Tobacco Research 45 (2) : 75 - 80, 2019

EXPRESSION ANALYSIS OF CYP82E GENES ASSOCIATED WITH TSNA IN *N. ATTENUATA* DATA HUB

T. SAROJA[1], K. PRABHAKARA RAO*[2], K. SARALA[2], T. DAVID SAMUEL RAJ[1] AND D. DAMODAR REDDY[2]

[1]V.S. KRISHNA COLLEGE, ANDHRA UNIVERSITY, VISAKHAPATANAM
[2]ICAR- CENTRAL TOBACCO RESEARCH INSTITUTE, RAJAHMUNDRY
* Corresponding author email: prabhakarabt@yahoo.co.in

(Recieved on 31st Aug, 2019 and accepted on 23rd Nov, 2019)

Tobacco is a high value commercial crop for varied traditional end uses smoking, chewing, hookah etc. apart from their medicinal values. Of late, the ill effects of tobacco use is becoming the major issue of concern. The health associated risks of tobacco consumption are majorly attributed to TSNAs, these are the nitrosated products of Nicotine and nornicotine, the major secondary metabolites reported in tobacco. The process of nitrosation is mediated by a key enzyme N-demethylase, which is encoded by CYP gene family. Among the various members of CYP family CYP82E genes are known to be the key functionaries in the regulation of TSNA. The expression these genes in the native state in various tissue and in response to various external cues will aid in understanding their regulation. In this regard an insilico expression analysis was carried out with the known CYP82E gene in the expression data base of Nicotiana attenuata, a close wild relative of *N._tabacum*. The NaDH currently hosts collections of predicted protein coding sequences of two recently sequenced Nicotiana species, and their functional annotations, 222 microarray datasets from 10 different experiments, a transcriptomic atlas based on 20 RNA-seq expression profiles and a metabolomic atlas. It provides a centralized platform for integrating and visualizing genomic, phylogenomic, transcriptomic and metabolomic data. The BLAST analysis of the candidate CYP genes with NaDH data hub revealed their corresponding gene identifiers which were further utilized for their expression analysis in transcriptome experiments. The CYP82E4 and E10 genes have identifier with more than 90% similarity (NIATv7_g20333.t1). In the micro array data, it was shown that the expression of CYP82E Id got slightly reduced in response to coronatine spray in treated plants compared to control. When plants were challenged with M.sexta neonates a lepidopteran pest the CYP82E Id expression got reduced in wild type. In herbivore attack it has shown strong expression in roots compared to leaves. The transcriptomic data in RNA sequencing experiments also suggests the prominent expression of genes in roots compare to other parts of the plants. In other expression datasets the CYP82E Id genes have shown differential expression. It indicates that CYP82E Id genes of TSNA shows tissue specific expression and the external cues like chemicals and microbes have relatively low effect on the regulation of these genes.

INTRODUCTION

Tobacco (*Nicotiana tabacum*) is an annual commercial crop cultivated all over the world with a long history of usage. It is being used for varied end uses including smoking, chewing, hookah etc. apart from the medicinal values. In the last few decades, there is an increasing global concern about the specific constituents of tobacco and their derivatives due to the possible detrimental effect on human health. Some of these are Tobacco-specific nitrosamines (TSNAs), derivatives of tobacco alkaloids considered to be the most important carcinogens in smokeless tobacco products and cigarette smoke. They are produced primarily during the curing and/or fermentation of the leaf (Bush *et al.,* 2001). Several TSNAs have been identified, but interest has focused on the four most important derivatives namely, 4-(methyl nitrosamino)-1-(3-pyridyl)-1-butanone (NNK), N-nitroso nornicotine (NNN), N-nitroso anabasine (NAB) and N2 -nitroso anatabine (NAT). Among these NNN and NNK have been shown to be strong

Key words: TSNA, Nornicotine, Tobacco

carcinogens in numerous animal studies, whereas NAT and NAB appear to be either weakly carcinogenic or benign (reviewed in Hecht, 1998 and Hecht, 2003). Previously it was predicted that majority of the NNN found in cigarette tobacco originated from nicotine (Hecht et al., 1978). Later investigations, however, demonstrated that NNN accumulation in cured tobacco leaves was correlated with nornicotine levels, (Bush *et al.,* 2001). Further, Ortholog searches revealed that highly expressed Nicotine N-Demethylase (NND) genes belong to CYP82E family and are responsible for its high level of nornicotine through conversion of nicotine (Sierro *et al.,* 2013). Concerted efforts are being made across the globe to regulate the levels of TSNA in cultivated tobacco (Lewis *et al.,* 2008; Li *et al.,* 2012). In the recent past, the wild relative of cultivated tobacco *Nicotiana attenuata*, has been developed as a model organism to study plant-environment interactions (Schuman *et al.,* 2012; Dinh *et al.,* 2013), and a large number of transcriptomic and metabolomic datasets have been generated with this plant. More than 230 transcriptomic data from *N. attenuata* have been submitted to the NCBI GEO database (Brockmoller *et al.,* 2017). The tools for centralizing, integrating and visualizing the omics data from sequenced and annotated genomes of *N. attenuata* and its close relative *N. obtusifolia*, are extremely useful in finding and characterizing the gene expression, tissue specificity, co expression and metabolite analysis. Utilising these transcriptomic data and tools, in the present investigation expression analysis of CYP82E genes were carried out to study the transcript accumulation and tissue specificity in different experimental datasets.

MATERIALS AND METHODS

The experimental— data sets available at *Nicotiana attenuata,* data hub (NaDH) were used for the gene expression, tissue specificity and co expression analysis. It comprises predicted protein coding sequences of 11 plant species, including two recently sequenced *Nicotiana* species, and their functional annotations, micro array data of 10 different field experiments of herbivores, pollinators and microbes attacks including 222 microarray datasets from a transcriptomic atlas based on 20 RNA-seq expression profiles and a metabolomic atlas based on 895 metabolite spectra analyzed by mass spectrometry data. Visualization tools including Electronic Fluorescent Pictograph (eFP) browser, co-expression networks, BLAST analysis were utilized in the expression analysis. For expression similarity Gini correlation coefficient (Ma *et al.,*2012) and for tissue specificity Shannon entropy (Gerstberger *et al.,*2014) were used.

RESULTS AND DISCUSSION

Identification of CYP82E4 homologues in NaDH

The expression of genes in different data sets of NaDH can be visualized only through the respective gene IDs or identifiers. Among the various *CYP82E* genes *CYP82E4* and *E1* were selected for expression analysis based on the research findings of Sierro *et al.,*(2013). The Search tool of NaDH was used with the available *CYP82E4* and *E1* gene sequences of *Nicotiana tabacum* as query against NaDH data. The database facilitates the similar sequences with gene Ids based on annotation, based on orthology or based on sequence similarity. The search analysis with query sequence resulted in the identification of several of gene ids with varied levels of similarity. Among these, the gene id with more than 92% similarity (NIATv7_g20333) having all the features of *CYP82E4* and *E1* denoted as *CYP82E* Id (Identifier) hereafter was selected for further analysis.

Expression analysis of *CYP82E* Id in micro array data

The Electronic Fluorescent Pictograph (eFP) browser tool of NaDH aids in visualizing the expression values of transcripts in a perceptive way._–It integrates the 10 micro array experiments, including the 256 microarrays and RNA -seq- data of 20 tissues and metabolite data of 14 tissues of *Nicotiana*._—The eFB browse tool has been used to study the expression of *CYP82E* Id in the micro array data generated through various on-field experiments. In the expression analysis, it identifies the probe sets for the primary gene and the mean of this probes was used as expression value.

Expression of *CYP82E* Id under Coronatine spray

Coronatine is a bacterial toxin which helps in re-opening of stomata after they close in response

to pathogen associated molecular patterns in some infections. Stitz *et al.*,2014 treated the *N. attenuata* plants with 1µM coronatine spray in every alternative day for a period of 10 days and the transcriptomic data was generated from the tissue samples of corolla at development stage of flower buds along with control. The available transcriptomic data was analyzed for the expression of *CYP82E* Id and it revealed that it has an expression value of 0.43 levels in control flowers, the expression level slightly reduced to 0.38 in treated samples of corolla (Table 1). It recorded slight decrease in the expression of *CYP82E* Id with Coronatine spray.

Expression of *CYP82E* Id in wild and engineered genotype for regulated cytokines

Cytokinins are regulators of a variety of developmental processes including delay of leaf senescence. A microarray dataset was generated with wild *N. attenauata* genotype and the engineered *N. attenauata* for delayed senescence (Senescence Associated Genes, Iso Pentynyl Transferase; SAG: IPT) challenged with *Manduca sexta* neonates for 3 days at early flowering stage along with controls. This data was analyzed for the expression of *CYP82E* Id in all the four lines. The expression levels of gene got reduced with *M.*

Table 1: Expression levels of *CYP82E* Id gene with respect to coronatine spray

S. No.	Treatment	Time	Developmental Stage	Expression Level
1	1 micro M Coronatine	10 day	flower buds the day before opening	0.38
2	control spray	10 day	flower buds the day before opening	0.43

Table 2: Expression levels of *CYP82E* Id gene in wild and genotype with regulated cytokines and senescence.

S. No.	Genotype	Treatment	Time	Tissue	Developmental Stage	Expression Level
1	SAG:IPT (higher cytokinins)	control	3d	young rosette leaf	Early flowering stage (6 weeks)	0.68
2	SAG:IPT (higher cytokinins)	*M. sexta* feeding	3d	young rosette leaf	Early flowering stage (6 weeks)	0.44
3	WT	control	3d	young rosette leaf	Early flowering stage (6 weeks)	2.2
4	WT	*M. sexta* feeding	3d	young rosette leaf	Early flowering stage (6 weeks)	0.22

Table 3: The signal intensities of *CYP82E* Id gene in herbivore attack

			TimeThe signal intensities of CYP gene			
	Root Control	Root Os	Systemic Leaf control	Systemic Leaf Os	Treated Leaf Control	Treated leaf Os
0h		7.92	1.11	1.11	1.29	1.29
1h	7.92	7.4	1.11	1.12	1.29	0.83
5h	8.18	8.93	0.72	1.33	0.49	0.87
9h	8.05	8.66	1.69	1.4	1.49	0.65
13h	8.05	7.86	1.64	1.17	1.2	1.11

sexta feeding in wild type and engineered tobacco plant compared to their respective untreated controls (Table 2).

Expression of *CYP82E* Id in response to herbivore attack

The feeding habit of herbivorous lepidopteron pest *Manduca sexta* on *N. attenuata* has become a model case for studying the various aspects of biotic stress. It modulates a diverse set of plant hormones like JA, methyl jasmonate and ethylene along with rapid induction of other genes responsive to wounding or herbivore attack (Von Dahl *et al.,* 2004) during the process of feeding. Inorder to study the expression of *CYP82E* Id under biotic stress conditions a micro array dataset was selected that was generated with challenging *N attenauta* plants by herbivores (Kim *et al.*,2011). Under this, the plants at rosettes stage were challenged with wounding followed by application of oral secretions of *M. sexta.* After the treatment the treated leaves, systemic leaves and roots along with controls were collected in the interval of every 4 hours and transcriptomic data was generated.

The analysis of the signal intensities revealed that the expression of *CYP82E* Id gene is much higher (more than 4 folds) in root compare to leaf (Table 3). In root slight induction of transcript was observed after 5 hours of treatment and subsequent reduction thereafter. The similar pattern of induced expression after 5 hours of treatment was also noted in systemic and treated leaves but in subsequent periods they showed differential expression. In total, the *CYP82E* Id has not shown much variation in expression in response to herbivore attack like other wound inducing genes with a sudden spurt in expression.

Expression analysis of *CYP82E* Id gene in RNA-seq data

The advances in molecular biology tools, lead to the RNA sequencing of various crops across the globe and the resultant data sets were available in the web resources. In *N. attenauta*, RNA sequencing data of 20 different tissues viz., leaf

Table 4: The RNA sequencing data and expression of of *CYP82E* Id gene in various tissues of *N. attenuata*

S. No.	Tissue	Treatment/development stage	No. of expressed genes	Expression of CYP gene (TPM)
1	Anther	Mature anther no treatment	11,550	0.0
2	Corolla	Late developmental stage, no treatment	13,486	0.69
3	Corolla	Early developmental stage, no treatment	13,662	0.32
4	Flower	Fully opened flowers, no treatment	14,390	0.59
5	Flower bud	Two early developmental stages of flowers, no treatment	14,543	0.0
6	Leaf	Rosette stage plants, no treatment	11,840	0.31
7	Leaf	Rosette stage plants, treated with 5 μL 1:1 diluted *M. sexta* oral secretion three times in leaves	12,179	0.16
8	Nectary	Mature nectary, no treatment	12,928	0.08
9	Ovary	Mature ovary, no treatment	13,960	0.0
10	Pedicel	Mature pedicel, no treatment	14,550	0.11
11	Pollen tube	No treatment	3,490	0.0
12	Root	Rosette stage plants, treated with 5 μL 1:1 diluted *M. sexta* oral secretion three times in leaves	15,499	5.37
13	Seed	Dry seeds	8,681	0.0
14	Seed	Treated with water	8,872	0.0
15	Seed	Treated with liquid smoke	9,227	0.0
16	Stem	Rosette stage plants, treated with 5 μL 1:1 diluted *M. sexta* oral secretion three times in leaves	14,682	0.21
17	Stigma	Mature stigma, no treatment	14,485	0.0
18	Style	Mature style, pollinated with pollens from different genotype	13,365	0.12
19	Style	Mature style without pollination	13,492	0.0
20	Style	Mature style, self-pollinated	13,533	0.0

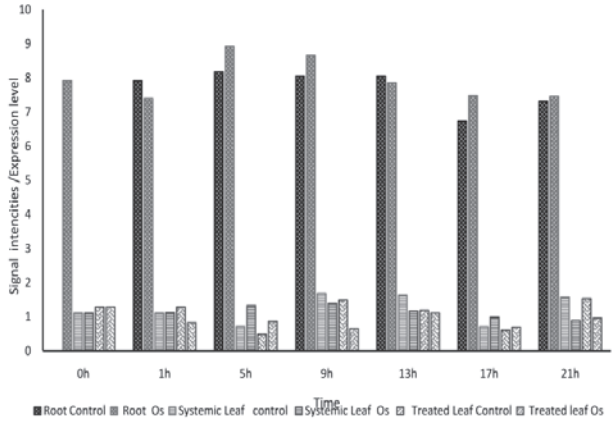

Fig. 1: The signal intensities denoting the expression of *CYP82E* Id gene under herbivore attack

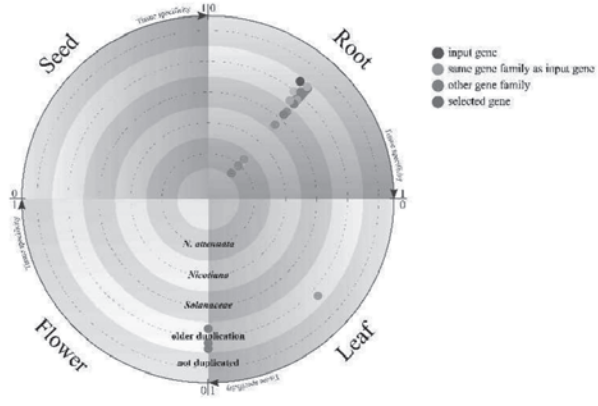

Fig. 2: Genes that are co-expressed with *CYP82E* Id gene (NIATv7_g20333) along with tissue specificity among the four major tissues.

control/treated, root treated, stem treated, flower buds, opening flower, corolla early/late, nectaries, ovary, style selfed/outcrossed/without pollination, pollen tubes, pedicels, stigma, anthers, seeds dry/watered/smoked with the expressed genes are available (Table 4). In order to analyze the tissue specific expression of *CYP82E* Id gene in various tissues all the data sets were analyzed with the gene probe set. The Expression of *CYP82E* Id was presented in log transformed TPM (Transcript Per Million) values. Among the various tissues, *Nioctiana* root has higher expression of *CYP82E* Id with 5.37 TPM compare to other tissues. No expression of *CYP82E* Id was found in Anther, Flower bud, Pollen tube, normal stigma, style and seed. It complements the previous data of microaaray experiments with herbivore attack, where root has maximum signal intensity compare to leaf. Among the floral parts corolla, has relatively higher expression of *CYP82E* Id gene.

Gene-to-gene co expression analysis of *CYP82E* Id

The gene-to-gene co expression analysis facilitates the search for potential co-regulated genes that may belong to the same pathway, have similar functions or are involved in similar biotic or abiotic responses. Further, the co expression analysis of gene generally aid in understanding the regulatory mechanisms and also predictive functions of the gene. Using the NaDH tools the gene to gene co expression analysis *CYP82E* Id (NIATv7_g20333) was carried out in 20 tissues, including different leaves, roots, flowers and seeds tissues adopting Gini correlation coefficient (Ma *et al.,* 2012) for expression similarity and Shannon entropy (Gerstberger *et al.*,2014) for tissue specificity. The similarity cutoff at 0.90 resulted in 1012 co expressed genes and when the cutoff was increased to 0.95 the co expressed genes were reduced to 318 and further increase of cut off to 0.99 the highest in the level the co expressed genes were reduced to 48. These co expressed genes majorly comprises stress related genes like transcription factor MYB39-like, small heat shock hsp G3-like, basic form of pathogenesis-related 1-like and some CYP genes. Among these expression of 44 genes are specific to root tissue (Fig.2) and are majorly stress and senescence specific. The conversion of nicotine to nor nicotine is mediated by a demythylase enzyme. Mann et al. (1958) reported the conversion of nicotine by a single dominant gene. Siminszky *et al.,* (2005) reported that a cytochrome P450 gene (CYP82E4v1) was involved in the conversion of nicotine to nornicotine. The *CYP82E* Id generally activates in senescent leaf for conversion of nicotine to nornictine and hence in the co expression also mostly senescence specific genes were observed.

In the micro array data, it was clearly shown that the expression of *CYP82E* Id got slightly reduced in response to coronatine spray in treated plants compared to control. When the plants were challenged with *M._sexta* neonates a lepidopteran pest the *CYP82E* Id expression got reduced in wild

type. In herbivore attack it has shown strong expression in roots compared to leaves. The transcriptomic data in RNA sequencing experiments also suggests the prominent expression of genes in roots compare to other parts of the plants. In other expression datasets the CYP82E Id genes have shown differential expression. It indicates that CYP82E Id genes of TSNA shows tissue specific expression and the external cues like chemicals and microbes have relatively low effect on the regulation of these genes.

REFERENCES

Bush, L.P., M. Cui, H. Shi, H.R. Burton, F.F. Fannin, L. Lei. and N. Dye. 2001. Formation of tobacco-specific nitrosamines in air-cured tobacco. **Rec. Adv. Tob. Sci**. 27: 23–46.

Brockmöller, T., L. Zhihao, L. Dapeng, G. Emmanuel, I. T. Baldwin and X. Shuqing. 2017. Nicotiana attenuata Data Hub (NaDH): an integrative platform for exploring genomic, transcriptomic and metabolomic data in wild tobacco. BMC Genomics 18:79.

Dinh, S.T., I.T. Baldwin and I. Galis. 2013. The HERBIVORE ELICITOR-REGULATED1 gene enhances abscisic acid levels and defenses against herbivores in *Nicotiana attenuata* plants. **Plant Physiol**.162 (4): 2106-2124.

Gerstberger, S., M. Hafner. and T. A. Tuschl. 2014. Census of human RNA-binding proteins. **Nat. Rev. Genet.** 15: 829–845.

Hecht, S.S., 1998. Biochemistry, biology and carcinogenicity of tobacco-specific Nitrosamines. **Chem. Res. Toxicol.** 11: 559–603.

Hecht, S.S., 2003. Tobacco carcinogens, their biomarkers and tobacco-induced cancer. **Nat. Rev. Cancer** 3: 733–744.

Hecht, S.S., C.B. Chen, N. Hirota, R.M. Ornaf, T.C. Tso. and D. Hoffmann. 1978. Tobacco specific nitrosamines: formation from nicotine in vitro and during tobacco curing and carcinogenicity in strain A mice. **J. Natl. Cancer Inst.** 60(4):819-24.

Kim, S.G., F. Yon, E. Gaquerel, J. Gulati. and I.T. Baldwin. 2011. Tissue specific diurnal rhythms of metabolites and their regulation during herbivore attack in a native tobacco, Nicotiana attenuata. **PLoS One**. 6(10): e26214.

Lewis, R.S., A.M. Jack, J.W. Morris, V.J.M. Robert, L. Gavilano, B. Siminszky, L.P. Bush, A.J. Hayes and R.E. Dewey. 2008. RNAi-induced suppression of nicotine demethylase activity reduces levels of a key carcinogen in cured tobacco leaves. **Plant Biotechnol. J**. 6: 346–354.

Li, D., R.S. Lewis, A.M. Jack, R.E. Dewey, S.W. Bowen and R.D. Miller. 2012. Development of CAPS and dCAPS markers for CYP82E4, CYP82E5v2, and CYP82E10 gene mutants reducing nicotine to nornicotine conversion in tobacco. **Mol. Breed.** 29: 589–599.

Ma, C and X. Wang. 2012. Application of the Gini correlation coefficient to infer regulatory relationships in transcriptome analysis. **Plant Physiol**. 160: 192–203.

Mann, T.J., J.A. Weybrew, D.F. Matzinger. and J.L. Hall. 1958. Inheritance of conversion of nicotine to nornicotine in varieties of *Nicotiana tabacum* and related amphiploids. **Crop Sci**. 4: 349-353.

Sierro, N., J.N. Battey, S. Ouadi, L. Bovet, S. Goepfert, N. Bakaher, M.C. Peitsch and N.V. Ivanov.2013. Reference genomes and transcriptomes of *Nicotiana sylvestris* and *Nicotiana tomentosiformis* **Genome Biol**. 14: R60.

Schuman, M.C., K. Barthel and I.T. Baldwin. 2012. Herbivory-induced volatiles function as defenses increasing fitness of the native plant *Nicotiana attenuata* in nature. **eLife.**1:00007.

Siminszky, B., L. Gavilano, S.W. Bowen. and R.E. Dewey. 2005. Conversion of nicotine to nornicotine in *Nicotiana tabacum* is mediated by CYP82E4, a cytochrome P450 monooxygenase. **Proc. Natl. Acad. Sci. U.S.A.** 102(41): 14919-14924.

Stitz, M., M. Hartl, I.T. Baldwin. and E. Gaquerel. 2014. Jasmonoyl-l-Isoleucine Coordinates Metabolic Networks Required for Anthesis and Floral Attractant Emission in Wild Tobacco (*Nicotiana attenuata*). **The Plant Cell**. 26 (10):3964-3983.

Von Dahl, C.C. and I.T. Baldwin. 2004. Methyl jasmonate and cis-jasmone do not dispose of the herbivore-induced jasmonate burst in *Nicotiana attenuata*. **Physiol Plant**.120:474–481.