

MOLECULAR CHARACTERIZATION AND EXPRESSION PATTERN OF TOBACCO (*Nicotiana tabacum*) APOPLASTIC INVERTASE GENE

CARACTERIZAÇÃO MOLECULAR E PADRÃO DE EXPRESSÃO DO GENE DA INVERTASE DO APOPLASTO EM *Nicotiana tabacum*

Nian FUZHAO; Zhao LEIFENG*

College of Tobacco Science, Yunnan Agricultural University, Kunming 650201, China

*Author to whom correspondence should be addressed; E-Mail: zhleifeng@163.com

ABSTRACT: The complete mRNA sequence of one tobacco (*Nicotiana tabacum*) gene- apoplastic invertase, was amplified using the rapid amplification of cDNA ends methods based on some tobacco ESTs. The full-length tobacco apoplastic invertase gene mRNA was 1,985bp containing an 1,740 bp open reading frame, which encodes a protein of 579 amino acids. Sequence analysis revealed that the apoplastic invertase of tobacco shares high homology with the apoplastic invertase of potato (82%), *Lycopersicon esculentum* (81%) and *Lycopersicon pennellii* (78%). Results also showed that tobacco apoplastic invertase gene has a closer genetic relationship with the apoplastic invertase gene of *Lycopersicon esculentum*. The prediction of transmembrane helices showed that tobacco apoplastic invertase might be a transmembrane protein. The expression profile was studied and the results indicated that tobacco apoplastic invertase gene was differentially expressed in detected tobacco tissues including leaf, stem, root and flower. Our experiment established the foundation for further research on this tobacco gene.

KEYWORDS: Tobacco. Apoplastic invertase. Tissue expression

INTRODUCTION

Apoplastic invertase is a member of glycosyl hydrolase family GH32. Glycosyl hydrolase family GH32 cleaves sucrose into fructose and glucose via beta-fructofuranosidase activity, producing invert sugar that is a mixture of dextrorotatory D-glucose and levorotatory D-fructose, thus named invertase. This family also contains other fructofuranosidases such as inulinase, exo-inulinase, levanase, and transfructosidases such as sucrose:sucrose 1-fructosyltransferase, fructan:fructan 1-fructosyltransferase, sucrose:fructan 6-fructosyltransferase, fructan:fructan 6G-fructosyltransferase and levan fructosyltransferases. These retaining enzymes catalyze hydrolysis in two steps involving a covalent glycosyl enzyme intermediate: an aspartate located close to the N-terminus acts as the catalytic nucleophile and a glutamate acts as the general acid/base; a conserved aspartate residue in the Arg-Asp-Pro (RDP) motif stabilizes the transition state. These enzymes are predicted to display a 5-fold beta-propeller fold as found for GH43 and CH68. The breakdown of sucrose is widely used as a carbon or energy source by bacteria, fungi, and plants. Invertase is used commercially in the confectionery industry, since fructose has a sweeter taste than sucrose and a lower tendency to crystallize. A common structural feature of all these enzymes is a 5-bladed beta-propeller domain, similar to GH43, that contains the catalytic acid and catalytic base. A long V-shaped groove,

partially enclosed at one end, forms a single extended substrate-binding surface across the face of the propeller[1-5] <http://www.ncbi.nlm.nih.gov/Structure/cdd/cd08996>.

Until today, the tobacco apoplastic invertase 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.

MATERIAL AND METHODS

Samples collection, RNA extraction and first-strand cDNA synthesis

Tobacco plants (Chinese local 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. Total RNA extraction and first-strand cDNA synthesis for these tissue samples were performed as the methods describe by Li et al. [6].

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 apoplastic invertase gene, the gene specific

primers (GSPs) were designed based on the coding sequence information from wine grape apoplastic invertase gene and its highly homologous tobacco EST sequence: FG186035. 5'-RACE GSP: 5'-ATGTCACCTCAACATCAGCCTGTGC-3' 3'-RACE GSP: 5'-TTAAAGGAATCACACCTGCACAGGC-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, 67°C 30 sec, 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.

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

qRT-PCR for evaluating the level of mRNA for apoplastic invertase gene 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, 56°C (optimal annealing temperature for each specific primer) for 15 s (Table 1), 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 [7].

Table 1. qRT-PCR primers for tobacco apoplastic invertase, actin genes and annealing temperature

Gene	Primer sequence	Ta/ °C	Length/(bp)
apoplastic invertase	Forward : 5'- CTGCCTTGTCCAAACTTGC3' Reverse: 5'- CGGATATTATGCCCGATGA-3'	56	255
Actin	Forward :5'- CCATTCTTCGTTTGGACCTT -3' Reverse: 5'- TTCTGGGCAACGGAACCT-3'	56	257

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 Clustalw software (<http://www.ebi.ac.uk/clustalw>).

RESULTS AND DISCUSSION

RACE results for tobacco apoplastic invertase gene

For tobacco apoplastic invertase gene, through 5'-RACE, one PCR product of 1,292bp was obtained. The 3'-RACE product was 735bp. These products were then cloned to T-vector and sequenced. Taken together, a 1,985 bp cDNA complete sequence was finally obtained (**Figure 1**).

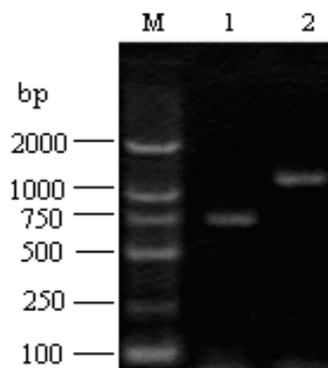


Figure 1. RACE results for tobacco apoplastic invertase gene. M DL2000 DNA markers; 1, 3'-RACE product for tobacco apoplastic invertase gene; 2, 5'-RACE product for tobacco apoplastic invertase gene.

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: KF308284).

The sequence prediction was carried out using the GenScan software and results showed that the 1,985-bp cDNA sequence represents one single

gene which encodes 579 amino acids. The theoretical isoelectric point (pI) and molecular weight (Mw) of the deduced proteins of these three tobacco genes were also computed using the Compute pI/Mw Tool (http://www.expasy.org/tools/pi_tool.html). The pI of tobacco apoplastic invertase is 8.87. The molecular weight of this putative protein is 65971.09.

```
TTCCTTCAGTTTCTTCTCCACATCACAAAAATTAAGGAAAAGAAAAACA AAAAGTGGAGTTTTTTAAGAAAAGCTCTTCTCTTTGGGAT
M E F F K K S S S L W D
ACTTTTGC AATTTCC TTAGTTAGTTTCATTGTAGTTTTTCCGAATATTAATGGAGTGTTTGGTCCCAAAAGTTTATATAGACTTACAA
T F A I S L V S F I V V F P N I N G V F G S H K V Y I D L Q
TCTTCAAGTCTGTTAATGTCAAAAATGCCACAGAAGTGGTTTCCATTTC AACCCACCAAGCATTGGATCAATGCTCCAATGTATTAC
S S S A V N V K N V H R T G F H F Q P P K H W I N A P M Y Y
AATGGAATCTATCATCTATTCTATCAATACAATCCAAAAGGATCAGTTGGGGT AATATTGTTGGGCTCATT CAGTTTCAAAAAGATTG
N G I Y H L F Y Q Y N P K G S V W G N I V W A H S V S K D L
ATTAATTTGGATTCATTAGAGCCTGCAATTTATCCATCTAAAAAATTTGACAAAATATGGAACATGGTCTGGATCAGCAACCAATTCCTCT
I N W I H L E P A I Y P S K K F D K Y G T W S G S A T I L P
AGTAAC AAGCCATTATTTTATATACTGGACTAGTAGATTCCAATAAACCCAAAGTCCAAAATTACGCCGTCCCGCAACTTATCCGAT
S N K P I I L Y T G L V D S N K T Q V Q N Y A V P A N L S D
CCATATCTCGTGAATGGATCAAGCCGATAACAACCCGTTAATCGTCCCCGATATCAGTATCAAC AAGACCC AATTTCCGACCCGCACA
P Y L R E W I K P D N N P L I V P D I S I N K T Q F R D P T
ACAGCTTGGATGGCCAAAGATGGTCATTGGAGAAATATAATGGGAAGTTTGAGAAAAAGAGGGGATTGGCAATAATGTATAGAAAGTAAG
T A W M G K D G H W R I I M G S L R K K R G L A I M Y R S K
GATTTTATGAGATGGATTAAGGCTAAACATCCATTCATTCACTGCTAATACAGGAAATGGGAAATGTCCTGATTTTTCCCTGTATCA
D F M R W I K A K H P L H S T A N T G N W E C P D F F P V S
TTACAAGGTACAATGGTTAGATAAATATGGTGAAGATTC TAAGTACGTACTTAAAAATAGTATGGATCTTACTAGTGTGAGTACTAT
L Q G T N G L D K Y G E D S K Y V L K N S M D L T R F E Y Y
ACTGTTGGTACGTACGATATTA AAAAAGACAGGTATATTCCAGATAACACTTCGGTTGATAGTTGGAAGGGATTGAGACTTGACTATGGA
T V G T Y D I K K D R Y I P D N T S V D S W K G L R L D Y G
AATTTCTACGCATCTAAATCATTTTATGATCCTAGCAAGAAATCGAAGAGTTACTTGGGGTGGGCTAACGAATCGGATATTATGCCCGAT
N F Y A S K S F Y D P S K N R R V T W G W A N E S D I M P D
GATGATATTAAAAAGGATGGGCTGGAATTC AAC TAATCCACGTAAATATGGCTCGACCTAGTGGTAACAAATGGTTCAATGCCCT
D D I K K G W A G I Q L I P R K L W L D P S G K Q L V Q W P
GTCGAAGAATTAGAACTCTAAGAAAACAAAAGTTCCAATTGAGCAACAAGAAGTTGAAC AATGGAGAAATGGTTGAGGGTTAAAGGATC
V E E L E T L R K Q K F Q L S N K K L N N G E M V E V K G I
ACACCTGCACAGGCTGATGTTGAAGTGACATTCTCTTTCGCAAGTTTGGACAAGGCAGAGCCATTGATCCTAATTTGGACTGATCTTTAT
T P A Q A D V E V T F S F A S L D K A E P F D P N W T D L Y
GCACAAGATGTTTGTCAATTAAGGGTTCAACTGTTC AAGGAGGGCTTGGCCATTGGTATGCAACATTGGCTTCTCAAAACTTGAAA
A Q D V C A I K G S T V Q G G L G P F G I A T L A S Q N L E
GAATACACACCTGTTTCTTTAGAGTTTTCAAAGCACAAACAAAATTAACAAGTTCTCATGTGCTCTGATGCCACAAGGTCAACTCTTAAG
E Y T P V F F R V F K A Q Q N Y K V L M C S D A T R S T L K
TACAACGAAACAATGTATAAACCTTCGTTGCTGATATGTGGATGTAGTTTACGACACAAAAGTTGCTCTCAGGAGTTTGATTGAT
Y N E T M Y K P S F A G Y V D V D L A D K K L S L R S L I D
AATTCAGTAGTGGAGAGTTTGTGCTGTGGTGA AAAACATGCATAACATCAAGAGTTTATCCAACATTGGCAATTTATGACAAGGCACAT
N S V V E S F G A G G K T C I T S R V Y P T L A I Y D K A H
TTATTTGCCTTTAAACAATGGTACGGAGCCAATTACAATTGAGACTCTAGATGCATGGAGTATGGGTAATGCTAAGATAGAAATAAAGTTGG
L F A F N N G T E P I T T I E T L D A W S M G N A K I E *
GAGAAAATATCACAAAAGGGGAAAAATAAGAATTTGATGGACATGGGAGTTTGTGCTAGATATGATTGTTACAACCTATTGTGCCAA
TAGACAAAAGCAAACCTGTAATATCCTCTATTTTAAATTTATGAGAATACTATGTTTTTTCCTTTCAAAAAAAAAAAAAAAAAAAAAA
```

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

Further BLAST analysis of this protein revealed that tobacco apoplastic invertase shares high homology with the apoplastic invertase of potato (Accession number: AEV46317, 82%), lycopersicon esculentum (Accession number: AAO45698, 81%) and lycopersicon pennellii (Accession number: CAB85898, 78%) (Figure 3).

Its conserved domains was identified as GH32_B_Fructosidase (Glycosyl hydrolase family 32, beta-fructosidase) (Figure 4).

The 3-D structural evidence of the putative conserved domain is also presented in Figure 5.

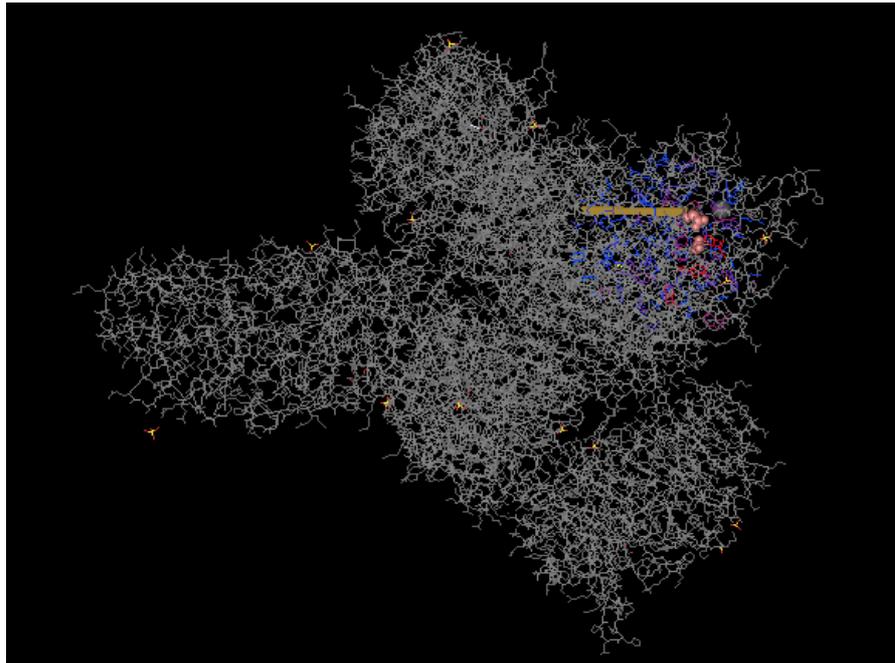


Figure 5. The 3-D structural evidence of the putative conserved domain of tobacco apoplastic invertase protein

The prediction of transmembrane helices in protein using the TMHMM Server v. 2.0 (<http://www.cbs.dtu.dk/services/TMHMM/>) showed

that tobacco apoplastic invertase protein might be a transmembrane protein (Figure 6).

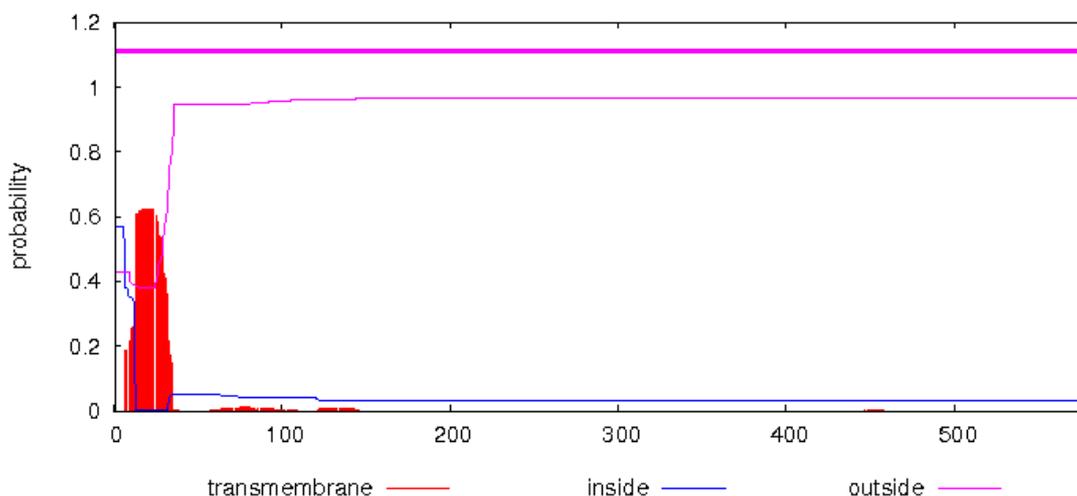


Figure 6. The transmembrane protein prediction of tobacco apoplastic invertase protein

Based on the results of the alignment of different species of apoplastic invertase proteins, a phylogenetic tree was constructed using the ClustalW software (<http://www.ebi.ac.uk/clustalw>),

as shown in Fig.7. The phylogenetic tree analysis revealed that the tobacco apoplastic invertase gene has a closer genetic relationship with that of *lycopersicon esculentum*.

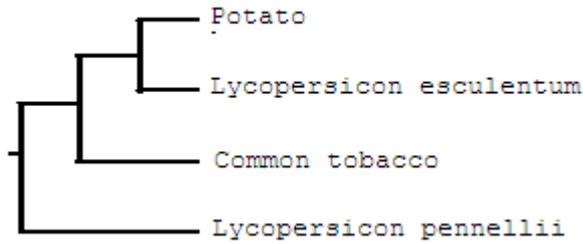


Figure 7. The phylogenetic tree for five kinds of apoplastic invertase genes

Tissue expression profile

Tissue expression profile analysis was carried out and results revealed that the tobacco

apoplastic invertase gene was highly expressed in leaf, but moderately expressed in root, stem and flower (Figure 8).

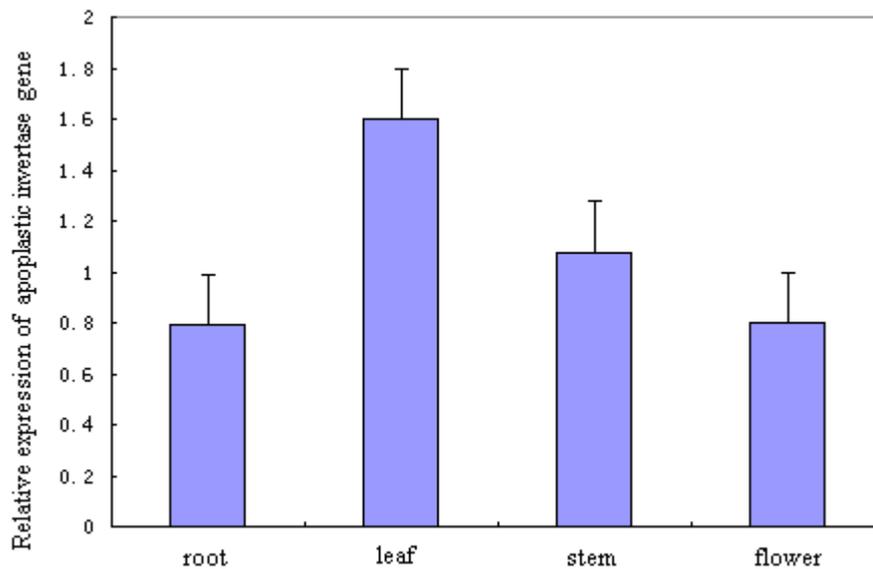


Figure 8. Expression analysis of apoplastic invertase gene mRNA in various tissues.

Comparative genomics research has revealed that virtually all (99%) of the protein-coding genes in humans align with homologs in mouse, and over 80% are clear 1:1 orthologs for human and mouse both belong to mammalian [8]. This extensive conservation in protein-coding regions implied that this conservation of protein-coding sequences may be expected in tobacco and lycopersicon esculentum for they are all plants. From the sequence analysis of apoplastic invertase genes, it can be seen that the coding sequences of apoplastic invertase genes were highly conserved in some plants. This implied that we can use the some plants as model organisms and isolate some tobacco genes based on the coding sequence information of these plants. Isolation of the tobacco apoplastic invertase gene in this experiment further validated that is an effective method.

The phylogenetic tree analysis revealed that the tobacco apoplastic invertase gene has a closer

genetic relationship with that of lycopersicon esculentum. This implied that we can use lycopersicon esculentum as model organism to study the tobacco apoplastic invertase gene or use tobacco as model organism to study the lycopersicon esculentum apoplastic invertase gene.

From the tissue distribution analysis in our experiment it can be seen that apoplastic invertase gene was obviously differentially expressed in some tissues. The suitable explanation for this under current conditions is that at the same time those biological activities related to the mRNA expression of this gene were presented diversely in different tissues.

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

RESUMO: A sequência completa do mRNA em fumo (*Nicotiana tabacum*) do gene da invertase apoplástica que foi amplificado usando a amplificação rápida da sua terminação pelo método usando cDNA. A sequência completa do mRNA foi de 1.985bp compreendendo uma ORF (“open reading frame”) de 1.740 pb que codificou 579 aminoácidos. Foi demonstrado uma homologia de 82 % nas sequências obtidas com a invertase apoplástica da batateira, 81 % com o tomateiro da espécie *Lycopersicon esculentum* e 78 % com *Lycopersicon pennellii*. Os resultados demonstraram uma homologia e relacionamento genético entre as espécies estudadas. A análise das hélices das transmembranas da invertase do apoplasto demonstraram que a mesma pode ser uma proteína da transmembrana. O padrão da expressão gênica estudado indicou que a invertase apoplástica dos tecidos do fumo incluindo, folha, caule, raiz e flor foi diferente. Novas pesquisas serão realizadas deste gene do fumo bem como sua fundamentação biológica.

PALAVRAS-CHAVE: Fumo. Invertase apoplástica. Expressão tecidual.

REFERENCES

- PARRENT, J. L. ; JAMES, T. Y. ; VASAITIS, R. ; TAYLOR, A. F. Friend or foe? Evolutionary history of glycoside hydrolase family 32 genes encoding for sucrolytic activity in fungi and its implications for plant-fungal symbioses. **BMC Evol Biol.** v. 30; n. 9, p. 148, Jun 2009. <http://dx.doi.org/10.1186/1471-2148-9-148>
- ALTENBACH, D.; RUDIÑO-PINERA, E.; OLVERA, C.; BOLLER, T.; WIEMKEN, A. RITSEMA, T. An acceptor-substrate binding site determining glycosyl transfer emerges from mutant analysis of a plant vacuolar invertase and a fructosyltransferase. **Plant Mol Biol.** v. 69, n. 1-2, p. 47-56, Jan 2009. <http://dx.doi.org/10.1007/s11103-008-9404-7>
- TSUJIMOTO, Y.; WATANABE, A.; NAKANO, K.; WATANABE, K.; MATSUI, H.; TSUJI, K.; TSUKIHARA, T.; SUZUKI, Y. Gene cloning, expression, and crystallization of a thermostable exo-inulinase from *Geobacillus stearothermophilus* KP1289. **Appl Microbiol Biotechnol.** v. 62, n. 2-3, p. 180-5, Aug 2003. <http://dx.doi.org/10.1007/s00253-003-1261-3>
- LAMMENS, W.; LE ROY, K.; SCHROEVEN, L.; VAN LAERE, A.; RABIJNS, A.; VAN DEN ENDE, W. Structural insights into glycoside hydrolase family 32 and 68 enzymes: functional implications. **J Exp Bot.** v. 60, n. 3, p. 727-40, 2009. <http://dx.doi.org/10.1093/jxb/ern333>
- YUAN, S.; LE ROY, K.; VENKEN, T.; LAMMENS, W.; VAN DEN ENDE, W.; DE MAEYER, M. pKa modulation of the acid/base catalyst within GH32 and GH68: a role in substrate/inhibitor specificity? **PLoS One.** v. 7, n. 5, p. 37453. 2012. <http://dx.doi.org/10.1371/journal.pone.0037453>
- LI, Y. MENG, F.; YIN, J.; LIU, H.; SI, Z.; NI, Z.; SUN, Q.; REN, J.; NIU, H. Isolation and comparative expression analysis of six MBD genes in wheat. **Biochim Biophys Acta**, 1779, n. 2, p. 90–98, 2008. <http://dx.doi.org/10.1016/j.bbagr.2007.09.004>
- LIVAK, K. J.; SCHMITTGEN, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) **Methods**, v. 25, n. 4, p. 402-8. Dec, 2001. <http://dx.doi.org/10.1006/meth.2001.1262>
- HARDISON, R. C (2003) **Comparative genomics.** PLoS Biol 1:E58. doi:10.1371/journal.pbio.0000058 <http://dx.doi.org/10.1371/journal.pbio.0000058>