International Coffee Genomics Network (ICGN)
Report 7th Coffee Genomics Workshop held at the XXII Plant and Animal Genome (PAG) Meeting San Diego, California January 11-15, 2014

Coffee Genomics Workshop Speakers

1. **Alexis Dereeper,** IRD, UMR RPB- France. The Coffee Genome Hub.
2. **Pablo Benavides**, CENICAFE, Colombia. The Genome of the Coffee Berry Borer, *Hypothenemus hampei*, the Major Insect Pest of Coffee Worldwide.
3. **Romain Guyot,** IRD-France. Transposable Elements in the Robusta Coffee Genome (*Coffea canephora*).
4. **Carmenza Góngora**, CENICAFE, Colombia. Induction of Defense Genes in Coffee Fruits from Different *Coffea* Species in Response to Attack by the Coffee Berry Borer, *Hypothenemus hampei.*
5. **Luiz Filipe Pereira** and **Douglas Domingues**, EMBRAPA/ IAPAR, Brazil. Diterpenes and Transcriptional Profile of CYPs genes during *Coffea arabica* Fruit Development.
6. **Victor Albert,** University of Buffalo, USA. The Role of Lineage-Specific Gene Family Expansions in Coffee Adaptation: The Case of N-Methyltransferases Involved in Caffeine Biosynthesis.

Coffee related abstracts presented at other PAG workshops- Bioinformatics

7. **Matthieu Rouard, Bioversity International, Montpellier**, France. The South Green Bioinformatics Platform.

Coffee Genomics Workshop at PAG

The Plant and Animal Genome (PAG) meeting is the largest international scientific conference reporting on animal and plant genomics developments in the world, this year with 3,397 participants from 62 countries. For those interested in participating in future meetings see http://www.intlpag.org. The XXIII Plant & Animal Genome Conference will be held in San Diego, January 10-14, 2015.

More than 50 scientists participated in our 7th coffee genomics workshop held as part of the PAG Meeting in San Diego on January 12, 2014. The co-organizers of the workshop, Marcela Yepes (Cornell University, my11@cornell.edu), Philippe Lashermes (IRD-CIRAD, France, philippe.lashermes@ird.fr), and Rod Wing (University of Arizona) thank the speakers for their participation and contributions. Abstracts of workshop and poster presentations on coffee are included as an appendix at the end of this report. The 8th Coffee Genomics Workshop will be held January 11, 2015 as part of the XXIII PAG meeting in San Diego, January 10-14, 2015. Please contact one of the organizers if interested in presenting a talk or poster, or with suggestions for new topics for workshop presentations or for round table discussion at the ICGN meeting. The coffee genomics workshop is an excellent opportunity to present advances in coffee genomics research to the International Plant and Animal Genomics Community and is helping our community explore new collaborations as well as funding opportunities.

ICGN survey and collaboration with the International Coffee Organization

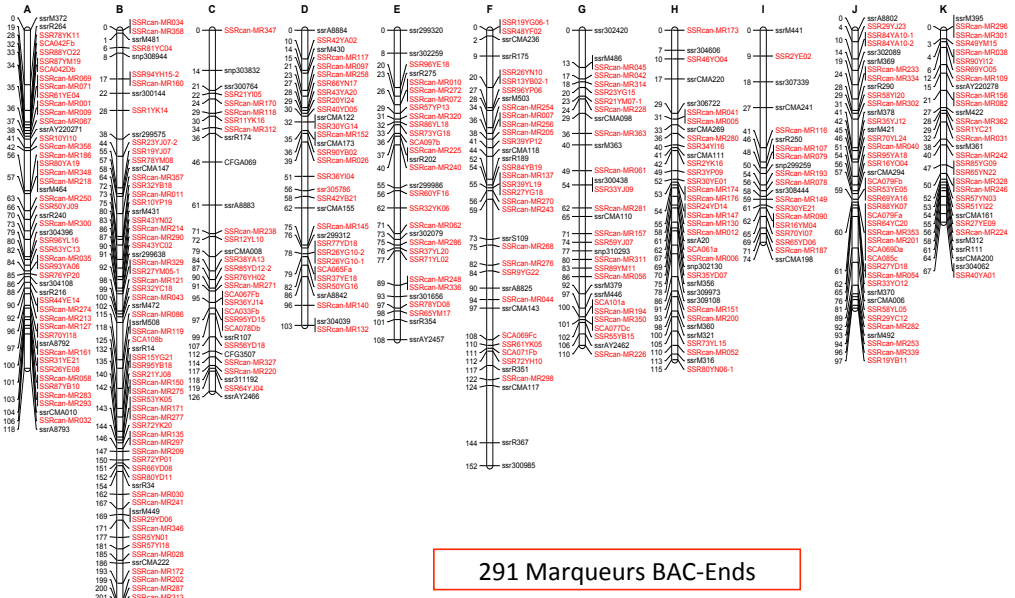
ICGN will conduct a survey in 2014 to help us update our mailing list, identify future priority projects for the community as well as new leadership to help secure funding for new proposals. ICGN members will be asked to help us contribute to this effort by completing and submitting the survey available at our www site (http://www.coffeegenome.org). Survey results will be discussed at the next ICGN meeting held in conjunction with the International Coffee Science ASIC 2014 Colombia meeting (http://www.asic2014colombia.org). We appreciate your feedback.

As the first *de novo* coffee genome reference and assembly become available (see report on the status of on going projects below), we would like to take advantage of the momentum to identify new priority projects of interest that ICGN can develop as a community to help mine the data generated and develop innovative tools and advanced resources in coffee genomics to address challenging issues for our community such as climate change adaptation and sustainability that could be accelerated with transforming genomic technologies and strategies. The African Coffee Research Network (ACRN) joined ICGN in 2011 as an institutional member, and its Director of Research and Development, Dr. Bayeta Bellachew helped us conduct the ICGN survey among ACNR members at several Coffee Research Institutions in Africa. We received through ACRN responses from scientists and scientific groups from the following countries: Ethiopia, Kenya, Rwanda, Uganda and Ghana with strong interest to work with ICGN on a global initiative to develop advanced genomic tools to speed up diversity characterization, enhanced utilization and conservation of *Coffea* germplasm in the context of climate change. In addition with support from the International Coffee Organization (ICO), ICO member countries have been contacted to discuss possible interest in developing a global initiative in collaboration with ICGN/ICO aiming at improving conservation and characterization of the world coffee gene pool for varietal development in a world of changing farming systems and climate. Other ICO member countries that have expressed strong interest in working on an ICGN/ICO collaborative proposal include, for Europe: France (IRD-CIRAD); for Latin America: Brazil, Colombia, Guatemala, Costa Rica, Mexico; for Africa: Cote D'Ivoire, Ethiopia, Kenya, Malawi, as well as the Inter-African Coffee Organization; and for Asia: India and Vietnam.

ICGN is grateful for the invitation by the ICO Executive Director Dr. Robeiro Oliveira Silva to participate as an observer in the ICO Council meetings in 2014, and we are looking forward to working closely with ICO officials on the preparation and submission of a first ICGN/ICO proposal, and to explore potential sources of finance for such joint initiative. Support from ICO and the private sector will be key for ICGN to secure future funding for diversity conservation efforts in *Coffea* with a broader funding base, and to promote coffee genomics research for coffee improvement targeting priority traits for different regions as well as for the coffee industry. Capacity building in developing countries to participate in coffee genomic research can be supported through ICGN networking to help us secure international funding for those efforts.

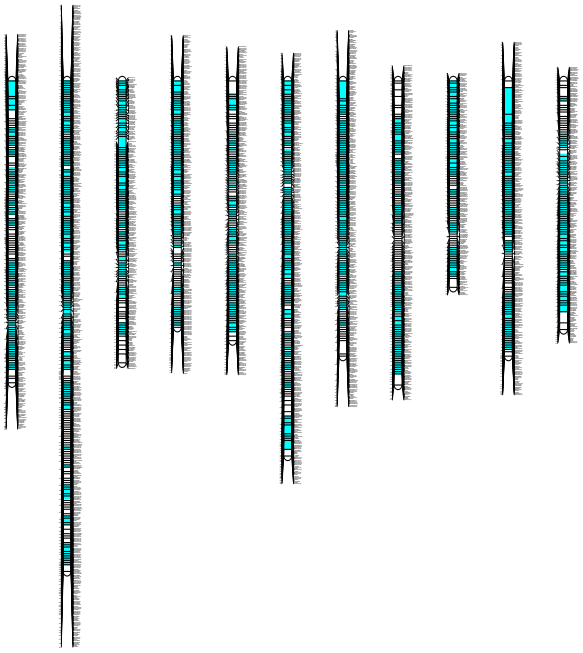
Update Status High-Density Mapping of the diploid species *Coffea canephora*

A high-density reference genetic map for *Coffea canephora* Pierre was constructed in collaboration with Nestlé R&D Centre and the Indonesian Coffee and Cocoa Research Institute. The population mapped was from a cross between two highly heterozygous genotypes, a Congolese group genotype (BP409) and a Congolese-Guinean hybrid parent (Q121). The segregating population is composed of 93 F1 individuals. DNA from the two parental clones and the segregating progeny were distributed to several ICGN members (on request). First, a high-density genetic map was constructed including 1481 loci covering 1400 cM markers, with a third of the SSR markers derived from BAC end sequences (see Figure below). The first set of markers mapped included: 360 RFLPs, >890 SSRs, and 213 SNPs, and were mined from genomic or EST libraries from different institutes (IRD, CIRAD, Trieste University, Cornell University, CENICAFE, and Nestlé).



As a second approach, Restriction Associated DNA sequencing (RADseq), which enables synchronous SNP marker discovery and genotyping using massively parallel sequencing, was used. The RAD libraries were made from digestion of DNA using two restriction enzymes, *Nsi*I (6 base cutter) and *Mse*I (4 base cutter). The fragments (150 - 500 bp) were selected to ligate to two adaptors, and one of them with tag for each progeny. Equal amount of amplicons from each individual were pooled to make Illumina RNAseq libraries with individual tags for each library. Co-segregating markers within 50 kb region (< 1 cM) based on the aligned template scaffold were sorted as bin. One marker from each bin was selected for mapping. The linkage analysis and map construction were performed using JoinMap software version 4.1 using LOD threshold of 5 and Kosambi's function to calculate genetic distance between two loci. The Robusta consensus genetic map was built using the F2 segregating loci as anchor markers in order to merge the two homologous parental linkage groups. Using RAD sequence data from the segregating population previously selected 1747 RADseq markers were added. The final high density Robusta map comprises 3230 loci, genetic size 1471 cM (1cM ~ 500 Kb), with an average density close to one marker every 220 Kb. The F1 high density genetic map will facilitate comparative genomic studies based on synteny, and provided the opportunity for anchoring and ordering the numerous scaffolds arising from the *Coffea canephora* genome sequencing (see report below). So far, the DNA sequences (scaffolds) anchored are covering approximately 75% of the genetic map (1023 cM). On going mapping efforts are focusing on the identification and mapping of SSRs from the *C. canephora* sequence scaffolds that are not or are insufficiently anchored. Both the high-density genetic map and the marker information will be freely available on a dedicated web-site once the construction of the map is completed. Please send information or comments to Dominique Crouzillat, Nestlé (dominique.crouzillat@rdto.nestle.com), Philippe Lashermes (philippe.lashermes@ird.fr), or Ray Ming (rming@life.illinois.edu).

The international *C. canephora* high density map is a highly valuable resource for different applications including transposition to other mapping populations, as genetic framework that can be used for various QTL studies, as well as genome structure comparisons. RAD sequencing is a powerful strategy for genotyping in coffee to provide access to high-throughput SNP detection.



Update status of the *Coffea canephora* genome sequencing project

With funding from the Agence Nationale de la Recherche (ANR; Genoplante ANR-08-GENM-022-001), France, several Institutes (Genoscope-CEA, IRD and CIRAD) combined their scientific resources and expertise to sequence, assemble, and annotate the entire genome of *C. canephora*. Additional partners include several ICGN members (EMBRAPA/Brazil, ENEA/Italy, University of Trieste/Italy, University of Queensland/Australia, CCRI/India, University of Illinois, Urbana/USA, Hawaii Agriculture Research Center HARC/USA, SUNY Buffalo/USA, University of Ottawa/Canada). A community effort for genome annotation is on going. The *C. canephora* genome consists of 11 chromosomes, is about 710 Mb in size, and was sequenced *de novo* with deep coverage using different sequencing platforms. Genoscope lead the sequencing and assembly of the *C. canephora* genome. Patrick Wincker, Head of Sequencing and Coordinator of Eukaryote Annotation and Analysis at Genoscope, presented the sequencing strategy and the status of the project during the 2011 coffee genomics workshop. In 2013 France Denoeud from Genoscope presented in our coffee genomics workshop an update on the first genome assembly, and this year (2014) Alexis Dereeper presented the **Coffee Genome Hub**, an integrative genome information system accessible through the South Green Bioinformatics Platform, developed to provide centralized access to all the coffee scientific community of the full *C. canephora* genome sequence, as well as genomics, genetics and breeding data and analysis tools to facilitate basic, translational and applied research in coffee. **Once the manuscript for the sequencing of the *C. canephora* is published in 2014, the community should be able to access this public database at http://coffee-genome.org.**

Summary of the sequencing strategy used for C. canephora:

C. canephora is one of the ancestral progenitors of the widely cultivated, *C. arabica,* a recent allotetraploid species formed from the merger of the diploid species *C. canephora* and *C. eugenioides*. The accession DH200-94, a doubled haploid genotype was selected for sequencing because of its homozygous nature to facilitate genome assembly. *De novo* genome sequencing with deep coverage was performed using both 454 Roche and Illumina next generation sequencing technologies. Direct whole genome shotgun (WGS) sequencing and paired-end sequencing of large insert libraries 8kb and 20 kb insert libraries was conducted. Furthermore, 73,728 BAC clones from two *C. canephora* BAC libraries *Hin*d III and *Bst*YI constructed in collaboration with Rod Wing at University of Arizona were BAC-end sequenced using Sanger technology. Average inserts for each BAC library were 166 and 121 kb in size, with 36,864 BAC clones per library for an estimated coverage of ~8.6X and 6.3 X per library, respectively. Both *C. canephora* BAC libraries are publicly available at Arizona Genomics Institute Resource Center (http://www.genome.arizona.edu.orders). BAC end sequences (BES) are also publicly available and were deposited in EMBL-EBI Bank (accession numbers FO535330, FO538768 to FO624989, and FO624992 to FO680656) (A. Dereeper *et al*. 2013. BAC-end sequences analysis provides first insights into coffee (*Coffea canephora* P.) genome composition and evolution. Plant Molecular Biology 83: 177-189 http://www.researchgate.net/publication/257120929_Dereeper_et_al_canephora_BAC_ends). The genome sequencing data generated for the *C. canephora* genome assembly included:

Roche/454 Titanium:

Whole Genome Shotgun (WGS) 454 sequencing: 28.9X coverage (assuming a genome size of 710 Mb) including:
23X single end 454 Titanium reads
– Reads single end: 14.83X , Mean size: 359 bp
– Long reads single end: 8.24X , Mean size : 462 bp
5.8X Paired-end sequencing of long insert libraries (8 and 20 Kb): 5.8X (2.2X for 8kb, 3.6X for 20kb), Mean Size: 252 bp.

Sanger end sequencing of Bacterial artificial chromosome (BACs):

Two BAC libraries (*Hind* III and *Bst*YI) were constructed in collaboration with Rod Wing at Arizona Genomics Institute. The BAC libraries have 73,728 clones (>11X coverage).
Sanger BAC-end sequencing: 131,412 BES were generated (73,728 BAC clones x 2 ends 5' and 3'): 0,27X
Mean BAC insert size : 135 Kb, range: 63,2Kb < insert < 253,6 Kb

Illumina sequencing was done at deep coverage (~70X) to correct 454 sequencing errors.

Single reads coverage 7.3X: read size 76 bp (4.8X) and read size 150 bp (2.5X)
Paired end reads coverage 62.4X: read size 76 bp (42.4X) and read size 108 bp (20X)

Assembler used Newbler and assembly statistics for the first assembly (data from Genoscope):

No. of Scaffolds 13,345
Size: 569 Mb (80% estimated genome size of 710 Mb)
Coverture 28.87 X (454/Sanger) and Illumina 69.7 X
N50: 1260 Kb (108 scaffolds)
N80: 65 kb (635 scaf)
Largest scaffold: 9.0 Mb

Automatic annotation by reconciliating proteic and EST hits, RNA-Seq data integrated into G-Mo-R-Se models, and *ab initio* predictions, was performed and provided 25,574 genes (with an average of 5.1 exons per gene). This relatively low gene number compared to other plant lineages reflects the fact that the *Coffea* lineage was less frequently subject to whole genome duplications

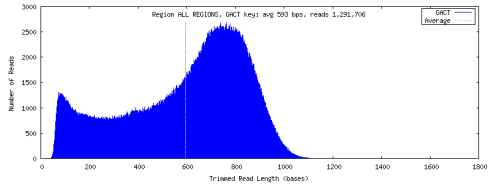
than other sequenced dicotyledon lineages. Its slow rate of genome evolution makes it a good model for paleogenomics in Asterids.

Update status of the *Coffea eugenioides* genome sequencing

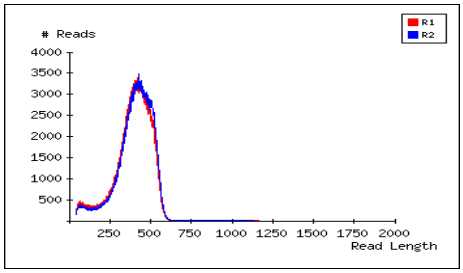
This project is funded by the InterAmerican Development Bank (FONTAGRO/SECCI). Genome sequencing was started at the end of 2012 and continued in 2013. The project is being developed collaboratively by the Colombian National Coffee Research Center (CENICAFE), and Cornell University. Funding for this project was secured jointly through a proposal prepared and submitted by Cornell University and CENICAFE.

We are mimicking the strategy used for the *C. canephora* sequencing to generate a high quality reference assembly for *C. eugenioides* using mixed next generation sequencing platforms: Roche 454 FLX+ and Illumina HiSeq 2500. We have sequenced a whole genome shotgun (WGS) library (fragment size >1,100 bases and <2,000 bases) using 454 FLX+ single end reads with mode length of 763 bases to generate a total of 6,082,341,937 bases for an estimated coverage of 9.2X for the *C. eugenioides* genome (~estimated genome size of 660 Mb*)*. See below quality control results of a typical 454 FLX+run:

Region	Passed Filter	Total Bases	Reads ≥ 1000 bases (% Passed Filter)	Read Length (Bases)				
				Mean	Median	Mode	Min	Max
1	634,065	373,033,536	4,712 (0.74%)	588.3	653	763	40	1,735
2	657,641	392,830,440	3,723 (0.57%)	597.3	660	782	40	1,692
ALL REGIONS	1,291,706	765,863,976	8,435 (0.65%)	592.9	657	763	40	1,735



We have completed sequencing of twelve 20 Kb long insert libraries (3.1X coverage) using paired end sequencing and Roche 454 FLX Titanium. Below is a graph of the paired end statistics profile for a run: Read length distribution of high quality reads. R1 = Coffee 20kb 1-1, R2 = Coffee 20kb 1-1. Linker positive displays statistics of reads with paired end linker sequence 71.47% and 73.62%. Linker Negative refers to reads with no paired end linker sequence 28.53% and 26.38%.



The overall data generated for the twelve 20 Kb-insert libraries paired end-sequenced using Roche 454 FLX Titanium includes:

Paired end reads 57% 5,876,463 reads x 170 bases x 2= 1,997,997,420 bases/660,000,000 = **3.03 X Coverage**
Non paired end reads 43% 4,225,283 reads x 298 bases = 1,259,134,334 bases/660,000,000 = **1.91 X Coverage**

We had technical issues with the construction of the 40 Kb long insert libraries for paired end sequencing using Roche 454 FLX Titanium. With the phasing out of 454 by Roche, we will use other state-of-the-art sequencing platforms to complete coverage. A first genome assembly for *C. eugenioides* was done by Roche in the Spring 2014 using the 454 data generated for the WGS library and the twelve 20 Kb insert libraries. With our long-read approach using Roche 454 FLX+ 1.0 Kb reads vs. Titanium, we increased by 50% the number of contigs (>500 bp) in the *C. eugenioides* first draft assembly (146,520) compared to the *C. canephora* assembly (96,182). We also had a larger number of scafolds (502) with N50 209.9 Kb. In order to connect and reduce the overall number of contigs and scaffolds in the assembly, as well as to increase the overall percent of genome assembled (currently 77-80%), we will continue sequencing *C. eugenioides* using other state-of-the-art sequencing platforms including PACBio and Illumina to complete deep coverage and generate a high quality assembly. In parallel, we are also doing transcriptome sequencing of *C. eugenioides* to validate assembly and annotate the genome.

Once the reference genomes of the diploid species *Coffea canephora* and *Coffea eugenioides* (parental diploid ancestors of the allotetraploid species *Coffea arabica)* become available in 2014, they will serve as frames for sequencing and assembly of *C. arabica*, the major cultivated coffee species in the world. By sequencing the coffee genome, we are building a solid foundation for deciphering the genetic and molecular bases of important biological traits in coffee that are relevant to growers, processors, and consumers. This knowledge will be fundamental to allow efficient use and conservation of coffee genetic resources, and for the development of improved cultivars in terms of quality and reduced economic and environmental costs, as well as, to advance efforts to adapt the crop to climate change. To ensure full benefit from the generated coffee genomic sequences and resources by the coffee sector, ICGN is exploring additional funding from International Funding Agencies for the development of friendly end-user tools as well as to organize training courses to promote community annotation efforts.

Acknowledgements:

ICGN is particularly grateful to all our workshops speakers who kindly accepted our invitation to participate in our 7th ICGN coffee genomics workshop. Abstracts of their presentations are enclosed below in the appendix, as well as posters presented on coffee genomics.

Upcoming Meetings of interest to the ICGN community

- 25th ASIC International Conference on Coffee Science, Colombia, September 8-13, 2014 http://www.asic2014colombia.org
- 11th Solanaceae Genomics Network SOL meeting, Brazil, November 2-6, 2014 http://solgenomics.net .
- 8th ICGN Coffee Genomics Workshop at XXIII Plant and Animal Genome (PAG) Meeting, San Diego, California, January 10-14, 2015 http://www.intlpag.org/

Pictures coffee genomics workshop participants and ICGN meeting participants:

Pictured ICGN meeting participants (from left to right):
Perla Hamon, IRD, France
Marcela Yepes, Cornell University, USA
Luiz Filipe Pereira, EMBRAPA, Brazil
Lukas Mueller, Boyce Thompson Institute, USA
Suzy Strickler, Boyce Thompson Institute, USA
Noriko Nakamura, Suntory, Japan
Yoshikazu Tanaka, Suntory, Japan
Douglas S. Domingues, IAPAR, Brazil

Other ICGN Coffee Genomics Workshop participants pictured below:

From left to right:
Pablo Benavides, CENICAFE, Colombia
Marcela Yepes, Cornell University, USA
Keithanne Mockaitis, Indiana University, USA
Carmenza Góngora, CENICAFE, Colombia

From left to right:
Marcela Yepes, Cornell University, USA
Keithanne Mockaitis, Indiana University, USA
Susan L. Ulanowicz, Roche, USA
Casey Matthews, Roche, USA

From left to right:
Alexis Dereeper, IRD, UMR RPB- France
Marcela Yepes, Cornell University, USA

Appendix

Abstracts 7th Coffee Genomics Workshop 2014

Workshop Co-Organizers:

Marcela Yepes, Cornell University (my11@cornell.edu)

Philippe Lashermes, L'Institut de Recherche pour le Développement (IRD), France (philippe.lashermes@ ird.fr)

Rod Wing, University of Arizona (rwing@Ag.arizona.edu)

(Program and abstracts also posted at:
https://pag.confex.com/pag/xxii/webprogram/Session2125.html

The Coffee Genome Hub

Alexis Dereeper *IRD, UMR RPB*, Gaëtan Droc *CIRAD, UMR AGAP*, Stéphanie Bocs *CIRAD, UMR AGAP*, Sebastien Ravel *IRD, UMR RPB*, Valentin Guignon *Bioversity International, Commodity Systems & Genetic*, Dominique This *Montpellier SupAgro, UMR AGAP*, Michelle Cotta *CIRAD, UMR AGAP*, Philippe Lashermes *Institut de Recherche pour le Développement*

The whole genome sequence of *Coffea canephora,* the perennial diploid species known as Robusta, has been recently obtained and portends to be published and released in the coming months. In the context of the *C.canephora* genome sequence project and to support post-genomics efforts, we developed the **Coffee Genome Hub**, an integrative genome information system accessible through the South Green Bioinformatics Platform, providing centralized access to coffee community to genomics, genetics and breeding data and analysis tools to facilitate basic, translational and applied research in coffee. Available data are the complete genome sequence of *C.canephora* along with gene structure, gene product information, metabolism, gene families, transcriptomics (ESTs, RNA-Seq), syntenic blocs, genetic markers and genetic maps. The hub provides also generic softwares (e.g. GMOD tools) for easy querying, visualizing and downloading research data. Notably, the Coffee Genome Hub includes a Genome Browser for the *C.canephora* genome sequence, enhanced by a Community Annotation System (CAS) enabling the improvement of automatic gene annotation through an annotation editor and facilitating the study of gene families, and a chromosome viewer displaying the distribution of genomics feature along the chromosomes. In addition, the hub aims at developing interoperability between other existing South Green tools managing Coffee data (phylogenomics resources, SNPs) and/or allowing data analyses (workflow manager).

The Coffee Genome Hub can be accessed at http://coffee-genome.org/.

The Genome of the Coffee Berry Borer, *Hypothenemus hampei*, the Major Insect Pest of Coffee Worldwide

Pablo Benavides *Centro Nacional de Investigaciones de Cafe, CENICAFE*, Lucio Navarro *Purdue University*, Flor Acevedo *Pennsylvania State University*, Ricardo Acuña *Centro Nacional de Investigaciones de Cafe, CENICAFE*, David O'Brochta *University of Maryland*, Stuart Jeffrey *Purdue University*, Jonathan Nuñez *Centro Nacional de Investigaciones de Cafe, CENICAFE*, Erick Hernandez *Centro Nacional de Investigaciones de Cafe, CENICAFE*, William Giraldo *Centro Nacional de Investigaciones de Cafe, CENICAFE*, Marco A. Cristancho *Centro Nacional de Investigaciones de Cafe, CENICAFE*, Marcela Yepes *Cornell University*, Herb Aldwinckle *Cornell University*, Alvaro Gaitan *Centro Nacional de Investigaciones de Cafe, CENICAFE*

Coffee supports the livelihood of millions of families in producing countries, and enriches the life of millions of coffee consumers worldwide. However, the most globally devastating insect pest of coffee: *Hypothenemus hampei*, the Coffee Berry Borer (CBB), threatens the sustainability of the coffee industry in the context of climate change, as increasing temperatures favor higher reproduction rates increasing rapidly insect populations. CBB infests the coffee beans, and the damage caused by the insect during feeding

makes the beans unmarketable impacting directly production and coffee quality. Control measures are hard to implement since the insect spends most of its life cycle inside the coffee berry, and insecticide resistance has been reported. Efforts to control this pest are underway at CENICAFE, the research branch of the Colombian National Coffee Growers Federation. We have sequenced the genome and transcriptome of *H. hampei* using next generation sequencing platforms (454 Roche and Illumina) to develop innovative control strategies that help reduce the impact of climate over CBB population dynamics, while gaining an in depth understanding of the biological basis for this insect's unique life history and habits, to use this knowledge to devise new strategies to control or eradicate this pest. In collaboration with Purdue University, the genome of CBB was assembled using more than 6 million FLX-454 reads obtained from both CBB males and females. The assembly was built with Newbler including single reads from WGS libraries (9.2X average coverage), and pair-end reads of long insert libraries (8Kb and 20Kb). The *H. hampei* genome assembly is 194 Mb, composed of 45,995 contigs (N50: 6.1Kb, ~165.6Mb total) and 9,932 scaffolds (N50: 437.7Kb). Although gene annotation is still in progress, initial analysis predicts ~20,500 genes, a similar number (~20,600) was obtained from whole transcriptome assembly (using RNA-seq), physical mapping using a CBB BAC library is on going. The identification of ubiquitous ultra-conserved and single-copy core genes was conducted, estimating a genome completeness of 95%. Genes involved in CBB metabolism are related to enzyme families involved with arabinoxylans and other metabolism polysaccharides, which possibly had the same evolutive dynamic as mannanases, an enzyme recently described in CBB as a key compound in the quick worldwide dispersal and strong specialization of CBB over *Coffea* berries. Putting together the biology of CBB, its genetics and the fact that the insect does outcross under field conditions, we are exploring control strategies to introduce deleterious genes in established CBB populations, as an autocidal genetic control strategy. Functional genomic strategies to characterize gene candidates have been developed in collaboration with University of Maryland. We have demonstrated that the transposons *Minos* and *piggyBac*are effective transgene vectors in this species. We successfully obtained CBB transgenic individuals using transposon-based germ-line transformation to advance functional genomics studies for candidate CBB control genes.

**Research Co-sponsored by the Colombian National Coffee Growers Federation and the Colombian Ministry of Agriculture.

Transposable Elements in the Robusta Coffee Genome (*Coffea canephora*)

Stephanie Bocs *CIRAD*, Alexis Dereeper *IRD, UMR RPB*, Thomas Gayraud *Institut de Recherche pour le Développement*, Véronique Jamilloux *INRA - URGI*, Philippe Lashermes *Institut de Recherche pour le Développement*, **Romain Guyot** *Institut de Recherche pour le Développement*

Coffee is one of the most important international trade commodities and is ranked as the second most valuable primary commodity exported by southern countries. Two species are mainly used in commercial production: *Coffea arabica*, known as Arabica and *Coffea canephora*, a perennial diploid species known as Robusta. Recently, 54.4 million of Roche 454 sequences, 131,412 Sanger BAC-end sequences and 60X Illumina coverage of the 710 Mb genome of a *C. canephora* Double Haploid accession (DH200-94) were generated, assembled and anchored to a genetic map. The *C. canephora* genome sequence represents a formidable resource to understand the chromosome structure and the genome evolution. It is now well established that plant genomes are dynamic structures submitted to a wide range of modifications via the activity of Transposable Elements (TEs). Transposable elements are mobile sequences that share several key properties such as the ability to move from one chromosome location to another, to amplify their copy number within the host genome and to contribute to the chromosome structure, organization and evolution. Particularly, TEs play a major role in creating structural variation and genetic diversity in plant genomes. Here we present the identification and classification of TEs in the 568 Mb genomic sequences of the *C. canephora* using a combination of *ab initio,* similarity and structure search approaches. We used mainly the REPET package V.2.1-RC (Flutre et al., 2011) to identify, classify and annotate TE. We found that 49% of the genomic sequences are composed of TEs similarly to other sequenced plant genomes such as banana, papaya, castor bean and soybean. Class I LTR retrotransposons represent the vast majority of identified elements, accounting to 42% of the genome assembly. *Gypsy* elements clearly outnumbering *Copia* elements since *Ty3-Gypsy* family covers 24.1% of the genome. Interestingly active non-autonomous LTR

retrotransposons elements were detected and classified into a new subgroup of non–autonomous elements containing a *capsid* domain but lacking the *polyprotein* region. Finally in an attempt to study conservation of LTR retrotransposons between coffee and reference plant genomes, we identified an outstanding conservation of several *Copia* groups across very distantly related plant species, suggesting that conservation of such elements or horizontal transfer events might be more frequent than recognized actually.

Induction of Defense Genes in Coffee Fruits from Different Coffea Species in Response to Attack by the Coffee Berry Borer, *Hypothenemus hampei*

Carmenza E. Góngora *Centro Nacional de Investigaciones de Cafe, CENICAFE*, Eliana Macea *Centro Nacional de Investigaciones de Cafe, CENICAFE*, Ana Maria Castro *Centro Nacional de Investigaciones de Cafe, CENICAFE*, Marco A. Cristancho *Centro Nacional de Investigaciones de Cafe, CENICAFE*, Pablo Benavides *Centro Nacional de Investigaciones de Cafe, CENICAFE*, David W. Galbraith *University of Arizona, BIO5 Institute &*, Cheryl Vanier *University of Nevada*, Marcela Yepes *Cornell University*, Herb Aldwinckle *Cornell University*, Alvaro Gaitan *Centro Nacional de Investigaciones de Cafe, CENICAFE*

Coffee Berry Borer (CBB), *Hypothenemus hampei* (Ferrari) (Coleoptera: *Curculionidae*) is the most important insect pest of coffee worldwide. No sources of natural resistance have been identified in *Coffea* species. However, *Coffea liberica* has shown some level of antibiosis against the insect. In order to understand the interactions between *C. arabica* and CBB, and to characterize the response of coffee plants under CBB attack, as well as, the bases of *C. liberica* antibiosis, functional genomics studies were carried out using cDNA libraries, ESTs, cDNA microarrays, an oligoarray containing 43,000 coffee sequences, and RNAseq with 25 million short reads (25-300bp). The results allowed the comparison of *C. liberica* vs. *C. arabica* fruits responses to insect infestation after 24h and 48h. Out of a set of 2,500 plant sequences that exhibited differential expression under insect attack, 900 were induced in *C. liberica*, at least 2 times more than in *C. arabica.* In order to validate some of the induced genes, quantifications through real-time PCR were done. At least 12 genes showed differential expression and four genes: an isoprene synthase gene, a patatin-like protein gene, a hevein-like protein sequence, and a trypsin inhibitor known also as miraculin-like gene, *CoMir*, were highly upregulated in *C. liberica* at 24 and/or 48 h after insect infestation compared to *C. arabica*. For each gene, further sequence characterization and comparison were carried out between both genotypes. Functional annotation indicate that they participate in separate defense plant processes such as volatiles synthesis, lipid or chitin degradation, and proteinase inhibition, suggesting the activation of different metabolic pathways and plant defense mechanisms in coffee plants in response to insect attack. One of those processes is the methyl-erythritol 4-phosphate (MEP), or non-mevalonate pathway, that leads to the production of isoprene. The effect that isoprene has on the CBB was measured by monitoring the development of the insect from egg to adult, on coffee-artificial diets amended with increasing concentrations of isoprene. Concentrations of isoprene above 25 ppm caused mortality and developmental delay in all insect stages from larva to adult, as well as the inhibition of larvae molting. In conclusion, comparative functional genomics studies allowed the identification of at least one possible mechanism of insect induced response in coffee, providing new tools to screen and utilize *Coffea*genetic resources, as well as, revealing unknown mechanisms of production of volatile substances in coffee plants with negative effects on CBB that may be applied for pest control purposes and possible interaction with other pests.

**Research Co-sponsored by the Colombian National Coffee Growers Federation and the Colombian Ministry of Agriculture

Diterpenes & Transcriptional Profile of CYPs Genes during *Coffea arabica* fruit Development

Luiz F. P. Pereira *Embrapa Café*, **Douglas S. Domingues** *Instituto Agronômico do Paraná*, Gonçalo Amarante Guimarães Pereira *Genomics and Expression Laboratory- UNICAMP*, Marcelo Falsarella Carazzolle *Genomics and Expression Laboratory- UNICAMP*, Osvaldo Reis Junior *Genomics and Expression Laboratory- UNICAMP*, Luiz G. E. Vieira *IAPAR (LBI-AMG)*, Maria Brigida dos Santos

Scholz *Instituto Agronômico do Paraná*, Cíntia Sorane Good-Kitzberger *Instituto Agronômico do Paraná*, Priscila Mary Yuyama *CIRAD*, Suzana Tiemi Ivamoto *Instituto Agronômico do Paraná*

Coffee oil is rich in kaurane family diterpenes, mainly cafestol (CAF) and kahweol (KAH), which are related with plant defense mechanisms, nutraceutical and sensorial beverage characteristics. In plants, the cytochrome P450s gene family (CYPs) is usually involved in the synthesis and interconversion of most plant secondary metabolites – which probably includes the diterpenes biosynthesis. Therefore, the aim of this study was to elucidate the CAF and KAH final biosynthesis steps combining biochemical and transcriptional analyses. We measured CAF and KAH by HPLC in flowers as well as fruit perisperm in several stages (30 to 210 days after flowering – DAF) to have the pattern of CAF and KAH tissue accumulation during fruit/grain development. CAF levels were detected mainly in flowers as well as in the perisperm decreasing after 120 DAF. On the other hand, KAH concentration increased with perisperm development reaching a peak at 120 DAF. Based on this HPLC analysis of diterpenes, 12 RNA-Seq libraries were obtained for *Coffea arabica* cv. IAPAR59: leaves, flowers and perisperm tissue from fruits along development. A total of 41.881.572 sequences were generated using Illumina, HiSeq2000. After clusterization, 127.600 contigs were formed with an average size of 1264bp. From those, 65480 were considered unique splicing variants (unigenes). With BLAST analysis we detected more than 250 CYPs which eight were used for further transcriptional analysis by qPCR in leaves, flowers and fruits in three developmental stages (90, 120 and 150 DAF). For five genes we observed a similar pattern between gene transcription and diterpenes concentration levels. Three CYPs (CaCYP76F2_1, CaCYP82C4, CaCYP74A1) had transcriptional patterns similar to CAF accumulation (most accumulated in flowers). On the other hand, two CYPs (CaCYP71A4_1 and CaCYP701A3) were related with KAH accumulation, lower in leaf and flower, but with increasing detection during fruit development. These five CYPs warrant further investigation as potential candidate genes involved in the final stages of CAF and KAH biosynthetic pathway providing us important clues and valuable information for future analysis of coffee diterpene synthesis. This is the first work with Illumina sequencing of coffee fruit tissue, which is providing important information on key genes related to enzymes and metabolites involved in fruit ripening and cup quality.

The Role of Lineage-Specific Gene Family Expansions in Coffee Adaptation: The Case of N-Methyltransferases Involved in Caffeine Biosynthesis

Victor A. Albert *University of Buffalo*, Pablo Librado *Universidad de Barcelona*, Christine Tranchant-Dubreuil *IRD, UMR DIADE*, Valerie Poncet *IRD, UMR DIADE*, Alexandre de Kochko *IRD, UMR DIADE*, Julio Rozas *Universidad de Barcelona*, Lorenzo Carretero-Paulet *University of Buffalo*

Using the new *Coffea canephora* genome sequence, we examined lineage-specific gene family expansions with potential adaptive value for coffee. We fit two different branch models implemented in the likelihood program BadiRate to 16,917 orthogroups identified in genome-wide comparisons with Arabidopsis, grape and tomato: i) a branch model accounting for coffee-specific expansions in size, and ii) a global ratio model, indicating stable sizes among species. To identify gene functions that may have played special roles during coffee evolution, we statistically examined differential representations of GO terms among expanded orthogroups. 98 out of 4,269 generic GO terms were differentially distributed, mostly corresponding to two main functional categories: defense response (including NBS and R genes) and metabolic process, the latter including different catalytic activities involved in secondary compound synthesis, e.g. flavonoids, terpenoids, phenylpropanoids or alkaloids. Particularly noteworthy was the enrichment in N-methyltransferases (NMTs), the main enzymes involved in the biosynthesis of caffeine, a purine alkaloid accumulated in coffee beans and the most characteristic secondary metabolite of the coffee plant. Phylogenetic analysis of 23 coffee NMTs together with NMTs from cacao and different tea species, including those known to be involved in caffeine biosynthesis, revealed the occurrence of four clades: caffeine biosynthetic NMTs are nested within three species-specific clades, while a fourth clade in coffee includes previously unreported NMTs. Selection analyses in these different NMT clades discounts the possibility of parallel molecular adaptations. Microsynteny analyses permitted unraveling the role of tandem duplications in the expansion of the coffee NMT gene family. Taken as a whole, these results indicate the independent acquisition of caffeine biosynthesis in all three caffeine producing plant species

examined, providing an outstanding example of convergent evolution of secondary metabolic pathways followed by coffee-specific gene family expansions through tandem duplication.

Abstracts Presented in Other Workshop Sessions:

Root Genomics

Unveiling Coffee Molecular Response Mechanisms to Drought: The Homeobox-Leucine Zipper I CAHB12expressed in Roots of *Coffea arabica* Is Able to Confer Water Deficit and Salt Stress Tolerance to Transgenic Plants

Márcio Alves-Ferreira, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil

Drought is one of the main abiotic stresses affecting plant growth, with serious negative consequences to crop yield worldwide. Among those crops, coffee is severely injured by water deficit. In the present work, 34 coffee homeobox (HB) genes were identified in the Brazilian Coffee Genome Project database, being subsequently classified through phylogenetic analyses together with *Arabidopsis* and rice HB genes. Three coffee HB genes (*CAZHD4*, *CAHB1-like2* and *CAHB12*) showed, by *in silico* analysis, an expression up-regulated by osmotic stress. qPCR analyses revealed that *CAHB12* is highly up-regulated in leaves and lateral roots of *Coffea arabica* plants, upon moderate and severe water deficit conditions. Functional characterization of transgenic *Arabidopsis* constitutively expressing *CAHB12* resulted in increased tolerance to drought and salt stresses, during distinct developmental stages. An insight into the gene set modulated by *CAHB12* ectopic expression was provided by massively parallel sequencing (RNA-Seq). Classical drought response genes were generally repressed, suggesting that other mechanisms are likely contributing to the more tolerant phenotype exhibited by over-expressing plants, such as the pathway signaled by heat shock proteins and heat shock transcriptional factors. Our data combined, suggest a possible role for *CAHB12* as a positive regulator of coffee stress response, and indicate this gene as a potential candidate for biotechnological approaches.

Plant Metabolic Network Resources and Applications

Curation in Metabolic Databases: Implementing Tools and Rules Across the Universal Metabolic Database MetaCyc to Species-Specific Plant/Genome Databases

Hartmut Foerster, Boyce Thompson Institute, Ithaca, NY; Lukas Mueller , Boyce Thompson Institute for Plant Research, Ithaca, NY

MetaCyc (http://metacyc.org) is a multi-species database providing non-redundant metabolic data from over 2400 species, particularly plants and microorganisms. The ongoing effort to improve, complement and expand the existing scope of curated pathways is crucial for the importance of the database as a significant reference and research tool for scientists studying plant metabolism. The primary goal of metabolic databases to capture the universe of primary and specialized (secondary) metabolism has been extended towards the presentation of plasticity and interconnectivity of biochemical networks. MetaCyc has been the principal matrix for many Pathway/Genome Database's (PGDB's) which have been predicted based on their sequenced genomes. The value of the database is defined by the accuracy, consistency and currentness of released data. Manual curation comprises the extraction, validation and presentation of structural and functional data from the body of scientific literature associated to pathways, compounds, enzymes, and genes. High data quality is assured by furnishing pathways and enzymes with evidence and citations linking back to the original source of research. Pathway tools represents the comprehensive software platform that allows curators to generate reactions, construct pathways and annotate enzymes and genes. We will present the scope of Pathway tools, explore the content of MetaCyc and explain how to access and query the database to retrieve information on pathways, enzymes and genes. We will also provide an

overview of SolCyc (http://solcyc.solgenomics.net/) a collection of species-specific databases of primarily *Solanacea* (tomato, pepper, eggplant, petunia) and close relatives such as **coffee** and how to compare species across available PGDB's.

Bioinformatics

The South Green Bioinformatics Platform

Mathieu Rouard, Bioversity International, Montpellier, France; Alexis Dereeper, IRD, UMR RPB, Montpellier cedex 5, France; Gaetan Droc , CIRAD, UMR AGAP, Montpellier, France; Gautier Sarah , INRA, UMR AGAP, Montpellier, France; Jean-François Dufayard, CIRAD, UMR AGAP, Montpellier, France; Valentin Guignon , Bioversity International, Montpellier, France; Chantal Hamelin, CIRAD, UMR AGAP, Montpellier, France; Felix Homa , CIRAD, UMR AGAP, Montpellier, France; Fred de Lamotte, INRA, UMR AGAP, Montpellier, France; Pierre Larmande , IRD, UMR DIADE, Montpellier, France; Delphine Lariviere , CIRAD, UMR AGAP, Montpellier, France; François Sabot , IRD, UMR DIADE, Montpellier, France; Guilhem Sempéré , CIRAD, UMR AGAP, Montpellier, France; Marilyne Summo, CIRAD, Montpellier, France; Bertrand Pitollat , CIRAD, Montpellier, France; Stephanie Pointet , CIRAD, UMR AGAP, Montpellier, France; Dominique This , Montpellier SupAgro, Montpellier, France; Stéphanie Bocs, CIRAD, UMR AGAP, Montpellier, France; Manuel Ruiz , CIRAD, UMR AGAP, Montpellier, France

The South Green platform (http://www.southgreen.fr/) is a local network of scientists gathering Bioinformatics skills based on the Agropolis campus that hosts research institutes such as CIRAD, IRD, INRA, SupAgro and Bioversity international. Based on this strong local community in the field of agriculture, food and biodiversity, various bioinformatics applications and resources dedicated to genomics of tropical and Mediterranean plants has been developed and published.

The objectives of South Green are to promote these original tools as well as their interoperability. Exchange and collaborative developments are also fostered through regular hands-on sessions on synergistic themes such as Galaxy, genome annotation or next generation genotyping. Finally, we provide access to computing facilities and hands-on training for both users and developers engaged in the network.

The South Green web portal contains currently 20 information systems and tools and targets about 30 plants. As a proof of concept for system interoperability, we recently released the Banana Genome Hub powered by major GMOD components (i.e. Chado, Cmap, Gbrowse, Tripal, Galaxy, Pathway tools) and South Green tools (e.g. GnpAnnot, GreenPhylDB, SNiPlay, TropGeneDB, ESTtik, OryGenesDB). This concept of hub can be extended to other crops as currently done for the **Coffee** genome.

Abstracts of Poster Presentations

Homeologous Genes Involved in Mannitol Synthesis Reveal Unequal Contributions in Responses to Abiotic Stress in *Coffea arabica*

Kenia Carvalho, Instituto Agronômico do Paraná, Londrina, Brazil; Carmen L.O. Petkowicz, Universidade Federal do Paraná, Curitiba, Brazil; Getulio T. Nagashima, Instituto Agronômico do Paraná, Londrina, Brazil; João C. Bespalhok Filho, Universidade Federal do Paraná, Curitiba, Brazil; Luiz G.E. Vieira, Universidade do Oeste Paulista, Presidente Prudente, Brazil; Luiz F. P. Pereira , Embrapa Café, Londrina, Brazil; Douglas S. Domingues , Instituto Agronômico do Paraná, Londrina, Brazil

Coffea arabica is an allotetraploid originating from a natural hybridization between *C. canephora* and *C. eugenioides*, and it accounts for 75% of the world's coffee production. Abiotic stresses, such as extreme temperatures, salinity, and drought, limit worldwide commercial coffee production. The accumulation of compatible solutes, such as mannitol, is known to be involved in abiotic stress tolerance in higher plants.

The aim of this study was to investigate the transcriptional responses of genes involved in mannitol biosynthesis and catabolism in *C. arabica* leaves under water deficit, salt stress and high temperature. Mannitol concentration was significantly increased in leaves of plants under drought and salinity but reduced by heat stress. Fructose content followed the level of mannitol only in heat-stressed plants, suggesting the partitioning of the former into other metabolites during drought and salt stress conditions. Transcripts of the key enzymes involved in mannitol biosynthesis, *CaM6PR*, *CaPMI* and *CaMTD,* were modulated in distinct ways depending on the abiotic stress. Our data suggest that changes in mannitol accumulation during drought and salt stress in leaves of *C. arabica* are due, at least in part, to the increased expression of the key genes involved in mannitol biosynthesis. In addition, the homeologs of the *C. canephora* subgenome did not present the same pattern of overall transcriptional response, indicating differential regulation of these genes by the same stimulus. Finally, this study adds new information on the differential expression of *C. arabica* homeologous genes under adverse environmental conditions.

Drought and Heat Stresses Have Distinct Impacts in Sub-Genomic Regulation Patterns of a Mannitol Biosynthesis Gene in Allotetraploid *Coffea arabica*

Douglas S. Domingues, Instituto Agronômico do Paraná, Londrina, Brazil; Kenia Carvalho, Instituto Agronômico do Paraná, Londrina, Brazil; Sandra M. B. Cação , Plant Biotechnology Laboratory, IAPAR, Londrina, Brazil; Nathalia V Silva, IAPAR; Luiz G. E. Vieira , IAPAR; Luiz-Filipe Protasio Pereira, Embrapa Café, Brazil

Coffea arabica is an allotetraploid originating from a natural hybridization between *C. canephora* and *C. eugenioides*, and it accounts for 75% of the world's coffee production. For allotetrapolyploids, it is expected that most genes are present in two homoeologous forms, highly similar but non-identical. Abiotic stresses, such as drought and extreme temperatures limit commercial coffee production. The accumulation of solutes, such as mannitol, is known to be involved in abiotic stress tolerance in plants. The key-enzyme involved in mannitol biosynthesis in plants is mannose-6-phosphate reductase (M6PR). Here, we report an inference of M6PR copies in *C. arabica* genome and the M6PR transcriptional profile under drought and heat stress. Based in allelespecific PCR and BAC library (~5X genome coverage) screening, we could infer that two homoeologous *CaM6PR* loci should exist in the tetraploid coffee genome. These observations are supported by analysis in sequenced diploid plant genomes, where M6PR is a single-copy gene. Using qPCR primers with specific amplification to *C. canephora* subgenome present in *C. arabica*, we could also evaluate transcriptional responses of "total" *CaM6PR* expression and "canephora-specific" pattern under water deficit and heat stress (37^{o}C). Both total and subgenomic M6PR transcripts, as well as mannitol concentration, are up-regulated upon drought in *C. arabica* leaves. Under high temperatures, we observed the opposite response: mannitol and M6PR transcripts are downregulated. Our results highlights that stress responses may trigger opposite molecular responses in plants, and guarantee further analysis in *Coffea arabica* homoelogous M6PR genes, in order to depict its response under abiotic stresses.

Finnancial support: FINEP, CNPq and Consórcio Pesquisa Café.

Coffee (*Coffea arabica*) Molecular Resistance Responses to the Root-Knot Nematode *Meloidogyne incognita*

Erika V. S. Albuquerque, Embrapa, Brasília, Brazil; Joseane Padilha , Embrapa, Brasília, Brazil; Sergio Alencar , Universidade Catolica de Brasilia, Brazil; Regina M. D. G. Carneiro , Embrapa Recursos Geneticos e Biotecnologia, Brazil; Gerogios Pappas , Universidade de Brasilia, Brazil; Anne-Sophie Petitot, Institut de Recherche pour le Développement, Montpellier, France; Anne-Claire Lecouls , Institut de Recherche Pour le Developement, France; Rodrigo R. Fragoso , Embrapa Cerrados, Brazil; Maria Fatima Grossi de Sa , Embrapa, Brasília, Brazil; Diana Fernandez , Institut de Recherche pour le Développement, Montpellier, France

Root-knot nematodes (*Meloidogyne* spp.) are major pests damaging the coffee culture (*Coffea arabica*). To gain knowledge about the molecular mechanisms of plant immunity in coffee, we investigated the

transcriptomic changes associated with the hypersensitive response (HR) to *Meloidogyne incognita* in resistant *C. arabica* 'UFV 408-28' plants. Deep sequencing (Illumina) data obtained from resistant roots at 6 days post inoculation (dpi), inoculated with nematodes (treated) or not (control), was tested by the Trinity method for *de novo* assembly of transcripts without a coffee reference genome forming contigs analyzed by BlastnNR or BlastnEST. With cutoffs to adjusted p-value ≤ 0,05 and 1≤log2FC≤-1 applied to the contigs, 3519 of them were considered downregulated and 868 upregulated regarding the treated sample (inoculated). We decided to check the presence of 92 selected genes from various pathways and with different functions, reported as related to biotic stress responses in plants. We found 89% of the selected genes among the contigs, being 53 differentially expressed. In the second part of our study, we compared the *in vitro* sequencing results to the expression profiles generated by RT-qPCR of these genes selected from literature during the coffee roots responses to *M. incognita*. The same samples sequenced by Illumina were analyzed together with roots collected at 4, 5 and 6 (dpi), corresponding to the previously determined onset of HR, in three independent experiments. We performed statistical analysis to verify the best model to fit each gene for inoculation, genotype, and the interaction inoculation x genotype effects. Coffee molecular responses were gradually increasing with infection time (dpi). Many expression differences were significant at 5% and 1% to the qPCR results. Transcriptomic data obtained by both approaches were consistent, and some specific novel cDNA sequences may be used to unveil the resistance mechanisms acting in *C. arabica*.

Cytosolic Glutamine Synthetase Is Modulated By Rust Infection and Nitrogen Status in *Coffea arabica*

João D. Soares, IAPAR, Londrina, Brazil; Viviane Y. Baba, IAPAR; Luiz-Filipe Protasio Pereira, Embrapa Café, Brazil; Douglas S. Domingues, Instituto Agronômico do Paraná, Londrina, Brazil

Nitrogen (N) is the mineral nutrient most required by plants, and its limitation directly influences plant development in natural and agricultural environments. In coffee, one of the most important agricultural commodities in the world, N- fertilization represents one of the major costs on its cultivation. Several reports have shown that nitrogen-limiting conditions can impact on plant disease development. The coffee rust, caused by *Hemileia vastratrix*, is considered the main disease in *Coffea arabica*. However, there are few molecular studies linking nutritional status of coffee plants and defense response to pathogens. Here, we evaluated the expression of cytosolic-glutamine-synthetase, CaGS1, in plants infected with rust and under different N availabilities. *C. arabica* plants that are resistant and susceptible to rust were challenged with *H. vastatrix*, under N-starvation and N-sufficiency. Young leaves were collected 0, 12, 24 and 48h after inoculation. qPCR analysis indicated a transcript peak in 12h after inoculation, that is more pronunciated in N-sufficient leaves. Susceptible plants had a second expression peak at 48 h in N-starved leaves. Enzyme activity followed the same pattern of transcriptional responses. In this way, we can infer that cytosolic glutamine synthetase has been modulated by coffee rust and may be influenced by pathogen-host interaction, since the response was different between resistant and susceptible cultivars. The results allow an initial understanding of the molecular interaction between mineral nutrition and coffee rust. Finnancial support: Consorcio Pesquisa Café.

Integrative System for Gene Family Gathering and Analysis in a Context of Crops' Stress Response Study

Delphine Lariviere, Cirad, UMR AGAP, Montpellier, France; Jean-François Dufayard , CIRAD, UMR AGAP, Montpellier, France; Stéphanie Bocs , CIRAD, Montpellier, France; Dominique This , CIRAD - Montpellier SupAgro, Montpellier, France

The study of plant tolerance to stress is a crucial issue for crops improvement and stability in constrained environments. Gene family analysis is an important way to understand complex processes underlying stress response in crops. Several tools exist to study families and propose automatically clustered families or curated published families. We observed that automatic clustering, efficient for global analyses, is rarely sufficient for precise studies: i) families can be spread in several cluster, or ii) intrusive sequences are

represented in clusters of interest. That is why biologists need most of the time to manually constitute their families. In response to this need, we propose to develop an integrative system that will allow to gather sequences from different sources for a customized family. This system will integrate several tools used for family construction and analysis. Currently, the prototype allows to query in-house Chado database (banana, **coffee**) and to import personal data with a web interface using the Drupal CMS. Multi-species analyzes are available using Galaxy workflows manager. Eventually, this system will be linked to a stress response oriented database. This database will propose a set of controlled vocabularies, some known stress resistance factors and other indicators to ease the identification of mechanisms implied in plant stress resistance. Beside that, this system will be generic for all type of gene families, and will propose a synthetic and dynamic view to offer to researchers an easy and intuitive way to curate their families.