Insights from the cDNA and EST analysis of *Antrodia cinnamomea*

Jing Li¹, Yanling Liu¹, Richard Yankey¹,², Dongmei Lin¹ & Zhanxi Lin¹*

¹National Engineering Research Center of JUNCAO Technology, Fujian Agriculture and forestry University, Fujian, Fuzhou 350002, PR China; ²Life Science College, Fujian Agriculture and Forestry University, Fujian, Fuzhou 350002, PR China; Zhanxi Lin - Email ID: lzxjuncao@163.com; *Corresponding author

Received July 21, 2020; Revised December 31, 2020; Accepted January 26, 2021, Published January 31, 2021

DOI: 10.6026/97320630017109

Abstract:
It is of interest to document the insights gleaned from the cDNA and EST analysis of *Antrodia cinnamomea* (a fungal species). Hence a library of sequences was constructed and analysed using standard procedures to gain new insights. Therefore, 65 ESTs, with size ranging from 300-2000 bp, were constructed. This included 46 ESTs with definite annotation, 18 ESTs were hypothetical and 1 new protein derived from BLAST analysis. We assigned 227 Gene Ontology terms linked to cell composition, transport, catalytic activity, and regulation functions in these sequences. Moreover, 56 matching genes were found in 8 Kyoto Encyclopedia of Genes and Genomes pathways. Data also showed 271 SSRs from *Antrodia cinnamomea* ESTs with an occurrence frequency of 96.82%. The STRING data analysis showed 29 genes encoded enzymes highly involved in protein-to-protein interactions linked to expression of regulation function. Thus, we documented some insights from the cDNA and EST analysis of *Antrodia cinnamomea* for further data mining.

Keywords: *Antrodia cinnamomea*; genome wide analysis; cDNA; EST; GO functional annotation;

Background:
*Antrodia cinnamomea* is a rare precious medical fungus found mainly in Taiwan China, which is well known for its immunomodulatory, anti-inflammation and anti-cancer pharmacological attributes [1]. Notably, triterpenoids and sterols are the most classical components and are biosynthesized via the mevalonic acid (MVA) pathway [2]. However, the relative genes and protein involved in the metabolite biosynthesis pathway of *A. cinnamomea* have rarely been studied. Less molecular information and EST sequences of *A. cinnamomea* were retrieved in the NCBI database. Complementary DNA (cDNA) library construction application for developing expressed sequence tags (ESTs) and transcriptome, was first reported in 1994 [3]. Gateway technology is a direct approach to speed-up functional open reading frame analysis [4], and the cDNA fragment is directly recombined into the destination vector, which is highly efficient, simplified cloning, chimeric clones reduced, less size bias and easily clone to a large extent, accurately reflecting the expression level of the original abundance of mRNA in the library. Simple sequence repeats (SSRs) are wide range extensive length polymorphisms used in selection of marker-assisted, genetic diversity print-finger, genetic mapping or breeding applications. SSRs are usually made of 1 to 6 repeat
nucleotides, which are composed in tandem and are widely distributed in the coding regions and non-coding regions of eukaryotic genes [5]. Screening of SSRs from the cDNA library is simple, inexpensive, and sequence-consistent, hence it is widely used in various plants and fungi such as rice, coffee, beans, rubber tree [6].

Methodology:
Fungi materials:
A cinnamomea AC001 (Genbank NO: KM925002) was donated by Taiwan Shennong Fungus Biotechnology Co., Ltd, and stored in the National Engineering Research Center of JUNCAO Technology of Fujian Agriculture and Forestry University. The strain was inoculated into a PDA plate and placed at a constant temperature of 28°C, then transferred to a liquid medium i.e., shaken at 28°C, 120 rpm in the dark, and cultured for 14, 21, 28, and 35 days. The mycelium was uniformly mixed after 35 days and used to construct a full-length cDNA library.

Isolation Kit (Thermo, USA). cDNA synthesis was performed by the CloneMiner II cDNA Library Construction Kit (Thermo, USA).

Full-length cDNA library construction:
The cDNA library of A. cinnamomea was constructed by using the Gateway method[9]. Briefly, the cDNA was used as a template to synthesize the first strand and the second strand of cDNA. The cDNA obtained by transcription was linked to three different reading frames of the adaptor, and Homologous Recombination performed the recombinant method. The primary cDNA library was prepared by cloning into the pDONR222 vector. Then, the primary cDNA library mixed plasmid was extracted, and a yeast two-hybrid cDNA library (secondary library) was prepared by LR recombination into the yeast vector pGADT7-DEST. Library identification and determination. The primary and secondary libraries of the obtained cDNA were diluted in a ratio of 1:1000 and coated with a plate. After incubating for 10 h at 37°C, single colonies were observed and counted. The titer of the amplified library was calculated as follows: (number of plaques × dilution factor × 10^3 μL/mL)/(diluted phage plated μL). The dilution factor is 1 × 10^4. The inserted sequences were randomly selected for PCR amplification using pDONR222-F primers: 5'-TCCCAATACGAGCTCTTAAAGCGAGCCGCGATGTT-3', pDONR222-R primers: 5'-AGAGCTGCGGAGAAACACGCTATGAC CGAGTATACGACTC-3'. The total volume of PCR was 20 μL, containing 1 μL template, 10 μL 2xPCR Master Mix (Thermo, USA), 1 μL each primer and 7 μL ddH2O. Cycling conditions were as follows: 94°C for 5 min, followed by 29 cycles of 95°C for 30 s, 57°C for 40 s and 72°C for 60 s, followed by 72°C for 5 min.

Bioinformatic analysis:
The clone transformants were randomly selected and sequenced by Bio-Sun Biotechnology Co., Ltd (Fujian, China). The adapter sequences were trimmed by Vecscreen tool (www.ncbi.nlm.nih.gov/tools/vecscreen) of National Center for Biotechnology Information (NCBI) and the inaccurate bases, poly-A tails, low-quality sequences (<100 bp) and other fragments [10] by CAP3 software removed to finally obtain more accurate sequencing. The NCBI BLAST program was used to compare the spliced single copy sequence with the nucleic acid library (NT) and protein (NR) databases (www.ncbi.nlm.nih.gov/) to complete the homology alignment of BLASTX. The database results with E value <1×10^-5 were generally regarded as a significant match [11]. NCBI's non-redundant protein database followed by the assignment of functionality via Gene Ontologies (GO) [12] using BLAST2GO (www.blast2go.com/) [13]. The GO analysis included functional classification of Molecular Function, Cellular Component, and Biological Process for EST sequence data by BLAST2GO software. Pathway assignments were mapped according to Kyoto Encyclopedia of Genes and Genomes (KEGG) database, while interaction protein and nucleic databases was conducted for biological functions by the STRING network (www.string-db.org/).

Figure 1: Electrophoresis of total RNA and mRNA of A. cinnamomea mycelium

Total RNA extraction and integrity assessment:
The total RNA of A. cinnamomea was extracted by TRIzol method [8], and the total RNA integrity and quality analysis detected by electrophoresis on a 1% agarose gel and Nanodrop2000C (Thermo, USA). The mRNA was isolated by the FastTrack MAG Maxi mRNA

Table 1: Electrophoresis of total RNA and mRNA of A. cinnamomea mycelium

| Time (s) | Sample |
|---------|--------|
| 18       | A      |
| 28       | B      |

Bioinformation 17(1): 109-118 (2021)
EST-SSR analysis:
Screening of EST-SSR was performed using the GRAMENE SSR online tool (https://archive.gramene.org/db/markers/ssrtool) with the standard length for searching for SSR including dinucleotide, trinucleotide, tetranucleotide, pentanucleotide, hexanucleotide with a minimum number of repetitions of 7, 5, 4, 4, 3. The SSR frequency and length were statistically analyzed.

Results & Discussion:
Total RNA assessment:
The total RNA extracted from the mycelium of A. cinnamomea was detected to be 963 ng/µL, the total amount was 481 µg, and the OD260/OD280 was 2.14, which indicated that the total RNA was integrated and stable enough for cDNA library construction. The integrity was measured by 1% agarose gel electrophoresis as shown in Figure 1A, and mRNA was further isolated by total RNA isolation (Figure 1B).

Construction of cDNA library:
The full-length cDNA primary library was transformed into a plate dilution of 1:1000, the library titer was 5.2×10⁶ CFU/mL, and the total number of clones was 1.0×10⁷ CFU (Figure 2A). The secondary library recombination rate was as high as 95%. The transformation plate dilution is 1:1000, the library titer is 6.1×10⁶ CFU/mL, Total number of clones is 1.2×10⁷ CFU (Figure 2C). The clones were selected randomly and amplified by PCR using pDONR222-F and pDONR222-R primers. The PCR products were detected by 1% agarose gel electrophoresis and the recombinant rate was 95%. The insert size is 300-2000 bp, and the average length is 1000 bp (Figures 2B and D), indicating that the library contains relatively long cDNAs that meet library capacity requirement. This method directly used cDNA as a template to synthesize cDNA double strands without any amplification process, and effectively kept the genetic information of full-length gene function. It also afforded an important resource for information about genes at the transcriptional level. The fragment is long enough to reflect the natural structure of the gene as much as possible, thus making it easier to obtain the complete sequence and functional information of the target gene in the library [14] and 95% of the recombinant cDNA library indicating the quality and providing a basis of the depth of the study.
EST analysis:
Several clones were randomly picked from the cDNA library for sequencing, and a total of 65 valid single sequences (unigene) were obtained after removing the vector sequence and low-quality sequences (<100 bp). As shown in Figure 3, 14 ESTs length were 500-999 bp, 21 ESTs length were 1000-1499 bp, 24 ESTs length were 1500-1999 bp and 2 ESTs length were >2000 bp. Among them, the longest EST was 3474 bp, and the shortest EST was 573 bp. The CAP3 program was spliced to obtain a unique sequence consisting of one contigs and 64 single singlets. A total of 65 individual ESTs were analyzed and submitted in GenBank (Table 1). The spliced sequence was subjected to homology alignment and gene function annotation with the NR database using NCBI BLASTX program. 65 single sequences except one (1.5%) had no significant homology; the remaining 64 (98.46%) had significant homologous sequences, including 46 known functional proteins and 18 unknown proteins. The homologous proteins were observed in Taiwanofungus camphoratus, Grifola frondosa, Wolfiporia cocos and fibroporia radiculosa etc. Obviously, 1 (Genbank NO.: MN205389) of the ESTs of the constructed library was noted to be novel and uncharacterized, and putatively expressed proteins with no homology matches of any public databases. Thus, its functional characterization and possible role in A. cinnamomea will be worthwhile to investigate.

Table 1: Comparison between ESTs with protein database of a cinnamomea mycelium cDNA library

| No. | Name                          | Accession No. | Max Score | E-value | Identities           |
|-----|-------------------------------|---------------|-----------|---------|----------------------|
| 1   | NAD-dependent formate dehydrogenase | MN205326      | 702       | 0       | 86.12%               |
| 2   | Mitochondrial carrier          | MN205327      | 484       | 5e-168  | 84.75%               |
| 3   | 1-Cys peroxiredoxin            | MN205328      | 462       | 7e-164  | 100%                 |
| 4   | NAD P-binding protein          | MN205329      | 253       | 6e-92   | 66.84%               |
| 5   | Phosphatidylinositol synthase   | MN205330      | 781       | 0       | 87.92%               |
| 6   | Carboxylic acid transporter    | MN205331      | 751       | 0       | 74.79%               |
| 7   | Predicted protein              | MN205332      | 103       | 3e-21   | 47.37%               |
| 8   | Alpha-ketoacid dehydrogenase    | MN205333      | 646       | 0       | 77.18%               |
| 9   | Acetyl-coenzyme a synthetase   | MN205334      | 1205      | 0       | 90.94%               |
| 10  | Dipeptidyl-peptidase 5         | MN205335      | 441       | 0       | 85.04%               |
| 11  | Zn-dependent exopeptidase      | MN205336      | 590       | 0       | 75.79%               |
| 12  | Methionine adenosyltransferase | MN205337      | 722       | 0       | 93.62%               |
| 13  | Ubiquitin-protein ligase       | MN205338      | 201       | 3e-56   | 53.37%               |
| 14  | Sorbitol dehydrogenase         | MN205339      | 709       | 0       | 91.01%               |
| 15  | Hypothetical protein LABUDRAFT_672451 | MN205340      | 372       | 9e-124  | 67.48%               |
| 16  | Hypothetical protein LABUDRAFT_306711 | MN205341      | 218       | 3e-68   | 68.92%               |
| 17  | Cytochrome P450 monoxygenase   | MN205342      | 1075      | 0       | 94.93%               |
| 18  | Hypothetical protein SCP_090480 | MN205343      | 114       | 2e-26   | 43.46%               |
| 19  | UMP-CMP kinase                 | MN205344      | 329       | 1e-109  | 71.03%               |
| 20  | Vacuolar amino acid transporter| MN205345      | 356       | 1e-116  | 87.50%               |
| 21  | U3 small nuclear RNA-associated protein  | MN205346      | 564       | 0       | 66.34%               |
| 22  | Hypothetical protein OBRIDRAFT | MN205347      | 181       | 7e-47   | 62.22%               |
| 23  | DUF1909-domain-containing protein | MN205348      | 131       | 8e-35   | 85.14%               |
| 24  | NAP-domain-containing protein  | MN205349      | 211       | 8e-10   | 80.93%               |
| 25  | Predicted protein              | MN205350      | 339       | 3e-112  | 64.48%               |
| 26  | Hypothetical protein POSPLADRAFT | MN205351      | 545       | 0       | 74.93%               |
| 27  | Hypothetical protein HYPPIDRAFT_29335 | MN205352      | 89        | 1e-17   | 70.31%               |
| 28  | Kynureninase                   | MN205353      | 701       | 0       | 77.85%               |
| 29  | Cytochrome P450                | MN205354      | 624       | 0       | 62.18%               |
| 30  | Mitochondrial genome maintenance| MN205355      | 282       | 1e-91   | 90.41%               |
| 31  | NADH dehydrogenase             | MN205356      | 595       | 0       | 78.36%               |
| 32  | Predicted protein              | MN205357      | 358       | 0       | 83.2%                |
| 33  | Indoleamine 2,3-dioxygenase    | MN205358      | 597       | 0       | 72.12%               |
| 34  | Pyruvate kinase                | MN205359      | 938       | 0       | 87.35%               |
| 35  | Predicted protein              | MN205360      | 282       | 2e-91   | 64.66%               |
| 36  | Subtilisin like protease       | MN205361      | 760       | 0       | 88%                  |
| 37  | Tubulin alpha-1a chain         | MN205362      | 795       | 0       | 98.07%               |
| 38  | 14-3-3-1 protein               | MN205363      | 508       | 3e-180  | 98.02%               |
| 39  | Dynin light chain, flagellar outer arm | MN205364      | 193       | 4e-60   | 91.89%               |
| 40  | ATP synthase II gamma          | MN205365      | 543       | 0       | 88.18%               |
| 41  | Kinase-like protein            | MN205366      | 664       | 0       | 81.00%               |
| 42  | Alpha/beta-hydrolase           | MN205367      | 600       | 0       | 74.35%               |
| 43  | Glycoside hydrolase family 5 protein | MN205368      | 323       | 3e-98   | 67.5%                |
| 44  | PpCYP74 mRNA                    | MN205369      | 590       | 0       | 62.29%               |
| 45  | Hypothetical protein MPER_04035 | MN205370      | 73.6       | 9e-13   | 85.17%               |
| 46  | Predicted protein              | MN205371      | 607       | 0       | 70.13%               |
| 47  | Calreticulin domain containing protein | MN205372      | 794       | 0       | 84.07%               |
Table 2: KEGG Pathway analysis of ESTs

| KEGG ID   | Term description                      | Gene count | False discovery rate | Matching proteins                                         |
|-----------|---------------------------------------|------------|----------------------|-----------------------------------------------------------|
| Sce0008   | Ribosom e biogenesis in eukaryotes     | 10         | 1.80E-08             | BMS1, IMP4, MPP10, NOP56, NOP58, PWP2, UTP13,              |
|           |                                       |            |                      | UTP14, UTP15, UTP4, AC52, AD01, ATP1, ATP6, ATP2, ATP4,   |
|           |                                       |            |                      | ATP5, ATP7, BNA5, CDC19, FDH1, GLN1, PDA1, PDB1,           |
|           |                                       |            |                      | SAM2, SHM2, URA6, XYL2                                   |
| Sce0110   | Metabolic pathways                    | 19         | 2.07E-05             | ATP1, ATP6, ATP2, ATP3, ATP4, AT5P, BNA5, CDC19, FDH1,   |
|           |                                       |            |                      | GLN1, PDA1, PDB1, SAM2, SHM2                             |
| Sce0190   | Oxidative phosphorylation             | 7          | 2.39E-05             | ATP1, ATP6, ATP2, ATP3, ATP4, AT5P, BNA5, CDC19, FDH1,   |
|           |                                       |            |                      | GLN1, PDA1, PDB1, SAM2, SHM2                             |
| Sce0200   | Carbon metabolism                     | 6          | 0.0023               | ACS2, CDC19, FDH1, SHM2                                  |
| Sce0260   | Pyruvate metabolism                   | 4          | 0.0029               | ACS2, CDC19, PDA1, PDB1                                  |
| Sce0607   | Glycolysis / gluconeogenesis          | 4          | 0.0058               | ACS2, CDC19, PDA1, PDB1                                  |
| Sce0680   | Methane metabolism                    | 3          | 0.0058               | ACS2, FDH1, SHM2                                         |
| Sce0690   | Glyoxylate and dicarboxylate metabolism| 3         | 0.0068               | FDH1, GLN1, SHM2                                         |

Supplementary Figure 1: Classification of identified genes based on the relevant biological process (A), Molecular Fucction (B) and cellular components (C).

Table 3: Repeat motifs in EST-SSRs of A. cinnamomea

| Items               | NO. | SSR                      |
|---------------------|-----|--------------------------|
| Di-nucleotide       | 12  | AG, AC, AT, TC, TA, TG, CT, CA, CG, GT, GA, GC |
| Tri-nucleotide      | 41  | AAG, ACA, ATC, AGC, AAC, TCA, TTA, TCC, TCG, TTG, TCG, CGG, CTG, CTC, CGC, CAT, CAG, CCA, GAA, GCC, GTC, GAT, GCC, GCT, GGT, GCA, CTA, CAT, CAG, CCA, GAA, GCC, GTC, GAT, GCC, GCT, GGT, GCA |
| Tetra-nucleotide    | 1   | ATAC                     |
| hexa-nucleotide     | 0   | --                       |
| hept-anucleotide    | 1   | CTGGGG,                  |
| Others              | 1   | CGTGGG                   |

ISSN 0973-2063 (online) 0973-8894 (print)
Bioinformation 17(1): 109-118 (2021)
©Biomedical Informatics (2021)
GO annotation analysis:
The single sequence of functional annotations was classified into functional categories in BLAST2GO software. A total of 227 GO terms were obtained, with an average of 6.837, divided into 34.36% biological processes, 27.75% cellular components, and 37.88% molecular functions, with the distribution of the ESTs (Supplementary Figure 1). According to the biological process, the component contained metabolic process (37%), cellular process (31%), biological regulation (9%), localization (8%), cellular component organization or biogenesis (8%) and regulation of biological process (7%), as shown in Supplementary Figure 1A. Regarding molecular function, 53% ESTs were associated with catalytic activity, 38% ESTs were associated with binding and only 8% ESTs were associated with transporter activity (Supplementary Figure 1B). The components of biological cellular process associated with the ESTs included membrane (19%), membrane part (17%), cell (16%), cell part (16%), organelle (13%), organelle part (8%), protein-containing complex (6%) and membrane-enclosed lumen (4%) (Supplementary Figure 1C). Supplementary Figure 2A and 3A shows the enrichment of gene functions. There are 5 GO terms in biological process that have highly enrichment value including cellular process (GO:0009987), phosphorylation (GO:0016310), cellular metabolic (GO:0044237), metabolic process (GO:0008152), oxidation-reduction process (GO:0055114), with 27, 6, 21, 32, 10 background genes respectively. According to molecular function GO term, there is catalytic activity (GO: 0003824) and oxidoreductase activity (GO:0016491) significantly enriched with 32 and 10 background genes (Supplementary Figure 2B and 3B). In cellular component GO term, cell nodescore (GO: 0005623), membrane nodescore (GO:0016020), intracellular part (GO: 0044424), integral component (GO:0016021), with 22, 26, 21 and 23 background genes, respectively (Supplementary Figure 2C and 3C).

**Table 4:** Di-nucleotide and tri-nucleotide repeat motifs in ESTs of *A. cinnamomea*

| Items | SSR | NO. | Rate of ESTs |
|-------|-----|-----|--------------|
| GC    | 40  | 14.76% |
| CG    | 36  | 13.28% |
| CT    | 26  | 9.59%  |
| TC    | 25  | 9.22%  |
| CA    | 18  | 6.64%  |
| TG    | 16  | 5.90%  |
| AT    | 13  | 4.80%  |
| GA    | 11  | 4.06%  |
| AC    | 10  | 3.69%  |
| GT    | 8   | 2.95%  |
| AG    | 6   | 2.21%  |
| TA    | 4   | 1.48%  |
| AGG   | 3   | 1.11%  |
| GTTC  | 3   | 1.11%  |
| GAA   | 3   | 1.11%  |
| CAG   | 2   | 0.74%  |
| CGT   | 2   | 0.74%  |
| GGG   | 2   | 0.74%  |
| GAT   | 2   | 0.74%  |
| GTG   | 2   | 0.74%  |
| TCC   | 2   | 0.74%  |
| TCG   | 2   | 0.74%  |
| TGC   | 2   | 0.74%  |
| TTG   | 2   | 0.74%  |
**Supplementary Figure 3**: The GO term ontology analysis of biological process (A), molecular function (B), cellular component (C).

**KEGG pathway and Interaction protein analysis**: KEGG is an approach to link genomic data with higher order functional sequence by computerizing current information on cellular processes and by standardizing gene annotations. It provides biochemical pathways for the annotations species in which the genome have been discovered. According to **Table 2**, a
total of 56 matching proteins were revealed to be involved in 8 KEGG pathways. 19 of 65 ESTs were annotated to the metabolic pathways (sce01100), which are the most represented pathways. While, the ribosome biogenesis in eukaryotes and oxidative phosphorylation are the second and third most represented pathways. In addition, carbon metabolism, pyruvate metabolism, glycolysis gluconeogenesis, methane metabolism and glycosylate and dicarboxylate metabolism were also represented. The interaction among the 65 ESTs were analysed by STRING database with the genome of Sachharomyces cerevisiae. The result showed a functional association network determined with 53 codes, 346 edges and PPI enrichment p-value of 0.00131 (Supplementary Table 1 see linked excel file). The 65 ESTs with the low confidence (0.150) minimum required interaction score is no more than 20 interactors. The predicted potential regulators with no clustering are shown in Figure 4. The STRING results revealed protein ATP2, PDB1, ATP1, ATP16, SAS10, PWP2, UTP13, ATP5, ATP4, UTP14, NOP58, NOP14, NOP56, ARC40, BMS1, UTP4, UTP15, ATP7, ARC18 were the functional partners. As Supplementary Figure 1 shows, there are a total of 11 GO terms which have high enrichment value including cellular process, phosphorylation, cellular metabolic, metabolic process, oxidation-reduction process, catalytic activity and oxidoreductase activity, cell nuclosde, membrane nucloscde, intracellular part, integral component of the A. cinnamomea cDNA library. In the present study, the STRING for A. cinnamomea is primarily based on the Sachharomyces cerevisiae genome. We sequenced among the matches, 29 genes which encoded enzymes highly involved in Ribosome biogenesis in eukaryotes, oxidative phosphorylation, carbon metabolism, glycosylate/ gluconeogenesis, glycoxylate and dicarboxylate metabolism, pyruvate metabolism and methane metabolism (Table 2).

EST-SSR analysis:
A total of 271 SSRs were isolated from 63 ESTs of A.cinnamomea with an SSR occurrence frequency of 96.82%. Among the EST-SSRs, there are 7 ESTs with 1 SSR, 12 ESTs with 2 SSRs, and 44 ESTs with 3 or more. The tri-nucleotides results shown in Table 3 accounts for the largest proportion of all SSRs. The di-nucleotide and tri-nucleotide repeat SSRs in the EST accounted for 95.31% of the total repeats, and the number of dinucleotide repeat motifs GC, CG, CT, and TC appeared more frequently with repeats of 14.76%, 13.28%, 9.59%, and 9.22%, respectively (Table 4). However, AG, GT, and TA were less frequent motifs accounting for less than 5%. In the tri-nucleotide repeat motif, AGG, GAA and GTC appear slightly higher, accounting for 3.33%, while CAC, CGC, AGC and ATC only occur once. The frequencies of tetra-nucleotides, hexa-nucleotides and hepta-nucleotides are all very low. However, di-nucleotide SSRs (18.75%) and tri-nucleotide SSRs (76.56%) presented higher polymorphic proportions than tetra-nucleotide in A. cinnamomea, which suggested that the SSRs which occurred within untranslated region were more polymorphic than those in exon regions.

Conclusion:
In this study, we document some insights from the cDNA and EST analysis of Antrodia cinnamomea for further data mining. This is helpful in the genome level understanding of Antrodia cinnamomea for application in biomedicine.

Acknowledgments:
Funding: The study was funded by the education department of Fujian Province, “the research on submerged fermentation and difference in terpene of Antrodia cinnamomea” (JT180121).

Conflicts of interest: All authors declare no conflict of interest.

References:
[1] Geethangili M & Tzeng YM. Evidence-Based Complementary and Alternative Medicine. 2011. [PMID: 19687189].
[2] Lu M-Y J et al. Proceedings of the National Academy of Sciences. 2014 111(44): E4743-E4752 [PMID: 2533675].
[3] Mao X-G, et al. Hereditas. 2006 28(7): 865-873 [PMID: 16825176].
[4] Weiste C, et al. Plant Journal. 2010 52(2): 382-390 [PMID: 17672839].
[5] Gupta P K&Varshney R. Euphytica. 2000 113(3): 163-185.
[6] Ellis J&Burke J. Heredity. 2007 99(2): 125 [PMID: 17519965].
[7] Li J, et al. Southwst China Journal of Agricultural Sciences. 2015 28(6): 2649-2654.
[8] Lee J T Y, et al. Journal of Biomaterials and Nanobiotechnology. 2011 2(01): 41.
[9] Cao S, et al. BMC biotechnology. 2011 11(1): 53 [PMID: 21595791].
[10] Huang X&Madan A. Genome research. 1999 9:868. [PMID: 10508846].
[11] Joshi A G&Pathak A R. Plant and Human Health, Volume 2. 2019 Springer. 207. [PMID: 19840835].
[12] Ashburner M, et al. Nature genetics. 2000 25(1): 25 [PMID: 10802651].
[13] Conesa A, et al. Bioinformatics. 2005 21(18): 3674-3676 [PMID: 16081474].
[14] Clarke L&Carbon J. Cell. 1976 9(1): 91-99 [PMID: 788919].

Edited by P Kanguene

License statement: This is an Open Access article which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited. This is distributed under the terms of the Creative Commons Attribution License
Articles published in BIOINFORMATION are open for relevant post publication comments and criticisms, which will be published immediately linking to the original article for FREE of cost without open access charges. Comments should be concise, coherent and critical in less than 1000 words.
