Plant Cell Rep (2008) 27:117–124
DOI 10.1007/s00299-007-0456-y

GENETICS AND GENOMICS

Guarana (*Paullinia cupana var. sorbilis*), an anciently consumed stimulant from the Amazon rain forest: the seeded-fruit transcriptome

Paula C. S. Ângelo · Carlos G. Nunes-Silva · Marcelo M. Brígido · Juliana S. N. Azevedo · Enedina N. Assunção · Alexandra R. B. Sousa · Fernando J. B. Patrício · Mailson M. Rego · Jean C. C. Peixoto · Waldesse P. Oliveira Jr. · Danival V. Freitas · Elionor R. P. Almeida · Andréya Márcya H. A. Viana · Ana Fabíola P. N. Souza · Edmar V. Andrade · Pablo O. A. Acosta · Jaqueline S. Batista · Maria E. M. T. Walter · Luciana Leomil · Daniel A. S. Anjos · Rodrigo C. M. Coimbra · Magda H. N. Barbosa · Eduardo Honda · Soraya S. Pereira · Artur Silva · José O. Pereira · Marcicleide L. Silva · Mozart Marins · Francisca J. Holanda · Rusleyd M. M. Abreu · Silvana C. Pando · José F. C. Gonçalves · Margarida L. Carvalho · Emygdia R. R. B. P. Leal-Mesquita · Márcio A. da Silveira · Weber C. Batista · André L. Atroch · Suzelei C. França · Jorge I. R. Porto · Maria Paula C. Schneider · Spartaco Astolfi-Filho · Brazilian Amazon Consortium for Genomic Research (REALGENE)

Received: 22 May 2007 / Revised: 30 August 2007 / Accepted: 10 September 2007 / Published online: 5 October 2007

Abstract Guarana (*Paullinia cupana* var. *sorbilis*) is a plant native to the central Amazon basin. Roasted seed extracts have been used as medicinal beverages since pre-Colombian times, due to their reputation as stimulants, aphrodisiacs, tonics, as well as protectors of the gastrointestinal tract. Guarana plants are commercially cultivated

Communicated by H. Judelson.

P. C. S. Ângelo · A. L. Atroch
Embrapa Western Amazon, Manaus, Amazonas, Brazil

C. G. Nunes-Silva · E. N. Assunção · D. V. Freitas · E. V. Andrade · L. Leomil · J. O. Pereira · M. L. Silva · S. Astolfi-Filho (✉)
Centro de Apoio Multidisciplinar, Universidade Federal do Amazonas, Av. Gal. Rodrigo Octávio Jordão, 3000, CEP 69.077-000 Manaus, Amazonas, Brazil
e-mail: sastolfi@ufam.edu.br

M. M. Brígido · M. E. M. T. Walter · D. A. S. Anjos · R. C. M. Coimbra
Departamento de Biologia Celular, Universidade de Brasília, Brasília, Distrito Federal, Brazil

J. S. N. Azevedo · A. Silva · M. P. C. Schneider
Laboratório de Polimorfismo de DNA, Universidade Federal do Pará, Belém, Pará, Brazil

A. R. B. Sousa · A. F. P. N. Souza · J. S. Batista · S. C. Pando · J. F. C. Gonçalves · J. I. R. Porto
Instituto Nacional de Pesquisa da Amazônia (INPA), Manaus, Amazonas, Brazil

F. J. B. Patrício · A. M. H. A. Viana · M. H. N. Barbosa · E. R. R. B. P. Leal-Mesquita
Centro de Ciência Biológica e Saúde, Universidade Federal de Maranhão, São Luiz, Maranhão, Brazil

M. M. Rego · P. O. A. Acosta
Departamento de Biologia, Universidade Federal de Roraima, Boa Vista, Roraima, Brazil

J. C. C. Peixoto
Departamento de Ciências Biológicas, Universidade Federal do Amapá, Macapa, Amapá, Brazil

W. P. Oliveira Jr. · M. A. da Silveira
Laboratório de Biotecnologia, Universidade Federal do Tocantins, Palmas, Tocantins, Brazil

E. R. P. Almeida
Embrapa Natural Resources and Biotechnology, Brasília, Distrito Federal, Brazil

E. Honda · S. S. Pereira · F. J. Holanda · W. C. Batista
Instituto de Pesquisa em Patologias Tropicais, IPEPATRO, Porto Velho, Rondonia, Brazil

M. Marins · S. C. França
Unidade de Biotecnologia, Universidade de Ribeirão Preto, UNAERP, Ribeirão Preto, São Paulo, Brazil

R. M. M. Abreu · M. L. Carvalho
Departamento de Ciências da Natureza, Universidade Federal do Acre, Rio Branco, Acre, Brazil

exclusively in Brazil to supply the national carbonated soft-drink industry and natural product stores around the world. In this report, we describe and discuss the annotation of 15,387 ESTs from guarana seeded-fruits, highlighting sequences from the flavonoid and purine alkaloid pathways, and those related to biotic stress avoidance. This is the largest set of sequences registered for the Sapindaceae family.

Keywords Guarana · Oxidative stress · Caffeine · Flavonoids · Disease resistance

Introduction

The genus *Paullinia* is classified as the Sapindaceae family, which includes maples (*Acer* spp.), horse chestnuts (*Aesculus* spp.), and lychee (*Litchi chinensis*). There are approximately 195 species of *Paullinia* (http://www.mobot.org/ 2006) distributed in the American tropics and subtropics. Of these, at least nine species are described as native to Brazil, including *P. cupana* (Kunth) var. *sorbilis* [(Mart.) Ducke]. This cultivated variety is a liana that becomes a shrub under cultivation (Erickson et al. 1984), and was recently demonstrated to possess 210 chromosomes, which accounts for 22 pg of DNA per nucleus (Freitas et al. 2007). Its fruits, called "guaraná", are dehiscent capsules that contain one to three dark brown seeds.

Guarana is commercially cultivated only in Brazil where Embrapa Western Amazon maintains a breeding program. Annual production is about 4,000 tons of roasted seeds (IBGE 2006), of which 70% are consumed by soft drink industries to produce non-alcoholic carbonated beverages. Drugstores and alternative natural products stores sell the remainder, mainly in the form of roasted seed powder.

The popular motivations for consuming guarana extracts as dietary supplements are: weight loss, energy boost, improvement of fitness and sexual performance, and increase of cognitive capacity (O'Dea 2003; Oliveira et al. 2005; Ray et al. 2005).

Scientific support is forthcoming for some of these motivations. Kennedy et al. (2004) reported an improvement in cognitive performance in humans, possibly due to the combination of caffeine with other components of the extract. Relaxation of isolated rabbit cavernous tissue was reported by Antunes et al. (2001) and increases in blood glucose level with decreases in liver glycogen content were observed in mice fed guarana water extracts (Miura et al. 1998). In addition, inhibitory effects on induced hepatocarcinogenesis were reported in mice (Fukumasu et al. 2006a), gastro-protective properties were demonstrated in rats (Campos et al. 2003), anti-oxidant effects associated with a reduction in lipid peroxidation were also observed in rats (Mattei et al. 1998), and platelet anti-aggregation action was observed in cells treated with water extracts of guarana seeds (Bydlowski et al. 1991).

We have screened more than 15,000 ESTs from guarana seeded-fruits to identify sequences involved in several metabolic pathways that may contribute to the knowledge of this plant, beginning to unveil its biological properties and mechanisms of protection against biotic stress.

Materials and methods

Construction of cDNA library and sequencing

Healthy fruits containing seeds were collected at three developmental stages (green immature fruit, intermediary, and mature) from BRS-Amazonas plants maintained at Embrapa Western Amazon (Manaus, Amazonas State, Brazil). Total RNA was extracted with Concert Plant RNA Reagent (Invitrogen) and 123 μg of total RNA from each developmental stage were used for Poly A+ RNA purification with Fast Track Kit (Invitrogen). cDNA synthesis and cloning were performed using the Super Script Plasmid System with Gateway Technology (Invitrogen). Fragments from 600 to 1,000 bp were selected for insertion into pCMV.SPORT6 vectors. *Escherichia coli* XL1-Blue cell transformation resulted in 1.6×10^6 transformants/μg cDNA. Recombinant plasmids were prepared by the alkaline lysis method (Birnboim 1983) and EST sequences (Sanger et al. 1977) were determined by the DYEnamic ET dye terminator® cycle sequencing procedure using a MegaBACE 1000 capillary sequencer (GE-Amersham Pharmacia Biotechnologies).

EST assembly and analysis

Electropherograms were base called using PHRED (Green 1996). Vector sequences were trimmed by Cross-Match. Accepted sequences had at least 100 bases with a PHRED quality equal to or better than 20, and these were used to feed the CAP3 program (Huang and Madan 1999). Assembled groups and singletons were submitted to an automatic EST annotation procedure (Brigido et al. 2005). Each group was compared to three databases using the Blast X (Altschul et al. 1997) program: GenBank NR (http://www.ncbi.nih.gov), KOG (http://www.ncbi.nih.gov/KOG), and Swissprot release 44 (ftp.ebi.ac.uk/pub/databases/swissprot/release/). The annotation pipeline automatically captured the gene descriptions and EC numbers from Swissprot best hits and the product name and functional category from KOG libraries using 10^{-10} as the *e*-value cut off. Annotation was inspected manually by

comparing Blast results and the automatically annotated EC numbers and functional classifications.

Due to the caffeine contribution to the stimulant, and most popular effect of guarana drinks, ESTs for caffeine synthases (CSs) were reassembled separately and reexamined in search of the best organization and putatively complete ORFs. The more complete ORFs found in the reassembled sequences were translated and aligned using clustal W program (Thompson et al. 1994). A neighbor-joining phylogenetic tree was developed, grouping these sequences and methyltransferases from other sources using PHYLYP with the Jones–Thornton substitution matrix (Felseinstein 1998).

Results and discussion

The sequencing effort generated 15,387 guarana ESTs from the seeded-fruit cDNA clones. These sequences were deposited in GenBank (http://www.ncbi.nih.gov) under the accession numbers EC763506–EC778393. The CAP3 assembly yielded 8,597 groups, with 5,969 singletons and 2,628 contigs composed of 9,418 ESTs. The average contig contained two to three sequences (73% of the total). KOG entries to *A. thaliana* sequences were found to be the best hits for 5,685 guarana EST groups (contigs and singletons) using 10^{-10} as the *e*-value cut off (overview in Table 1).

The classification of 3,380 sequences into KOG categories (Fig. 1) indicated that 14% of the classified EST groups were categorized as involved in posttranslational modification, protein turnover, and chaperones. The next category was ribosomal structure, biogenesis and translation, and signal transduction mechanisms. Carbohydrate and energetic metabolism were also significantly found.

The contigs assembled from 20 or more ESTs are shown in Table 2, representing the most abundant transcripts.

Table 1 Summary of guarana seeded-fruit EST project results

Descriptive category	No. of ESTs
Total cDNAs sequenced	15,387
Number of clusters (contigs)	2,628
Number of singletons	5,969
Redundancy*	61%
Average EST length	773 bp
EST group matches with *e*-value $<10^{-10}$	
All organisms	5,956
Arabidopsis thaliana	5,685
No matches	2,641
KOG classified transcriptional units	3,380
G + C content	45.1

* (number of ESTs assembled in clusters/total number of ESTs) × 100

Most of them are distributed in either carbon fixation or carbohydrate related functions and storage protein metabolism.

Metallothioneins produced two of the most populated contigs holding 65 ESTs (Table 2). A high frequency of metallothioneins and other anti-oxidant related sequences was similarly found in the transcriptomes of other maturing fruits (Moyle et al. 2005; Davies and Robinson 2000; Grimplet et al. 2005). These enzymes have been implicated in reducing hydroxyl radical induced-DNA damage in watermelons (Akashi et al. 2004). A similar DNA protective effect was pointed out for guarana extracts (Fukumasu et al. 2006b).

Orthologues of late embryogenesis abundant (LEA) proteins were also found among the most abundant sequences. LEA proteins have been related to drought and salt resistance (Oraby et al. 2005).

Transcripts related to secondary metabolism and protection against biotic stress is highlighted in Table 3 and discussed below. These transcripts are targets for further investigation into the unique guarana biological properties and desirable agronomical characteristics.

Secondary metabolism transcripts

Secondary metabolism accounted for 4% of the EST groups according to KOG classification (Fig. 1). The *P. cupana* transcriptome included at least 129 sequences related to flavonoid metabolism, including all the key enzymes regulating the central pathway of its biosynthesis (Table 3), which indicates that this is a very active pathway in seeded-fruits. This is in agreement with the phytochemical evidence for this genus (Abourashed et al. 1999).

Flavonoids are powerful anti-oxidants. The health benefits of tea consumption are imputed to its high concentrations of flavonol dimers and polymers (Frei and Higdon 2003). Orthologues to enzymes involved in catechin and epicatechin biosynthesis, dihydroflavonol 4-reductase and anthocyanidin synthase, respectively, were found (Table 3). Catechins are believed to contribute to weight reduction (Murase et al. 2002) and to increase, in concert with caffeine, energy expenditure, as observed in men who received encapsulated guarana (Berube-Parent et al. 2005). Isolated tannins were found to exert anti-mutagenic and anti-inflammatory effects (Tobi et al. 2002; Sugisawa et al. 2004).

The stimulant properties of guarana drinks are reputedly due to the seeds' richness in caffeine, 2.7–5.8% of the dry weight compared to 1.0–2.4% reported for coffee seeds (Escobar et al. 1985; Weckerle et al. 2003). The high number of orthologues to enzymes of the caffeine

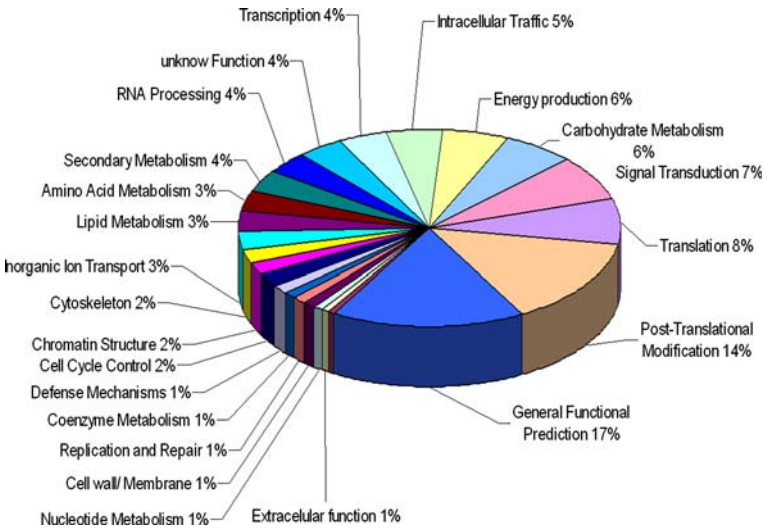

Fig. 1 Distribution of guarana ESTs based on the KOG functional categories

Table 2 The most populated contigs found in the guarana seeded-fruit transcriptome

ESTs	Annotation	*e*-value	Organism
83	Chlorophyll *a*/*b*-binding protein (cab proteins)	$1.0e^{-139}$	*Nicotiana sylvestris*
75	2S albumin precursor	$3.0e^{-17}$	*Ricinus communis*
57	RUBISCO small chain, chloroplast precursor	$1.0e^{-68}$	*Fagus crenata*
49	Oleosin isoform B	$9.0e^{-30}$	*Glycine max*
43	Metallothionein 2b	$1.0e^{-16}$	*Populus trichocarpa*
42	Mannose/glucose-specific lectin	$3.0e^{-16}$	*Parkia platycephala*
36	Fructose-bisphosphate aldolase	$1.0e^{-119}$	*Manihot esculenta*
34	unspecific monooxygenase	$1.0e^{-75}$	*Nicotiana tabacum*
28	RUBISCO activase 1	0.0	*Larrea tridentate*
27	Legumin type B precursor	$1.0e^{-38}$	*Anacardium occidentale*
26	Late embryogenesis abundant (LEA) M17 protein	$1.0e^{-10}$	*Arabidopsis thaliana*
26	Seed tetraubiquitin	$1.0e^{-123}$	*Helianthus annuus*
25	ADP-ribosylation factor	$5.0e^{-99}$	*Gossypium hirsutum*
24	Plant invertase/pectin methylesterase inhibitor	$1.0e^{-48}$	*Nicotiana tabacum*
22	Metallothionein-like protein	$2.0e^{-7}$	*Carica papaya*
21	Putative deoxycytidine deaminase	$1.0e^{-79}$	*Oryza sativa*
21	Predicted protein	$4.0e^{-12}$	*Magnaporthe grisea*
21	Aquaporin	$1.0e^{-110}$	*Ricinus communis*
20	Putative histone H2A	$5.0e^{-38}$	*Oryza sativa*

The contigs are represented by the number of reads, the annotation, *e*-value, and the orthologue source species

biosynthesis pathway (94 ESTs, Table 3) is in accordance with the abundance of caffeine in guarana seeds. This pathway is better understood in *Coffea arabica* (coffee; Mazzafera et al. 1994; Ogawa et al. 2001; Uefuji et al. 2003) and *Camellia sinensis* (tea, Kato et al. 1999, 2000), where it involves three *N*-methylation steps (xanthine→7-methylxanthosine→theobromine→caffeine), all of them performed by *S*-adenosine-L-methionine (SAM)-dependent methyltransferases or caffeine synthases (CSs). Guarana ESTs showed the highest similarity to two *C. sinensis* enzymes named TCS1 and TCS2 (Yoneyama et al. 2006). TCS1 is a dual-function methylxanthosine/theobromine methyltransferase (Kato et al. 1996; 1999).

A phylogenetic tree based on the distances among different CSs from *C. arabica*, *C. sinensis*, *T. cacao*, and CS contigs from guarana is shown in Fig. 2. Caffeine pathway methyltransferases have been demonstrated to group preferably according to the source plant rather than to their catalysis (Yoneyama et al. 2006) and this was confirmed in this work.

Table 3 Guarana ESTs related to secondary metabolism and biotic stress

Annotated product	GI	Organism	ESTs	Identity (%)
Secondary metabolism				
Flavonoid pathways				
Flavonol synthase	2,465,434	*Petunia hybrida*	34	47–81
Flavonoid hydroxylase	5,921,647	*Petunia hybrida*	30	46–72
Chalcone synthase	567,937	*Camellia sinensis*	17	70–94
Isoflavone reductase	3,243,234	*Pyrus communis*	17	53–87
Anthocyanidin synthase	296,844	*Malus* sp	13	47–82
Dihydroflavonol 4-reductase	33,313,474	*Malus domestica*	8	68–80
UDP-glucose: flavonoid 3-*o*-glucosyltransferase	13,620,873	*Vitis vinifera*	6	51–61
UTP-glucose glucosyltransferase	453,251	*Manihot esculenta*	3	52–56
UDP rhamnose: anthocyanidin-3-glucoside rhamnosyltransferase	397,567	*Petunia hybrida*	1	55
Purine alkaloid pathways				
Caffeine synthase (TCS2)	51,968,288	*Camellia sinensis*	73	45–62
Caffeine synthase (TCS1)	9,967,143	*Camellia sinensis*	21	43–60
Carotenoid pathways				
Beta-carotene hydroxylase	11,245,486	*Citrus unshiu*	6	51–79
9-*cis*-Epoxycarotenoid dioxygenase	38,112,198	*Vitis vinifera*	5	42–84
Carotenoid cleavage dioxygenase 1	49,659,726	*Lycopersicon esculentum*	5	65–92
Zeta-carotene desaturase ZDS2	33,313,474	*Malus domestica*	1	88
Biotic stress				
Pathogenesis-related proteins	7,269,429	*Arabidopsis thaliana*	86	47–61
Endochitinase precursors	168,440	*Zea mays*	28	53–75
Cystein protease inhibitor	288,188	*Vigna unguiculata*	19	51–74
WRKY transcription factor	48,686,707	*Vitis vinifera*	17	34–85
Avr9/Cf-9 rapidly elicited protein	53,857,147	*Nicotiana tabacum*	9	37–75
Wound-induced protein WIN1 precursor	21,617	*Solanum tuberosum*	8	69–70
Hevamine A precursor	3,452,146	*Hevea brasiliensis*	4	55–71
Basic endochitinase precursor	166,665	*Arabidopsis thaliana*	3	51
Wax synthase	5,020,219	*Simmondsia chinensis*	3	50–55

Gene index number and source species for the best hit in the NR GenBank are listed as well as total guarana EST counts (including ESTs grouped in contigs and singletons) and amino acid identity percentages to the GenBank product (the lowest and the highest identities are displayed for each EST group). A minimum *e*-value of 10^{-10} was set as the threshold to identify putative orthologous sequences

The metabolic pathway for carotenoid biosynthesis is also activated in guarana fruits since 17 reads, including those involved in the synthesis of carotenoids, such as lycopene and zeta-carotene (Sun et al. 1996), were identified (Table 3). Class IV small heat shock proteins, which support carotenoid deposition into the cell membrane during the chloroplast (green plastids) to chromoplast (red plastids) transition (Lawrence et al. 1997), were found to represent 14% of the identified chaperones.

Plant resistance ESTs

At least 177 guarana reads showed significant similarity to plant resistance proteins. From these, 86 ESTs for pathogenesis related (PR) proteins which biological functions in defense reactions could not be clearly defined, followed by chitinases that are known to degrade *N*-acetyl-glucose polymers in the chitin of insect exoskeletons and fungal cell walls. Biotic stress control is apparently enhanced by a cysteine-proteinase inhibitor known as cystatin that exhibits anti-fungal activity (Soares-Costa et al. 2002). WRKY factors that induce PR protein transcription (Eulgem et al. 1999) were found. Finally, some ESTs corresponding to a dual-function chitinase/lysozyme hevamine (Potter et al. 1993) were identified.

Final considerations

EST analyses have proven to be a cost-effective method to reveal targets for future works attempting to access the

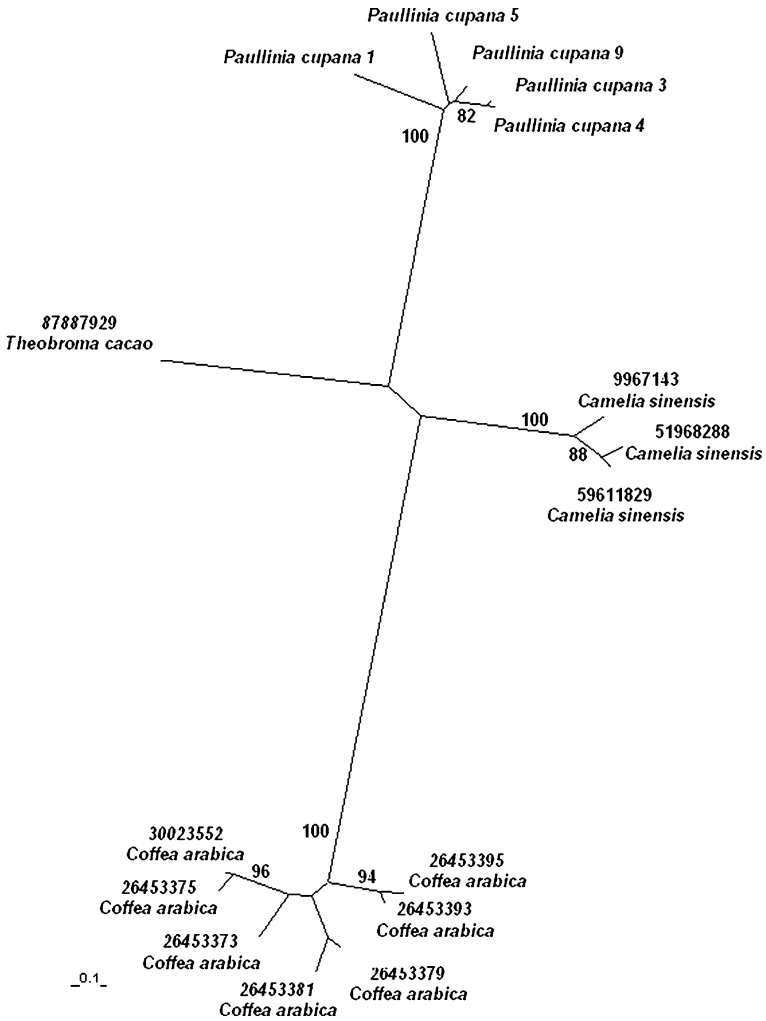

Fig. 2 Phylogenetic relationships among caffeine synthase (CS) sequences annotated from guarana seeded-fruits, *Coffea arabica* (coffee), *Camellia sinensis* (tea) and *Theobroma cacao* (cacao). CS related ESTs from guarana transcriptome were grouped and translated and the predicted polypeptides were aligned to various CSs identified in the GenBank. Maximum likelihood tree was obtained with Phylip (Felsenstein 1998) and relevant neighbor joining bootstraps (1,000 replicates) were shown. Numbers above species names refer to the GI (gene index), except for *Paullinia cupana* that refers to grouped ESTs

complete structure and expression patterns of genes that contribute to the biological properties of guarana extracts (Henman 1982; Espinola et al. 1997).

Additionally, many ESTs are related to valuable agronomical characteristics, such as the PR related transcripts, which can be considered a starting point for a large range of investigations.

A GenBank search using Sapindaceae as the key word resulted in 548 hits for proteins and 2,087 for nucleotide sequences. More than 8,000 transcriptional units (Unique Sequences) were analyzed by the REALGENE consortium, comprising the largest set of sequences for this botanical family.

Acknowledgments This work was supported by grants from the Ministry of Science and Technology and CNPq, process no 48.0120/2002-1. We are thankful to Ana Lúcia Assad, Flávio Neves Bittencourt, and Juçara Parra for the invaluable help and support on this project. We are grateful to Maria do Perpetuo Socorro Lira and Firmino José do Nascimento Filho for access to the Guarana Germoplasm Bank at Embrapa Western Amazon. The authors are thankful for critical and constructive suggestions for the article by Prof. Charles R. Clement and Rosa de Nazaré Silva Clement. We are also grateful for help in the sequencing and/or annotation by Andréa Vernes Fernandes, Jonso V. Souza, Maria Silvanira R. Barbosa, Marta Sofia Peixe Carepo, Rafaela Carvalho de Souza, and Viviane Basso Chiesa.

References

Abourashed EA, Toyang NJ, Choinski J Jr, Khan IA (1999) Two new flavone glycosides from *Paullinia pinnata*. J Nat Prod 62:1179–1181

Akashi K, Nishimura N, Ishida Y, Yokota A (2004) Potent hydroxyl radical-scavenging activity of drough-induced type-2 metallothionein in wild watermelon. Biochem Biophys Res Commun 323:72–78

Altschul SF, Madden TL, Schaffer AA, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402

Antunes E, Gordo WM, de Oliveira JF, Teixeira CE, Hyslop S, De Nucci G (2001) The relaxation of isolated rabbit corpus cavernosum by the herbal medicine Catuama and its constituents. Phytother Res 15:416–421

Berube-Parent S, Pelletier C, Dore J, Tremblay A (2005) Effects of encapsulated green tea and Guarana extracts containing a

mixture of epigallocatechin-3-gallate and caffeine on 24 h energy expenditure and fat oxidation in men. Br J Nutr 94:432–436

Birnboim HC (1983) A rapid alkaline extraction method for the isolation of plasmid DNA. Methods Enzymol 100:243–255

Brígido MM, Walter ME, Oliveira AG, Inoue MK, Anjos DS, Sandes EF, Gondim JJ, Carvalho MJ, Almeida NF Jr, Felipe MS (2005) Bioinformatics of the *Paracoccidioides brasiliensis* EST Project. Genet Mol Res 4:203–215

Bydlowski SP, D'Amico EA, Chamone DA (1991) An aqueous extract of guarana (*Paullinia cupana*) decreases platelet thromboxane synthesis. Braz J Med Biol Res 24:421–424

Campos AR, Barros AI, Santos FA, Rao VS (2003) Guarana (*Paullinia cupana* Mart.) offers protection against gastric lesions induced by ethanol and indomethacin in rats. Phytother Res 17:1199–1202

Davies C, Robinson SP (2000) Differential screening indicates a dramatic change in mRNA profiles during grape berry ripening. Cloning and characterization of cDNAs encoding putative cell wall and stress response proteins. Plant Physiol 122:803–812

Erickson HT, Corrêa MPF, Escobar JR (1984) Guaraná (*Paullinia cupana*) as a commercial crop in Brazilian Amazonia. Econ Bot 38:273–286

Escobar JR, Costa PRC, Corrêa MPF (1985) Variação do teor de cafeína na semente de guaraná em progênies de polinização aberta. Manaus: EMBRAPA-UEPAE de Manaus (Boletim de Pesquisa, 5)

Espinola EB, Dias RF, Mattei R, Carlini EA (1997) Pharmacological activity of Guarana (*Paullinia cupana* Mart.) in laboratory animals. J Ethnopharmacol 55:223–229

Eulgem T, Rushton PJ, Schmelzer E, Hahlbrock K, Somssich IE (1999) Early nuclear events in plant defence signalling: rapid gene activation by WRKY transcription factors. EMBO J 18:4689–4699

Felsenstein J (1998) PHYLIP—phylogeny inference package (Version 3.2). Cladistics 5:164–166

Frei B, Higdon JV (2003) Antioxidant activity of tea polyphenols in vivo: evidence from animal studies. J Nutr 133:3275S–3284S

Freitas DV, Carvalho CR, Nascimento-Filho FJ, Astolfi-Filho S (2007) Karyotype with 210 chromosomes in guarana (*Paullinia cupana* 'Sorbilis'). J Plant Res 120:399–404

Fukumasu H, Avanzo JL, Heidor R, Silva TC, Atroch A, Moreno FS, Dagli ML (2006a) Protective effects of guarana (*Paullinia cupana* Mart. var. sorbilis) against DEN-induced DNA damage on mouse liver. Food Chem Toxicol 44:862–867

Fukumasu H, Silva TC, Avanzo JL, Lima CE, Mackowiak II, Atroch A, Spinosa H, Moreno FS, Dagli ML (2006b) Chemopreventive effects of *Paullinia cupana* Mart var. sorbilis, the guarana, on mouse hepatocarcinogenesis. Cancer Lett 233:158–164

Green P (1996) PHRAP documentation. University of Washington, Seattle. http://bozeman.mbt.washington.edu/phrap.docs/phrap.html

Grimplet J, Romieu C, Audergon J-M, Marty I, Albagnac G, Lambert P, Bouchet J-P, Terrier N (2005) Transcriptomic study of apricot fruit (*Prunus armeniaca*) ripening among 13006 expressed sequence tags. Physiol Plant 125:281–292

Henman AR (1982) Guaraná (*Paullinia cupana* var. sorbilis): ecological and social perspectives on an economic plant of the central Amazon basin. J Ethnopharmacol 6:311–338

Huang X, Madan A (1999) CAP3: a DNA sequence assembly program. Genome Res 9:868–877

IBGE, Instituto Brasileiro de Geografia e Estatística (2006) Sistema IBGE de Recuperação Automática (SIDRA). http://www.ibge.gov.br

Kato M, Kanehara T, Shimizu H, Suzuki T, Gillies FM, Crozier A, Ashihara H (1996) Caffeine biosynthesis in young leaves of *Camellia sinensis*: in vitro studies on *N*-methyltransferase activity involved in the conversion of xanthosine to caffeine. Physiol Plant 98:629–636

Kato M, Mizuno K, Fujimura T, Iwama M, Irie M, Crozier A, Ashihara H (1999) Purification and characterization of caffeine synthase from tea leaves. Plant Physiol 120:579–586

Kato M, Mizuno K, Crozier A, Fujimura T, Ashihara H (2000) Caffeine synthase gene from tea leaves. Nature 406:956–957

Kennedy DO, Haskell CF, Wesnes KA, Scholey AB (2004) Improved cognitive performance in human volunteers following administration of guarana (*Paullinia cupana*) extract: comparison and interaction with Panax ginseng. Pharmacol Biochem Behav 79:401–411

Lawrence SD, Cline K, Moore GA (1997) Chromoplast development in ripening tomato fruit: identification of cDNAs for chromoplast-targeted proteins and characterization of a cDNA encoding a plastid-localized low-molecular-weight heat shock protein. Plant Mol Biol 33:483–492

Mattei R, Dias RF, Espinola EB, Carlini EA, Barros SB (1998) Guarana (*Paullinia cupana*): toxic behavioral effects in laboratory animals and antioxidants activity in vitro. J Ethnopharmacol 60:111–116

Mazzafera P, Wingsle G, Olsson O, Sandberg G (1994) *S*-adenosyl-L-methionine: Theobromine 1-*N*-methyltransferase an enzyme catalyzing the synthesis of caffeine in coffee. Phytochemistry 37:1577–1584

Miura T, Tatara M, Nakamura K, Suzuki I (1998) Effect of guarana on exercise in normal and epinephrine-induced glycogenolytic mice. Biol Pharm Bull 21:646–648

Moyle R, Fairbairn DJ, Ripi J, Crowe M, Botella JR (2005) Developing pineapple fruit has a small transcriptome dominated by metallothionein. J Exp Bot 56:101–112

Murase T, Nagasawa A, Suzuki J, Hase T, Tokimitsu I (2002) Beneficial effects of tea catechins on diet-induced obesity: stimulation of lipid catabolism in the liver. Int J Obes Relat Metab Disord 26:1459–1464

O'Dea JA (2003) Consumption of nutritional supplements among adolescents: usage and perceived benefits. Health Educ Res 18:98–107

Ogawa M, Herai Y, Koizumi N, Kusano T, Sano H (2001) 7-Methylxanthine methyltransferase of coffee plants. Gene isolation and enzymatic properties. J Biol Chem 276:8213–8218

Oliveira CH, Moraes ME, Moraes MO, Bezerra FA, Abib E, De Nucci G (2005) Clinical toxicology study of an herbal medicinal extract of *Paullinia cupana*, Trichilia catigua, Ptychopetalum olacoides and Zingiber officinale (Catuama) in healthy volunteers. Phytother Res 19:54–57

Oraby HF, Ransom CB, Kravchenko AN, Sticklen MB (2005) Barley HVA1 gene confers salt tolerance in R3 transgenic oat. Crop Sci 45:2218–2227

Potter S, Uknes S, Lawton K, Winter AM, Chandler D, DiMaio J, Novitzky R, Ward E, Ryals J (1993) Regulation of a hevein-like gene in *Arabidopsis*. Mol Plant Microbe Interact 6:680–685

Ray S, Phadke S, Patel C, Hackman RM, Stohs S (2005) Short-term and long-term in vivo exposure to an ephedra- and caffeine-containing metabolic nutrition system does not induce cardiotoxicity in B6C3F1 mice. Arch Toxicol 79:330–340

Sanger F, Nickelen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74:5463–5467

Soares-Costa A, Beltramini LM, Thiemann OH, Henrique-Silva F (2002) A sugarcane cystatin: recombinant expression, purification, and antifungal activity. Biochem Biophys Res Commun 296:1194–1199

Sugisawa A, Kimura M, Fenech M, Umegaki K (2004) Anti-genotoxic effects of tea catechins against reactive oxygen species in human lymphoblastoid cells. Mutat Res 559:97–103

Sun Z, Gantt E, Cunningham FX Jr (1996) Cloning and functional analysis of the beta-carotene hydroxylase of *Arabidopsis thaliana*. J Biol Chem 271:24349–24352

Thompson J, Higgins D, Gibson T (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680

Tobi SE, Gilbert M, Paul N, McMillan TJ (2002) The green tea polyphenol, epigallocatechin-3-gallate, protects against the oxidative cellular and genotoxic damage of UVA radiation. Int J Cancer 102:439–444

Uefuji H, Ogita S, Yamaguchi Y, Koizumi N, Sano H (2003) Molecular cloning and functional characterization of three distinct *N*-methyltransferases involved in the caffeine biosynthetic pathway in coffee plants. Plant Physiol 132:372–380

Weckerle CS, Stutz MA, Baumann TW (2003) Purine alkaloids in *Paullinia*. Phytochemistry 64:735–742

Yoneyama N, Morimoto H, Ye CX, Ashihara H, Mizuno K, Kato M (2006) Substrate specificity of *N*-methyltransferase involved in purine alkaloids synthesis is dependent upon one amino acid residue of the enzyme. Mol Genet Genomics 275:125–135