https://doi.org/10.1051/e3sconf/202129203101

Initial Characterization and Expression Pattern Analysis of Tobacco (*Nicotiana Tabacum*) Sucrose Synthase Gene

Meiwei Zhao[1], Juan Zou[3], Sixuan Leng[3], Ying Tao[3], and Zhengxiong Zhao[2*]

[1]College of Agronomy and Biotechnology, Yunnan Agricultural University, Kunming 650201, China
[2]College of Resources and Environment, Yunnan Agricultural University, Kunming 650201, China
[3]Research and Development Center, China Tobacco Yunnan Industrial Co.,Ltd, Kunming 650106, China

Abstract. A novel tobacco(nicotiana tabacum) sucrose synthase genewas characterized using the rapid amplification of cDNA ends methods based on one tobacco EST. Thissucrose synthase gene mRNA was 2,954bp in length containing a 2,418bp open reading framewhich encodes a protein of 805 amino acids. Sequence analysis revealed that thissucrose synthase of tobacco shares high homology with the sucrose synthase of Nicotianaattenuata(98%), Nicotianasylvestris(98%), Nicotianatomentosiformis(98%), Capsicumannuum(95%),Capsicumbaccatum(95%), Solanumtuberosum(94%),Solanumlycopersicum(94%) and Solanumpennellii(94%). Results also showed that tobacco sucrose synthase gene has a closer genetic relationship with the sucrose synthase gene of Capsicumannuum.Genomic DNA analysis indicated that tobacco sucrose synthase gene is structured in 13 exons and 12 introns.Tissue expression profile analysis revealed that this tobacco sucrose synthase gene was highly expressed in leaf, but moderately expressed in root, stemand flower. These established the foundation for further research on the tobacco sucrose synthase gene.

1 Introduction

Sucrose serves as carbon and energy sources for plant metabolism. Whensuffering low temperature or drought stress, plant cells can accumulate sucrose to stabilize membranes and proteins[1-3]. Sucrose synthase functions in UDP-glycosyltransferase activityand sucrose synthase activity[4-6] and catalyzes the reversible conversion of sucrose and a nucleoside diphosphate into the corresponding nucleoside diphosphate-glucose and fructose[7]. More than ten sucrose synthaseisoforms have been found in the many plant species and they display different developmental expression patterns[1,8,9]. For the involvement of sucrose synthase in the synthesis of UDP-glucose and ADP-glucose linked to plant cellulose and starch biosynthesis,that mutations of the sucrose synthase gene have beenreported to impaired the sucrose synthaseactivity and the accumulation levels of ADP-glucose, UDP-glucose, cellulose and starch[10]. However, in 2012, Baroja-Fernández et al.disprovedthe Barratt et al.[10] claims and demonstrated that sucrose synthase activity of mutats in leaves and stems can support normal cellulose and starch biosynthesis in Arabidopsis.

Although sucrose synthasegene play important roles in the plant carbon metabolism, until today, the full-length mRNA of tobacco sucrose synthase gene has not been reported yet. In present experiment, we will isolate the complete mRNA sequences of this tobacco gene, subsequently perform some necessary sequence analysis and tissue expression analysis for this gene. These will establish the primary foundation of understanding this tobacco gene.

2 Material and mehods

2.1 Samples collection, RNA extraction and first-strand cDNA synthesis

Tobacco plants (Chinese commercial variety Yunyan 85) were grown in a naturally lit glasshouse with normal irrigation and fertilization. The tissues including leave, stem, root, flower were harvested and immediately frozen in liquid nitrogen and stored at -80°C[11]. Total RNA extraction and first-strand cDNA synthesis for these tissue samples were performed as the methods describe by Li et al.[12].

2.2 5’and 3’-RACE

5’- and 3’-RACE were performed as the instructions of BD SMARTTM RACE cDNA Amplification Kit (BD science, USA). For the tobacco sucrose synthase gene, the gene specific primers (GSPs) were designed based on the coding sequence information from potato sucrose synthase gene and its highly homologous tobacco EST sequence:FG169638.

5’-RACE GSP:
5’- GTGCTGGTGATGATAAAATCGGTGT-3’
3’-RACE GSP:

* Corresponding author: zhaozx0801@163.com

© The Authors, published by EDP Sciences. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (http://creativecommons.org/licenses/by/4.0/).

5'- CGTAACACAGTGCACCATAGCTCAT-3'.

RACE touchdown PCRs were carried out with 5 cycles of 94°C: 30 sec and 72°C: 3 min, followed by 5 cycles of 94°C: 30 sec, 65°C: 30 sec and 72°C: 3 min, finally with 25 cycles of 94°C: 30 sec, 64°C: 30 sec and 72°C: 3 min to terminate reaction. These RACE PCR products were then cloned into PMD18-T vector (TaKaRa, China) and sequenced bidirectionally with the commercial fluorometric method. At least five independent clones were sequenced for each PCR product.

2.3 Quantitative real time PCR (qRT-PCR) for tissue expression profile analysis

qRT-PCR for evaluating the level of mRNA for sucrose synthasegene was performed on the ABI Prism 7300 Sequence Detection Systems (Applied Biosystems, Foster City, CA, USA). PCR reactions for each sample were carried out in 25µl reaction volume containing 1µl SYBR Green real-time PCR Master Mix, 100 ng cDNA template and 200 nM each primer. Conditions for real-time PCR were: an initial denaturation at 95 °C for 3 min, 40 cycles of 95 °C for 15 s, 52°C for 15 s (Table 1) and 72°C for 20 s. For each sample, reactions were set up in triplicate to ensure the reproducibility of the results. The gene relative expression levels were quantified relative to the expression of the reference gene, actin (GenBank Accession No. GQ339768) by employing the $2^{-\Delta\Delta Ct}$ value model [11,13].

2.4 Sequence analysis

The cDNA sequence prediction was conducted using GenScan software (http://genes.mit.edu/GENSCAN.html).

The protein prediction and analysis were performed using the Conserved Domain Architecture Retrieval Tool of BLAST at the National Center for Biotechnology Information (NCBI) server (http://www.ncbi.nlm.nih.gov/BLAST) and the Clustal Omega software (http://www.ebi.ac.uk/Tools/msa/clustalo/).

The theoretical isoelectric point (pI) and molecular weight (Mw) of the deduced protein of the tobacco gene was computed using the Compute pI/Mw Tool (http://www.expasy.org/tools/pi_tool.html).

3 Results

3.1 RACE results for tobacco sucrose synthase gene

For tobacco sucrose synthase gene, through 5'-RACE, one PCR product of 1,654bp was obtained. The 3'-RACE product was 1,456bp. These products were then cloned to T-vector and sequenced. Taken together, a 2,954-bp cDNA complete sequence was finally obtained (Figure 1).

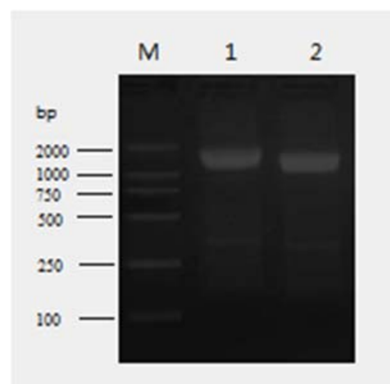

Figure 1. RACE results for tobacco sucrose synthase gene. M DL2000 DNA markers; 1, 5'-RACE product for tobacco sucrose synthase gene; 2,3'-RACE product for tobacco sucrose synthase gene

3.2 Sequence analysis

These cDNA nucleotide sequence analysis using the BLAST software at NCBI server (http://www.ncbi.nlm.nih.gov/BLAST) revealed that this gene was not homologous to any of the known tobacco gene and it was then deposited into the Genbank database (Accession number: KF977579).

The sequence prediction was carried out using the GenScan software and results showed that the 2,954-bp cDNA sequence represents one single gene which encodes 805 amino acids (Figure 2). The pI of tobacco sucrose synthase is 5.94. The molecular weight of this putative protein is 92531.43.

Further BLAST analysis of this protein revealed that tobacco sucrose synthase has high homology with the sucrose synthase of nicotianaattenuata(Accession number: XP_01922761, 98%), nicotianasylvestris (Accession number: XP_009760458, 98%), nicotianatomentosiformis(Accession number: XP_009589756, 98%), capsicumannuum(Accesession number: XP_016540310, 95%),capsicumbaccatum(Accesession number:PHT52244, 95%), solanumtuberosum(Accesession number:AAA97571, 94%),solanumlycopersicum(Accesession number:NP_001234655, 94%) and solanumpennellii(Accesession number:XP_015059625, 94%)(Figure 3). Its conserved domain was identified as Sucrose_synth .

The 3-D structural evidence of the putative conserved domain is also presented in figure 5. Based on the results of the alignment of different species of sucrose synthase proteins, a phylogenetic tree was constructed using the Clustal Omega software, as shown in Figure 6. The phylogenetic tree analysis revealed that the tobacco sucrose synthase gene has a closer genetic relationship with that ofcapsicumannuum.

Table1. qRT-PCR primers for tobacco sucrose synthase, actin genes and annealing temperature

Gene	Primer sequence	Ta/ °C	Length/(bp)
Sucrose synthase	Forward : 5'- GAGCCGAGCACTCACATA-3' Reverse: 5'-AGCCACAAGATTTCCCTC-3'	51	194
Actin	Forward :5'- CCATTCTTCGTTTGGACCTT -3' Reverse: 5'- TTCTGGGCAACGGAACCT-3'	56	257

```
GCGGCGCGGGGATCACCATTCATAAACAAACCTCTTTCATTTGCTTCTTTGATTTCCATTCCTTTCGTTCATTTCATTTGCTGCTTATTCTTGTTCTCTTTTTCTTCCAT
TAGTTTCATTATTCCATTCCTTATTCCTCTCATTTTTGCTCTCTCTTCTTTTTCTTTAAAAGTTGAAAGTCATCTGAGAATTTCCAGCTGCTGAATTCAACTGCAATG
                                                                                                              M
GCCGAACGTGTGCTAACTCGTGTTCACAGCCTTCGCGAACGTCTTGATGCTACTTTGGCTGCTCATCGCAATGAGATTTTGCTGTTTCTTTCAAGGATTGAAAGCCAC
 A  E  R  V  L  T  R  V  H  S  L  R  E  R  L  D  A  T  L  A  A  H  R  N  E  I  L  L  F  L  S  R  I  E  S  H
GGAAAAGGGATATTGAAACCTCACCAGTTGCTGGCTGAGTTTGAATCAATTCACAAAGAAGACAAAAACAAACTGAATGATCATGCTTTTGAAGAAGTCCTGAAATCT
 G  K  G  I  L  K  P  H  Q  L  L  A  E  F  E  S  I  H  K  E  D  K  N  K  L  N  D  H  A  F  E  E  V  L  K  S
ACTCAGGAAGCAATTGTCTTGTCCCCTTGGGTTGCGCTTGCCATTCGTCTGCGGCCTGGTGTGTGGGAATATGTTCGTGTGAATGTCAATGCACTTATTGTCGAGGAG
 T  Q  E  A  I  V  L  S  P  W  V  A  L  A  I  R  L  R  P  G  V  W  E  Y  V  R  V  N  V  N  A  L  I  V  E  E
CTGACTGTGCCTGAATATTTGCAATTCAAGGAAGAACTTGTTAATGGAACCTCGAACGATAACTTTGTTCTTGAGCTGGATTTTGAGCCCTTCACTGCATCATTTCCC
 L  T  V  P  E  Y  L  Q  F  K  E  E  L  V  N  G  T  S  N  D  N  F  V  L  E  L  D  F  E  P  F  T  A  S  F  P
AAACCAACCCTCACCAAATCAATTGGAAATGGAGTTGAATTCCTCAATAGGCACCTTTCTGCGAAAATGTTCCATGACAAGGAAAGCATGACCCCGCTTCTTGAATTT
 K  P  T  L  T  K  S  I  G  N  G  V  E  F  L  N  R  H  L  S  A  K  M  F  H  D  K  E  S  M  T  P  L  L  E  F
CTTCGGGTTCACAATTATAAGGGCAAGACAATGATGCTGAATGACAGAATACAGAATTTAACCACTCTGCAAAATGTCCTAAGGAAGGCAGAGGAATACCTTATTATG
 L  R  V  H  N  Y  K  G  K  T  M  M  L  N  D  R  I  Q  N  L  T  T  L  Q  N  V  L  R  K  A  E  E  Y  L  I  M
CTTCCCCCTGAAACTCCATTTTCCGAATTCGAACACAAGTTCCAAGAAATTGGATTGGAGAAGGGATGGGGCGACACTGCGGAGCGCGTGCTAGAGATGATATGCATG
 L  P  P  E  T  P  F  S  E  F  E  H  K  F  Q  E  I  G  L  E  K  G  W  G  D  T  A  E  R  V  L  E  M  I  C  M
CTTCTTGATCTCCTCGAGGCTCCCGATTCCTGTACTCTTGAGAAGTTCTTGGGGAGAATTCCTATGGTGTTCAATGTGGTTATCCTTTCCCCCCACGGATATTTCGCC
 L  L  D  L  L  E  A  P  D  S  C  T  L  E  K  F  L  G  R  I  P  M  V  F  N  V  V  I  L  S  P  H  G  Y  F  A
CAGGAAAATGTCTTGGGTTATCCCGACACTGGTGGCCAGGTTGTCTATATATTAGATCAAGTTCCAGCCTTGGAGCGTGAAATGCTTAAGCGCCTAAAGGAGCAAGGA
 Q  E  N  V  L  G  Y  P  D  T  G  G  Q  V  V  Y  I  L  D  Q  V  P  A  L  E  R  E  M  L  K  R  L  K  E  Q  G
CTTGATATCACACCGCGTATTCTTATTGTTACTCGTCTGCTACCTGATGCAGTTGGAACGACTTGTGGTCAGCGGCTTGAGAAGGTGTATGGAGCCGAGCACTCACAT
 L  D  I  T  P  R  I  L  I  V  T  R  L  L  P  D  A  V  G  T  T  C  G  Q  R  L  E  K  V  Y  G  A  E  H  S  H
ATTCTGAGGGTCCCCTTTAGGACTGAGAAGGGCATTGTTCGTAAATGGATCTCTCGCTTTGAAGTGTGGCCATATATGGAGACTTTCACTGAGGATGTCGCAAAAGAA
 I  L  R  V  P  F  R  T  E  K  G  I  V  R  K  W  I  S  R  F  E  V  W  P  Y  M  E  T  F  T  E  D  V  A  K  E
CTTGCTGCAGAATTGCAGGCCAAGCCAGATTTGATAATAGGCAACTATAGCGAGGGAAATCTTGTGGCTTCATTGCTCGCTCATAAGTTAGGCGTAACACAGTGCACC
 L  A  A  E  L  Q  A  K  P  D  L  I  I  G  N  Y  S  E  G  N  L  V  A  S  L  L  A  H  K  L  G  V  T  Q  C  T
ATAGCTCATGCATTGGAGAAAACAAAGTATCCTGATTCTGACATCTACTGGAAAAAATTCGATGAAAAATACCATTTCTCGTCCCAGTTTACCGCTGATCTTATTGCA
 I  A  H  A  L  E  K  T  K  Y  P  D  S  D  I  Y  W  K  K  F  D  E  K  Y  H  F  S  S  Q  F  T  A  D  L  I  A
ATGAATCACACCGATTTTATCATCACCAGCACTTTCCAGGAGATAGCAGGAAGCAAGGACACTGTCGGACAGTACGAGAGTCATCAGGCATTCACAATGCCCGGATTG
 M  N  H  T  D  F  I  I  T  S  T  F  Q  E  I  A  G  S  K  D  T  V  G  Q  Y  E  S  H  Q  A  F  T  M  P  G  L
TACAGAGTTGTTCACGGCATTGATGTGTTCGACCCCAAATTCAACATTGTCTCACCTGGAGCTGACATAAACCTCTATTTCCCATATTCCGAGAAGGAAAAGAGACTG
 Y  R  V  V  H  G  I  D  V  F  D  P  K  F  N  I  V  S  P  G  A  D  I  N  L  Y  F  P  Y  S  E  K  E  K  R  L
ACAGCACTTCACCCTGAAATCGAGGAGCTGCTGTACAGTGACATTGAGAACGAGGAACATCTGTGTGTGCTAAAGGACAGGAATAAGCCAATTATATTCACCATGGCT
 T  A  L  H  P  E  I  E  E  L  L  Y  S  D  I  E  N  E  E  H  L  C  V  L  K  D  R  N  K  P  I  I  F  T  M  A
AGATTGGATCGAGTGAAGAACTTAACTGGACTTGTGGAGTTGTACGCCAAGAACCCACGGCTAAGGGAGTTGGTTAACCTTGTCGTGGTTGGTGGAGACCGAAGGAAG
 R  L  D  R  V  K  N  L  T  G  L  V  E  L  Y  A  K  N  P  R  L  R  E  L  V  N  L  V  V  V  G  G  D  R  R  K
GAATCCAAAGATTTGGAAGAACAGGCAGAGATGAAGAAGATGTATGAACTTATAAAGACGCACAATTTAAACGGCCAATTCCGATGGATTTCTTCCCAGATGAACCGC
 E  S  K  D  L  E  E  Q  A  E  M  K  K  M  Y  E  L  I  K  T  H  N  L  N  G  Q  F  R  W  I  S  S  Q  M  N  R
GTGAGGAATGGCGAACTCTACAGGTACATTGCCGATACTAGGGGAGCTTTTGTGCAGCCTGCATTTTACGAGGCTTTTGGTTTGACTGTTGTTGAGGCCATGACCTGT
 V  R  N  G  E  L  Y  R  Y  I  A  D  T  R  G  A  F  V  Q  P  A  F  Y  E  A  F  G  L  T  V  V  E  A  M  T  C
GGTTTGCCTACATTTGCAACTAATCATGGCGGTCCAGCTGAGATCATCGTTAACGGAAAATCTGGCTTCCACATCGATCCATATCACGGTGAGCAAGCTGCTGATCTG
 G  L  P  T  F  A  T  N  H  G  G  P  A  E  I  I  V  N  G  K  S  G  F  H  I  D  P  Y  H  G  E  Q  A  A  D  L
CTAGCTGATTTCTTTGAGAAATGTAAGACAGAACCTTCTCATTGGGAAACCATTTCAACGGGTGGCCTGAAGCGCATCCAAGAGAAGTACACGTGGCAAATCTACTCG
 L  A  D  F  F  E  K  C  K  T  E  P  S  H  W  E  T  I  S  T  G  G  L  K  R  I  Q  E  K  Y  T  W  Q  I  Y  S
GAGAGGCTATTGACATTGGCTGCTGTTTACGGGTTCTGGAAACATGTTTCTAAGCTTGATCGTCTAGAAATCCGTCGATATCTTGATATGTTTTATGCTCTCAAATAC
 E  R  L  L  T  L  A  A  V  Y  G  F  W  K  H  V  S  K  L  D  R  L  E  I  R  R  Y  L  D  M  F  Y  A  L  K  Y
CGCAAGATGGCTGAAGCTGTTCCATTGGCTGCTGAGTGAAGCGAGAGATACCATTTGAGATGACCAAAAAAAACTGCCAATTTGGTTGAATAAAATGCTGTGACAAGAGG
 R  K  M  A  E  A  V  P  L  A  A  E  *
CTTCTATTCTCTATTTTTAGAAAATTATTACTGTGTAGTGATCCTCTTTTCCCTATTTGTATGCCAACTACCCCTCCCCCCTCCACCCCTTTTAATCTCTCTTACTTTTT
TTTTGTCTGCATTTTATCCTGTATTCCTGCCAATATTGGTTGGAGAAATTGCAAATTCGATATTCGGAGCTGTTAGTGTGAATAAAAGATCAAATTTCAATCTGTCCAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
```

Figure 2. The complete mRNA of tobacco sucrose synthase gene and its encoding amino acids *indicates the stop codon

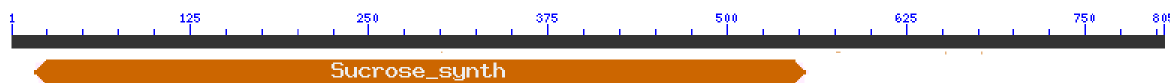

Figure 3. The putative Sucrose_synth domain of the protein encoded by tobacco sucrose synthase gene

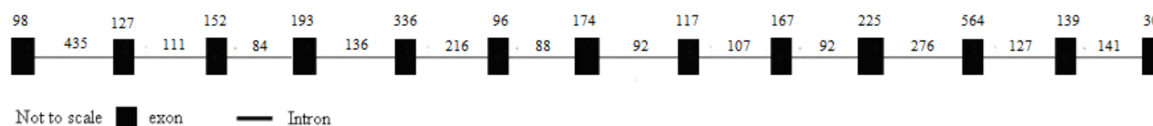

Figure 4. The genomic sequence organizations representing the coding region of the sucrose synthase gene

https://doi.org/10.1051/e3sconf/202129203101

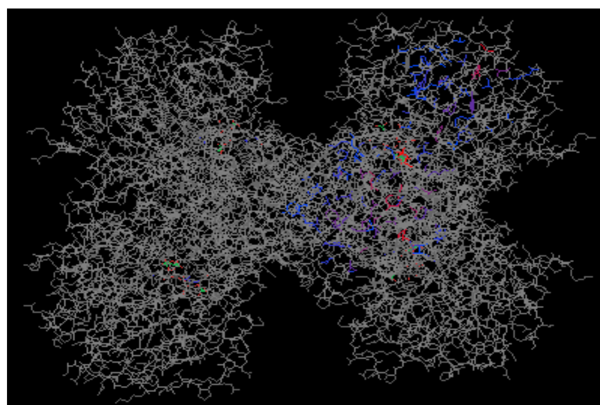

Figure 5. The 3-D structural evidence of the putative conserved domain of tobacco sucrose synthase protein

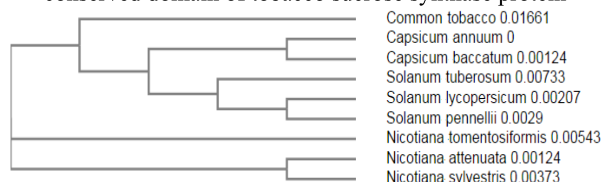

Figure 6. The phylogenetic tree for six kinds of sucrose synthase genes

To obtain the genomic DNA of coding region of common tobacco sucrose synthase gene, the publicly available common tobacco whole genome shotgun sequence database was screened using the coding sequences of common tobacco sucrose synthase geneas seeds by BLAST analysis. Twooverlapping genome shotgun sequences (GenBank accession no.NW_015916701.1 and NW_015909541.1) were found to encompass the entire sucrose synthase gene. The common tobacco sucrose synthase gene is 4,323-bp in length and consists of 13 exons and 12 introns. All exon–intron splice junction sequences conform to the GT–AG rule (Figure 4).

3.3 Tissue expression profile

Tissue expression profile analysis was carried out and results revealed that the tobacco sucrose synthase gene was highly expressed in leaf, but moderately expressed in root, stemand flower (Figure 7).

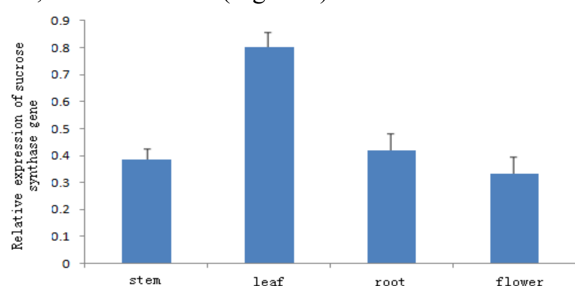

Figure 7. Expression analysis of sucrose synthase gene mRNA in various tissues

4 Conclusions

More than fourteen sucrose synthaseisoforms have been found in common tobacco plants and they display different developmental expression patterns[1]. Comparative analyses of gene structure and mRNA sequence indicated that this tobacco sucrose synthase gene we isolated is different with allcommon tobacco sucrose synthaseisoform genes. So that this is a novel tobacco sucrose synthase gene.

The phylogenetic tree analysis revealed that the tobacco sucrose synthase gene has a closer genetic relationship with that of capsicumannuum. This implied that we can use capsicumannuum as model organism to study the tobacco sucrose synthase gene or use tobacco as model organism to study the capsicumannuum sucrose synthase gene.

From the tissue distribution analysis in our experiment it can be seen that this sucrose synthase gene was obviously deferentially expressed in some tissues and was highly expressed in leaf. The expression patterns of genes can often partly reveal their likely physiological functions[1]. For sucrose synthase functions in UDP-glycosyltransferase activity and sucrose synthase activity[4-6], this indicated that the sucrosebiosynthesis of this novel tobaccosucrose synthase gene is higher in leaf.

In conclusion, we first isolated a novel tobacco sucrose synthase gene and performed necessary sequence analysis and tissue expression profile analysis. This established the primary foundation for further research on this tobacco gene.

Acknowlegement

This work was supported by The Research and Development Foundation of Yunnan Agricultural University (KX900187000) and The Basic Research Foundation of Yunnan Tobacco Industry Co. Ltd."Construction and application of R&D information platform based on market demand orientation"(2019xx01).

References

1. Wang Z., Wei P., Wu M., Xu Y., Li F., Luo Z., Zhang J., Chen A., Xie X., Cao P., Lin F., Yang J.., Planta. 242, 14 (2015)
2. Bertrand A., Bipfubusa M., Claessens A., Rocher S., Castonguay Y., Plant Sci. 264, 7 (2017)
3. Takehara K., Murata K., Yamaguchi T., Yamaguchi K., Chaya G., Kido S., Iwasaki Y., Ogiwara H., Ebitani T., Miura K., Breed Sci. 68, 7 (2018)
4. Zrenner R., Salanoubat M., Willmitzer L., Sonnewald U., Plant J. 7, 11 (1995)
5. Tang G.Q., Sturm A., Plant Mol Biol. 41, 14 (1999)
6. Thévenot C., Simond-Côte E., Reyss A., Manicacci D., Trouverie J., Le Guilloux M., Ginhoux V., Sidicina F., Prioul J.L., J Exp Bot. 56, 15 (2005)

7. Baroja-Fernández E., Muñoz F.J., Li J., Bahaji A., Almagro G., Montero M., Etxeberria E., Hidalgo M., Sesma M.T., Pozueta-Romero J., Proc Natl Acad Sci U S A. 109, 6 (2012)
8. Baud S., Vaultier M.N., Rochat C., J Exp Bot. 55, 13 (2004)
9. Fallahi H.et al., J Exp Bot. 59, 13 (2008)
10. Barratt D.H., Derbyshire P., Findlay K., Pike M., Wellner N., Lunn J., Feil R., Simpson C., Maule A.J., Smith A.M., Proc Natl Acad Sci U S A. 106, 9 (2009)
11. Nian F., Zhang Y., Su X., Zou J. and Zhao L., IJAST. 1,5 (2013)
12. Li Y., Meng F., Yin J., Liu H., Si Z., Ni Z., Sun Q., Ren J., Niu H., Biochim Biophys Acta. 1779, 9 (2008)
13. Livak K.J., Schmittgen T.D., Methods. 25, 7 (2001)