Gene 506 (2012) 202–206

Contents lists available at SciVerse ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Short Communication

Construction of cDNA library and preliminary analysis of expressed sequence tags from tea plant [*Camellia sinensis* (L) O. Kuntze]

Munmi Phukon, Richa Namdev, Diganta Deka, Mahendra K. Modi, Priyabrata Sen *

Department of Agricultural Biotechnology, Assam Agricultural University, Jorhat, 785 013, India

ARTICLE INFO

Article history:
Accepted 19 June 2012
Available online 1 July 2012

Keywords:
Camellia sinensis
Tender shoots
cDNA library
ESTs

ABSTRACT

Tea is the most popular non-alcoholic and healthy beverage across the world. The understanding of the genetic organization and molecular biology of tea plant, which is very poorly understood at present, is required for quantum increase in productivity and efficient use of germplasm for either cultivation or breeding program. Single-pass sequencing of randomly selected cDNA clones is the most widely accepted technique for gene identification and cloning. In the present study, a good quality cDNA library was constructed and preliminary analysis of ESTs was carried out. The titers of unamplified and amplified libraries were 1.4×10^6 pfu/ml and 5.27×10^8 pfu/ml respectively. A total of 210 cDNA clones from the constructed cDNA library were sequenced and analyzed. A total of 84 high quality Expressed Sequence Tags (ESTs) were generated, among which 71 ESTs had significant homology with sequences in NCBI non-redundant protein database by BLAST X analysis. About 80% ESTs had poly (A) tail at 3′ end indicating that the cDNAs were full length. The database-matched ESTs were classified into putative cellular roles, viz. energy-related category (corresponding to 20% of total BLAST X matched ESTs), Transcription (14.2%), protein synthesis (14.2%) cell growth and division (8.6%), cell structure (5.7%), signal transduction (5.7%), transporters (2.9%), disease and defenses (2.9%), secondary metabolism (2.9%) and gene regulation (2.9%). This study provides an overview of the mRNA expression profile and first hand information of gene sequence expressed in tender leaves and apical buds of tea plant.

1. Introduction

The Indian tea is an epitome of taste, warmth and happiness. Being the largest producer of tea in the world India also consumes large amounts of tea. "Chai" as tea beverage is known in India has become an essential part of daily life. The tea plant (*Camellia sinensis*) has received much attention for its aroma, taste and medicinal benefits (Lin et al., 2003) and has great value as a source of secondary metabolic products (Yamamoto et al., 1998). A number of barriers hinder the production of tea, of which biotic or abiotic stress is predominant.

Expressed Sequence Tags (ESTs) (Adams et al., 1991) are short and single-pass sequence reads from mRNA. They represent a snapshot of genes expressed in a given tissue and/or at a given developmental stage (Mekhedov et al., 2000). ESTs have been reported both in the literature and public databases for rice cDNAs (Sasaki et al., 1994; Uchimaya et al., 1992), *Arabidopsis* cDNAs (Hofte et al., 1993; Newman et al., 1994) and Maize (*Zea mays* L.) cDNAs (Keith et al., 1993). Untill 2005, only eight *C. sinensis* ESTs were deposited to dbEST (Boguski et al., 1993) of the GenBank (Benson et al., 2003) as reported by Chen et al. (2005). The construction of cDNA library, generation and analysis of ESTs has also been reported in tea (Chen et al., 2005; Park et al., 2004).

Because ESTs represent a copy of just the business part of a genome, that which is expressed, to sequence the functional portions using ESTs strategy was believed to be a more effective way for discovering more functional genes (Ablett et al., 2000; Bausher et al., 2003; Carson and Botha, 2000; Delseny et al., 1997; Liu et al., 1995, 2010; Thanh et al., 2011; Yamamoto and Sasaki, 1997). In this study, we report the construction of a high quality of cDNA library from tender shoots of *C. sinensis*, the preliminary analysis of expressed sequence tags, the putative function analysis of ESTs and the gene expression pattern in tender shoots.

2. Materials and methods

2.1. Plant materials

Tender leaves and apical buds from TV-1 clone, an Assam hybrid of *C. sinensis* (L) O. Kuntze collected from the Tea Experimental Field,

Abbreviations: %, Percent; °C, Degree; dbEST, database EST; DMSO, Dimethyl sulphoxide; cDNA, Complementary Deoxyribonucleic acid; EST, Expressed Sequence Tag; et al, Et alia; Fig., Figure; ml, Milliliter; mRNA, Messenger RNA; NCBI, National Center for Biotechnology Information; pfu, Plaques forming unit; RNA, Ribonucleic acid; X-gal, 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside; Kb, Kilobases; Bp, Base pairs; EtBr, Ethidium bromide; BLAST X, Algorithm that searches protein databases using a translated nucleotide query; PCR, polymerase chain reaction; LB media, Luria–Bertani media; IPTG, Isopropyl-beta-thio galactopyranoside; CBL-interacting protein kinase, Calcineurin B-like -interacting protein kinase.

* Corresponding author.
E-mail address: pbsen14@yahoo.co.in (P. Sen).

0378-1119/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2012.06.055

Assam Agricultural University, Jorhat, Assam, India were used as plant materials. Healthy tender leaves were collected and immediately frozen in liquid nitrogen and put in −80 °C until further use.

2.2. Total RNA isolation

The total RNA from leaves and apical buds was extracted using Guanidine HCl (without using liquid nitrogen). This procedure based on Guanidine hydrochloride developed by Logemann et al. (1987) and modified by Lal et al. (2001) was adapted with little modification. The RNA thus extracted was subjected to spectrophotometric analysis.

2.3. cDNA library construction

Double-strand cDNA was synthesized using the SMART™ cDNA Library Construction Kit (Clontech laboratories Inc., CA, USA) following the manufacturer's instruction in an ABI GeneAmp PCR System 9700 (Applied Biosystem, USA). PCR product was run in agarose gel electrophoresis to check the double stranded cDNA quality. Double-stranded cDNA was then digested by SfiI enzyme (New England Biolabs Inc., USA) and fractionated by using CHROMA SPIN-400 columns (Clontech Laboratories Inc.) in order to filter the cDNA fragments shorter than 400 bp. The cDNAs larger than 400 bp were collected and ligated to λ TriplEx2 vector performing overnight ligation at 16 °C. The ligated product was packaged with Gigapack III Gold Packaging Extract (Stratagene, USA).

2.4. Titration of the primary library

The cDNA inserts having a length of more than 400 bp were used for cloning into λ TriplEx2 vectors. Successful transformation was done by using the XL1-Blue strain of *E. coli*. Transformed culture was plated on LB/$MgSO_4$/IPTG/X-Gal (with 50 μg/ml ampicillin) plates for blue-white screening. The number of clones was counted to calculate the library titer using the formula: pfu/ml = number of plaques × dilution factor × 10^3 μl/ml (μl of diluted phage plated). The size of the insert fragment was confirmed by the PCR method using random selection of 10–15 clones as described by Gao and Hu (2001). After amplification, the completed cDNA libraries were stored in 7% dimethyl sulfoxide at −80 °C.

2.5. Sequencing and analysis

cDNA clones were randomly picked and subjected to 5′ end single-pass sequencing on an ABI 3130xl Genetic Analyzer (Applied Biosystems). Sequences less than 200 bp were excluded, the remaining were subjected to data analysis. The leading vector, tailing and poor quality sequences were removed by the software "SeqScanner". The repeated ESTs were deleted. The primarily edited EST sequences were subjected to BLAST X analysis against the non-redundant protein database of NCBI (National Center for Biotechnology Information) to search for similarity (http://www.ncbi.nlm.nih.gov/blast). The result of the EST analysis is grouped for functional annotation.

3. Results

3.1. Total RNA isolation and LD-PCR

The A_{260}/A_{280} ratio of isolated RNA was found to be 1.88 confirming good quality of isolated RNA. Denaturing agarose gel electrophoresis of isolated RNA sample showed clear distinct band of 23S and 16S corresponding to a molecular weight of 5.0 kb and 1.8 kb respectively. Double-stranded cDNA (5 μl) was analyzed on a 1.1% agarose/EtBr gel. The ds cDNA appeared as a 0.1–4 kb smear on the gel, which confirmed the successful synthesis of ds cDNA.

3.2. Constructon of cDNA library

The cDNA size fractionation was carried out in Chromaspin-400 columns (Fig. 1). The titers of unamplified and amplified libraries were 1.4×10^6 pfu/ml and 5.27×10^8 pfu/ml respectively with recombination percentage of 80.71%. To check the size of the inserts, screening of inserts was done by PCR amplification. The PCR amplification was carried out with T3 and T4 primers to determine the presence and size of the cDNA insert. The results showed that out of 340 randomly selected cDNA clones, 210 clones contained inserts with size ranging from 300 bp to 1100 bp (Fig. 2). These data indicated that the tea cDNA library thus constructed had high titer, high recombination rate and large insert.

3.3. Generation of expressed sequence tags and sequence analysis

A total of 210 clones were sequenced of which sequences less than 200 bp were excluded. The leading vector, tailing and poor-quality sequences were removed by software "SeqScanner". The repeated ESTs were deleted. Finally, a total of 84 ESTs were generated which were of more than 200 bp long and subjected to BLAST X analysis against the non-redundant protein database of NCBI to search for similarity. Most of the ESTs were 300–700 bp long, with an average of 415 bp. The ESTs, after BLAST X analysis, were classified into four groups (Fig. 3). The generated ESTs were submitted to NCBI (http://www.ncbi.nlm.nih.gov/blast).

About 80% ESTs (67 nos.) had poly (A)-tail at the 3′ end, indicating that the cDNAs were full length and the remaining cDNAs (17 nos.) had no poly (A)-tail. BLAST X analysis revealed a wide distribution of homologous genes ranging from algae to angiosperm. Eight ESTs (9.5% of total ESTs) represented genes that had been previously identified in non-plant organisms such as *E. coli*, *Drosophila*, human etc. This indicates that these genes are highly conserved from prokaryotes to vertebrates to higher plants (Chen et al., 2005). An overview of the ESTs BLAST X analysis is in Table 1.

About 20% of database-matched 71 ESTs were found to match proteins from model plants such as *Arabidopsis* (7.0%, 5 ESTs), rice (8.4%, 6 ESTs), tobacco (5.6%, 4 ESTs), etc.

3.4. Abundance of ESTs with know on putative functions by BLAST X analysis

Abundance of ESTs with known or putative functions by BLAST X analysis is shown in Table 2. The most abundant ESTs were cyclophilin, 60S ribosomal protein and histones.

These ESTs are mostly expressed genes in tender shoots of *C. sinensis*. A total of 22 ESTs were found to have more than 2 copies, 13 single copies of ESTs and 20 ESTs showed homology with hypothetical or predicted protein in different plants.

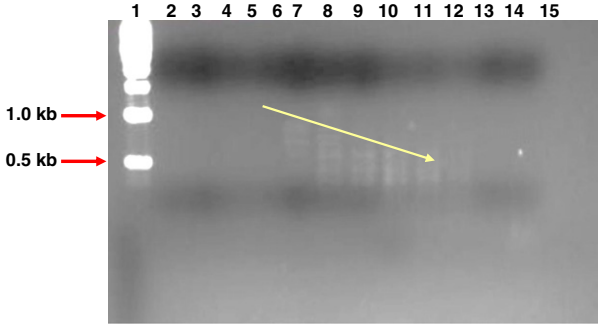

Fig. 1. Electrophoresis banding pattern of fractionated ds cDNA Lane 1: 500 bp DNA size marker, Lanes 2–15: fractionated cDNA.