.1 |
| | | ec:3.6.1.9- diphosphatase | DC892585.1 |
| | | ec:2.7.7.6- RNA polymerase | DC891552.1, DC891730.1,DC891710.1 |
| | | ec:DNA polymerase | DC891710.1 |
| | | ec:6.3.5.5- synthase (glutamine hydrolysing) | DC891710.1 |
| | | ec:1.17.4.1- reductase | DC891710.1 |
| | | ec:2.7.4.14- kinase | DC892087.1 |
| 2 | Purine metabolism | ec:3.6.1.15- phosphatase | DC891710.1, GO256710.1, DC891552.1 |
| | | ec:kinase | DC891710.1 |
| | | ec:uridine 5'-nucleotidase | DC891710.1 |
| | | ec:3.6.1.3-adenylpyrophosphatase | DC891710.1, GO256710.1,DC892585.1, DC8991552.1 |
| | | ec:3.6.1.9-diphosphatase | DC892585.1 |
| | | ec:2.7.7.6- RNA polymerase | DC891552.1, DC891730.1, DC891710.1 |
| | | ec:2.7.7.7- DNA polymerase | DC891710.1 |
| | | ec:6.3.5.3- synthase | DC891710.1 |
| | | ec:2.4.2.14-phosphoribosyldiphosphate 5-amidotransferase | DC891710.1 |
| | | ec:1.17.4.1- reductase | DC891710.1 |
| | | ec:2.4.2.7-phosphoribosyltransferase | DC891710.1 |
| 3 | Oxidative phosphorylation | ec:1.9.3.1- oxidase | DC891918.1 |
| | | ec:1.6.99.3- dehydrogenase | DC892585.1, DC891552.1 |
| | | ec:1.6.5.3- reductase (H+-translocating) | DC891552.1 |
| | | ec:3.6.1.1- diphosphatase | DC891710.1 |
| | | ec:3.6.3.6- ATPase | DC891710.1 |
| 4 | Thiamine metabolism | ec:3.6.1.15- phosphatase | DC891710.1, GO256710.1, DC892585.1, DC891552.1 |
| | | ec:2.7.4.3- kinase | DC891710.1 |
| | | ec:2.2.1.7- synthase | DC891710.1 |
| 5 | Drug metabolism- other enzyme | ec:3.1.1.1- all-esterase | DC892585.1, DC891710.1 |
| | | ec:1.17.4.1- reductase | DC891710.1 |
| | | ec:2.4.1.17-1-napthanol glucuronyltransferase | DC891710.1 |
| | | ec:2.7.4.14- kinase | DC892087.1 |
| 6 | Glyoxylate and dicarboxylate metabolism | ec:2.1.2.1-hydroxymethyltransferase | DC891710.1 |
| | | ec:4.1.1.39- carboxylase | DC891552.1 |
| | | ec:1.11.1.6- equilase | DC891710.1 |
| 7 | Pyruvate metabolism | ec:1.8.1.4- dehydrogenase | DC891710.1 |
| | | ec:6.4.1.2-carboxylase | DC891552.1 |
| | | ec:4.4.1.5- lyase | DC891710.1 |
| | | ec:1.1.1.39- dehydrogenase (decarboxylating) | DC891710.1 |
| | | ec:1.1.1.38- dehydrogenase (oxaloacetate-decarboxylating ) | DC891710.1 |
| | | ec:3.1.2.6- hydrolase | DC891710.1 |
| | | ec:1.2.4.1- dehydrogenase (acetyl-transfering) | DC891710.1 |

| | | | |
|---|---|---|---|
| 8 | Biosynthesis of antibiotics | ec:1.8.1.4- dehydrogenase | DC891710.1 |
| | | ec:6.4.1.2- carboxylase | DC891552.1 |
| | | ec:2.1.2.1-hydroxymethyltransferase | DC891710.1 |
| | | ec:2.7.1.11-phosphohexokinase | DC891710.1 |
| | | ec:4.1.2.13- aldolase | DC891710.1 |
| | | ec:2.7.4.3-kinase | DC891710.1 |
| | | ec:2.7.4.2- kinase | DC891710.1 |
| | | ec:5.3.1.9- isomerase | DC891710.1 |
| | | ec:4.1.1.39- carboxylase | DC891552.1 |
| | | ec:3.1.3.25- phosphatase | DC891710.1 |
| | | ec:1.1.1.49- dehydrogenase (NADP+) | DC891710.1 |
| | | ec:6.3.5.3- synthase | DC891710.1 |
| | | ec:1.8.3.5- oxidase | DC891710.1 |
| | | ec:2.4.2.18-phosphoribosyltransferase | DC891710.1 |
| | | ec:2.2.1.6- synthase | DC891710.1 |
| | | ec:1.11.1.6- equilase | DC891710.1 |
| | | ec:2.4.2.14-phosphoribosyldiphosphate | DC891710.1 |
| | | 5-amidotransferase | |
| | | ec:2.5.1.54- synthase | DC891710.1 |
| | | ec:2.2.1.7- synthase | DC891710.1 |
| | | ec:2.7.2.8- kinase | DC891710.1 |
| | | ec:4.3.2.1- lyase | DC891710.1 |
| | | ec:1.2.4.1- dehydrogenase (acetyl- transferring) | DC891710.1 |
| | | ec:1.4.1.14- synthase (NADH) | DC891710.1 |
| | | ec:2.1.3.3-carbamoyltransferase | DC891710.1 |
| | | ec:3.1.1.31-phosphogluconolactonase | DC891710.1 |
| | | ec:5.4.99.5- mutase | DC891710.1 |
| | | ec:2.6.1.42- transaminase | DC891710.1 |
| 9 | Phenylalanine metabolism | ec:4.3.1.24- ammonia-lyase | DC892361.1 |
| | | ec:4.3.1.25- ammonia-lyase | DC892361.1 |
| | | ec:3.5.1.4- acylamidase | DC891710.1 |
| | | ec:1.4.3.21- oxidase | DC891710.1 |
| 10 | Aminobenzoate degradation | ec:3.1.3.41- nitrophenyl phosphatase | DC892585.1, DC891710.1 |
| | | ec:3.5.1.4- acylamidase | DC891710.1 |
| 11 | Starch and sucrose metabolism | ec:3.2.1.4- endo-1,4-beta-D-glucanase | DC891710.1 |
| | | ec:5.3.1.9- isomerase | DC891710.1 |
| | | ec:3.1.3.12- trehalose 6- phosphatase | DC891710.1 |
| | | ec:3.6.1.9-diphosphatase | DC892585.1 |
| | | ec:2.4.1.34-synthase | DC891710.1 |
| | | ec:2.7.7.27-adenylyltransferase | DC891710.1 |
| | | ec:2.4.1.18- branching enzyme | DC891710.1 |
| | | ec:2.4.1.13- synthase | DC891710.1 |
| | | ec2.4.1.12- synthase (UDP-forming) | DC891710.1 |
| | | ec:2.4.1.11- synthase | DC891710.1 |
| 12 | Th1 and Th2 zcell Differentiation | ec:3.1.3.16- phosphatase | DC892585.1, DC891710.1 |
| 13 | T cell receptor signaling pathway | ec:3.1.3.16- phosphatase | DC892585.1, DC891710.1 |
| 14 | Porphyrin and Chlorophyll | ec:1.2.1.70- reductase | DC891710.1 |
| | | ec:1.14.13.122- oxygenase | DC891710.1 |
| | Metabolism | ec:5.4.3.8-2,1-aminomutase | DC891710.1 |
| | | ec:42.1.75- synthase | DC891710.1 |
| | | ec:1.14.13.81--IX monomethyl ester (oxidative) cydase | DC1556.1 |
| | | ec:4.99.1- ferrochelatasse | DC891710.1 |
| | | ec:3.1.1.14- CLH | DC891710.1 |
| | | ec:4.2.1.24- synthase | DC891710.1 |
| | | ec:2.4.1.17- 1-naphthol glucuronyltransferase | DC891710.1 |
| | | ec:1.3.3.3- oxidase | DC891710.1 |

| | | | |
|---|---|---|---|
| 15 | Phenylpropanoid biosynthesis | ec:1.11.1.7-lactoperoxidase | DC891710.1 |
| | | ec:4.3.1.24- ammonia-lyase | DC89236.1 |
| | | ec:4.3.1.25-ammonia-lyase | DC89236.1 |
| 16 | Pantothenate and CoA biosynthesis | ec:3.6.1.9- diphosphatase | DC892585.1 |
| | | ec:2.2.1.6- synthase | DC891710.1 |
| | | ec:2.6.1.42- transaminase | DC891710.1 |
| 17 | Nicotinase and nicotinamide metabolism | ec:3.1.3.5- uridine 5'-nucleotidase | DC891710.1 |
| | | ec:3.6.1.9- diphosphatase | DC892585.1 |
| | | ec:2.7.1.23- kinase | DC891710.1 |
| 18 | Glycerophospholipid metabolism | ec:2.7.1.107- kinase (ATP) | DC891710.1 |
| | | ec:1.1.5.3- dehydrogenase | DC892585.1 |
| | | ec:3.1.4.4- D | DC891710.1 |
| 19 | Carbon fixation in photosynthetic organisms | ec:4.1.2.13- aldolase | DC891710.1 |
| | | ec:4.1.1.39- carboxylase | DC891552.1 |
| | | ec:1.1.1.39- dehydrogenase (decarboxylating) | DC891552.1 |
| 20 | Propanoate metabolism | ec:1.8.1.4- dehydrogenase | DC891710.1 |
| | | ec:6.4.1.2- carboxylase | DC891552.1 |
| | | ec:1.2.1.27- dehydrogenase (CoA-acylating) | DC891710.1 |
| 21 | Glycolysis/ gluconeogenesis | ec:1.8.1.4- dehydrogenase | DC891710.1 |
| | | ec:2.7.1.11-phosphohexokinase | DC891710.1 |
| | | ec:5.3.1.9- isomerase | DC891710.1 |
| | | ec:5.3.1.9- isomerase | DC891710.1 |
| | | ec:1.2.4.1- dehydrogenase (acetyl-transferring) | DC891710.1 |
| 22 | Mannose type O-glycan biosynthesis | ec:2.4.1.135- 3-beta-glucuronosyltransferase | Dc891710.1 |
| 23 | Glutathione | ec:1.1.1.49- dehydrogenase | DC891710.1 |
| | Metabolism | (NADP+) | |
| | | ec:4.3.2.9- gamma-giutamyl-amino acid cyclotransferase | DC891710.1 |
| | | ec:1.17.4.1- reductase | DC891710.1 |
| 24 | Tryptophan Metabolism | ec:1.11.1.6- equilase | DC891710.1 |
| | | ec:3.5.1.4- acylamidase | DC891710.1 |
| 25 | Glycosaminoglycan biosynthesis-heparan sulfate / Heparin | ec:2.4.1.135- 3-beta- glucuronosyltransferase | DC891710.1 |
| 26 | Cyanoamin acid metabolism | ec:2.1.2.1-hydroxymethyltransferase | DC891710.1 |
| 27 | Steroid degradation | ec:1.1.1.145-dehydrogenase | DC891710.1 |
| 28 | Valine, lucine and isolucine biosynthesis | ec:2.2.1.6- synthase | DC891710.1 |
| | | ec:2.6.1.42- transaminase | DC891710.1 |
| | Glycospingolipid biosynthesis- ganglio series | ec:3.2.1.23- lactase (ambigous) | DC891710.1 |
| 29 | Styrene degradation | ec:3.5.1.4- aclamidase | DC891710.1 |
| 30 | Valine, lucine and isolucine degradation | ec:1.8.1.4- dehydrogenase | DC891710.1 |
| | | ec1.2.1.27- dehydrogenase (CoA-acylating) | DC891710.1 |
| | | ec:2.6.1.42- transaminase | DC891710.1 |
| 31 | Galactose metabolism | ec:2.7.1.11-phosphohexokinase | DC891710.1 |
| | | ec:3.2.1.22- melibiase | DC891710.1 |
| | | ec:3.2.1.23- lactase (ambigous) | DC891710.1 |
| 32 | Amino sugar and nucleotide sugar metabolism | ec:5.3.1.9- isomerase | DC891710.1 |
| | | ec:2.7.7.27-adenyltransferase | DC891710.1 |
| | | ec:2.4.1.43- 4-alpha-galacuronosyltransferase | DC891710.1 |
| | | ec:3.2.1.14- ChcC | DC891710.1 |
| 33 | Biosynthesis of unsaturated fatty Acids | ec:1.14.19.3- 6- desaturase | GO256710.1 |
| 34 | Metabolism of xenobiorics by cytochrone P450 | ec:2.4.1.17- 1-naphthol glucuronyltransferase | DC891710.1 |

| | | | |
|---|---|---|---|
| 35 | mTOR signaling pathway | ec:2.7.11.24- protein kinase | DC891710.1 |
| 36 | Sphingolipid metabolism | ec:3.2.1.22- melibiase | DC891710.1 |
| | | ec:3.2.1.23- lactase (ambigous) | DC891710.1 |
| 37 | Aminoacyl-tRNA biosynthesis | ec:6.1.1.16- ligase | DC891710.1 |
| | | ec:6.1.1.11- ligase | DC891710.1 |
| | | ec:6.1.1.7- ligase | DC891710.1 |
| | | ec:6.1.1.2- ligase | DC891710.1 |
| 38 | Tyrosine metabolism | ec:1.44.3.21- oxidase | DC891710.1 |
| 39 | Histidine metabolism | ec:3.1.3.15- histidinol phosphate phosphatase | DC891710.1 |
| | | ec:5.3.1.16- isomerase | DC891710.1 |
| 40 | Tropane, piperidine and pyridine alkaloid biosynthesis | ec:1.4.3.21- oxidase | DC891710.1 |
| 41 | Phosphatidylinositol signaling system | ec:2.7.1.107- kinese (ATP) | DC891710.1 |
| | | ec:3.1.3.25- phosphatase | DC891710.1 |
| | | ec:2.7.1.68- 5-kinase | DC891710.1 |
| | | ec:2.7.1.158- 2-kinase | DC891710.1 |
| | | ec:2.7.1.159- 5/6-kinase | DC891710.1 |
| 42 | Cysteine and methionine metabolism | ec:5.3.1.23- isomerase | DC891710.1 |
| | | ec:2.6.1.42- transaminase | DC891710.1 |
| 43 | Folate biosynthesis | ec:1.5.1.3- reductase | DC891710.1 |
| 44 | Inositol phosphate metabolism | ec:3.1.3.25-phosphatase | DC891710.1 |
| | | ec:27.1.68- 5-kinase | DC891710.1 |
| | | ec:2.7.1.134- 1-kinase | DC891710.1 |
| | | ec:2.7.1.158- 2-kinase | DC891710.1 |
| | | ec:2.7.1.159- 5/6-kinase | DC891710.1 |
| 45 | Beta-alannine metabolism | ec:1.4.3.21- oxidase | DC891710.1 |
| 46 | Glycine, serine and threonine metabolism | ec:1.8.1.4- dehydrogenase | DC891710.1 |
| | | ec:2.1.2.1-hydroxymethyltransferase | DC891710.1 |
| | | ec:1.4.4.2- dehydrogenase (aminomethyl-transferring) | DC891710.1 |
| | | ec:1.4.3.21- oxidase | DC891710.1 |
| 47 | Linoleic acid metabolism | ec:1.14.19.3- 6-desaturase | GO256710.1 |
| 48 | Alanine, aspartate and glutamate metabolism | ec:6.3.5.5- synthase (glutamine-hydrolysing) | DC891710.1 |
| | | ec:2.4.2.14-phosphoribosyldiphosphate 5-amidotransferase | DC891710.1 |
| | | ec:4.3.2.1- lyase | DC891710.1 |
| | | ec:1.4.1.14- synthase(NADH) | DC891710.1 |
| 49 | Fatty acid biosynthesis | ec:6.4.1.2- carboxylase | DC891552.1 |
| 50 | Glycosphingolipid biosynthesis - globo and isogloboseries | ec:2.4.1.69- 1 galactoside alpha-(1,2)-fucosyltransferase | DC891710.1 |
| | | ec:3.2.1.22-melibiase | DC891710.1 |
| 51 | Fructose and mannose metabolism | ec:2.7.1.11-phosphohexokinase | DC891710.1 |
| | | ec:4.1.2.13- aldolase | DC891710.1 |
| 52 | Arginine and purine metabolism | ec:4.1.1.19- decarboxylase | DC891710.1 |
| | | ec:3.5.1.4- acylamidase | DC891710.1 |
| 53 | Aflatoxin biosynthesis | ec:6.4.1.2- carboxylase | DC891552.1 |
| 54 | Isoquinoline alkaloid biosynthesis | ec:1.4.3.21- oxidase | DC891710.1 |
| 55 | Indole alkaloid biosynthesis | ec:4.3.3.2- synthase | DC891710.1 |
| 56 | Lysine degradation | ec:2.1.1.43- N-methyltransferase | DC891710.1 |
| 57 | Biotin metabolism | ec:6.3.4.15- carboxyl-carrier protein] ligase | DC891710.1 |
| 58 | Phenylalaline, tyrosine and tryptophan biosynthesis | ec:2.4.2.18-phosphoribosyltransferase | DC891710.1 |
| | | ec:2.5.1.54- synthase | DC891710.1 |
| | | ec:5.4.99.5- mutase | DC891710.1 |

| | | | |
|---|---|---|---|
| 59 | Steroid hormone biosynthesis | ec:1.1.1.145-dehydrogenase | DC891710.1 |
| | | ec:2.4.1.17- 1-naphthol glucuronyltransferase | DC891710.1 |
| 60 | Glycosaminoglycan biosynthesis-chondroitin sulfate / dermatan sulfate | ec:2.4.1.135- 3-beta- glucuronosyltransferase | DC891710.1 |
| 61 | Vitamin B6 metabolism | ec:1.4.3.5- 5'-phosphate synthase | DC891710.1 |
| 62 | Glycerolipid metabolism | ec:2.7.1.107- kinase (ATP) | DC891710.1 |
| | | ec:3.2.1.22- melibiase | DC891710.1 |
| | | ec:2.3.1.20- O-acyltransferase | DC891710.1 |
| 63 | Arginine | ec:2.7.2.8- kinase | DC891710.1 |
| | Biosynthesis | ec:4.3.2.1- lyase | DC891710.1 |
| | | ec:2.1.3.3-carbamoyltransferase | DC891710.1 |
| 64 | Arachidonic acid metabolism | ec:5.3.99.3-- synthase | DC891710.1 |
| 65 | Drug metabolism- cytochrome P450 | ec:1.14.13.8-monooxygenase | DC891710.1 |
| | | ec:2.4.1.17- 1-naphthol glucuronyltransferase | DC891710.1 |
| 66 | Retinol metabolism | ec:2.4.1.17- 1-naphthol glucuronyltransferase | DC891710.1 |
| 67 | Methane metabolism | ec:2.1.2.1-hydroxymethyltransferase | DC891710.1 |
| | | ec:2.7.1.11-phosphohexokinase | DC891710.1 |
| | | ec:4.1.2.13- aldolase | DC891710.1 |
| 68 | Ether lipid metabolism | ec:3.1.4.4- D | DC891710.1 |
| 69 | One carbon pool by folate | ec:6.3.3.2- cyclo-ligase | DC891710.1 |
| | | ec:2.1.2.1-hydroxymethyltransferase | DC891710.1 |
| | | ec:2.1.1.45- synthase | DC891710.1 |
| | | ec:1.5.1.3- reductase | DC891710.1 |
| 70 | Zeatin biosynthesis | ec:1.5.99.12-dehydrogenase | DC891710.1 |
| 71 | Nitrogen metabolism | ec:4.2.1.1- anhydrase | DC891710.1 |
| | | ec:1.7.1.3- reductase (NADPH) | DC891710.1 |
| | | ec:1.4.1.14- synthase (NADH) | DC891710.1 |
| 72 | C5-branched dibasic acid metabolism | ec:2.2.1.6- synthase | DC891710.1 |
| 73 | Pentose and glucuronate interconversions | ec:4.2.2.2- lyase | DC891710.1 |
| | | ec:3.1.1.11- pectin demethoxylase | DC891710.1 |
| | | ec:2.4.1.17- 1-naphthol glucuronyltransferase | DC891710.1 |
| | | ec:3.2.1.15- pectin depolymerase | DC891710.1 |
| 74 | Terpinoid backbone biosynthesis | ec:2.7.4.2- kinase | DC891710.1 |
| | | ec:1.8.3.5- oxidase | DC891710.1 |
| | | ec:2.2.1.7- synthase | DC891710.1 |
| 75 | Butanoate metabolism | ec:2.2.1.6- synthase | DC891710.1 |
| 76 | Glucosinolate | ec:2.6.1.42- transaminase | DC891710.1 |
| | Biosynthesis | | |
| 77 | Riboflavin metabolism | ee:3.6.1.9- diphosphatase | DC892585.1 |
| 78 | Cutin, suberine and wax biosynthesis | ec:2.3.1.20- O-acyltransferase | DC891710.1 |
| 79 | Steroid biosynthesis | ec:5.5.1.9- cycloisomerase | DC891710.1 |
| 80 | Carbon fixation pathways in prokaryotes | ec:6.4.1.2- carboxylase | DC891552.1 |
| 81 | Glycosaminoglycan degradation | ec:3.2.1.23- lactase (ambiguous) | DC891710.1 |
| 82 | Ascorbate and aldarate metabolism | ec:2.4.1.17- 1-naphthol glucuronyltransferase | DC891710.1 |

| | | | |
|---|---|---|---|
| 83 | Streptomycin biosynthesis | ec:3.1.3.25- phosphatase | DC891710.1 |
| 84 | Pentose phosphate pathway | ec:2.7.1.11- phosphohexokinase | DC891710.1 |
| | | ec:4.1.2.13- aldolase | DC891710.1 |
| | | ec:5.3.1.9- isomerase | DC891710.1 |
| | | ec:1.1.1.49- dehydrogenase (NADP+) | DC891710.1 |
| | | ec:3.1.1.31- phosphogluconolactonase | DC891710.1 |
| 85 | Other glycan degradation | ec:3.2.1.24- alpha-D-mannosidase | DC891710.1 |
| | | ec:3.2.1.23- lactase (ambiguous) | DC891710.1 |
| 86 | Citrate cycle (TCA cycle) | ec:1.8.1.4- dehydrogenase | DC891710.1 |
| | | ec:1.2.4.1- dehydrogenase (acetyl-transferring) | DC891710.1 |
| 87 | Glycosphigolipid biosynthesis- lacto and neolacto series | ec:2.4.1.69- 1 galactoside alpha-(1,2)-fucosyltransferase | DC891710.1 |

process (3 SSR-EST), Metabolic process (3 SSR- EST), Oxidation-reduction process (3 SSR-EST), Response to water (3 SSR-EST), ATP hydrolysis couple proton transport (2 SSR-EST), Cell redox homeostasis (2 SSR-EST), Ceramide metabolic process, Denovo pyrimidine nucleobase biosynthetic process (2 SSR-EST), DNA replication (2 SSR-EST), Electron transport chain (2 SSR- EST), Glutamine metabolic process (2 SSR-EST), Lipid catabolic process (2 SSR-EST), Lipid transport (2 SSR- EST), Negative regulation of transcription, DNA-templated (2 SSR-EST), Phosphorelay signal transduction system (2 SSR-EST), Photosynthetic electron transport in photosystem II (2 SSR-EST), Protein glycosylation (2 SSR-EST), Proton transport (2 SSR-EST), Protein ubiquitination (2 SSR-EST), Protein-chromophore linkage (2 SSR-EST), Regulation of transcription, DNA template (2 SSR-EST), Response to abscisic acid (2 SSR-EST), Response to stress (2 SSR-EST), RNA processing (2 SSR-EST). The remaining markers were involved in a less amount of Biological process and also those processes were occurred in less number, so the most frequently occurred biological processes were taken in to consideration (Figure 3a).

### Molecular functions

Molecular function describes the actions or activities that a gene product (or a complex) performs. Here, in molecular function, the most frequent resulted functions were as follows; Structural constituent of ribosome (13 SSR-EST), DNA binding (7 SSR-ESTs) ATP binding (6 SSR-ESTs), RNA (rRNA, tRNA) binding (5 SSR-EST), Transmembrane transporter activity (4 SSR-EST), heme binding (4 SSR-EST), Zinc ion binding (3 SSR-EST), Transferase activity (3 SSR-EST), Monooxygenase activity (3 SSR-EST), Lipid binding (3 SSR-EST), Iron ion binding (3SSR-EST), GTP binding (3 SSR-EST), Electron transfer activity (3 SSR-EST), DNA binding transcription factor activity (3 SSR-EST) (Figure 3b).

### Cellular components

Cellular component is a component of cell, but with the provision that it is part of some larger object. This study meet the most frequently observed cellular components were, Integral component of membrane (9 SSR-ESTs), Ribosome (6 SSR-ESTs), Nucleus (6 SSR-ESTs), Small ribosomal subunit (5 SSR-ESTs), Proton transporting ATP synthase complex, Catalytic core F (1) (3 SSR-ESTs), Photosystem I (3 SSR-ESTs), Golgi membrane (3 SSR-ESTs) (Figure 3c).

The SSR-ESTs after FDM assessment had been further analyzed in Blast2Go for EC mapping and then figuring out its KEGG pathways. The EC mapping and KEGG pathway enrichment assessment resulted 87 numbers of metabolic pathways and the enzyme codes for 146 numbers of EST-SSR sequences i.e. Multiple quantity of sequences were involved in exclusive metabolic pathways and additionally one sequence have a couple of number of enzyme codes. This analysis summarized that the enzyme code ec: 3.6.1.15-phosphatase corresponds to Thiamine metabolism pathway with involvement of maximum 4 numbers of SSR-ESTs (Table 2). This prediction may leads to gather the information regarding the involvement of selected primers from SSR-ESTs of *C. limon* with different metabolic pathways.

## CONCLUSION

*Citrus limon* is an ever inexperienced plant with high rate of medicinal value. Microsatellites or SSRs play a prime function in polymorphism analysis and in marker assisted selection. *In silico* approach for predicting SSRs within the complete genome, was observed to be both cost and time effective and additionally helps to increase a novel generation of molecular markers as well. So, this study exhibits 420 EST-SSR sequences which give 128 tremendous primers which will probably beneficial for genetic mapping, gene populace examine and so on. Also the functional domain analysis or GO annotation of resulted EST-SSRs can provide statistics concerning the putative functions of transcribed genetic markers, which might have the way for future studies in the aspect of breeding and genetic studies of *Citrus limon* plant and its functional characterization.

## ACKNOWLEDGEMENT

We really thanks to Tectona Biotech Resource Centre (TBRC), Odisha, for providing necessary facilities, as well as we present our heartily gratitude to Dr. Shovan Kumar Mishra, Director of TBRC, for his endless support throughout this work.

## REFERENCES

1. Aslam MS, Ahmad MS. Worldwide Importance of Medicinal Plants: Current and Historical Perspectives. Recent Adv Biol Med. 2016; 2: 88-93.
2. Panda S, Naik D, Kamble A. Population structure and genetic diversity of the perennial medicinal shrub Plumbago. AoB Plants. 2015; 7.
3. Mohamed M, Jemai R, Kriaa W, Chtourou Y, Fetoui H. Citrus Limon

from Tunisia: Phytochemical and Physicochemical Properties and Biological Activities. BioMed Res Int. 2018; 1-10.

4. Bhattacharyya U, Pandey SK, Dasgupta T. Identification of EST-SSRs and FDM in sesame (Sesamum indicum L.) through data mining. Scholarly J Agri Sci. 2014; 4: 60-69.

5. Kaliswamy P, Vellingiri S, Nathan B, Selvaraj S. Microsatellite analysis in the genome of Acanthaceae: An in silico approach. Pharmacogn Mag. 2015; 11: 152-156.

6. Hynniewta M, Malik SK, Rao SR. Genetic diversity and phylogenetic analysis of Citrus (L) from north-east India as revealed by meiosis, and molecular analysis of internal transcribed spacer region of rDNA. Meta Gene. 2014; 2: 237-251.

7. Sorkheh K, Prudencio AS, Ghebinejad A, Dehkordi MK, Erogul D, Rubio M, et al. In silico search, characterization and validation of new EST-SSR markers in the genus Prunus. BMC Res Notes. 2016; 9: 336.

8. Usdin K. The biological effects of simple tandem repeats: Lessons from the repeat expansion diseases. Genome Res. 2008; 18: 1011-1019.

9. Ranjan A, Kumari A, Pandey DM. Annotation of stress responsive candidate genes in peanut ESTs. Interdiscip Sci Comput Life Sci. 2014; 1-9.

10. Huang X, Madan A. CAP3: A DNA Sequence Assembly Program. Genome Res. 1999; 9: 868-877.

11. Beier S, Thiel T, Munch T, Scholz U, Mascher M. MISA-web: a web server for microsatellite prediction. Bioinfor. 2017; 33: 2583-2585.

12. Sahu J, Sarmah R, Dehury B, Sarma K., Sahoo S, Sahu M. Mining for SSRs and FDMs from expressed sequence tags of Camellia sinensis. Bioinformation. 2012; 8: 260-266.

13. Guo Q, Wang JX, Su LZ, Lv W, Sun YH, Li Y. Development and Evaluation of a Novel Set of EST-SSR Markers Based on Transcriptome Sequences of Black Locust (*Robinia pseudoacacia* L.). Genes. 2017; 8: 177.

14. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, et al. Primer3-new capabilities and interfaces. Nucleic Acids Res. 2012; 40: e115.

15. Götz S, García-Gómez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ. High-throughput functional annotation and data mining with the Blast2GO suite. Nucleic Acids Res. 2008; 36: 3420-3435.

16. Flores-Bautista E, Cronick CL, Fersaca AR, Martinez-Nuñez MA, Perez-Rueda E. Functional Prediction of Hypothetical Transcription Factors of *Escherichia coli* K-12 Based on Expression Data. Comput Struct Biotechnol J. 2018; 16: 157-166.

**Cite this article**

*Ray M, Pradhan SP, Sahoo S (2019) Computational in Sight into Identification and Analysis of SSR-FDM in Citrus limon. Int J Plant Biol Res 7(1): 1111.*