International Coffee Genome Network (ICGN)

http://www.coffeegenome.org

Celebrating 10 years: 2005-2015

- ICGN has been committed to sequence the coffee genome and to advance coffee genomic research through international partnerships for sustainable coffee production.
- ICGN networks scientific groups (113 members in >56 countries) around the world in Africa, the Americas, Europe, and Asia (for comprehensive list of members see http://www.coffeegenome.org/about/members.php).

Participating Research Institutes
113 members/>56 countries

Brazilian Consortium for Coffee Research/ EMBRAPA, IAPAR, UNICAMP (**Brazil**), CENICAFE (**Colombia**), PROMECAFE/ IICA/CATIE (**Mexico, Guatemala, Peru, Honduras, Costa Rica, El Salvador, Nicaragua, Ecuador, República Dominicana, Venezuela, Haiti, Bolivia, Panama, Jamaica**, and others)

CCRI/CCMB/CBC/CFTRI (**India**),

Inter-African Coffee Organization (IACO)/ African Coffee Research Network (ACRN) 25 countries (**Ethiopia, Kenya, Uganda, Ivory Coast, Tanzania, Congo DRC, Ghana, Rwanda, Burundi, Madagascar,** and others)

Cornell University + HARC/UIUC + AGI + BTI (**USA**)

IRD/CIRAD/ENSAM, Nestle R&D Tours (**France**),

University of Trieste/ Illycafe/ ENEA (**Italy**)

University of Bohn/Max Plank/Botanical Garden Berlin (**Germany**)

Swedish Agricultural University (**Sweden**)

Bioversity International

(for comprehensive list see http://www.coffeegenome.org/about/members.php).

ICGN Founding Members

Founding Meetings
Paris- 4&5 April 2005

ICGN History

- Initial discussions 20th ASIC International Coffee Science Conference Bangalore, India (Oct 11-15, 2004)
- ICGN launching meeting in Paris, France (April 4-5, 2005) organized by IRD/CIRAD, with support from USDA to fund travel for scientists from coffee producing countries
- Follow up meeting in Trieste, Italy (April 28-29, 2006)
- Meetings in conjunction with the ASIC International Coffee Science Conferences in Montpellier, France (Sep 16, 2006); Campinas, Brazil (Sep 14-19, 2008); Bali, Indonesia (Oct 3-8, 2010); San José, Costa Rica (Sep 12-16, 2012), and Armenia, Colombia (Sep 8-13, 2014)
- Annual meetings in conjunction with the Coffee Genomics Workshops organized by ICGN as part of the Plant and Animal Genome Conference in San Diego, California (2008-2015) (see http://int-pag.org)- co-organizers: Marcela Yepes, Cornell; Philippe Lashermes IRD; Rod Wing AGI, U. Arizona

ICGN International Coffee Genome Network

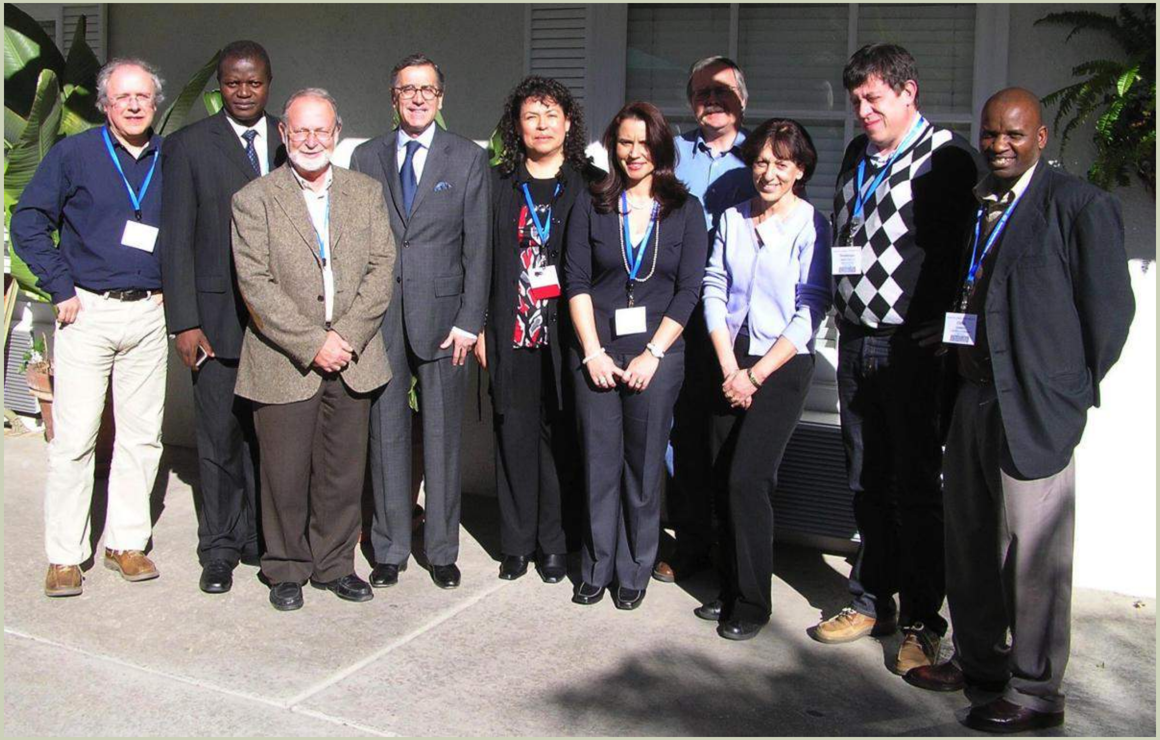
ICGN meeting in conjunction with the 22nd ASIC International Conference on Coffee Science in Campinas, Brazil, September 16th, 2008

A total of 37 scientists representing 11 countries attended the ICGN meeting.

ICGN meeting at the Plant and Animal Genome Conference and 1st Coffee Genomics Workshop, Jan 13, 2008

International Coffee Genome Network

ICGN meeting at the Plant and Animal Genome Conference, and Coffee Genomics Workshop, Jan 12, 2009, San Diego, CA



With the participation of Nestor Osorio, Executive Director, International Coffee Organization

Coffee Importance

- Coffee is the first Agricultural commodity in world trade, second in value only to oil (Global market US $90 billion/yr).
- Coffee is produced mainly by small holders. A sustainable coffee sector is needed to meet the millennium goals in terms of poverty eradication (>125 million people dependent on coffee for their livelihood).
- US is the first coffee consuming country in the world. Increased coffee quality responsible for enhanced consumption and consumer satisfaction (market increase 2% per year).

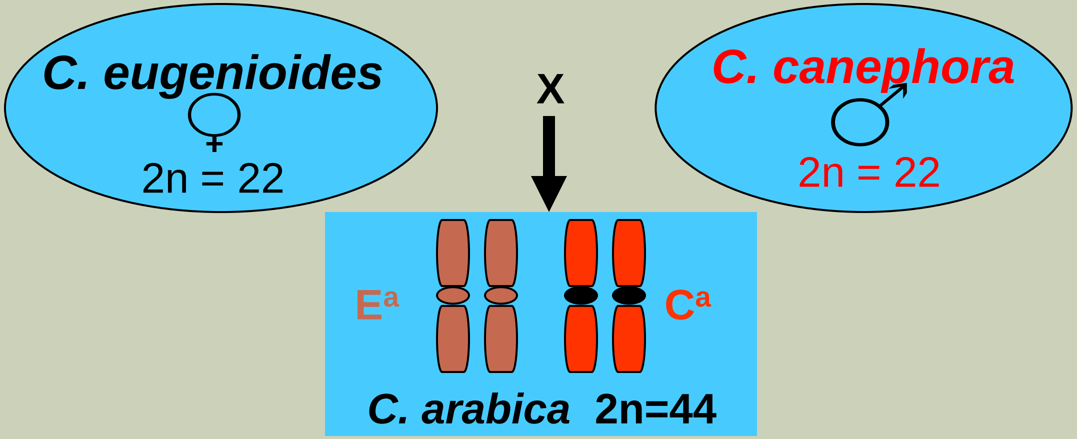
Coffee and Climate change

- Sustainable coffee production is very vulnerable to climate change at a global scale due to increased insect and disease incidence, as well as abiotic stress.
- A major constrain is the very narrow genetic pool of cultivated varieties. Urgent need to coordinate efforts at a global scale to promote *Coffea* germplasm characterization, enhanced utilization, and conservation in the context of climate change adaptation.
- Advanced genomic tools are needed to accelerate progress in coffee breeding and selection of resistance varieties through enhanced knowledge of gene function and allelic diversity.

Coffea diversity studies

- Facilitate positional cloning and candidate gene discovery, characterize gene families and metabolic pathways of interest.
- Enhance understanding of genome organization and evolution in *Coffea* and develop more efficient strategies to capture diversity represented in the >100 species of the genus.
- Unveil the genetic and molecular bases of important biological traits in coffee tree species that are relevant to **growers, processors and coffee consumers**.

Coffee genome sequencing



C. eugenioides ♀

2n = 22

X

C. canephora ♂

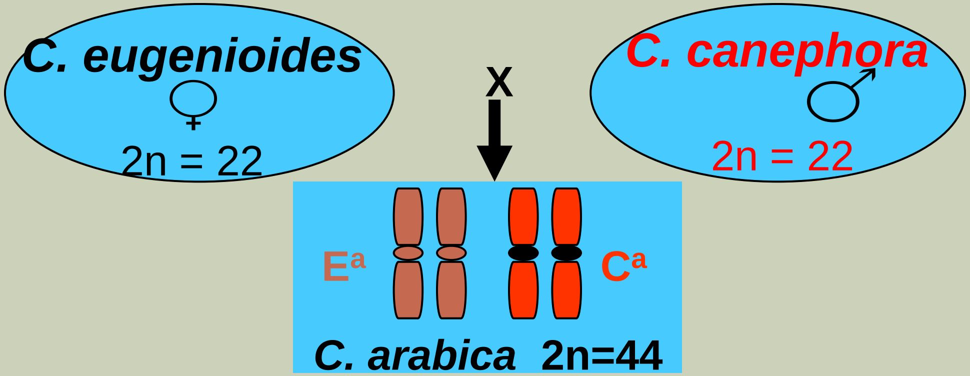
2n = 22

E^a C^a

C. arabica 2n=44

- *Coffea arabica* (1,300 Mb) recent allotetraploid formed from the merge of two diploid species: *C. canephora* (710 Mb) and *C. eugenioides* (660 Mb).
- Decoding the coffee genome will contribute to understand the functional and agronomic significance of polyploidy.

Coffee genome sequencing



- *C. canephora* sequencing, assembly, and annotation project was co-lead by Genoscope-CEA and IRD-CIRAD funded by the Agence Nationale de la Recherche (ANR), France, as well as several ICGN member countries that joined the effort for mapping, sequencing, and annotation.
- The first genome assembly was published recently Denoeud *et al.* 2014. Science v. 345: 1181-1184.

C. canephora sequencing/assembly

1. **Sequencing was done using 454 Roche GX FLX Titanium, Sanger and Illumina:**

Roche/454:

WGS 454 sequencing : 28.87 X (assuming a genome size of 710 Mb) including

- Reads single end : 14.83 X , Mean size: 359 bp
- Long reads single end : 8.24 X , Mean size : 462 bp

Paired-end sequencing of long insert libraries (8 and 20 Kb) : 5.8 X (2.2X for 8kb, 3.6X for 20kb), Mean Size : 252 bp

Sanger:

Two BAC libraries (*Hind III* and *Bam HI*) were constructed in collaboration with Rod Wing at Arizona Genomics Institute.

Sanger BAC-end sequencing: 144,000 BES were generated (72,000 BAC clones x 2 ends 5' and 3'): 0,27X

Mean BAC insert size : 135 Kb, range: 63,2Kb < insert < 253,6 Kb

C. canephora sequencing/assembly strategy

Illumina Sequencing (coverage 69.7 X)
- Single end reads: 7.3X , size: 76 bp (4,8X) or 150 bp (2,5X)
- Paired end reads : 62.4X, size: 76 bp (42,4 X) or 108 bp (20 X)

2. For the assembly, a first assembly was done using the 454 and Sanger data, and the Illumina data was used to correct errors due to homopolymers

568.6 Mb total assembled genome (80% of 710 Mb genome)

C. canephora assembly

Assembly statistics

Jan 2011 (**Patrick Wincker**'s presentation from Genoscope at PAG)

Size: 540 Mb (76% estimated size)

N50 contigs : 6.2 kb

Coverture 15x

N50 scaffolds : 443 kb

Largest scaffold : 12.8 Mb

Jan 2013 (**France Denoeud**'s presentation from Genoscope at PAG)

Nb of Scaffolds 13,345

Size: 568.6 Mb (80% estimated genome size)

Coverture 28.87 X (454/Sanger) and Illumina 69.7 X

N50: 1.2 Mb (108 scaf)

N80: 65 kb (635 scaf)

Largest scaffold: 9.0 Mb

C. canephora assembly

Denoeud *et al.* 2014. Science v. 345: 1181-1184

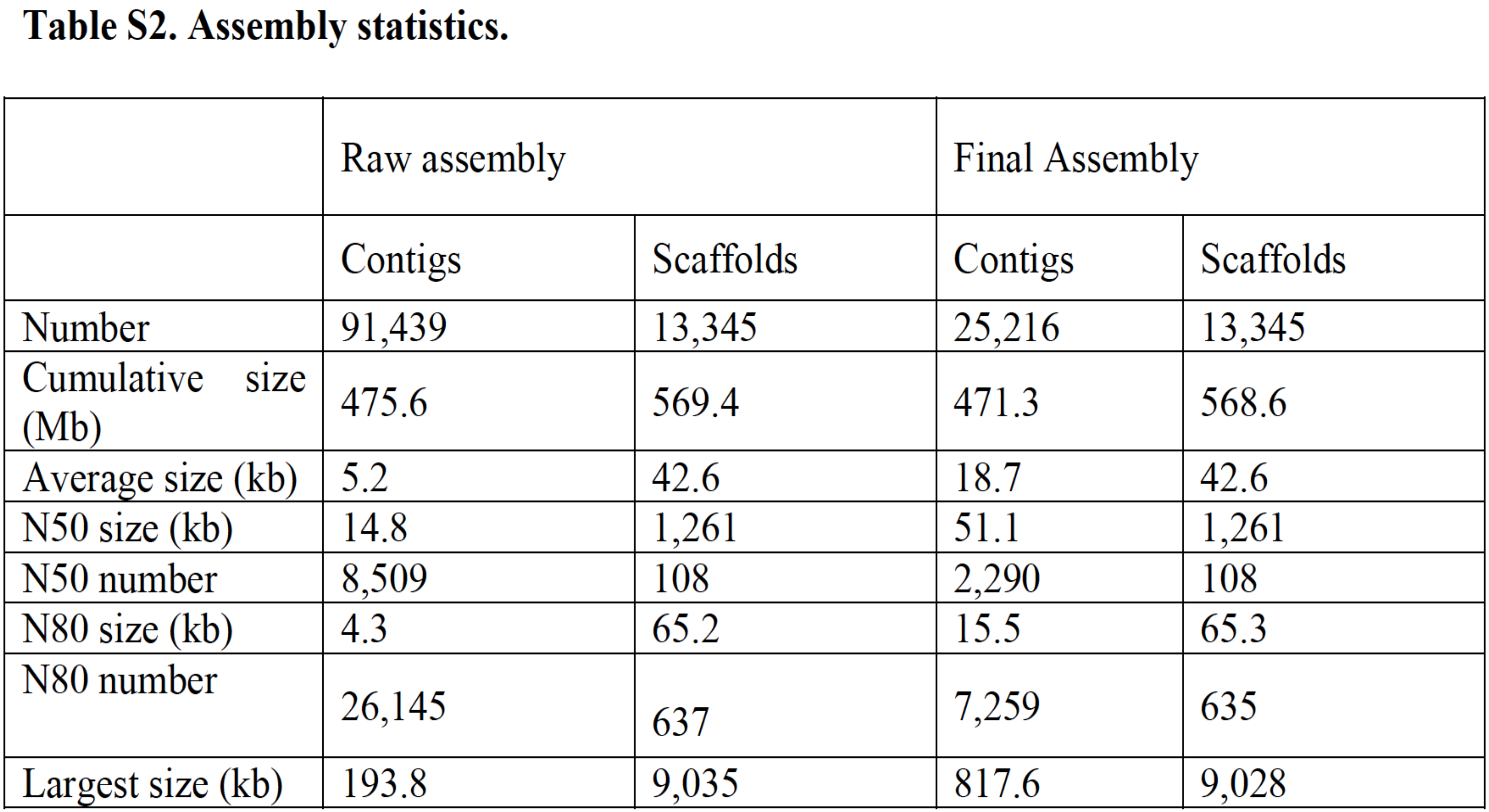
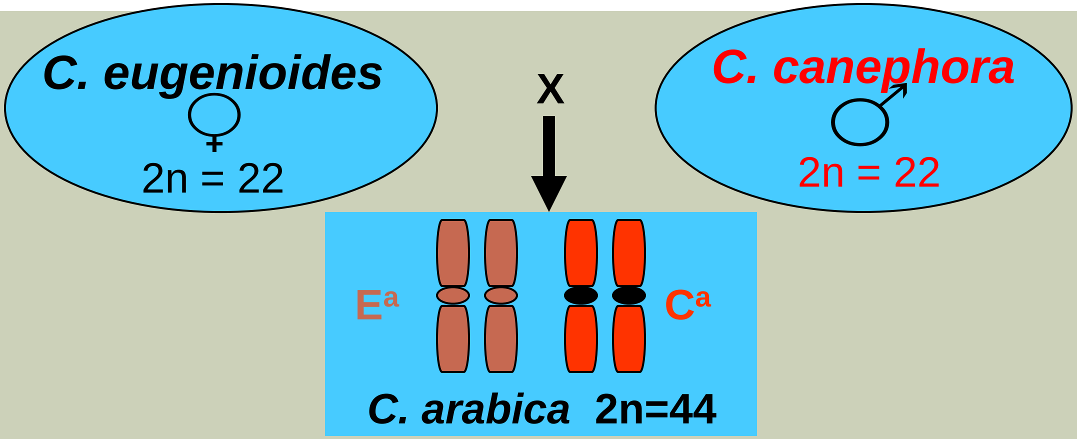


Table S2. Assembly statistics.

	Raw assembly		Final Assembly	
	Contigs	Scaffolds	Contigs	Scaffolds
Number	91,439	13,345	25,216	13,345
Cumulative size (Mb)	475.6	569.4	471.3	568.6
Average size (kb)	5.2	42.6	18.7	42.6
N50 size (kb)	14.8	1,261	51.1	1,261
N50 number	8,509	108	2,290	108
N80 size (kb)	4.3	65.2	15.5	65.3
N80 number	26,145	637	7,259	635
Largest size (kb)	193.8	9,035	817.6	9,028

Coffee genome sequencing



C. eugenioides ♀
2n = 22

X

C. canephora ♂
2n = 22

E^a C^a

C. arabica **2n=44**

Coffea eugenioides and *Coffea arabica* sequencing and assembly: Cornell University and the Colombian National Coffee Research Center (CENICAFE) were funded for this project by the InterAmerican Development Bank/Fondo Regional de Tecnología Agropecuaria (IDB/FONTAGRO)

25th ASIC Int. Coffee Conference Armenia Colombia

500 participants from 47 countries

25th ASIC Int. Coffee Conference Armenia Colombia

500 participants from 47 countries

CENICAFE: Colombian National Coffee Research Center

Cenicafé

1938

Federación Nacional de Cafeteros de Colombia

CENICAFE: Colombian National Coffee Research Center

Federación Nacional de Cafeteros de Colombia

CORNELL

IDB

Inter-American Development Bank

540 × 304 - export.gov

Cenicafé

Libertad y Orden

Federación Nacional de Cafeteros de Colombia

Ministerio de Agricultura y Desarrollo Rural

8th Coffee Genomics Workshop

1. **Alvaro Gaitán and Marco Cristancho**, CENICAFE, **Colombia**, Long-Read deep sequencing and assembly of the allotetraploid *Coffea arabica* cv. Caturra and its maternal ancestral diploid species *Coffea eugenioides*.
2. **Michele Morgante**, Universitá di Udine and Istituto di Genomica Applicata, **Italy**, Progress report on the sequencing and assembly of the allotetraploid *Coffea arabica* var. Bourbon genome.
3. **Alexandre de Kochko**, IRD UMR DIADE, **France**, Dihaploid *Coffea arabica* genome sequencing and assembly.
4. **Romain Guyot**, IRD UMR DIADE, **France**, Transposable element distribution, abundance and impact in evolution in the genus *Coffea*.
5. **Elaine Silva Dias**, UNESP, **Brazil**, Impact of transposable elements on the evolution of *Coffea arabica* (Rubiaceae)

ICGN International Coffee Genome Network

8th Coffee Genomics Workshop

Alvaro Gaitan

Marco Cristancho

Michele Morgante

Alexandre de Kochko

Romain Guyot

Elaine Dias

ICGN International Coffee Genome Network

8th Coffee Genomics Workshop

Thank you to the ICGN Coffee Genomics Workshop
speakers and attendants at PAG January 10-14, 2015

ICGN International Coffee Genome Network

8th Coffee Genomics Workshop

Thank you to the ICGN Coffee Genomics Workshop speakers and attendants at PAG January 10-14, 2015