



**UNIVERSIDADE FEDERAL DO AMAZONAS**  
**PROGRAMA MULTI-INSTITUCIONAL**  
**DE PÓS-GRADUAÇÃO EM BIOTECNOLOGIA**

**CARACTERIZAÇÃO GENÔMICA E MOLECULAR DO GUARANAZEIRO**  
**(*Paullinia cupana* var. *sorbilis*)**

**DANIVAL VIEIRA DE FREITAS**

**MANAUS**  
**AMAZONAS – BRASIL**  
**2009**



**DANIVAL VIEIRA DE FREITAS**

**CARACTERIZAÇÃO GENÔMICA E MOLECULAR DO GUARANAZEIRO**

**(*Paullinia cupana* var. *sorbilis*)**

**Orientador: Dr. Spartaco Astolfi Filho**

**Co-orientador: Dr. Sérgio Ricardo Nozawa**

Tese apresentada a Universidade Federal do Amazonas, como parte dos requisitos do Programa Multi-Institucional de Pós-Graduação em Biotecnologia, para obtenção do título de *Doctor Scientiae*.

**MANAUS  
AMAZONAS – BRASIL  
2009**

## Ficha catalográfica

F866c Freitas, Danival Vieira de  
Caracterização genômica e molecular do guaranazeiro  
(*Paullinia cupana* var. *sorbilis*) / Danival Vieira de Freitas  
– Manaus, UFAM, 2009.  
70f.: il., color

Tese (Doutorado em Biotecnologia) – Universidade  
Federal do Amazonas.  
Orientador: Astolfi-Filho, Sapartaco  
Co-orientador: Nozawa, Sérgio Ricardo  
Área de concentração: Agroflorestal

1. Citogenética. 2. Ploidia. 3. Citometria de Fluxo. 4.  
Genoma funcional. 5. Cafeína. 6. Expressão Gênica. 7.  
qRT PCR. Título.

CDU 582.772.4 (043.5)

**Contatos e informações:** danivalfreitas@hotmail.com



***Dedico a Deus,***

***aos meus Pais Antônio José de Freitas e Delfina Borges Vieira de Freitas***

***e meus irmãos Ana Paula, Jáder e Adelson,***

***e a Daniela meu grande Amor...***

***... eles sabem porque.***

## AGRADECIMENTOS

A Deus por tudo.

À minha família pelo amor e apoio incondicionais.

A Universidade Federal do Amazonas em especial ao Centro de Apoio Multidisciplinar (CAM), e ao Centro Universitário Nilton Lins pela oportunidade e disposição de sua infraestrutura para realização do doutorado.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), à Fundação de Amparo à Pesquisa do Estado do Amazonas (FAPEAM), a Superintendência da Zona Franca de Manaus (SUFRAMA) pelos recursos concedidos para realização desta pesquisa.

Ao Professor Spartaco, pela orientação, pela amizade, pelas oportunidades, pelos incentivos, e pelo convívio durante esses anos.

Ao Professores/Pesquisadores Charles Clement, Paula Ângelo, José Odair, Jorge Porto, Eliana Feldberg, Luis Antônio S. Contim e Edmar pelos conhecimentos transmitidos e pelos aconselhamentos.

À Universidade Federal de Viçosa (UFV), nas pessoas do Professor Carlos Carvalho do Departamento de Biologia Geral e a EMBRAPA Amazônia Ocidental, nas pessoas do Firminio, Paula Angelo, Nelcimar e André, pela parceria na condução de diversas atividades de pesquisas.

Ao Pesquisador Sérgio Ricardo Nozawa, pelas sugestões e pelo aconselhamento.

Aos membros da banca examinadora pelas sugestões.

Aos colegas do Programa de Pós Graduação em Biotecnologia e do Centro Universitário Nilton Lins pela convivência e amizade.

Aos funcionários da Coordenação do Programa de Pós-graduação em Biotecnologia, que, direta ou indiretamente, contribuíram para a realização deste trabalho.

A todos os amigos e familiares não citados pelo nome, mas que de alguma forma contribuíram para a realização deste trabalho.

São os meus sinceros agradecimentos!

## Caracterização genética e molecular do guaranazeiro (*Paullinia cupana* “*sorbilis*”)

Autor: Danival Vieira de Freitas

Orientador: Dr. Spartaco Astolfi Filho

Co-orientador: Sérgio Ricardo Nozawa

### RESUMO

A espécie *Paullinia cupana* originária da Amazônia, é dividida em duas variedades, *P. cupana* var. *cupana* (Venezuela e Colômbia) e *P. cupana* var. *sorbilis* (Brasil). Os poucos estudos com o guaranazeiro concentraram, principalmente, nos seus aspectos botânicos, no seu melhoramento e em testes experimentais sobre o efeito de seu extrato para fins terapêuticos. O objetivo principal do presente trabalho foi ampliar o conhecimento sobre as informações básicas do ponto de vista genético e molecular do guaranazeiro. Deste modo foi proposto: (1) determinar o cariótipo e o tamanho do genoma da planta (*P. cupana* var. *sorbilis*); (2) caracterizar e validar seqüências expressas em seus frutos, e em especial as das enzimas relacionadas com a síntese de cafeína (N-metiltransferases); (3) determinar o seu número de cópias; (4) determinar o padrão de expressão em diferentes órgãos/tecidos. A var. *sorbilis* apresentou elevado número de cromossomos ( $2n = 210$ ) e conteúdo de DNA ( $2C = 22,5$  pg DNA/células). A análise do cariograma mostrou que os cromossomos estão agrupados em dois conjuntos, um tetraplóide e outro hexaplóide. O seqüenciamento de transcritos gerou um banco com 15.387 ESTs, distribuídas em 2.628 contigs e 5.969 singletons. Entre estas estão as enzimas que participam das vias de síntese da cafeína sintase, sendo o melhor *hit* com as seqüências de *N-metiltransferases*, aqui denominadas *GNMT*. As análises indicam

que a sequência *GNMT* é homóloga as *N-metiltransferases* da via biossintética da cafeína de outras espécies tais como café (*Coffea arabica* and *C. canephora*), cacau (*Theobroma cacao*), e chá verde (*Camellia sinensis*). O número de cópias gênicas da N-metiltransferase da *GNMT* presentes no genoma do guaranazeiro foi cerca de três vezes mais cópias que os espécimes selvagens (*guaranarana*, *Paullinia* sp). O fruto do guaranazeiro com 70 dias de desenvolvimento tem o maior nível de transcritos da *GNMT*, sendo que o número de transcritos aumentou gradualmente com o processo de maturação do fruto do guaranazeiro. Os resultados aqui discutidos são especialmente interessantes e podem colaborar com a experimentação científica que estão, aos poucos, comprovando as propriedades medicinais atribuídas ao guaraná e seu extrato obtido do pó de suas sementes.

## Genetic and molecular characterization of guaranazeiro (*Paullinia cupana* “*sorbilis*”)

Author: Danival Vieira de Freitas

Advisores: Dr. Spartaco Astolfi Filho & Sérgio Ricardo Nozawa

### ABSTRACT

*Paullinia cupana* is originary from Amazon and is divided in two different varieties, *P. cupana* var. *cupana* (Venezuela and Colombia) and *P. cupana* var. *sorbilis* (Brazil), this last one is the true guarana plant. The few studies on this plant are concentrated, mainly, in its bothanical aspects and effects of its seed extract over some medical experimentation. The main objective of this work was to amplify knowledgment over those informations about the guarana plant, moreover by the means of genetic molecular. Aiming the following propositions: (1) to determine the guarana plant (*P. cupana* var *sorbilis*) caryotype and its genome size; (2) to characterize and validate sequences expressing in the guaraná berries, gathering especial information on those related with caffeine synthesis (N-methyl-transferases); (3) to determine its (N-methyl-transferases) copy number in the plant genome; (4) to determine its (N-methyl-transferases) expression patterns in different organs/tissues. The results showed an elevated number of chromosomes (210) and DNA content (2C = 22.5 pg DNA/cell). The caryogram analysis presented two sets of chromosomes, one tetraploid and other hexaploid. The transcripts sequencing yield a 15,387 ESTs databank, grouped in 2,628 contigs and 5,969 singletons. Among these are enzymes genes involved in caffeine synthesis, the so called caffeine sintases. The analysis of these guarana sequences (denominated *GNMT*) shows it is homologous to N-methyltransferases of other species such as coffee (*Cofea arábica* and *C. canephora*), cocoa (*Theobroma*

*cacao*), and green tea (*Camellia sinensis*). Regarding to the number of genome copies, the guarana plant presented three times more than its native wild plants (the guaranaranas, *Paullinia* sp) correlates. Fruits posses the highest level of *GNMT* transcript expression when it reach 70 days of development age, increasing the rate as berries ripening process occurs. These results are especially interesting and can collaborate with the experimentations in progress course which are proving the healing properties attributed to this plant and its seed powder extract.

## SUMÁRIO

<b>1. INTRODUÇÃO GERAL.....</b>	<b>- 12 -</b>
<b>2. OBJETIVOS.....</b>	<b>- 15 -</b>
2.1. GERAL.....	- 15 -
2.2. ESPECÍFICOS .....	- 15 -
<b>3. CAPÍTULO I – OS SEGREDOS MOLECULARES DE UMA PLANTA LENDÁRIA ....</b>	<b>- 16 -</b>
3.1. SINOPSE .....	- 16 -
<b>4. CAPÍTULO II – CARIÓTIPO COM 210 CROMOSSOMOS NO GUARANAZEIRO (PAULLINIA CUPANA VAR. SORBILIS).....</b>	<b>- 22 -</b>
4.1. SINOPSE .....	- 22 -
<b>5. CAPÍTULO III – TRANSCRIPTOMA DO FRUTO DO GUARANAZEIRO (PAULLINIA CUPANA VAR. SORBILIS), UM ESTIMULANTE ORIGINÁRIO DA FLORESTA AMAZÔNICA CONSUMIDO TRADICIONALMENTE.....</b>	<b>- 27 -</b>
5.1. SINOPSE .....	- 29 -
<b>6. CAPÍTULO IV – N-METILTRANSFERASE DO GUARANAZEIRO (PAULLINIA CUPANA VAR. SORBILIS) ENVOLVIDA NA BIOSÍNTESE DA CAFEÍNA.....</b>	<b>- 38 -</b>
6.1. SINOPSE .....	- 38 -
6.2. INTRODUÇÃO.....	- 39 -
6.3. MATERIAL E MÉTODOS.....	- 43 -
6.3.1. <i>Material vegetal e isolamento de ácidos nucleicos.....</i>	<i>- 43 -</i>
6.3.2. CARACTERIZAÇÃO IN SILICO DE UMA N-METILTRANSFERASE DO GUARANAZEIRO .....	- 44 -
6.3.3. <i>Análise do número de cópias e transcritos do gene GNMT por PCR em Tempo Real usando o sistema SYBR Green.....</i>	<i>- 45 -</i>
6.4. RESULTADOS .....	- 48 -
6.4.1. <i>Caracterização in silico das N-metiltransferases.....</i>	<i>- 48 -</i>
6.4.2. <i>Estimativa do número de cópias do gene GNMT .....</i>	<i>- 51 -</i>
6.4.3. <i>Quantificação do número de transcritos da GNMT.....</i>	<i>- 53 -</i>
<b>7. REFERÊNCIAS.....</b>	<b>- 59 -</b>
<b>8. CONSIDERAÇÕES FINAIS E PESPECTIVAS .....</b>	<b>- 68 -</b>

## 1. INTRODUÇÃO GERAL

A espécie *Paullinia cupana* originária da Amazônia é dividida em duas variedades, *P. cupana* var. *cupana* (Venezuela e Colômbia) e *P. cupana* var. *sorbilis* (Brasil) (Lleras, 1992). As duas variedades foram cultivadas pelos indígenas, sendo a var. *sorbilis* cultivada no Baixo Rio Amazonas pelas tribos Maués e Andirás, e a var. *typica* pelos Barés no Alto Rio Negro (Lleras, 1984). As duas variedades diferenciam-se principalmente no tamanho, forma e cor do fruto. No alto Rio Negro, a variedade *cupana* foi incorporada aos hábitos alimentares dos colonizadores e viajantes da região, não sendo difundida fora de seu habitat natural. No baixo Rio Amazonas, a variedade *sorbilis* teve ampla aceitação e disseminação de suas sementes, desde a chegada dos primeiros colonizadores, tornando-se uma das espécies nativas com maior importância socioeconômica para a região (Lleras, 1992).

*P. cupana* var. *sorbilis* (guaranazeiro) é uma planta perene arbustiva, trepadeira e lenhosa que pode atingir 10 metros de altura, embora quando cultivada em áreas abertas, cresce no máximo até três metros (Erickson et al., 1984). O guaranazeiro é uma espécie com crescente demanda devido às múltiplas e comprovadas propriedades medicinais do extrato do seu fruto (Henman, 1982). O extrato pode ser utilizado para diminuição da massa corpórea (Boozer et al., 2001; Andersen e Fogh, 2001), como estimulante do sistema nervoso central (Espinola et al., 1997), para melhora do desempenho cognitivo (Espinola et al., 1997; Kennedy et al., 2004), como afrodisíaco (Antunes et al., 2001), como bactericida e antioxidante (Mattei et al., 1998; Basile et al., 2005). Além disso, o extrato pode também inibir a hepatocarcinogênese (Fukumasu et al., 2006b), diminuir a agregação de plaquetas

(Bydlowski et al., 1991) e, proteger contra lesões gástricas induzidas por etanol (Campos et al., 2003).

Os poucos estudos celulares e moleculares realizados com o guaranazeiro concentraram, principalmente, na análise da diversidade genética através da avaliação morfométrica e molecular. Os resultados reportam reduzida variabilidade molecular (Sousa, 2003) e alta diversidade morfométrica (Corrêa, 1989). Estudos citogenéticos de *P. cupana* var *sorbilis* reporta que a espécie apresenta 210 cromossomos (Freitas et al., 2007). A poliploidia pode conduzir às diferenças imediatas no tamanho das células, no tamanho dos estômatos, no tamanho da planta. A aparência imediata destas novas características é indicativa de mudanças na expressão dos genes, embora as causas desta nova variação não sejam bem compreendidas (Osborn et al., 2003).

O estabelecimento de um banco de dados contendo informações de estudos de transcriptomas tem proporcionado o desenvolvimento de métodos de análises globais para o estudo de mudanças na expressão gênica, associadas com importantes processos celulares nos organismos (Fukuda et al., 2008; Cánovas et al. 2004). Devido a grande importância socioeconômica e a carência de estudos genéticos e moleculares, *P. cupana* tem atraído considerável interesse científico, por exemplo, pela “Rede da Amazônia legal de Pesquisas Genômicas – REALGENE” formado por um Consórcio de Instituições da Amazônia, apoiado pelo CNPq/MCT, que vem estudando o genoma funcional (transcriptoma do fruto) do guaranazeiro (Ângelo et al., 2008).

Muitos eventos relacionados aos fatores de crescimento e desenvolvimento de plantas ou até mesmo em resposta a vários estímulos são resultantes da alteração na expressão gênica. A análise da função do gene necessita,

conseqüentemente, da análise temporal e espacial dos padrões de expressão gênica. A informação mais conclusiva sobre mudanças nos níveis da expressão gênica pode ser obtida através da análise de mudanças qualitativas e quantitativas dos RNAs mensageiros, proteínas e metabolitos. A determinação qualitativa e quantitativa dos níveis de transcritos de células vegetais permite que genes, diferencialmente expressos, possam ser identificados e, conseqüentemente, sua função metabólica possa ser investigada. Portanto, o conhecimento da localização celular dos transcritos é essencial, a fim de avaliar a função do gene em um contexto integrado (Pesquet *et al.*, 2004).

Durante muitos anos, a vasta maioria dos estudos de expressão gênica foi realizada utilizando análises com *gel-blots* não quantitativas ou semi-quantitativas (Bustin, 2002). A maximização da sensibilidade dos métodos de quantificação tem levado ao desenvolvimento de técnicas cada vez mais avançadas, dentre elas o PCR Quantitativo em Tempo Real (qRT-PCR), porém apenas nos últimos anos esta técnica tem sido utilizado como ferramenta nos estudos de expressão gênica e quantificação de sequências específicas com vegetais (Gachon, 2004). Comparado aos outros métodos de PCR, o qRT-PCR é mais específico, sensível e reprodutível (Freeman *et al.*, 1999; Raeymaekers, 2000).

O objetivo do presente trabalho foi analisar o cariótipo e o conteúdo de DNA do guaranazeiro (*Paullinia cupana* var. *sorbilis*), identificar os genes expressos no seu fruto e, estimar o número de cópias da N-metiltransferase no genoma, principal enzima na síntese de cafeína, do guaranazeiro (*GNMT*) e o nível da sua expressão em diferentes tecidos.

## 2. OBJETIVOS

### 2.1. Geral

- ✓ Analisar o cariótipo e o conteúdo de DNA do guaranazeiro (*Paullinia cupana* var. *sorbilis*), identificar os genes expressos no seu fruto, além de estimar o número de cópias de uma N-metiltransferase no genoma do guaranazeiro e o nível de expressão gênica em diferentes tecidos.

### 2.2. Específicos

1. Desenvolver um protocolo para isolamento de núcleos do guaranazeiro (*Paullinia cupana* var *sorbilis*);
2. Estimar o conteúdo de DNA de *Paullinia cupana* var *sorbilis*;
3. Determinar o cariótipo de *Paullinia cupana* var *sorbilis* e apresentar um cariograma com alta resolução;
4. Seqüenciar, validar, ordenar e anotar as seqüências expressas (EST's) no fruto de *Paullinia cupana* var *sorbilis*;
5. Identificar e caracterizar as EST's com homologia para *N-metiltransferase* envolvidas na biossíntese de cafeína do guaranazeiro (*GNMT*);
6. Estimar a variação do número de cópias da *GNMT* no genoma do guaranazeiro (*Paullinia cupana* var. *sorbilis*) e de espécimes selvagens (guaranaranas - *Paullinia* sp);
7. Quantificar os níveis de transcritos da *GNMT* expressos em diferentes órgãos/tecidos do guaranazeiro (*Paullinia cupana* var. *sorbilis*) e de espécimes selvagens (guaranaranas - *Paullinia* sp);

### **3. Capítulo I – Os segredos moleculares de uma planta lendária**

**Artigo de apresentação do guaranazeiro publicado na *Ciência Hoje* (2009) 43: 31-35.**

*Indramara Lobo de Araújo (Programa de Pós-graduação em Desenvolvimento Regional), Danival Vieira de Freitas (Laboratório de Expressão Gênica, Centro Universitário Nilton Lins e Programa - Multi-institucional de Pós-graduação em Biotecnologia, Universidade Federal do Amazonas), Carlos Gustavo Nunes-Silva (Laboratório de Expressão Gênica, Centro Universitário Nilton Lins), Spartaco Astolfi-Filho (Laboratório de Tecnologia de DNA, Centro de Apoio Multidisciplinar - Universidade Federal do Amazonas)*

#### **3.1. Sinopse**

Quando os primeiros europeus chegaram ao Brasil, os índios já consumiam o guaraná como alimento e em seus rituais religiosos. O guaraná, natural da Amazônia, era utilizado pelos povos indígenas como estimulante e revigorante. Hoje em dia o uso da semente do Guaraná se difundiu como fitoterápico rico em cafeína e estimulante do sistema nervoso central. Além da cafeína, a semente do Guaraná contém amido, proteínas, óleo fixo e matérias aromáticas, resinosas e pépticas. O Guaraná também é usado como tônico geral e no combate ao estresse. As propriedades terapêuticas e estimulantes do guaraná descrito na farmacopéia baseada no conhecimento indígena tradicional estão sendo comprovadas cientificamente por pesquisadores atraídos pela grande importância socioeconômica e medicinal dessa planta.

# Os segredos moleculares de uma planta lendária

**O guaraná (*Paullinia cupana*) é uma planta originária** da bacia amazônica.

Seu nome popular é uma derivação da palavra indígena *wara'ná*, que significa 'árvore que sobe apoiada em outra', já que, na floresta, o guaraná desenvolve-se como uma trepadeira. A espécie tem duas variedades: *P. cupana* variedade *typica*, que ocorre na Venezuela, na Colômbia e no Brasil (na região do alto rio Negro), e *P. cupana* variedade *sorbilis*, conhecida como o guaraná verdadeiro e encontrada na área central da Amazônia brasileira, principalmente na região de Maués, município no leste do Amazonas. As diferenças principais entre as duas estão no tamanho, na forma e na cor do fruto. Ambas foram domesticadas por grupos indígenas: a *sorbilis* pelos Maués e Andirá, no baixo rio Amazonas, e a *typica* pelos Barés, no alto rio Negro. ▶

BIOTECNOLOGIA

Em seu estado natural, nas matas, o guaranazeiro cresce como uma trepadeira, apoiando-se nas árvores para atingir a parte superior da cobertura florestal, a cerca de 30-40 m do solo. Quando cultivado em campo aberto, porém, é arbustivo, alcançando no máximo 3 m de altura. A planta tem folhas grandes, de um verde escuro acentuado, e frutos arredondados, que surgem em cachos vermelhos ou amarelos, parecidos com as ramas de café maduro. A semente única, preta e brilhante, é o produto utilizável da planta. Quando o fruto amadurece, a casca rompe-se espontaneamente e a semente, envolta em uma polpa branca (o arilo), fica exposta, dando aos frutos a aparência de um olho. Os frutos, depois de colhidos, são amontoados ou acondicionados em sacos por três dias, para que fermentem, o que facilita a retirada da casca. Após o despolpamento, as sementes são torradas e moídas.

O Brasil é o único produtor comercial de guaraná do mundo, embora existam pequenas áreas plantadas, para subsistência, na Venezuela e no Peru. A produção brasileira de amêndoas (sementes) é atualmente de cerca de 3 mil toneladas anuais, e há possibilidade real de expansão, o que repre-

sentaria um importante reforço para a economia regional. O guaraná é exportado em forma de xarope, pó, refrigerantes e outros subprodutos, e entre os países compradores estão Japão, Alemanha, Itália, Inglaterra, Estados Unidos e Espanha. A perspectiva de aumento nas exportações é otimista, porque é um produto natural e exótico, apelos fortes no mercado atual, somando-se a isso a comprovação científica de suas propriedades terapêuticas e a determinação de empresas brasileiras de difundir o consumo do refrigerante sabor guaraná em outros países.

### **A lenda do guaraná**

Conta a tradição oral dos índios da etnia Saterê-Mawê que um casal de índios maués vivia há muitos anos sem ter filhos, mas desejava muito uma criança. Um dia, o casal pediu que Tupã, o 'rei' dos deuses, lhe desse um filho, para completar sua felicidade. Tupã, sabendo que o casal era cheio de bondade, atendeu o desejo, e nasceu um lindo menino. O tempo passou e o menino cresceu bonito e generoso.

Plantação de guaranazeiros no Amazonas



No entanto, Jurupari, o deus da escuridão, sentia uma extrema inveja do menino e da paz e felicidade que ele transmitia. Um dia, o menino foi coletar frutos na floresta e Jurupari aproveitou a ocasião para lançar sua vingança. Transformou-se em uma serpente venenosa e mordeu o menino, matando-o instantaneamente. A triste notícia espalhou-se rapidamente. Nesse momento, trovões ecoaram e fortes relâmpagos caíram sobre a aldeia. A mãe, que chorava em desespero, entendeu que os trovões eram uma mensagem de Tupã, dizendo que ela deveria plantar os olhos da criança, dos quais uma nova planta cresceria dando saborosos frutos. Os índios obedeceram aos pedidos da mãe e plantaram os olhos do menino. Nesse lugar cresceu o guaraná, cujas sementes são negras, cercadas pelo arilo branco, lembrando olhos humanos.

### Uma planta 'milagrosa'

O primeiro relato escrito sobre essa planta e sua importância é do jesuíta João Felipe Bettendorf (1625-1698) por volta de 1669. Bettendorf chegou ao Pará em 1661, enviado pelo também jesuíta Antonio Vieira (1608-1697) para fundar uma missão na tribo dos tapajós, na confluência dos rios Tapajós e Amazonas. Essa missão deu origem à cidade de Santarém (PA). Ele dizia que o povo Andirá fazia uso da planta, tida como 'milagrosa', secando e pisando o fruto, e fazendo com ele bolas, pelas quais tinham o mesmo apreço que os brancos por ouro. Com o auxílio de uma pedra, eles desfaziam a bolinha (amêndoa seca e pisada) em uma cuia com água e ingeriam o líquido. Essa bebida, nas palavras do jesuíta, "dá tanta força que, indo à caça, um dia até outro, não sentem fome, além do que tiram febres, câibras e dores de cabeça".

Por volta de 1810, dois naturalistas europeus, o alemão Friedrich Alexander von Humboldt (1769-1859) e o francês Aimé Jacques Bonpland (1773-1858), foram os pioneiros a coletar o guaraná para estudos, ao passarem pela Venezuela. A descrição e a classificação científicas da espécie, porém, só foram feitas em 1823 pelo botânico alemão Karl S. Kunth (1788-1855), com base em plantas coletadas por Humboldt e Bonpland. Kunth incluiu o guaraná na família Sapindaceae e deu a ele o nome científico de *Paullinia cupana*. Em 1818, em suas andanças pelo baixo Amazonas, na região de Maués, outro alemão, o botânico Carl Friedrich von Martius (1794-1868), coletou novamente o guaraná e o batizou, em livro também de 1823, como *Paullinia sorbilis*, mas depois reconheceu como pioneiro o nome dado por Kunth.

### GENÉTICA VEGETAL



Sementes  
(amêndoas)  
de guaraná

Martius isolou das sementes uma substância cristalina, amarga e branca, com propriedade fisiológica marcante. Essa substância foi denominada 'guaranina' e, mais tarde, renomeada como 'cafeína'. O guaraná, porém, contém grande variedade de substâncias. O Instituto de Tecnologia de Alimentos, do governo do estado de São Paulo, analisou a semente (amêndoa) do guaraná e determinou sua composição química (ver 'Alto teor de cafeína').

### A ciência confirma a tradição

O conhecimento acumulado pelos índios a respeito do guaraná foi transmitido oralmente aos colonizadores e seus descendentes (comunidades ribeirinhas) e posteriormente relatados em obras etnobotânicas, como, por exemplo, o *Dicionário das plantas úteis do Brasil*, lançado em 1926 pelo botânico português Manuel Pio Correia (1844-1934). Essas informações serviram de base para a realização de inúmeros estudos científicos sobre o guaraná, que comprovaram os conhecimentos tradicionais dos povos indígenas da Amazônia sobre as propriedades de *Paullinia cupana* variedade *sorbilis*.

A suposição de que o sucesso na descoberta de novos fármacos seria maior a partir de pesquisas etnodirigidas, ou seja, que levassem em conta os saberes e práticas tradicionais, levou a um aumento significativo nos estudos sobre o guaraná. Já foram comprovadas cientificamente propriedades antioxidantes (redução de radicais tóxicos de oxigênio), antimutagênicas (inibição de mutações) e antineoplásicas (inibição do surgimento de tumores malignos). Estudos confirmaram ainda que o guaraná melhora a atividade motora, retardando a fadiga e, em consequência, aprimorando o desempenho físico, sobretudo em atividades de longa duração; ajuda a melhorar a performance cognitiva e a reduzir a perda de memória em portadores do mal de ▶

## ALTO TEOR DE CAFEÍNA

A análise de uma amostra de amêndoa de guaraná (matéria seca) feita pelo Instituto de Tecnologia de Alimentos, de Campinas (SP), mostrou a seguinte composição química: amido (40%), proteína bruta (12,4%), tanino (10,8%), umidade (6,82%), açúcares totais (6,1%), açúcares redutores (4,4%), cafeína (2,97%), fibras (2,87%), matéria graxa (2,48%) e cinzas (1,55%), além das vitaminas A (375 UI por 100 g) e B1 (ou tiamina, 0,57 mg/100 g), e dos elementos potássio (609 mg/100 g), fósforo (144 mg/100 g), cálcio (92,3 mg/100 g) e ferro (6,2 mg/100 g). Nas sementes, a cafeína pode variar de 2,5% a 5% da matéria seca. O extrato do guaraná contém, além da cafeína, a teobromina e a teofilina, alcalóides com moléculas similares e também com efeitos estimulantes (as sementes de cacau têm maior teor de teobromina e o chá verde e o chá preto têm maior teor de teofilina). O guaraná contém ainda polifenóis (substâncias com ação antioxidante, antiviral e bactericida), além de outros tipos de compostos químicos de interesse farmacológico. A cafeína é encontrada em muitas plantas, várias delas consumidas pelos humanos, como café, erva-mate, chá, coca e cacau, mas o guaraná é a que apresenta a maior concentração desse composto – praticamente 3,5 vezes maior do que nos grãos de café da espécie arábica.

Alzheimer (propriedades neuroprotetoras); e reduz em até 31% a agregação de plaquetas que causam trombose, o que ajuda a proteger os vasos sanguíneos da aterosclerose. A cafeína presente no fruto neutraliza o efeito da adenosina no metabolismo das células de gordura, inibindo a geração de mais gordura (o que a torna muito útil em produtos, por exemplo, contra a celulite). Já o extrato é capaz de reduzir o apetite, contribuindo para a perda de peso. A maioria desses efeitos era descrita, de forma indireta, pelo conhecimento indígena.

As pesquisas vêm sendo realizadas, no Brasil e no exterior (Estados Unidos, Reino Unido, Itália e outros locais), nos últimos 20 anos. Em 1988 e 1991, por exemplo, o brasileiro Sérgio P. Bydlowski e equipe revelaram que o extrato de guaraná previne a aterosclerose e tem ação anticoagulante em animais de laboratório. Outros estudos utilizaram voluntários humanos, como os que confirmaram os efeitos contra a fadiga (em 1995, pelo norte-americano Lawrence Spriet e equipe), contra a perda de memória (em 2003, pela brasileira Maria Teresa S. Trevisan e equipe, e em 2004, pelo inglês David O. Kennedy e equipe) e de inibição do apetite (em 2005, pelos norte-americanos Aleda M. Hess e Donald L. Sullivan). Culturas de células também foram empregadas em outras investigações sobre o guaraná. Mais recentemente, em 2006, o brasileiro Heidge Fukumasu e equipe descobriram, em animais de laboratório, que o guaraná previne o câncer de fígado – propriedade não prevista pelo conhecimento tradicional.

RODOGRUPPO KEYSSTONE

Produção de guaraná no Brasil e nos três principais estados produtores, de 1999 a 2007, segundo dados do IBGE e do Ministério da Agricultura

## Os segredos genéticos

Os maiores desafios para o aumento na produção são as doenças que atingem o guaranazeiro e que, de tempos em tempos, assolam a região de cultivo no Amazonas. Essas doenças, que constituem um grande gargalo para a expansão da cultura na região, atraíram a atenção de centros de pesquisa, principalmente a Embrapa Amazônia Ocidental, vinculada à Empresa Brasileira de Pesquisa Agropecuária (Embrapa). A Embrapa Amazônia Ocidental coordena o programa de pesquisa dedicado ao guaraná, estruturado em duas linhas básicas (recursos e melhoramento genético e manejo cultural) e desenvolvido na sede da instituição, no Campo Experimental de Maués (AM) e em outras estações experimentais. Esses projetos, incluídos no Programa Nacional de Pesquisa, visam desenvolver sistemas sustentáveis de produção de matérias-primas, originárias de culturas regionais não irrigadas. A ênfase é dada na criação de variedades com características genéticas que permitam melhor adaptação aos ecossistemas amazônicos, visando reduzir custos de produção e controlar doenças e pragas.

A grande importância socioeconômica do guaranazeiro na região atraiu também o interesse científico da Rede da Amazônia Legal de Pesquisas Genômicas (Realgene), formada por um consórcio de instituições de pesquisa e apoiada pelo Ministério da Ciência e da Tecnologia e pela Fundação de Amparo a Pesquisa do Estado do Amazonas. Recentemente, a Realgene finalizou os estudos sobre o tamanho do genoma (quantidade de DNA) de *P. cupana sorbilis*, o número de cromossomos (os 'pacotes' em que o genoma se divide) e seu genoma funcional (o conjunto de genes expressos no fruto, ou seja, que contém as informações necessárias para a produção de proteínas e outras moléculas – também chamado de transcriptoma).

O objetivo principal da Realgene, com o sequenciamento do genoma funcional do guaraná, foi o de ampliar o conhecimento básico sobre essa planta, do ponto de vista genético e molecular, incluindo a descoberta de genes responsáveis por processos metabólicos importantes. O trabalho, do qual participaram três dos autores deste artigo, foi publicado na revista científica *Plant Cell Reports*, e identifica os genes expressos em três diferentes fases de maturação do fruto. Esse estudo possibilitou a descoberta da expressão de genes importantes, como os de resistência a insetos e ao estresse hídrico, e os



que contêm informações para a produção de enzimas-chave de rotas metabólicas relevantes, como as de síntese da cafeína, do betacaroteno (composto precursor da vitamina A) e da chalcona (substância, segundo alguns autores, com efeito cicatrizante e anti-inflamatório), entre outras.

O conhecimento da estrutura de genomas é cada vez mais importante para a compreensão da evolução dos genes e para a sua modificação, visando obter plantas mais resistentes e/ou produtivas. O estudo sobre a estrutura do genoma do guaraná revelou um conjunto de cromossomos (cariótipo) com 210 unidades e cerca de 23 milhões de pares de bases – quatro vezes o tamanho do genoma humano. Esse grande número de cromossomos do guaraná, no entanto, contrasta com o cariótipo das outras plantas do gênero *Paullinia*, pois todas as sete espécies já estudadas têm apenas 24 cromossomos. O guaraná, portanto, apresenta poliploidia (multiplicação do número de cromossomos), fenômeno que talvez seja resultado do longo processo de domesticação dessa planta pelos nativos.

Os antigos conhecimentos a respeito do guaraná, muitos deles comprovados em estudos científicos, somados às novas informações sobre seu genoma, podem direcionar a pesquisa no sentido de descobrir os princípios ativos de *Paullinia cupana* e possibilitar a descoberta ou o desenvolvimento (a partir dos compostos presentes na planta) de novos fitofármacos. Também podem facilitar o melhoramento genético da espécie, tanto para aumentar sua produtividade quanto para reforçar sua resistência a doenças, pragas e estresses não biológicos (períodos de falta de água, por exemplo). A descoberta de novas propriedades e a comprovação das propriedades descritas na farmacopéia popular nos fazem entender por que os indígenas consideram a planta 'mágica', a ponto de utilizá-la em alguns de seus principais rituais e de, ao longo do tempo, terem promovido a domesticação e o melhoramento genético da espécie. ■

### Sugestões para leitura

- ÂNGELO, P. C. S. e outros. 'Guaraná (Paullinia cupana var. sorbilis), an anciently consumed stimulant from de Amazon rain forest: the seeded fruit transcriptome', in *Plant Cell Reports*, v. 27(1), p. 117, 2008.
- TAVARES, A. M. e outros. *Cultura do guaranazeiro no Amazonas*. Manaus, Embrapa Amazônia Ocidental, 2005.
- WATZEL, J. 'O guaraná, seu valor industrial e medicinal', in *Boletim do Ministério da Agricultura*, v. 26 (4/6), p. 25, 1937.

## **4. Capítulo II – Cariótipo com 210 cromossomos no guaranazeiro** **(*Paullinia cupana* var. *sorbilis*)**

**Artigo publicado na *Journal Plant Research* (2007) 120:399-404.**

### **4.1. Sinopse**

Este trabalho reporta o kariograma e a análise do tamanho do genoma de *Paullinia cupana* var. *sorbilis*, fornecendo a caracterização morfológica dos cromossomos e o conteúdo de DNA. A quantificação do tamanho e estrutura do genoma são cruciais para a compreensão global de um organismo e, diante dos avanços biotecnológicos, estes estudos têm sido importantes para avaliar a cobertura de uma biblioteca genômica, estimar o número de cópias de um determinado gene no genoma, e desenvolver estratégias para a clonagem de genes baseado no mapeamento do genoma. Os resultados mostraram elevado número de cromossomos ( $2n = 210$ ) e conteúdo de DNA ( $2C = 22,5$  pg DNA/célula). Estudos citológicos na família Sapindaceae têm concentrado, principalmente, na diversidade do número ( $2n = 14$  a  $96$ ) e tamanho dos cromossomos. No gênero *Paullinia*, sete espécies tiveram o kariótipo estudado e todas apresentaram  $2n = 24$  cromossomos. Estes dados permitiram considerar que o kariótipo do guaranazeiro, em função do número e do tipo de cromossomos, é de origem complexa, que incluiu em sua gênese eventos de poliploidização e rearranjo do conjunto de cromossomos.

## Karyotype with 210 chromosomes in guaraná (*Paullinia cupana* ‘Sorbilis’)

Danival Vieira de Freitas · Carlos Roberto Carvalho ·  
Firmino José do Nascimento Filho ·  
Spartaco Astolfi-Filho

Received: 8 August 2006 / Accepted: 30 December 2006  
© The Botanical Society of Japan and Springer 2007

**Abstract** The genus *Paullinia* includes the economically important *P. cupana*, known as guaraná in Brazil and more recently in the world market. Native Americans of the Maué and Andirá tribes cultivated *P. cupana* ‘Sorbilis’ in central Amazon, and the Barés cultivated the ‘Typica’ variety in the upper Negro River (Brazil). Cytological studies in the Sapindaceae family have concentrated on the diversity in number (from  $2n = 14$  to 96) and size of the chromosomes. In *Paullinia*, seven species have been karyotyped and all show  $2n = 24$ . Meristem maceration, cellular dissociation and air-drying techniques were used for cytogenetic preparations and DNA content was determined by flow cytometry. Chromosome characterization and DNA content of *Paullinia cupana* Kunth ‘Sorbilis’ (Mart.) Ducke (Sapindaceae) were studied. The high chromo-

some number ( $2n = 210$ ) fall into two cytomorphological groups: (a) a metacentric and submetacentric group showing 25 sets of three pairs of chromosomes (2–76); (b) a group containing only acrocentric showing 12 sets of two pairs of chromosomes (82–105), a homologous submetacentric pair (1) and an acrocentric pair (81). Mean nuclear DNA content of guaraná was  $2C = 22.8$  pg. A karyogram was set up showing a high chromosome number complement.

**Keywords** Chromosome number · *C* value ·  
Flow cytometry · Guaraná · Polyploidy ·  
*Paullinia cupana* ‘Sorbilis’

### Introduction

Sapindaceae is a tropical and subtropical family with 140 genera and approximately 1,800 species (Judd et al. 2002). Radlkofer (1931–1934) recognized 14 tribes, among which *Paullinieae* (6 genera with 430 species) is the only one with a climbing habit. The genus *Paullinia* includes 180 species of neotropical lianas, except for *Paullinia pinnata*, which also occurs in the African tropics. This genus contains the economically important *P. cupana* Kunth, known as guaraná in Brazil and more recently in the world market. The species originated in the Amazon and is divided in two varieties, which differ mainly in fruit size, shape and colour: *P. cupana* ‘Typica’ Ducke (Venezuela and Colombia) and *P. cupana* ‘Sorbilis’ (Martius) Ducke (Brazil), the true guaraná (Radlkofer 1895; Lleras 1992). Native Americans of the Maué and Andirá tribes cultivated ‘Sorbilis’ in central Amazon, and the Barés cultivated ‘Typica’ in the upper Negro River in

D. V. de Freitas · S. Astolfi-Filho  
Instituto de Ciências Biológicas,  
Laboratório de Tecnologia de DNA,  
Universidade Federal do Amazonas,  
69.077-000 Manaus, AM, Brazil  
e-mail: danivalfreitas@hotmail.com

S. Astolfi-Filho  
e-mail: sastolfi@ufam.edu.br

C. R. Carvalho (✉)  
Departamento de Biologia Geral,  
Laboratório de Citogenética e Citometria de Plantas,  
Universidade Federal de Viçosa, Av. PH Rolfs, s/n,  
CEP 36.570-000 Viçosa, Minas Gerais, Brazil  
e-mail: ccarvalho@ufv.br

F. J. do N. Filho  
Setor de Genética e Melhoramento de Plantas,  
Embrapa/CPAA, Caixa Postal 319, 69048-660  
Manaus, AM, Brazil  
e-mail: firmino@cpaa.embrapa.br

northwestern Amazon. In the latter area, the variety 'Typica' was incorporated into the food habits of colonizers and travellers, without spreading outside of its original distribution. In central Amazon, the 'Sorbilis' variety became much more widely accepted, and is now a native species with great social and economic importance (Lleras 1992).

*Paullinia cupana* 'Sorbilis' (guaraná) is a climbing, woody, perennial vine that can reach 10 m in height, although it grows only to 3 m in the form of a spreading shrub when cultivated in open areas. The native people of central Amazon prepare a stimulating beverage from the caffeine-rich seeds (Henman 1982). Nowadays, guaraná is increasingly on demand for the multiple and medically proven properties of its seed extract, used as a weight reducer (Boozer et al. 2001), a memory enhancer and stimulant of the central nervous system (Espinola et al. 1997), an aphrodisiac (Antunes et al. 2001), a bactericide, an antioxidant (Basile et al. 2005) and for protection against gastric injuries caused by ethanol (Campos et al. 2003). The short-term medicinal effects of guaraná in humans and other mammals are generally due to its high concentration of caffeine and other alkaloids, as well as considerable amounts of tannin (Sombra et al. 2005).

To date, cytological studies in the Sapindaceae family have concentrated on the diversity in number and size of the chromosomes (Ferrucci 2000). Reported numbers vary from  $2n = 14$  in *Cardiospermum integerrimum* Radlk. (Ferrucci 1989) to  $2n = 96$  in *Melicoccus lepidopetalus* Radlk. (Ferrucci and Solís Neffa 1997). The *Paullinieae* tribe presents an almost complete aneuploid series from  $x = 7$  to 13, and some higher basic numbers (e.g.,  $x = 15$  and 16). In *Paullinia*, seven species have been karyotyped and all show  $2n = 24$ . The genome size of these *Paullinia* species was not determined. In the Sapindaceae family only four species have been analyzed: *Blighinia sapida* Kon. ( $2C = 1.8$  pg DNA), *Sapindus trifoliatus* L. ( $2C = 1.1$  pg DNA), *Hemigyrosa canescens* Thew. ( $2C = 1.8$  pg DNA) and *Litchi chinensis* Sonn. ( $2C = 1.4$  pg DNA). In this study we report the karyogram of *Paullinia cupana* 'Sorbilis', provide a morphological characterization of its chromosomes and its DNA content.

## Materials and methods

### Plant material

Seeds from clones of *P. cupana* 'Sorbilis' kept at the Embrapa Amazônia Ocidental germplasm bank in

Manaus (Amazonas, Brazil) were used. At the Federal University of Viçosa (Minas Gerais, Brazil), the seeds were sown in a greenhouse at 28°C with appropriate irrigation routine until germination. Seeds of *Vicia faba* ssp. *faba* 'Inovec' ( $2C = 26.90$  pg) and *Pisum sativum* L. cv Citrad ( $2C = 9.09$  pg), kindly donated by Dr. Jaroslav Dolezel (Experimental Institute of Botany, Czech Republic), were germinated in conditions similar to those of guaraná and used as standards for the DNA content analysis.

### Cytogenetic preparations

Roots of guaraná seedlings were treated with 3  $\mu$ M Oryzalin solution for 3 h at 30°C plus 15 min at -20°C. They were then washed in distilled water for 15 min and fixed in methanol/acetic acid (Merck®) solution (3:1) at -20°C. After 24 h, the roots were washed and digested in a Flaxzyme (NOVO Ferment®)/distilled water solution (1:10) at 35°C for 120 min. Root tips were washed in distilled water for 20 min and again fixed and stored at -20°C. The slides were prepared by the cellular dissociation of apical meristems and air-drying techniques. The slides were stained with Giemsa 5% solution in phosphate buffer (pH 6.8) for 5 min, washed twice in distilled water, and air-dried at 50°C, as described by Carvalho and Saraiva (1997).

### Image analysis

The images were captured with a video camera (CoolSnap-PRO™) connected to an Olympus™ BX60 microscope, using a  $\times 100$  immersion objective, and processed with the software IMAGE PRO-PLUS™ version 4.51 (Media Cybernetics®). Chromosome morphology analyses were also carried out using the Image SXM program (Rasband 1997), and the karyogram was assembled in a nonconventional way, with two cytomorphological groups: a metacentrics/submetacentrics group and another containing only acrocentrics separated in the karyogram, in order to enhance the visual aspects of the unusual karyotype.

### Flow cytometry

Tests with *Pisum sativum* L. cv Citrad ( $2C = 9.09$ ) as internal standards were carried out with guaraná. Vigorous young leaves were rinsed in distilled water at 4°C and simultaneously cut into 2 cm<sup>2</sup> fragments in 1 mL of modified NIB (Nucleus Isolation Buffer, Saxena and King 1989) solution (0.14 M Dextran

**Fig. 1** Mitotic metaphase of *Paullinia cupana* ‘Sorbilis’ with  $2n = 210$  chromosomes, prepared by enzymatic maceration, cellular digestion, air-drying and staining with Giemsa 5% solution. White spots mark the chromosomes that were involved in overlaps. Bar = 5  $\mu\text{m}$



Pharmacia<sup>®</sup>, 10 mM MES Sigma<sup>®</sup>, 10 mM NaCl, 10 mM KCl, 0.1 mM spermine Sigma<sup>®</sup>, 2.5 mM EDTA Sigma<sup>®</sup>, 2.5 mM PVP Sigma<sup>®</sup>, 0.01% Triton-X 100 Sigma<sup>®</sup>) at 4°C. The suspension was filtered through a nylon filter with a 40  $\mu\text{m}$  mesh diameter, then centrifuged (100 g for 5 min) and the pellet transferred to 2 mL tubes, and stained with 1.5 mL OTTO II solution (Otto 1990) 0.4 M  $\text{Na}_2\text{HPO}_4 \cdot 12\text{H}_2\text{O}$  + 15  $\mu\text{M}$  propidium iodide (PI) + 50  $\mu\text{M}$  RNase (Sigma<sup>®</sup>), for 10–15 min in the dark, and analyzed with a Partec PAS<sup>®</sup> flow cytometer (Partec<sup>®</sup> GmbH, Munster, Germany), equipped with a Laser source and a series of filters (TK 420, TK 560 and RG 610). The equipment was calibrated and aligned using microbeads and standard solutions according to the manufacturer’s recommendations (Partec<sup>®</sup>). The FlowMax<sup>®</sup> software (Partec<sup>®</sup>) was used for the data analyses. More than 5,000 nuclei in the foliar samples were analyzed and those with coefficients of variation above 3% were discarded. Three independent replications were analyzed on different days and average values of the  $2C$  DNA content are reported in picograms (pg). The histogram peak of *P. sativum* nuclei  $G_0/G_1$  was set up for channel 100.

## Results

### Cytogenetic analysis

Associating cytogenetic techniques of enzymatic digestion, cellular dissociation and air-drying with digital image-analysis tools, metaphases with appropriate quality for morphological characterization and karyogram assembly were obtained. Considering the

large number of chromosomes in this species, dozens of slides were prepared, of which 13 metaphases allowed unambiguous counting of 210 chromosomes, even though most of them showed 11–14 overlapping chromosomes, and two exceptionally good metaphases

**Table 1** Morphological characterization of the  $2n = 210$  mitotic chromosomes of guaraná (*Paullinia cupana* ‘Sorbilis’)

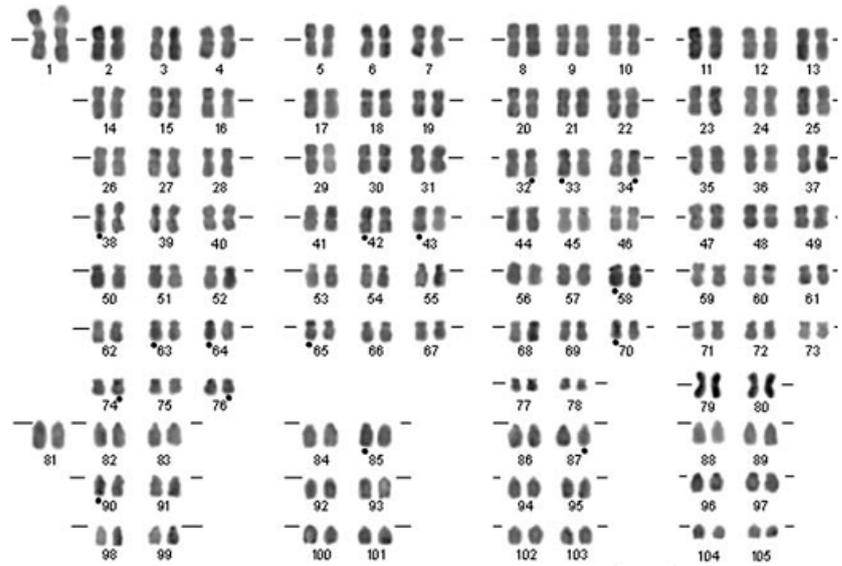
Chromosomes							
Pairs	Length <sup>a</sup>	Arm ratio <sup>b</sup>	Class <sup>c</sup>	Pairs	Length <sup>a</sup>	Arm ratio <sup>b</sup>	Class <sup>c</sup>
1	5.08	1.77	SM	62–64	1.91	2.22	SM
2–4	3.18	1.55	SM	65–67	1.89	2.07	SM
5–7	3.07	1.16	M	68–70	1.89	1.84	SM
8–10	3.03	1.20	M	71–73	1.70	1.56	SM
11–13	3.03	1.55	SM	74–76	1.51	2.19	SM
14–16	2.93	1.58	SM	77–78	1.20	1.71	SM
17–19	2.82	1.88	SM	79–80	2.39	1.24	M
20–22	2.81	1.24	M	81	2.43	19.48	A
23–25	2.80	1.02	M	82–83	2.08	19.87	A
26–28	2.79	1.56	SM	84–85	2.05	6.93	A
29–31	2.67	1.12	M	86–87	1.97	10.18	A
32–34	2.64	1.95	SM	88–89	1.94	7.10	A
35–37	2.61	1.10	M	90–91	1.88	9.11	A
38–40	2.61	1.02	M	92–93	1.84	12.13	A
41–43	2.43	1.02	M	94–95	1.83	9.14	A
44–46	2.34	1.54	SM	96–97	1.75	3.61	A
47–49	2.30	1.06	M	98–99	1.72	7.06	A
50–52	2.18	2.24	SM	100–101	1.63	5.31	A
53–55	2.11	2.24	SM	102–103	1.61	5.66	A
56–58	2.01	2.13	SM	104–105	1.20	6.41	A
59–61	1.97	1.64	SM				

<sup>a</sup> Mean value of pairs in micrometers

<sup>b</sup> Arm ratio (long/short)

<sup>c</sup> Class: *M* metacentric, *SM* submetacentric, *A* acrocentric

**Fig. 2** Karyogram of *P. cupana* 'Sorbilis' assembled from metaphase chromosomes shown in Fig. 1. The assembly was done so as to highlight relations between a metacentrics/submetacentrics group (2–76) and an acrocentrics group (82–105). Several pairs (1, 77, 78, 79, 80 and 81) of chromosomes did not fit into these two main groups. *Black spots* below the chromosomes mark those that were involved with overlaps. *Bar* = 5  $\mu$ m



showing seven and another with only five minor overlaps that was disjointed digitally in the image (Fig. 1).

The chromosomes were characterized according to the arm ratio (long/short) criteria described by Levan et al. (1964) and revised by Guerra (1986a, b) (Table 1). The karyogram was assembled so as to highlight morphological relations and composition among groups of chromosomes (Fig. 2). The 210 chromosomes appear to be divided into two main groups: (a) a group containing long (1.5–3.2  $\mu$ m) metacentrics and submetacentrics (25 sets of three pairs of chromosomes, pairs 2–76 in Figs. 2, 3); (b) a group containing only acrocentrics (12 sets of 2 pairs of chromosomes, pairs 82–105). A homologous submetacentric pair (1) and an acrocentric pair (81) were distinguished by their length and for not participating in these two main groups. Two smaller submetacentric pairs (77 and 78) and two densely stained metacentric pairs (79 and 80) were distinguished for forming two pairs of chromosomes groupings without being acrocentric (Figs. 2, 3).

#### Flow cytometry analysis

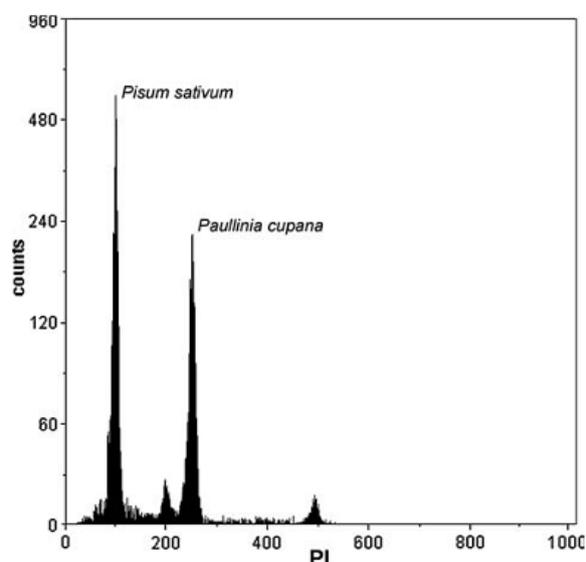
Flow cytometry of nuclei stained with propidium iodide generated a histogram of the relative DNA contents of guaraná and the comparative standard. The histogram presented two dominant peaks corresponding to nuclei in  $G_1/G_0$  of *P. cupana* cells (Channel 251) and *Pisum sativum* (Channel 100, internal standard  $2C = 9.09$  pg) (Fig. 4). The  $2C$  DNA content of guaraná was determined by the peaks relative positions to standard. The three repetitions using *Vicia faba* as standard yielded a mean nuclear DNA content of  $2C = 22.8$  pg for this guaraná species.

#### Discussion

The  $2n = 210$  chromosome number found in this study is the first report for this species and contrasts dramatically with other species of *Paullinia* reported in the



**Fig. 3** Idiogram showing groups with representative sets of three pairs of chromosomes (2–76) and sets of two pairs of chromosomes (82–105), and the non-grouped chromosomes (1, 77–80 and 81). *Bar* = 5  $\mu$ m



**Fig. 4** Histograms of the relative DNA content analyzed by flow cytometry of nuclei stained with propidium iodide (PI). *Paullinia cupana* ( $2C = 22.8$  pg) and *Pisum sativum* (internal standard,  $2C = 9.09$  pg). Note that the peak corresponding to *Pisum sativum* nuclei  $G_0/G_1$  was calibrated for channel 100

literature, all of which presented  $2n = 24$  chromosomes (Mangenot and Mangent 1958; Semple 1974; Ferrucci 1981; Guerra 1986b; Ferrucci and Solís Neffa 1997; Ferrucci 2000). Although the cytology of the Sapindaceae is poorly known, the presence of plants with high chromosome number is thought to be rare (Ferrucci 2000). Polyploidy is a key feature in plant evolution, especially in cultivated plants (Hancock 2005). The Native Americans who discovered the medicinal and stimulant properties of guaraná would have noticed these characteristics in a new plant during domestication and propagated it, creating the ‘Sorbilis’. In fact, the legend of the origin of guaraná describes exactly this episode of interest by natives (Pereira 1974).

The number and types of chromosome groups found in guaraná suggest a complex origin for this karyotype with high chromosome number, followed by numerical chromosomal rearrangements. For example, the first pair of chromosomes (1), which have secondary constrictions and a submetacentric morphology, may suggest an origin by fusion between metacentric and acrocentric sources. It has long been recognized that chromosomal fusion, i.e., the occurrence of Robertsonian translocations, plays a major role in karyotype evolution, however, in plants this changes have been detected in relatively few taxa (Jones 1998). Admitting that this fusion may have occurred, it is possible to argue that this event happened after the formation of

the complements with high chromosome number, since polyploidization can act as a trigger and/or facilitator to speed up evolution (Feldman and Levy 2005).

The DNA content of guaraná ( $2C = 22.8$  pg DNA) is much higher than the  $2C$  values of other species of the Sapindaceae family, which vary from 1.1 to 1.8 pg, but was expected given the large number of chromosomes. The DNA content of guaraná can be classified as a high  $C$  value, since Levin (2002) considers that the majority of the species present less than 10 pg of DNA. The very large chromosome complement of guaraná suggests that future research should concentrate on identifying the *Paullinia* species that participated in the hybridization(s) at the high-chromosome-number origin event, as well as determining the chromosome complement of the ‘Typica’ variety in northwestern Amazon. This variety and the species involved in the origin of guaraná are part of the primary gene pool of this crop and need to be collected and evaluated for use in guaraná improvement.

**Acknowledgments** We thank Prof. Charles R. Clement, Instituto Nacional de Pesquisas da Amazônia, for his critical and constructive suggestions on the article. This research is part of the REALGENE work supported by MCT/CNPq Brazil.

## References

- Antunes E, Gordo WM, Oliveira CE, Hyslop S, Nucci G (2001) The relaxation of isolated rabbit corpus cavernosum by the herbal medicine Catuama and its constituents. *Phytother Res* 15:416–421
- Basile A, Ferrara L, Del Pezzo M, Mele G, Sorbo S, Bassi P, Montesano D (2005) Antibacterial and antioxidant activities of ethanol extract from *Paullinia cupana* Mart. *J Ethnopharmacol* 102:32–36
- Boozer CN, Nasser JA, Heymsfield SB, Wang V, Chen G, Solomon JL (2001) An herbal supplement containing Ma Huang-Guaraná for weight loss: a randomized, double-blind trial. *Int J Obes* 25:316–324
- Campos AR, Barros AIS, Santos FA, Rao VSN (2003) Guarana (*Paullinia cupana* Mart.) offers protection against gastric lesions induced by ethanol and indomethacin in rats. *Phytother Res* 17:1199–1202
- Carvalho CR, Saraiva LS (1997) High-resolution HKG-banding in maize mitotic chromosomes. *J Plant Res* 110:417–420
- Espinola EB, Dias RF, Mattei R, Carlini EA (1997) Pharmacological activity of guaraná (*Paullinia cupana* Mart.) in laboratory animals. *J Ethnopharmacol* 55:223–229
- Feldman M, Levy AA (2005) Allopolyploidy—a shaping force in the evolution of wheat genomes. *Cytogenet Genome Res* 109:250–258
- Ferrucci MS (1981) Recuentos cromosômicos en Sapindáceas. *Bonplandia* 5:73–81
- Ferrucci MS (1989) Cromosomas en *Cardiospermum* y *Diplokeleba* (Sapindaceae), significado taxonômico y evolutivo. *Bonplandia* 6:151–164
- Ferrucci MS (2000) Cytotaxonomy of Sapindaceae with special reference to the tribe Paullinieae. *Genet Mol Biol* 23:941–946

- Ferrucci MS, Solís Neffa VG (1997) Citotaxonomía de Sapindaceae sudamericanas. *Bol Soc Argent Bot* 33:77–83
- Guerra MS (1986a) Reviewing the chromosome nomenclature of Levan et al. *Rev Bras Genet* 9:741–743
- Guerra MS (1986b) Citogenética de Angiospermas coletadas em Pernambuco. I. *Rev Bras Genet* 9:21–40
- Hancock JF (2005) Contributions of domesticated plant studies to our understanding of plant evolution. *Ann Bot* 96:953–963
- Henman AR (1982) Guaraná (*Paullinia cupana* ‘Sorbilis’): ecological and social perspectives on an economic plant of the central Amazon basin. *J Ethnopharmacol* 6:311–338
- Jones K (1998) Robertsonian fusion and centric fission in karyotype evolution in higher plants. *The Bot Rev* 64:273–289
- Judd WS, Campbell CS, Kellogg EA, Stevens PF (2002) Plant systematics, a phylogenetic approach. Sinauer Associates, Sunderland
- Levan A, Fredga A, Sanderberg AA (1964) Nomenclature for centromeric position in chromosome. *Hereditas* 52:201–220
- Levin DA (2002) The role of chromosomal change in plant evolution. Oxford University Press, New York
- Lleras E (1992) Espécies de *Paullinia* com potencial econômico. In: Bermejo H, León JE (eds) Cultivos Marginados: Otra Perspectiva de 1492. Food and agriculture organization (FAO) and Jardín Botánico de Córdoba. FAO Plant Prod Prot Pap (Roma) 26:221–226
- Mangenot S, Mangenot G (1958) Deuxième liste de nombres chromosomiques nouveaux chez diverses Dicotylédones et Monocotylédones d’Afrique occidentale. *B Jardin Bot l’état* 28:315–329
- Otto F (1990) High-resolution DNA flow cytometry in malignant melanoma. *Cytometry Suppl* 4:55
- Pereira N (1974) Panorama da alimentação indígena: Comidas, bebidas and tóxicos na Amazônia brasileira, Editora São José, Rio de Janeiro
- Pikaard CS (1999) Nucleolar dominance and silencing of transcription. *Trends Plant Sci* 4:478–483
- Radkofer L (1895) Monographie der Sapindaceen-Gattung Paullini. Verlag der Konigliche Akademie, München
- Radlkofer L (1931–1934) Sapindaceae. In: Engler A (ed) (1965) Das Pflanzenreich (IV). Verlag von HR Engelmann/J Cramer, Weinheim 165:1–1539
- Rasband W (1997) Image SXM 1.61. A public domain software for image analysis, written at the US National Institute of Health, extensions by Steve Barrett and available from the Internet by anonymous ftp from <http://www.zippy.nimh.nih.gov>
- Saxena PK, King J (1989) Isolation of nuclei and their transplantation into plant protoplast. In: Bajaj YPS (ed) Biotechnology in agriculture and forestry. Plant protoplast and genetic engineering. Springer, Heidelberg, pp 328–342
- Semple JC (1974) Chromosome numbers of phanerogams. *Ann Mo Bot Gard* 61:902–903
- Solís Neffa VG, Ferrucci MS (1998) Cariótipos de Sapindaceae sudamericanas. *Bol Soc Argent Bot* 33:185–190
- Sombra LL, Gomez MR, Olsina R, Martýnez LD, Silva MF (2005) Comparative study between capillary electrophoresis and high performance liquid chromatography in “guarana” based phytopharmaceuticals. *J Pharmaceut Biomed* 36:989–994

## **5. Capítulo III – Transcriptoma do fruto do guaranazeiro (*Paullinia cupana* var. *sorbilis*), um estimulante originário da Floresta Amazônica consumido tradicionalmente**

Artigo publicado na *Plant Cell Report* (2008) 27:117–124.

### **5.1. Sinopse**

O presente estudo permitiu identificar as seqüências gênicas envolvidas nas vias metabólicas que conduzem à biossíntese de compostos de elevado interesse econômico no guaranazeiro. O seqüenciamento de transcritos de três fases de desenvolvimento dos frutos (verdes imaturos, estágio intermediário e maduros) do guaranazeiro da cultivar BRS-Amazonas, gerou um banco com 15.387 EST's (*expressed sequence tags*) válidas distribuídas em 2.628 contigs e 5.969 singletons, com comprimento médio de 773 pares de bases. Algumas das ESTs são interessantes e podem explicar, pelo menos em parte, as propriedades medicinais atribuídas ao extrato do pó das sementes torradas e que estão sendo aos poucos comprovadas por experimentação científica. Entre estas estão às enzimas que participam das vias de síntese e catálise de flavonóides e carotenóides e as cafeína sintases. Foi identificado também um grupo de seqüências relacionadas com genes de vias de interação planta-patógeno, incluindo proteínas PR (*pathogenesis related*), inibidores de proteases de cisteína e seqüências relacionadas a endoquitinases.

## 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  
© Springer-Verlag 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, Macapá, 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  $1.6 \times 10^6$  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<sup>®</sup> 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 (Felsenstein 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 <i>e</i> -value $<10^{-10}$	
All organisms	5,956
<i>Arabidopsis thaliana</i>	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)  $\times$  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.

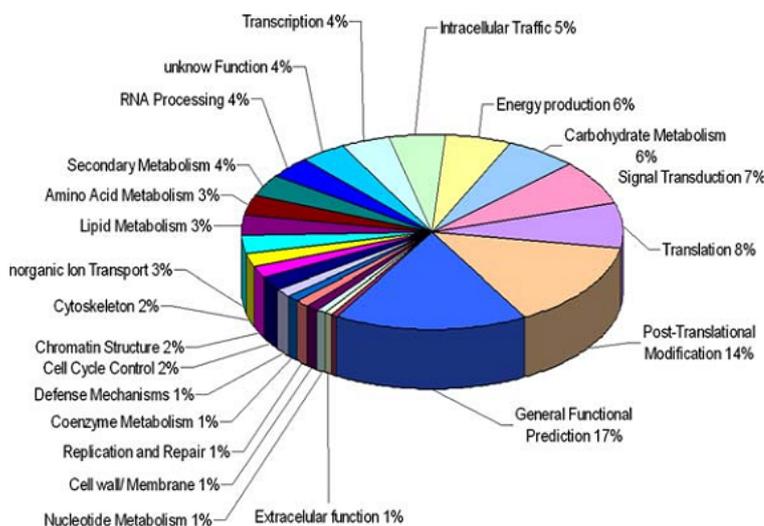
## Secondary metabolism transcripts

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

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

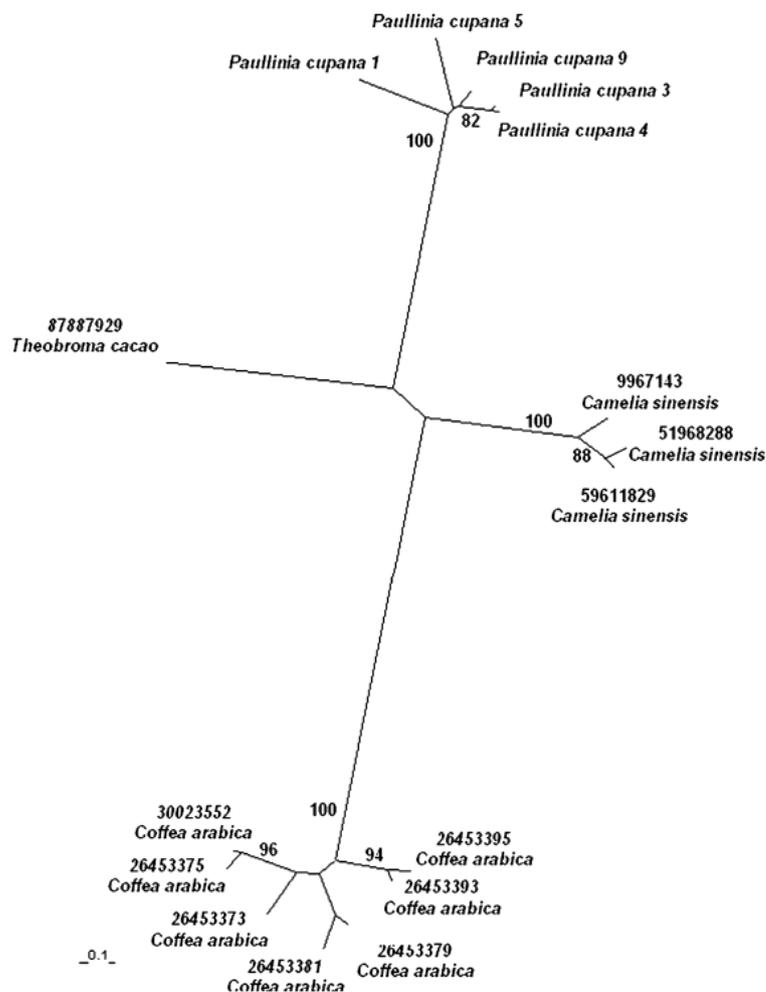
At least 177 guarana reads showed significant similarity to plant resistance proteins. From these, 86 ESTs for pathogenesis related (PR) proteins which biological functions in

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 agro-nomical 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, Tataru 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* (Catuaema) 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) Antigenotoxic 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

## 6. Capítulo IV – N-metiltransferase do guaranazeiro (*Paullinia cupana var. sorbilis*) envolvida na biossíntese da cafeína

### 6.1. Sinopse

O Brasil é o único produtor mundial de guaraná e atende ao mercado nacional e internacional. O guaranazeiro é essencialmente uma planta em processo de melhoramento genético que possui propriedades medicinais e estimulantes. A partir da biblioteca de cDNA do fruto do guaranazeiro foram caracterizados 89 EST's com homologia com N-metiltransferase (CS) do cacau, 68 com a do chá da Índia (tanto TCS1 quanto TCS2) e 60 com a do café (*Coffea arabica*). A sequência de cDNA do fruto do guaranazeiro com maior similaridade com as N-metiltransferases das outras espécies estudadas foi a analisada neste estudo, tem 755 pb e foi denominada *GNMT*. Esta sequência foi utilizada para a síntese química de iniciadores para a análise do número de cópias no genoma e dos níveis de transcritos da *GNMT* de vários órgãos/tecidos do guaranazeiro e espécimes selvagens. O número de cópias gênica no genoma do guaranazeiro apresentou cerca de 3,0 vezes mais cópias que os espécimes selvagens (guaranarana). O fruto do guaranazeiro com 70 dias de desenvolvimento apresentou o maior nível de transcritos (3,287) da *GNMT*, o qual aumentou gradualmente com o processo de maturação do fruto do guaranazeiro.

## 6.2. Introdução

A cafeína é um dos principais alcalóide purina em *Paullinia cupana*. A primeira análise química no guaranazeiro foi realizada no século XVII pelo botânico alemão Theodore Von Martius, que isolou das sementes uma substância cristalina, amarga e branca, denominada “guaranina”, que foi mais tarde renomeada cafeína.

A cafeína é conhecida por ter efeitos sensoriais e estimulantes quando consumida nas bebidas como o guaraná (*Paullinia cupana*), o café (*Coffea arábica*, *Coffea canephora*), o chá (*Camellia sinensis*) e a erva-mate (*Ilex paraguayensis*). As ações farmacológicas provavelmente são relacionadas com o bloqueio dos receptores de adenosina A<sub>1</sub> (Cauli e Morelli, 2005). O papel biológico exato da cafeína e dos alcalóides púricos relacionados é ainda obscuro, embora haja duas hipóteses principais: a) propõe-se que a cafeína realiza a função de proteção das folhas e frutos contra predadores (teoria da defesa química [Frischknecht, 1985; Hollingsworth et al., 2002]) e; b) a cafeína liberada pelo revestimento da semente impede a germinação de outras sementes (teoria alelopática ou autotóxica [Friedman e Waller, 1983]).

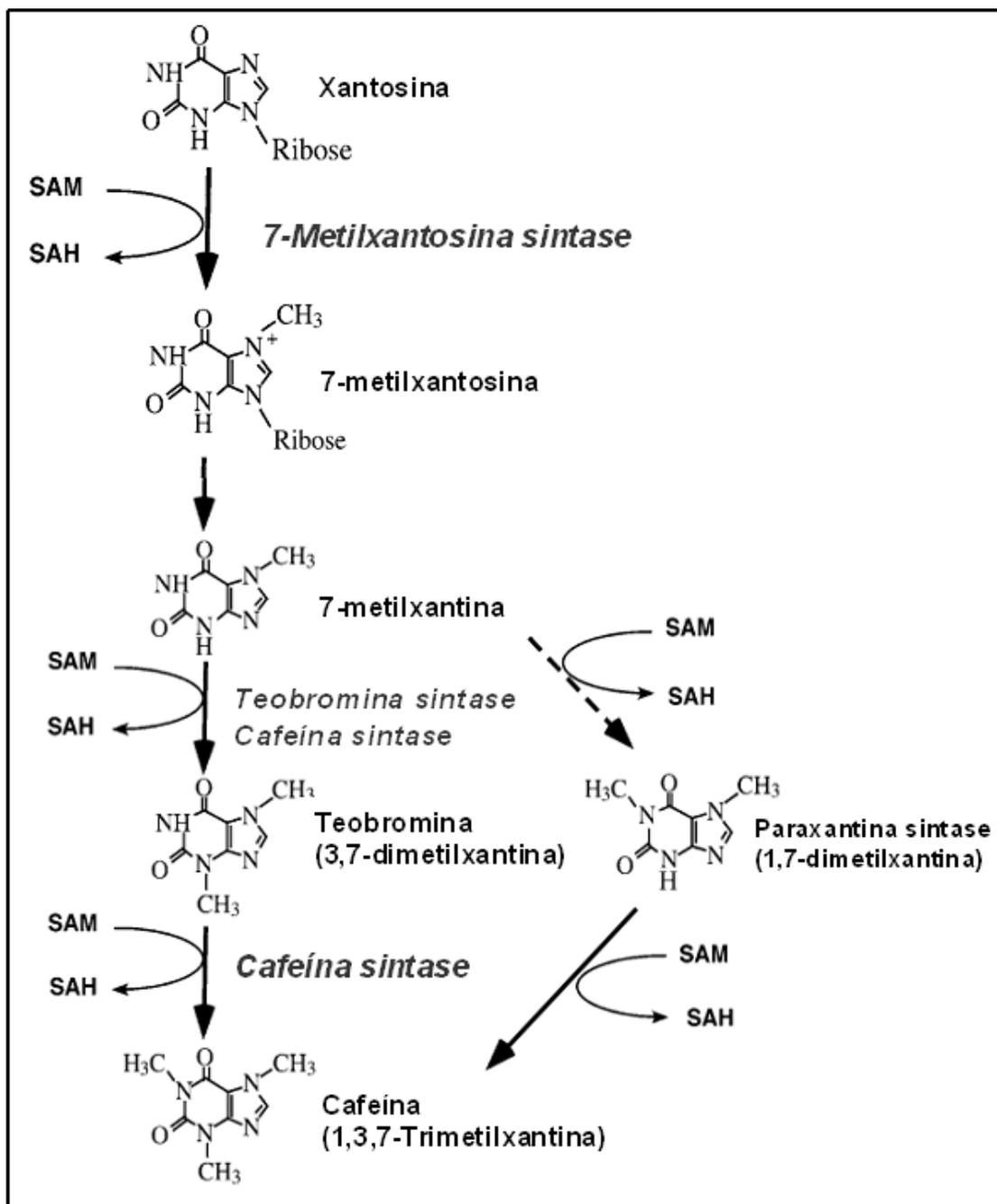
A cafeína pode ser produzida por duas vias (Figura 1), a principal envolve a 7-metilxantina → teobromina → cafeína (Suzuki 1972; Ashihara e Kubota 1987; Ashihara et al. 1996, 2004; Kato et al. 1996). Esta via ocorre no chá e no café, e é apoiada por experimentos utilizando N-metiltransferases (NMTs) recombinantes e naturais destas espécies (Kato et al. 1999, 2000; Mizuno et al. 2003a, b; Uefuji et al. 2003). A via principal parece ser universal nos alcalóides púricos de plantas (Ashihara 1993; Koyama et al. 2003; Yoneyama et al. 2006). A outra via metabólica envolve a xantosina → 7-metilxantosina → 7-metilxantina → paraxantina → cafeína, sendo esta mais elucidada no café. A taxa de biossíntese de cafeína parece ser

regulada, primeiramente, pela atividade da *N*-metiltransferase. O primeiro, o segundo e o quarto passo da principal via são catalisados pela *N*-metiltransferase que utilizam a *S*-adenosil-*L*-metionina (*SAM*) como o doador de metil (Kato,1996). Os dois passos finais são catalisados pela *CS*, uma enzima bifuncional no chá (*TCS1* e *TCS2*) que compreende duas atividades *N*-metiltransferases dependentes de *SAM* (Yoneyama et al., 2006).

As NMT que estão envolvidas na biossíntese da cafeína podem ser divididas em três tipos, de acordo com a especificidade do substrato. Estes três tipos de enzimas e os genes correspondentes já foram caracterizados no café e no chá. A primeira etapa de metilação na via biossintética da cafeína do xantosina a 7-metilxantosina é catalisado pelo 7-metilxantosina sintase (Mizuno et al., 2003b; Uefuji et al., 2003). O único substrato que é disponível para a 7-metilxantosina sintase é a xantosina. Recentes estudos da estrutura química desta proteína têm mostrado que ela é uma enzima bifuncional: realiza a transferência de metil e cliva os nucleosídeos (McCarthy e McCarthy, 2007). As últimas duas etapas da biossíntese da cafeína são catalisadas pela cafeína sintase (Kato et al., 1999, 2000; Mizuno et al., 2003a; Uefuji et al., 2003). A primeira NMT da via de biossíntese da cafeína a ser caracterizada bioquimicamente foi clonada de folhas novas do chá (*C. sinensis*); esta descoberta permitiu o isolamento subsequente de muitos genes de NMT que são envolvidos na biossíntese da cafeína (Kato et al., 1999, 2000).

Por muitos anos, *Paullinia cupana* foi o foco das atenções devido à presença da cafeína (Henman, 1982; Belliardo et al., 1985; Bempong & Houghton, 1992). Nas sementes do guaranazeiro são encontradas elevadas concentrações de cafeína, assim como “teofilina” e o “teobromina” (em quantidades bem menores) (Weckerle et al., 2003). Embora o guaranazeiro seja a planta que mais acumula cafeína, no que

se refere a via biossintética desse alcalóide é a menos estudada. O presente trabalho apresenta os resultados da caracterização molecular *in silico* de EST`s com homologia para *N-metiltransferases* do guaranazeiro (*GNMT*), a quantificação do seu número de cópias no genoma e os níveis de transcritos expressos em diferentes tecidos do guaranazeiro (*Paullinia cupana* var. *sorbilis*) e de espécimes selvagens (guaranaranas - *Paullinia* sp) pelo método qRT-PCR.



**Figura 1:** Rota biossintética da Cafeína. As setas sólidas indicam a principal rota de biossíntese. 7-metilxantósina sintase e teobromina sintase catalizam apenas a conversão de xantósina para 7-metilxantósina e metilxantina para teobromina, respectivamente. A cafeína sintase possui amplo substratos específicos. A primeira (1), a terceira (3), e a quarta (4) etapa da via principal são reações de N-metilações (Ishida et al., 2009).

## 6.3. MATERIAL E MÉTODOS

### 6.3.1. Material vegetal e isolamento de ácidos nucleicos

Material vegetal de diferentes acessos e tecidos foi coletado no Banco de Germoplasma de *Paullinia* sp., existente na sede da Empresa Brasileira de Agropecuária - Amazônia Ocidental (EMBRAPA – CPAA) no município de Manaus, Amazonas, Brasil (Tabela 1). Os tecidos coletados foram transportados em gelo seco e armazenados a  $-80^{\circ}\text{C}$  para extração de ácidos nucleicos. A extração do DNA foi feita pelo método CTAB (Doyle e Doyle, 1987) utilizando o protocolo modificado por Ferreira e Grattapaglia (1998). O material vegetal (folíolo expandido do guaranazeiro (acesso BRS 300) e de duas espécimes selvagem - guaranaranas) foi macerado com nitrogênio líquido. O RNA total foi extraído a partir de aproximadamente 100 mg de diferentes tecidos utilizando o kit de extração Concert Plant RNA da Invitrogen (Invitrogen, San Francisco, CA), de acordo com as recomendações do fabricante. Ao final deste processo foi gerado um volume de 40  $\mu\text{L}$  de solução (RNA total + 40  $\mu\text{L}$  de água RNase free). As amostras de DNA e RNA extraídas foram quantificadas pelo espectrofotômetro Nanodrop (Thermo Scientific NanoDrop™).

Uma quantidade de aproximadamente 5  $\mu\text{g}$  de RNA total de cada tecido estudado foi tratada com DNase I utilizando o kit da Ambion, em uma reação com volume final de 20  $\mu\text{L}$ , segundo o protocolo do fabricante. Posteriormente, foi realizado PCR convencional utilizando os iniciadores para o gene da *GNMT* para confirmar a ausência de DNA.

A síntese de cDNA foi realizada a partir de 1  $\mu\text{g}$  de RNA total pela utilização da transcriptase reversa com primers hexâmeros randômicos (5  $\mu\text{M}$ ) utilizando o kit

First Strand cDNA Synthesis<sup>®</sup> (Fermentas, Hanover, MD) em uma reação com volume final de 20 µL, conforme o protocolo do fabricante.

### 6.3.2. Caracterização *in silico* de uma N-metiltransferase do guaranazeiro

Os experimentos foram conduzidos utilizando clones da biblioteca de cDNA de três fases de maturação do fruto do guaranazeiro (15, 45 e 70 dias) (Angelo et al., 2008). Os *contigs* e *singletons* (GeneBank: EC 763506 até EC 778393, [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)) que apresentaram homologia com *cafeínas sintase* e *catalase* (gene normalizador da qRT PCR), assim definidos pela anotação automática, foram selecionados para caracterização *in silico*. As sequências da *catalase* e de outras N-metiltransferases da via de biossíntese da cafeína do chá, do cacau e do café foram acessadas no banco de dados de plantas (Uniprot - [www.uniprot.gov](http://www.uniprot.gov)) e utilizadas como *drivers* na análise filogenética. As sequências foram alinhadas usando o programa ClustalW (Thompson et al., 1994). O alinhamento da *catalase* e da *GNMT* foi manualmente ajustado para preservar subdomínios conservados. O alinhamento foi ajustado para remover *gaps* na maioria das sequências. EST's com sequências inferiores a 100 pb e com *e-value* menores que  $1e^{-10}$  foram descartadas. O alinhamento foi analisado pelo programa Phylip (Felsenstein, 1995) e pelo programa de livre domínio (CLC Sequence Viewer – 6.1 / [www.clcbio.com](http://www.clcbio.com)), utilizando o algoritmo neighbor-joining (NJ) e a distância genética como critérios. O *bootstrap* foi executado com 1.000 repetições usando NJ/UPGMA.

Os iniciadores para a quantificação do número de cópias no genoma e do número de transcritos da CS e os da *catalase* foram construídos a partir da sequência com melhor *hit* no alinhamento acima citado. Os iniciadores para *GNMT* (*GNMT-F/1* CTCTTCAAAGAAAATCACCACCAT e *GNMT-R/1*

CCTGAAGCACACCCCAAATCT) e para *catalase* (*Cat-F/1*  
TGCCGAATCTTTGCTTATGGT e *Cat-R/1* GAGCATCAGGTAGTTTGGTCCAA)  
foram desenhados utilizando-se o software Primer Express (Applied Biosystems). A  
especificidade dos iniciadores foi avaliada verificando a formação de banda única  
em gel de agarose e pela curva de dissociação na qRT-PCR. Para verificar a  
especificidade dos iniciadores, todos os DNA/cDNA foram submetidos a PCR  
convencional para verificar o fragmento amplificado. A reação ocorreu em  
termociclador Veriti<sup>®</sup> (Applied Biosystems) com 35 ciclos, iniciando com uma  
temperatura de 94°C por 1 minuto, anelamento de 60°C por 45 segundos e  
alongamento a 72°C por 1 minuto. Os *amplicons* foram purificados utilizando o kit  
Wizard PCR Preps DNA Purification System (Promega, Southampton, UK), de  
acordo com as recomendações do fabricante. Parte do amplicon purificado foi  
clonado para construção da curva padrão para quantificação do número de cópias  
gênicas.

### **6.3.3. Análise do número de cópias e transcritos do gene *GNMT* por PCR em Tempo Real usando o sistema SYBR Green**

Para as análises do número de cópias no genoma e do número de transcritos  
utilizou-se a metodologia de quantificação por PCR tempo Real (qRT PCR),  
utilizando-se o equipamento ABI Prism 7300 Sequence Detection System (Applied  
Biosystems). As reações foram realizadas com volume de 5 µl de “SYBR<sup>®</sup> Green  
PCR Master Mix” (Applied Biosystems), 1 µl do iniciador *GNMT* – F/1 (1 pmol/ µl) e 1  
µl do iniciador *GNMT* – R/1 (1 pmol/µl), 1 µl de DNA/cDNA e água (RNase free). As  
condições de termociclagem foram: um passo inicial de 95°C por 10 minutos,  
seguidos por 40 ciclos de 95°C por 15 segundos, 60°C por 30 segundos e 72°C por

20 segundo. As reações foram realizadas em triplicata, permitindo a detecção de possíveis erros. Quando os resultados das triplicatas apresentaram uma variação maior que 1 ciclo, novas repetições foram feitas. Reações sem cDNA também foram utilizadas como uma forma de detectar possíveis contaminações (controle negativo).

A quantificação do número de cópias de cafeína sintase presentes nos genomas estudados foi realizada pelo método da curva padrão externa (Hernandez et al., 2004; Lee et al., 2006). A construção desta curva foi feita a partir da clonagem do produto da amplificação da *cafeína sintase* e da *catalase* no plasmídeo pGEM®-T Easy Vector. Os plasmídeos com os insertos foram obtidos por meio de mini-preparação, linearizados com a enzima *Pst*I e quantificados utilizando espectrofotometria (Nanodrop). A partir de alíquotas de concentrações conhecidas dos plasmídeos recombinantes, com cinco diluições seriadas do DNA (50 ng/μl até 5 pg/μl), foi realizada a qRT PCR. O ciclo no qual foi detectado um aumento significativo da fluorescência (*Crossing point* - *CP*), foi considerado para quantificar o produto inicial em cada amostra e utilizado como referência para a curva-padrão. A curva padrão foi obtida plotando os valores de *Cp* versus o logaritmo da concentração. Os valores dos pontos de amplificação do *Cp* do *GNMT* foram utilizados para calcular o número de cópias dos genomas, estimado pela equação:  $N=N_0 \times E^n$ , onde *N* é o número de cópia após *n* ciclos e *E* é a eficiência de amplificação. *N<sub>0</sub>* pode ser calculado pela equação assumindo que o número de ciclos equivale ao *Cp*.  $N_0 = 10^{-(Cp/slope) - 1}$  (Pfaffi, 2002). Os resultados do *Cp* obtidos da qRT-PCR para análise do número de cópias no genoma foram analisados no Microsoft Excel. A Análise de Variância – *single-factor* (ANOVA) foi realizado pelo método estatístico descritivo no programa SSP versão 2.75.

O método  $\Delta\Delta C_t$  (Pfaffl et al., 2002) foi utilizado para determinar a expressão do gene da *GNMT*. A taxa de expressão deste transcrito de interesse foi estimada em diferentes tecidos (Tabela 1), sendo a expressão da *catalase* foi utilizada como calibrador e o número de transcritos da *GNMT* do tecido foliar do guaranazeiro foi utilizado como amostra referência. Os valores  $\Delta\Delta C_t$ , a média e o erro padrão foram calculados pelo software de domínio livre Rest 2008 V2.0.7 (Pfaffl et al., 2002). A regressão linear dos valores de  $C_p$  também foi utilizada para avaliar a expressão da *catalase* entre os tecidos e os experimentos de qRT-PCR.

## 6.4. RESULTADOS

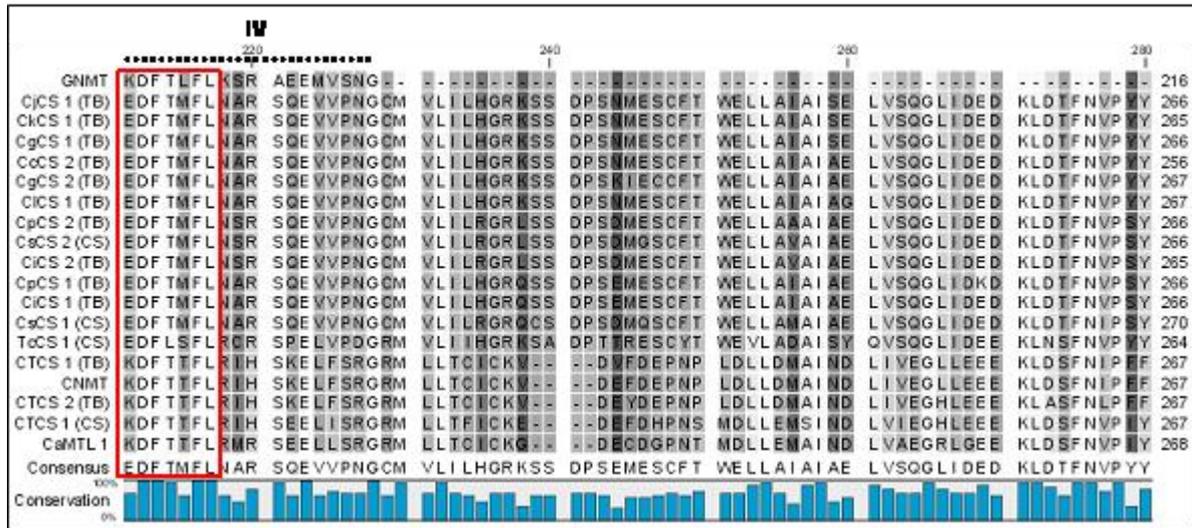
### 6.4.1. Caracterização *in silico* das N-metiltransferases

A partir da biblioteca de cDNA do fruto do guaranazeiro foram caracterizados 89 EST's com homologia para a CS do cacau, 68 para a do chá da Índia (tanto TCS1 quanto TCS2) e 60 para a do café (*Coffe arabica*). A sequência de cDNA do guaranazeiro que apresentou melhor *hit* com as sequências de *N-metiltransferases* depositada no Uniprot possui 755 bp e foi denominada *GNMT*. Esta sequência foi traduzida pelo programa Clustal W (Thompson et al., 1994) e utilizada para o alinhamento e análise filogenética com as sequências de aminoácidos da *teobromina sintase* e *cafeína sintase* classes I e II do café (*Coffea arábica* e *C. canephora*), cacau (*Theobroma cacao*), e chá da Índia (*Camellia* sp), todas já caracterizadas (Figura 2). Os resultados mostraram que a *GNMT* é mais estreitamente relacionada a enzima de cacau classe I (Figura 2).

A sequência gênica da *GNMT* compartilha com as demais N-metiltransferases conhecidas quatro regiões altamente conservadas, *motif A*, *motif B'*, *motif C* e a região de YFFF. As seqüências consenso dos *motifs* da região A e C, e da região YFFF da *GNMT* foram idênticas as das espécies comparadas. Entretanto, o *motif B'* situado entre o *motif A* e o *motif C*, que apresenta a seqüência 'LNDLPSNDFN' com elevado número de aminoácidos hidrofílicos, apresentou variações.

Em seu conjunto, as análises indicam que a sequência escolhida (*GNMT*) é homóloga as N-metiltransferase da via biossintética da cafeína de outras espécies.

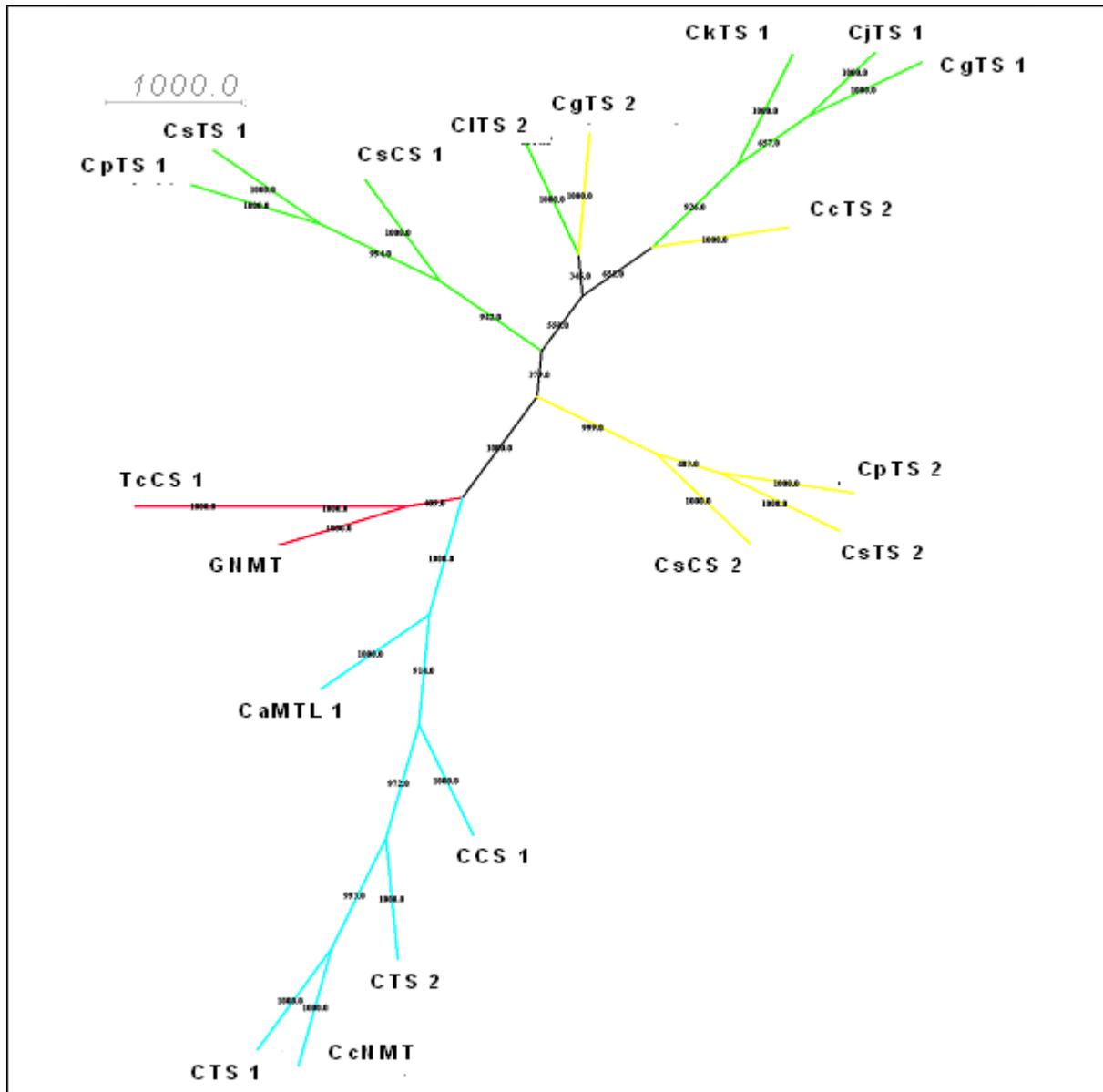




**Figura 2:** Comparação de aminoácidos da *GNMT*, da *teobromina sintase* e *cafeína sintase* classes I e II do café [*Coffea arábica* (cafeína sintase classe 1 - CaMTL 1) e *C. canephora* (cafeína sintase classe 1 – CCS1, N-metiltransferase – CcNMT, teobromina sintase classe 1 – CTS 1, teobromina sintase classe 2 – CTS 2)], cacau [*Theobroma cacao* (cafeína sintase classe 1 - TcCS1)- ], e espécies do gênero *Camellia* [*C. cinensis* (cafeína sintase classe 1 – CsCS 1, cafeína sintase classe 2 – CsCS 2, teobromina sintase classe 1 – CsTS 1, teobromina sintase classe 2 – CsTS 2), *C. granthamiana* (cafeína sintase classe 1 – CgTS 1, cafeína sintase classe 2 – CgTS 2), *C. lutchuensis* (teobromina sintase classe 2 – CITS 2), *C. kissi* (teobromina sintase classe 1 – CkTS 1), *C. japonica* (teobromina sintase classe 1 – CjTS 1), *C. chrysantha*. Os boxes representam resíduos de aminoácidos idênticos. Os *motifs* que interagem com a SAM (A, B' e C) e da região conservada denominada de “região YFFF” são apresentados pelos boxes vermelhos fechados. As seqüências conservadas na família das metiltransferases são indicadas pelas linhas pontilhadas (I,II,III e IV).

#### **6.4.2. Estimativa do número de cópias do gene *GNMT* no genoma do guaranazeiro**

Para determinar o número de cópias do gene da *GNMT* do guaranazeiro e de outras duas espécimes selvagens (guaranaranas), utilizou-se PCR em tempo real. A especificidade foi checada pela análise da curva de melting e gel de agarose. A curva de melting apresentou apenas um pico e o tamanho estimado do amplicon foi de cerca de 70 pb correspondendo ao tamanho esperado. A curva padrão foi linear na escala testada ( $R^2 \geq 0,9800$ ) para as reações em triplicatas. A partir do valor do *slope* da curva padrão (-2,8849) foi possível verificar uma elevada eficiência da amplificação, que foi determinada na linha investigada e utilizada para quantificação da taxa de expressão da *GNMT*. O número médio de cópias gênicas da *GNMT* no genoma do guaranazeiro foi estimado em cerca de 3,0 vezes o número de cópias da guaranarana - II e 1,5 vezes o da guaranarana - I (Tabela 1).



**Figura 3:** Análise filogenética das sequências de aminoácidos da *GNMT*, da *teobromina sintase* e *cafeína sintase* classes I e II do café [*Coffea arábica* (cafeína sintase classe 1 - CaMTL 1) e *C. canephora* (cafeína sintase classe 1 – CCS1, N-metiltransferase – CcNMT, teobromina sintase classe 1 – CTS 1, teobromina sintase classe 2 – CTS 2)], cacau [*Theobroma cacao* (cafeína sintase classe 1 - TcCS1)-), e espécies do gênero *Camellia* [*C. cinensis* (cafeína sintase classe 1 – CsCS 1, cafeína sintase classe 2 – CsCS 2, teobromina sintase classe 1 – CsTS 1, teobromina sintase classe 2 – CsTS 2), *C. granthamiana* (cafeína sintase classe 1 – CgTS 1, cafeína sintase classe 2 – CgTS 2), *C. lutchuensis*

(teobromina sintase classe 2 – CITS 2), *C. kissi* (teobromina sintase classe 1 – CkTS 1), *C. japonica* (teobromina sintase classe 1 – CjTS 1), *C. chrysantha* (teobromina sintase classe 2 – CjTS 2).

**Tabela 1:** Estimativa da expressão gênica da *cafeína sintase* (CS). A estimativa da variação do número de cópias no genoma e o nível de transcritos de CS foram estabelecidos por PCR em Tempo Real no guaranazeiro (acesso BRS) e de espécies afins (guaranaranas).

Acesso	RNCEG <sup>1</sup>	Tecido (amostra)	Expressão relativa	Std. Error
Guaranazeiro	2,98 X (3)	(I) Folíolo expandido	1,0	0,95 - 1,05
		(II) Pecíolo	0,10	0,06 - 0,17
		(III) Inflorescência	0,74 *	0,50 - 1,10
		(IV) Fruto 15 dias	0,13	0,07 - 0,22
		(V) Fruto 45 dias	2,07	1,37 - 3,13
		(VI) Fruto 70 dias	3,29	2,98 - 3,64
Guaranarana - I	1,79 X (2)	(VII) Folíolo expandido	0,21	0,13 - 0,32
Guaranarana - II	1,00 X (1)	(VIII) Inflorescência	0,36	0,24 - 0,52
		(IX) Fruto início maturação	1,11 *	0,88 - 1,40
		(X) Fruto final maturação	0,53 *	0,32 - 0,89
		(X) Folíolo expandido	0,70 *	0,47 - 1,03

**RNCEG<sup>1</sup>** – Razão do número de cópias da *GNMT* estimada no genoma dos acessos estudados por qRT PCR. Entre parênteses os valores arredondados.

\* Não diferem estatisticamente da amostra de referência (folíolo expandido do acesso BRS) a 5% de probabilidade.

### 6.4.3. Quantificação do número de transcritos da N-metiltransferase do guaranazeiro

Para analisar a expressão do gene *GNMT* a acumulação relativa de mRNA foi quantitativamente examinada em vários tecidos do guaranazeiro e espécimes selvagens, incluindo folíolo, inflorescência, fruto em diferentes fases de desenvolvimento (Tabela 1). O gene escolhido para normalizar (controle interno) as

análises da expressão da *GNMT* foi o da *catalase*. Os níveis de mRNA foram calculados cruzando os valores de cada tecido com os níveis de expressão do tecido de referência (folíolo do guaranazeiro - padrão). A expressão do controle interno, *catalase*, foi relativamente estável nos seis tecidos do guaranazeiro examinados, com uma variação de até 58% do número de transcritos detectada entre os diferentes tecidos. Os valores de *Cp* variaram entre 27 e 35 ciclos nos diferentes órgãos/tecidos, indicando um poder de detecção de até quatro ordens de magnitude para os níveis de expressão do gene.

A análise da expressão da *GNMT* dos tecidos do guaranazeiro apresentaram elevados níveis de transcritos quando comparado com os mesmos tecidos das guaranaranas (Tabela 1). Por outro lado, o nível de transcritos na inflorescência do guaranazeiro, nos frutos e nos folíolos expandido da guaranarana - II não apresentaram diferenças estatísticas. O pecíolo e o fruto (15 dias) do guaranazeiro, o folíolo da guaranarana - I, e a inflorescência da guaranarana - II apresentaram números menores de transcritos quando comparados com o padrão. Já os frutos com 45 e 70 dias do guaranazeiro apresentaram valores superiores de transcritos comparado com o padrão. O fruto do guaranazeiro com 70 dias de desenvolvimento apresentou o maior número de transcritos (3,287). Os valores de transcritos aumentaram gradualmente com o processo de maturação do fruto do guaranazeiro. Este quadro não foi observado nas guaranaranas, que diminuíram os níveis de transcritos com o avanço do processo de maturação.

## 6.5. DISCUSSÃO

O Brasil é o único produtor mundial de guaraná e atende ao mercado nacional e internacional. O guaranazeiro é uma planta em processo de melhoramento que possui propriedades medicinais e estimulantes (O'Dea, 2003; Oliveira et al., 2005; Ray et al., 2005; Campos et al., 2003; Fukumasu et al., 2006a), sendo considerada a maior fonte natural de cafeína na semente seca. *P. cupana* var. *sorbilis* apresenta valores entre 2,7 a 5,8 % de cafeína (Mazzafera 1994a; Baumann et al., 1995); o chá da Índia (*C. sinensis* var. *sinensis*) 2,8% (Nagata and Sakai, 1984); o café (*C. arábica* e *C. canephora*) cerca de 1,0 a 2,4 % (Koshiro et al., 2006); e o cacau (*Theobroma cacao*) 2,5% (Senanayake e Wijesekera, 1971; Athayde et al., 2000).

A identificação no fruto do guaranazeiro de um elevado número de EST's que codificam N-metiltransferase da via de biossíntese da cafeína é coerente com o alto conteúdo de cafeína na planta (Tabela 1). A comparação da *GNMT* com outras N-metiltrasferase está sumarizada na Figura 2. A análise da sequência parcial da *GNMT* sugere que este gene é bastante similar ao gene da *CS* do cacau, do chá e do café. Os resultados foram, também, consistentes com o relacionamento taxonômico das espécies analisadas, sendo a *GNMT* mais similar a enzima do cacau (cafeína sintase classe I).

Os quatro *motifs* caracterizados e conservados [(A, B` e C) e YFFF] na *GNMT* podem ter papéis importantes tanto na reação catalítica de transferência de *N*-metil, quanto no posicionamento da *N*-metilação do anel da purina (Anaya et al., 2006). Acredita-se que o *motif* A seja um sitio de ligação de SAM (Kagan e Clarke, 1994; Joshi e Chiang, 1998), que comumente inicia um importante papel na síntese de cafeína como doador de metil (Kishiishi et al., 2001). Yoneyama et al. (2006)

reportaram duas classes de genes envolvidos na síntese da cafeína. As enzimas da classe I são genes que participam da síntese da cafeína e teobromina, enquanto que as funções das enzimas da classe II ainda não foram reportadas (Ishida et al., 2009). A *GNMT* deve pertencer a classe I, uma vez que apresentou alta similaridade com a CS classe do cacau.

Um protocolo foi estabelecido para quantificação do número de cópias no genoma e da expressão em nível transcricional do gene *GNMT* por PCR em tempo real. O elevado número de cópias do gene *GNMT* no genoma do guaranazeiro em relação as de guaranaranas (espécimes selvagens) já era esperado, visto que o guaranazeiro apresenta elevada ploidia, com 210 cromossomos, estruturados em um conjunto hexaplóide e um grupo tetraplóide de cromossomos (Freitas et al., 2007), enquanto *Paullinia* tem 24 cromossomos como número base.

Os níveis de expressão de transcritos da via da biossíntese da *cafeína* do guaranazeiro foi quantificado nas folhas, na inflorescência e nos frutos de diferentes estágios de desenvolvimento, tanto no guaranazeiro, quanto nos espécimes selvagens (guaranaranas) por qRT PCR. A atividade de biossíntese da cafeína depende acima de tudo, da quantidade e atividade das enzimas, da disponibilidade de substratos, tais como o xantossina e o SAM, do estágio de desenvolvimento e do tecido, e da variação sazonal (Kato et al., 2000). Os transcritos de *GNMT* foram detectados em todos os tecidos usados no experimento, sendo a expressão gênica mais marcante nos frutos, seguida pelo folíolo, inflorescência e por fim no pecíolo. Estes resultados foram similares tanto no guaranazeiro quanto nas guaranaranas. No gênero *Paullinia*, os alcalóides purinas são preferencialmente alocados no fruto e outros tecidos mais velhos, e não em brotações e tecidos novos como na planta de chá (Weckerle et al., 2003). Estudos da expressão dos genes que codificam a CS no

chá (*TCS1* e *TCS2*) nas folhas novas e maduras, caule e raízes de mudas de chá (*Camellia sinensis*), com quatro meses, apresentaram quantidades mais elevadas de transcritos de *TCS1* e *TCS2* nas folhas novas do que em outras partes da planta (Li *et al.*, 2008). A rota biossintética da cafeína do café, envolvendo a adenina e guanina foi encontrada apenas nas folhas jovens, mas a conversão de teobromina e cafeína foi encontrada em folhas maduras e velhas (Ashihara *et al.*, 1996). Por este motivo, a atividade da CS parece estar presente nas folhas de café até mesmo após a maturação. Isto é diferente das folhas de chá, em que a atividade da CS desapareceu após completo desenvolvimento das folhas (Fujimori *et al.*, 1991).

Os níveis de expressão da *GNMT*, possivelmente, relacionada com a síntese de cafeína no guaranazeiro e guaranaranas foram mais elevados no endosperma, assim como no café. Os elevados níveis de expressão quantificados nos frutos do guaranazeiro apresentaram padrões sincronizados de aumento da expressão, acompanhando o desenvolvimento dos frutos. No endosperma do café os genes que participam da síntese da *cafeína* também apresentaram semelhantes padrões sincronizados de expressão, com alta atividade de síntese de cafeína nas sementes em desenvolvimento (Mazzafera *et al.*, 1991; Mazzafera *et al.*, 1994).

No tecido foliar do guaranazeiro e das guaranaranas o índice de expressão da *GNMT* não apresentou diferença estatística dos níveis de expressão da inflorescência, embora em ambas as espécies os valores estimados nos tecidos foliares foram duas vezes maiores. Nos estudos fitoquímicos foi reportado que o folíolo do guaranazeiro apresenta maior acúmulo de teobromina (0,139 % do peso seco de teobromina) (Weckerle *et al.*, 2003). No café, as folhas também apresentaram elevada atividade de síntese de teobromina (*teobromina sintase*) e reduzida atividade da CS (Mizuno *et al.*, 2003a).

Os valores observados do conteúdo de alcalóides purinas nas folhas (1,27 %), no pecíolo (0,14 %), inflorescência (0,160 %) e no fruto (4,55 %) são correspondentes as taxas de expressão da *GNMT* nos respectivos órgãos/tecidos do guaranazeiro (Weckerle et al., 2003). As guaranaranas apresentaram baixos níveis de expressão de transcritos da *GNMT* comparado com guaranazeiro. Uma das possíveis explicações para este fato poderia ser devido ao menor número de cópias do gene *GNMT* no genoma das guaranaranas.

O alto número de EST`s identificadas no transcriptoma e o elevado nível de transcritos de *GNMT* detectados no fruto do guaranazeiro podem ser consequência do grande número de genes da *GNMT* e reponsável pelo alto teor de cafeína acumulado.

No guaranazeiro, dando continuidade a este trabalho, há necessidade de obter-se a sequência completa dos cDNAs (mRNAs) das *GNMTs*, compará-las *in silico* com as sequências desse tipo de enzima das outras espécies e, com isso, inferir o número e classe dessas enzimas do guaranazeiro. Por outro lado, pesquisas referentes ao isolamento e caracterização bioquímca dessas enzimas são necessárias para compreensão mais precisa da(s) via(s) metabólica(s) de síntese da cafeína e dos outros alcalóides púricos do guaranazeiro. Estudos de biologia molecular envolvendo a clonagem e seqüenciamento dos genes da *GNMTs* e estudos das forças de seus promotores nos diferentes tecidos do guaranazeiro também serão de relevância para a melhor compreensão da atividade de síntese de cafeína nos diferentes órgãos/tecidos do guaranazeiro.

## 7. Referências

- Anaya AL, Cruz-Ortega R, Waller GR. 2006. Metabolism and Ecology of Purine Alkaloids. *Frontiers in Bioscience* 11:2354-2370.
- Andersen T, Fogh J. 2001. Weight loss and delayed gastric emptying following a South American herbal preparation in overweight patients. *Journal of Human Nutrition and Dietetics* 14:243–250.
- Angelo PCS, Nunes-Silva CG, Brígido MM, Azevedo JS, Assunção EN, Sousa ARB, Patrício FJB, Rego MM, Peixoto JCC, Oliveira-Junior WP, Freitas DV, Almeida ERP, Viana AMHA, Souza AFPN, Andrade EV, Acosta POA, Batista JS, Walter MEMT, Leomil L, Anjos DAS, Coimbra RCM, Barbosa MHN, Honda E, Pereira SS, Silva A, Pereira JO, Silva ML, Marins M Holanda FJ, Abreu RMM, Pando SC, Gonçalves JFC, Carvalho ML, Leal-Mesquita ERRBP, Silveira MA, BATISTA WC, Atroch AL, França SC, Porto JIR, Schneider MPC, Astolfi-Filho S - Brazilian Amazon Consortium for Genomic Research (REALGENE). 2008. Guarana (*Paullinia cupana* var. *sorbilis*), an anciently consumed stimulant from the Amazon rain forest: the seeded-fruit transcriptome. *Plant Cell Rep* 27:117–124.
- 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. *Phytotherapy Research* 15:416-421.
- Ashihara H. 1993. Purine metabolism and the biosynthesis of caffeine in mate leaves. *Phytochemistry* 33:1427–1430.
- Ashihara H and Kubota H. 1987. Biosynthesis of purine alkaloids in *Camellia* plants. *Plant Cell Physiol* 28:535–539.

- Ashihara H, Monteiro AM, Gillies FM, Crozier A. 1996. Biosynthesis of caffeine in leaves of coffee. *Plant Physiol.* 111: 747-753.
- Ashihara H and Suzuki T. 2004. Distribution and biosynthesis of caffeine in plants. *Front Biosci* 9:1864–1876.
- Athayde ML, Coelho GC, Schenkel EP. 2000. Caffeine and theobromine in epicuticular wax of *Ilex paraguariensis* A. St.-Hil. *Phytochemistry* 55:853-857.
- Basile A, Ferrara L, Del Pezzo M, Mele G, Sorbo S, Bassi P, Montesano D. 2005. Antibacterial and antioxidant activities of ethanol extract from *Paullinia cupana* Mart. *J Ethnopharmacol* 102:32–36 .
- Baumann TW, Schulthess BH, Hanni K. 1995. Guarana (*Paullinia cupana*) rewards seed dispersers without intoxicating them by caffeine. *Phytochemistry* 39:1063–1070.
- Belliardo F, Martelli A, Valle MG. 1985. HPLC determination of caffeine and theophylline in *Paullinia cupana* Kunth (guaraná) and *Cola* spp. samples. *Z. Lebensm. Unters. Forsch.*180: 398-401.
- Bempong DK, Houghton PJ. 1992. Dissolution and absorption of coffee from guarana. *J.Pharm. Pharmacol.* 44:769-771.
- Boozer CN, Nasser JA, Heymsfield SB, Wang V, Chen G, Solomon JL. 2001. An herbal supplement containing Ma Huang-Guarana for weight loss: a randomized, double-blind trial. *International Journal of Obesity and Related Metabolic Disorders* 25:316-324.
- Bustin SA. 2002. Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems. *Journal of Molecular Endocrinology*, 29:23-39.

- Bydlowski SP, D'Amico EA, Chamone DA. 1991. An aqueous extract of guarana (*Paullinia cupana*) decreases platelet thromboxane synthesis. *Brazilian Journal of Medical and Biological Research* 24: 421-424.
- Campos AR, Barros AI, Santos FA, Rao VS. 2003. Guaraná (*Paullinia cupana* Mart.) offers protection against gastric lesions induced by ethanol and indomethacin in rats. *Phytother Res* 17:1199–1202.
- Cánovas FM, Gaudot ED, Recobert G, Jorrin J, Mock HP, Rossignol M. 2004. Plant proteome analysis. *Proteomics* 4: 285-298.
- Cauli O, Morelli M. 2005. Caffeine and the dopaminergic system. *Behav Pharmacol* 16: 63–77.
- Corrêa MPF. 1989. Caracteres quantitativos e qualitativos para descrição morfológica do guaraná (*Paullinia cupana* var. *sorbilis* (Mart.) Ducke). 186p. Tese (Doutorado em Botânica) Instituto Nacional de Pesquisas da Amazônia/Fundação Universidade do Amazonas, Manaus, Amazonas.
- Doyle JJ and Doyle JL. 1987. A rapid DNA isolation procedure to small amounts of fresh leaf tissue. *Phytochemical Bulletin* 19:11-15.
- Erickson HT, Corrêa MPF, Escobar JR. 1984. Guaraná (*Paullinia cupana*) as a commercial crop in Brazilian Amazonia. *Economic Botany* 38: 273-286.
- Espinola EB, Dias RF, Mattei R, Carlini EA. 1997. Pharmacological activity of guarana (*Paullinia cupana* Mart.) in laboratory animals. *Journal of Ethnopharmacology* 55:223–229.
- Felsenstein J. 1998. PHYLIP—phylogeny inference package (Version 3.2). *Cladistics* 5:164–166.
- Ferreira ME and Grattapaglia D. 1998. Introdução ao uso de marcadores moleculares em análise genética. 3. ed. Brasília:EMBRAPA-CENARGEN. 220p.

- Freeman WM, Walker SJ, Vrana KE. 1999. Quantitative RT-PCR: pitfalls and and potential. *Biotechniques* 26:112–122.
- Freitas DV, Carvalho CR, Nascimento-Filho FJ, Astolfi-Filho S. 2007. Karyotype with 210 chromosomes in guaraná (*Paullinia cupana* 'sorbilis'). *Journal of Plant Research* 120:399-404.
- Friedman J, Waller GR. 1983. Caffeine hazards and their prevention in germinating seeds of coffee. *J Chem Ecol* 9:1099–1106.
- Frischknecht PM. 1985. Purine alkaloid formation in buds and developing leaflets of *Coffea arabica*: expression of an optimal defense strategy. *Phytochemistry* 3: 613–616
- Fujimori N, Suzuki T, Ashihara H. 1991. Seasonal variations in biosynthetic capacity for the synthesis of caffeine in tea leaves. *Phytochemistry* 30: 2245-2248.
- Fukuda S, Mikami K, Uji T, Park Eun-Jeong, Ohba T, Asada K, Kitade Y, Endo H, Kato I, Saga N. 2008. Factors influencing efficiency of transient gene expression in the red macrophyte *Porphyra yezoensis*. *Plant Science* 174:329–339.
- 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.
- Gachon C, Mingam A, Charrier B. 2004. Real-time PCR: what relevance to plant studies? *Journal of Experimental Botany* 55:1445–1454.

- Henman AR. 1982. Guarana (*Paullinia cupana* var. *sorbilis*): ecological and social perspectives on an economic plant of the central Amazon basin. *J. Ethnopharmacol.* 6:311–338.
- Hernandez M, Duplan MN, Berthier G, Vaitilingom M, Hauser W, Freyer R, Pla M, Bertheau Y. 2004. Development and comparison of four real-time polymerase chain reaction systems for specific detection and quantification of *Zea mays* L. *Journal of Agricultural and Food Chemistry* 52:4632-4637.
- Hollingsworth RG, Armstrong JW, Campbell E. 2002. Caffeine as a repellent for slugs and snails. *Nature* 417: 915–916.
- Ishida M, Kitao N, Natsu M, Tanikawa K, Kato M. 2009. Occurrence of theobromine synthase genes in purine alkaloid-free species of *Camellia* plants. *Planta*, 229:559–568.
- Joshi CP and Chiang VL. 1998. Conserved sequence motifs in plant S-adenosyl-L-methionine-dependent methyltransferases. *Plant Mol Biol* 37:663–674.
- Kagan RM and Clarke S. 1994. Widespread occurrence of three sequence motifs in diverse S-adenosylmethionine-dependent methyltransferases suggests a common structure for these enzymes. *Arch. Biochem. Biophys.* 310:417–427;1994.
- 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, Crozier A, Fujimura T, Ashihara H. 2000. Caffeine synthase gene from tea leaves. *Nature* 406: 956–957.

- 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.
- 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*. *Pharmacology, Biochemistry, and Behavior* 79: 401-411.
- Kishiishi C, Kato A, Yama S, Crozier A, Ashihara H. 2001. A new caffeine biosynthetic pathway in tea leaves: utilisation of adenosine released from the S-adenosyl-L-methionine cycle, *FEBS Lett.* 499:50–54.
- Lee C, Kim J, Shin SG, Hwang S. 2006. Absolute and relative QPCR quantification of plasmid copy number in *Escherichia coli*. *Journal of Biotechnology* 123:273-280.
- Li Y, Ogita S, Keya CA, Ashihara H. 2008. Expression of caffeine biosynthesis genes in tea (*Camellia sinensis*). *Journal of Biosciences* 63:267–270.
- Lleras E, Leite AMC, Scariot AS, de Sa Brandao JE. 1992. Final Report to UN Food and Agriculture Organization (FAO). Brasília, 65pp.
- Lleras E. 1984. Considerações sobre a distribuição geográfica e taxonomia do guarana (*Paullinia cupana var sorbilis*) e taxas afins na Amazônia. In: Simpósio Brasileiro do guaraná. I. Anais EMBRAPA/UEPAE – Manaus.
- Mattei R, Dias RF, Espinola EB *et al.* 1998. Guarana (*Paullinia cupana*): toxic behavioral effects in laboratory animals and antioxidant activity in vitro. *J Ethnopharmacol* 60:111-116.

- Mazzafera P, Wingsle G, Olsson O, Sandberg G. 1994. S-adenosylmethionine: Theobromine 1-N-methyltransferase an enzyme catalysing the synthesis of caffeine in coffee. *Phytochemistry* 37:1577-1584.
- Mazzafera P. 1991. Trigonelline in coffee. *Phytochemistry* 30:2309–2310.
- McCarthy AA, McCarthy JG. 2007. The structure of two N-methyltransferases from the caffeine biosynthetic pathway. *Plant Physiol* 144:879–889.
- Mizuno K, Kato M, Irino F, Yoneyama N, Fujimura T, Ashihara H. 2003a. The first committed step reaction of caffeine biosynthesis: 7-methylxanthosine synthase is closely homologous to caffeine synthases in coffee (*Coffea arabica* L.). *FEBS Lett* 547: 56–60.
- Mizuno K, Okuda A, Kato M, Yoneyama N, Tanaka H, Ashihara H, Fujimura T. 2003b. Isolation of a new dual-functional caffeine synthase gene encoding an enzyme for the conversion of 7-methylxanthine to caffeine from coffee (*Coffea arabica* L.). *FEBS Lett* 534: 75–81.
- Mohanpuria P, Kumar V, Joshi R, Gulati A, Singh P; Ahuja SK Y. 2009. Caffeine Biosynthesis and Degradation in Tea [*Camellia sinensis* (L.) O. Kuntze] is under Developmental and Seasonal Regulation. *Mol Biotechnol* 43:104–111.
- Nagata T and Sakai S. 1985. Purine base pattern of *Camellia irrawadiensis*. *Phytochemistry* 24:2271-2272.
- O’Dea JA. 2003. Consumption of nutritional supplements among adolescents: usage and perceived benefits. *Health Educ Res* 18:98–107.
- Ogita S, Uefuji H, Morelli M, Sano H. 2005. Metabolic engineering of caffeine production. *Plant Biotechnol* 22: 461–468.
- 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
- Osborn TC, Pires JC, Birchler JA, Auger DL, Chen ZJ, et al. 2003. Understanding mechanisms of novel gene expression in polyploids. *Trends Genet* 19:141–147.
- Pesquet E, Barbier O, Ranocha P, Jauneau A, Goffner D. 2004. Multiple gene detection by in situ RT-PCR in isolated plant cells and tissues. *The Plant Journal* 39: 947–959.
- Pfaffi MW, Horgan GW, Dempfle L. 2002. Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *Nucleic Acids Res.* 30: e36.
- Radkofer L. 1895. Monographie der Sapindaceen-Gattung *Paullinia*. Verlagder Konigliche Akademie, Munchen.
- Raeymaekers L. 2000. Basic principles of quantitative PCR. *Mol. Biotechnol.* 15: 115–122.
- 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
- Senanayake UM and Wijesekera ROB. 1971. Theobromine and caffeine content of the cocoa bean during its growth. *J Sci Food Agric* 22:262-263.
- Sousa NR. 2003. Variabilidade genética e estimativas de parâmetros genéticos em germoplasma de guaranazeiro. Lavras: UFLA. 99p. il. (Tese – Doutorado em Genética e Melhoramento de Plantas).
- Suzuki T. 1972. The participation of S-adenosylmethionine in the biosynthesis of caffeine in the tea plants. *FEBS Lett* 24:18–20.

- Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTALW: 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.
- Uefuji H, Shinjiro O, 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.

## 8. CONSIDERAÇÕES FINAIS E PERSPECTIVAS

Atualmente, o guaranazeiro (*Paullinia cupana*) é, indubitavelmente, uma das espécies com características medicinais que tem despertado interesse mundial. Importantes propriedades fitoterápicas foram descritas nesta espécie, que vão desde um simples estimulante (físicos e cognitivos) até antitumorais (antihepatocarcinogênico). Devido a sua grande importância socioeconômica, o guaranazeiro tem atraído a atenção tanto da comunidade técnico-científica quanto da empresarial. A possibilidade de se obterem novas moléculas do guaranazeiro, ou mesmo, em grandes quantidades, purificadas, destinadas ao desenvolvimento de bioprodutos torna a biotecnologia um instrumento imprescindível para o desenvolvimento econômico-social da região Amazônica. Este trabalho teve como objetivo principal ampliar o conhecimento sobre as informações do ponto de vista genético e molecular desta espécie, pois genomas decifrados são abundantes fontes de genes e informações de seus produtos de expressão com valor biotecnológico.

As análises do tamanho do genoma e da organização cromossômica do guaranazeiro são importantes: 1) para avaliar a cobertura de uma biblioteca genômica; 2) na estimativa do número de cópias de um gene no genoma; 3) no desenvolvimento de estratégias futuras para a clonagem e expressão de genes baseado no seu mapeamento; 4) no apoio a programas de melhoramento genético.

A análise do cariótipo de outras espécies de *Paullinia*, especialmente as que poderiam ter contribuído para a origem da variedade *sorbilis*, bem como da *cupana* observada por Humboldt e Bonpland, poderia contribuir para a compreensão da origem evolutiva do guaranazeiro cultivado. Estes resultados podem, ainda, permitir

outros tipos de análises genético-moleculares para apoiar o programa de melhoramento.

Entre as estratégias oferecidas pela biotecnologia para a produção de compostos bioativos oriundos de espécies amazônicas, a genômica funcional tem sido aplicada com sucesso na identificação de genes que codificam para as enzimas envolvidas no metabolismo secundário. Muitas vias metabólicas que conduzem à biossíntese de compostos de elevado interesse econômico são freqüentemente mal caracterizadas, principalmente em função da limitada informação genômica das plantas. A identificação dos genes expressos no fruto do guaranazeiro, como os genes que codificam enzimas envolvidas na biossíntese da cafeína, resultou na geração de informações, seja do genoma ou do transcriptoma. Este estudo tem permitido acelerar as investigações que buscam compreender o metabolismo celular em tecidos especializados ou o seu papel no desenvolvimento do guaranazeiro. O desenvolvimento prévio da biblioteca de cDNA permite, ainda, a utilização de tecnologias como os *microarrays* para a análise da alteração na expressão gênica relacionada à produção de metabólitos secundários, e a estresse biótico e abiótico em resposta a vários estímulos. As análises da função dos genes caracterizados no fruto do guaranazeiro necessitam, conseqüentemente, serem complementados pela análise temporal e espacial dos padrões de sua expressão.

Em seu conjunto, as análises indicam que a sequência escolhida (*GNMT*) é homóloga as N-metiltransferases da via biossintética da cafeína de outras espécies. A informação mais conclusiva sobre os níveis de expressão gênica da N-metiltransferase do guaranazeiro envolvida na biossíntese de cafeína foi obtida por meio da análise de mudanças quantitativas do mensageiro RNAs em diferentes órgãos/tecidos. A determinação quantitativa dos níveis de transcritos da *GNMT*

permitiu conhecer que o gene é diferencialmente expresso, sendo marcante sua atividade no fruto em desenvolvimento. O conhecimento da localização celular da *GNMT* foi essencial, a fim de avaliar a função do gene em um contexto integrado. Diante dos resultados parciais faz-se necessário: (1) o seqüenciamento completo do(s) cDNA das N-metiltransferase; (2) o desenvolvimento de trabalhos que decifrem de forma mais detalhada a expressão gênica e o estudo dos padrões de expressão espacial de todos os genes envolvidos na biossíntese da cafeína; (3) expressão funcional das N-metiltransferases do guaranazeiro em plantas modelos, como *Arabidopsis* para elucidar suas funções biológicas.