Computational Biology and Chemistry 64 (2016) 1–8

Contents lists available at ScienceDirect

Computational Biology and Chemistry

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

Research article

Computational analysis of *atp*B gene promoter from different Pakistani apple varieties

Tariq Mahmood*, Najeeb Ullah Bakht, Ejaz Aziz

Department of Plant Sciences, Quaid-i-Azam University, Islamabad 45320, Pakistan

ARTICLE INFO

Article history:
Received 17 April 2015
Received in revised form 27 April 2016
Accepted 5 May 2016
Available online 7 May 2016

Keywords:
Apple
Genetic diversity *atp*B promoter
cis-acting regulatory elements

ABSTRACT

Apple is the fourth most important fruit crop grown in temperate areas of the world belongs to the family *Rosaceae*. In the present study, the promoter (~1000 bp) region of *atp*B gene was used to evaluate the genetic diversity and phylogeny of six local apple varieties. *atp*B gene is one of the large chloroplastic region which encodes β-subunit of ATP synthase and previously it had been used largely in phylogenetic studies. During the present study, *atp*B promoter was amplified, sequenced and analyzed using various bioinformatics tools including Place Signal Scan, MEGA6 and BLASTn. During the phylogenetic analysis, obtained phylogram divided the studied varieties into two clusters revealing the monophyletic origin of studied apple varieties. Pairwise distance revealed moderate genetic diversity that ranges from 0.047–0.170 with an average of 0.101. While identifying different *cis*-acting elements present in the *atp*B promoter region, results exhibited the occurrence of 56 common and 20 unique *cis*-regulatory elements among studied varieties. The identified *cis*-acting regulatory elements were mapped as well. It was observed that Kala Kulu has the highest unique features with reference to the availability of *cis*-acting elements. Moreover, the possible functions of all regulatory elements present on the promoter sequence of *atp*B gene were predicted based on already reported information regarding their *in vivo* role.

1. Introduction

Apple (*Malus domestica*), which belongs to the family *Rosaceae*, is one of the significant cultivated fruit crops of the world temperate regions (Janick and Moore, 1996). The total number of species in the genus *Malus* varies between different studies as Robinson et al. (2001) recognized seventy eight while Harris et al. (2002) reported 55 species. Global apple production exceeds 75.4 million tonne in 2013 (Sonmezuglu and Kutuk, 2014). Apple is the fourth major fruit crop of Pakistan and mainly cultivated in northern areas of Khyber Pukhtunkhwa, Punjab and Baluchistan. However, Baluchistan province is the major contributor of apple production. Various varieties of apple being grown in Pakistan includes Kala Kulu, Golden Delicious, Super, Gaja, Red Delicious, Mashadi, etc. (Mukhtar et al., 2010; Muhammad et al., 2011). Jasra et al. (2001) reported severe decline in apple yield in Baluchistan (Pakistan) due to selection of variety as pollinizer as well as attack of insects. So, accurate identification and assessment of genetic polymorphism of the existing cultivar is of utmost importance for successful apple breeding programs.

Malus domestica is a hybrid species from a number of progenitor species, including *M. sylvestris*, *M. prunifolia*, *M. dasyphyllus*, *M. orientalis* and *M. praecox* (Korban and Skirvin, 1984; Cornille et al., 2013; Nikiforova et al., 2013). The domestication of apple orchards started in Central Asia along with secondary contribution from the European crabapple dated back in near four thousand years after the first use of grafting practices (Cornille et al., 2014). The road named Old Silk that connects Eastern China with Black Sea has been considered the mainstream way in the evolution of *Malus* (Juniper et al., 1999).

Complete genome sequence of *Malus domestica* Borkh has been published (Velasco et al., 2010). All members in genus *Malus* contain equal number of x = 17 chromosomes (Evans and Campbell, 2002). Genetic markers collectively selected from nuclear DNA (nrDNA), mitochondrial DNA (mtDNA) and chloroplast DNA (cpDNA) have been widely used to trace the origin of domesticated apples. Moreover, extensive studies have been reported on apple to examine its genetic variability using technique such as random amplified polymorphic DNA (RAPD) (Dantas et al., 2000; Goulao and Cristina, 2001), amplified fragment length polymorphism (AFLP) (Goulao et al., 2001; Hokanson et al., 2001), restriction

* Corresponding author.
E-mail addresses: tmahmood.qau@gmail.com, tmahmood@qau.edu.pk (T. Mahmood).

http://dx.doi.org/10.1016/j.compbiolchem.2016.05.002
1476-9271/

fragment length polymorphism (RFLP) (Nybom and Schaal, 1990; Gardiner et al., 1996; Khadivi-Khub et al., 2014), trnH-psbA intergenic spacer (Volk et al., 2015), inter-simple sequence repeat (ISSR) (Kashyap et al., 2010) and simple sequence repeat (SSR) (Gharghani et al., 2009; Zhang et al., 2012; Gross et al., 2014).

cpDNA sequences are traditionally used as staple data source for studying the molecular phylogeny of plants (Patwardhan et al., 2014). This is because cpDNA is relatively conserved as well as smallest as compared to nuclear and mitochondrial genome along with the ease of PCR amplification and sequencing of chloroplast genes (Dong et al., 2012). cpDNA as a molecular marker can be used efficiently to determine genetic diversity as well as phlyogenetic lineage (Coart et al., 2003). Several cpDNA genes including *mat*K, *rbc*L, *atp*B, *rpl*16, *rps*16, *rps*11 and *ndh*F have been used for inferring phylogeny in plants at different taxonomic levels (Zhang, 2000; Gao et al., 2008; Li, 2008). Previously phylogenetic analyses based on different regions of the chloroplast genome have been conducted in the family *Rosaceae* (Lo and Donoghue, 2012).The *atp*B gene, located on the cpDNA, encodes the β (beta) subunit of ATP synthase. In many studies scientist used *atp*B gene to elucidate phylogenetic relationship for different plant species because it has a relatively slow rate of nucleotide substitution compared to most genes encoded by the cpDNA (Savolainen et al., 1995; Magee et al., 2010). On the other hand no elaborative studies have been undertaken for *atp*B gene characterization in any single aspects for apple chloroplast. Therefore, *atp*B based analysis can be a good candidate to study the phylogeny of *Malus* species. The present study has therefore been conducted with an aim of evaluating the genetic diversity and phylogeny of six apple varieties based on *atp*B gene promoter sequences and to analyze the functional importance and identification of its regulatory elements.

2. Materials and methods

2.1. Plant material and extraction of genomic DNA

Fresh leaves samples of the selected apple varieties were collected from Pishin, Baluchistan (Pakistan) and stored at 4 °C for further processing. Total genomic DNA was extracted using CTAB-based method (Richard, 1997) with slight modification by incubating the homogenized mixture at 70 °C for 30 min.

2.2. Primer designing

Primers was designed for *atp*B gene promoter using Primer 3 (version 4.0) (http//primer3.sourceforage.net/) based on cpDNA sequences of tobacco available at NCBI GenBank (www.ncbi.nih.gov). The sequence of primer is as follow;

*atp*B Forward: 5′-CCAGAAGTAGTAGGATTGATTCTCA-3′

*atp*B Reverse: 5′-TCTTCAGGTGGAACTCCAGGTT-3′.

Table 1
Accession numbers of sequences of *atp*B gene promoter from selected apple varieties.

Apple Variety	Accession No.
Malus domestica var. Gaja	KP176610
Malus domestica var. Kala Kulu	KP176611
Malus domestica var. Mashadi	KP176612
Malus domestica var. Super	KP176613
Malus domestica var. Sur Kulu	KP176614
Malus domestica var. Shin Kulu	KP176615

2.3. PCR amplification and sequencing of atpB promoter region

The amplification of *atp*B gene promoter region was carried out in a total volume of 25 μL containing 16.2 μL of nano pure water, 2.5 μL of 10 X PCR buffer, 1.5 μL of 2 mM dNTPs, 1.5 μL of 25 mM $MgCl_2$, 1 μL of each primer, 0.3 μL of *Taq* polymerase (5 U) and 1 μL of template DNA using PCR MultiGene Thermal Cycler (Labnet). The PCR cycling profile was as follows: 5 min at 94 °C, 35 cycles of 30 s at 94 °C, 1 min at 60 °C, 1 min at 72 °C and final extension of 20 min at 72 °C. Amplification was confirmed by running it on 1.5% agarose gel. After the purification of amplified products by JET quick (Genomed) PCR Product Purification Kit, the amplified samples were sent for sequencing from Macrogen Inc. (Seoul, Korea).

2.4. Sequence analysis

The obtained sequences were aligned by using ClustalW. Homology with already reported sequences of *atp*B promoter in nucleotide databases were detected using nucleotide blast (BLASTn) (http://blast.ncbi.nlm.nih.gov/Blast.cgi). Sequenced data was submitted to Genbank and accession numbers were allotted by the Genbank staff (Table 1). The neighbor-joining (NJ) method from the MEGA6 (Tamura et al., 2013) program was used to construct phylogenetic tree. The statistical significance of tree branching was tested by performing 500 bootstrap replications. Finally, the *cis*-regulatory elements were identified using PLACE web site (http://www.dna.affrc.go.jp/PLACE) and maps were constructed using DOG 2.0 (Domain Graph, version 2.0) (http://dog.biocuckoo.org/).

3. Results and discussion

The obtained sequence data of *atp*B gene promoter was analyzed. Sequences were aligned and compared with nucleotide sequences in the GenBank using ClustalW and BLASTn to determine conserved regions and similarities respectively. Blast results showed higher than 90% sequence identity with *Malus* cpDNA.

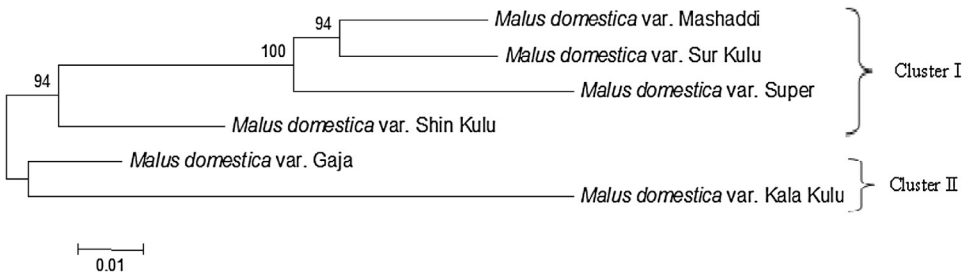

Fig. 1. Phylogram constructed by MEGA6 for *atp*B gene promoter sequences indicating genetic relationship among selected apple varieties with bootstrap values.