



Plant & Animal Genomes XIX Conference

January 15-19, 2011
Town & Country Convention Center
San Diego, CA

W152: Coffee Genomics

Sequencing The Coffee Genome

[Patrick Wincker](#)¹ , [Victor A. Albert](#)² , [Alan A. Andrade](#)³ , [Xavier Argout](#)⁴ , [Benoit Bertrand](#)⁴ ,
[Alexandre de Kochko](#)⁵ , [Giovanni Giuliano](#)⁶ , [Giorgio Graziosi](#)⁷ , [Robert Henry](#)⁸ , [Jayarama Jayarama](#)⁹
, [Philippe Lashermes](#)⁵ , [Ray Ming](#)¹⁰ , [Chifumi Nagai](#)¹¹ , [Steve Rounsley](#)¹² , [David Sankoff](#)¹³

¹ Genoscope, Institut de Génomique du CEA, CP5706 91057 Evry cedex France

² Department of Biological Sciences, University at Buffalo (SUNY), USA

³ EMBRAPA, Final Av. W/5 Norte Brasília-DF 70770-917, Brasil

⁴ CIRAD 34398 Montpellier Cedex 5, France

⁵ IRD Centre de Montpellier BP 64501 34394 Montpellier cedex 5, France

⁶ ENEA, Post Box 026, via Anguillarese 301 - 00123 S.M. di Galeria, Roma, Italy

⁷ University of Trieste. P.le Valmaura . 34100 Trieste, Italy

⁸ University of Queensland. Hartley Teakle Building [#83] St Lucia 4072 Australia

⁹ CCRI, Coffee Board, Chikmagalur Dist., Karnataka State, India

¹⁰ University of Illinois at Urbana-Champaign. Urbana, IL 61801 USA

¹¹ HARC PO Box100, Kunia HI 96759-0100 USA

¹² Dow Agrosiences, Indianapolis, IN, USA

¹³ University of Ottawa, 585 King Edward Avenue, K1N 6N5 Ottawa, Canada

Commercial coffee production relies mainly on two closely related species: *Coffea arabica* and *C. canephora*, which account respectively for 70 and 30% of the coffee production. All coffee species are diploid ($2n=2x=22$) and generally self-incompatible, except for *C. arabica* which is the only tetraploid ($2n=4x=44$) and self-fertile. Molecular analyses (Lashermes et

al. 1999) have indicated that *C. arabica* is a recent allotetraploid (CE genome) formed by hybridisation between two related diploid species: *C. canephora* (C genome) and *C. eugenioides* (E genome). In spite of the close relationship between the two constitutive sub-genomes, *C. arabica* displays diploid-like meiotic behavior with bivalent formation (Krug and Mendes 1940, Lashermes et al. 2000). The genomes of coffee species (Cros et al. 1995; Noirod et al. 2003) appear to be of rather low size (i. e. about 660, 710 and 1300 Mb for *C. eugenioides*, *C. canephora* and *C. arabica*, respectively). Several institutes are combining their scientific resources and expertise to sequence, assemble, and annotate the entire genome of *C. canephora*. The *C. canephora* genome consists of 11 chromosomes, is about 710 Mb in size, and is being sequenced de novo with deep coverage using 454 paired-end and single reads, and 50x coverage with Illumina GAIIx data to obtain a reference genome for *Coffea*. The overall sequencing strategy as well as progress of the project will be described.