

International Coffee Genomics Network (ICGN) Report Coffee Genomics Workshop

Plant and Animal Genome (PAG-32) Meeting,
San Diego, California, January 10-15, 2025

<https://plan.core-apps.com/pag32/event/ccb7e3a674d26e5273e1670285c21e9a>

Abstracts

Sunday, January 12, 2025: 4:00 PM-6:10 PM

Celebrating 20 years of the International Coffee Genomics Network (ICGN): Advancing coffee genomics research with Telomere-to-Telomere (T2T) genome assemblies

Marcela Yepes¹, Marco Aurelio Cristancho², Aleksey Zimin^{3,4}, Carmenza E. Góngora²,
Claudia Flórez², Alvaro Gaitán² and Herb Aldwinckle¹, ⁽¹⁾ Cornell University/ School of Integrative
Plant Sciences/ Plant Pathology and Plant Microbe Biology Section, Geneva, NY, ⁽²⁾ Federación Nacional
de Cafeteros de Colombia (FNC)/ Centro Nacional de Investigaciones de Café (CENICAFE) (Colombian
National Coffee Growers Federation/Colombian National Coffee Research Center, FNC/CENICAFE),
Chinchiná, Colombia, ⁽³⁾ University of Maryland, College Park, MD, ⁽⁴⁾ Johns Hopkins University,
Department of Biomedical Engineering, Baltimore, MD

The International Coffee Genomics Network (ICGN) (<https://www.coffeegenome.org/>) celebrates 20 years in 2025. On behalf of the coffee research community worldwide, ICGN is committed to advancing coffee genomics research with the first Telomere-to-Telomere (T2T) coffee genome assemblies. Our ICGN network was launched in Paris, France, in 2005. This first meeting was attended by 29 scientists from 12 countries world-wide, including participants from coffee producing countries and from the public and private sectors. The Institut de Recherche pour le Développement (IRD) and the Centre de Coopération Internationale en Recherche Agronomique pour le Développement (CIRAD) hosted and organized the meeting. Cornell University and Bioversity International are founding members, and have continuously served on ICGN's Steering Committee and Secretariat. ICGN has been successful at increasing both private and public investment in Coffee Genomics research, securing funding from: the French Agence Nationale de la Recherche (ANR), the InterAmerican Development Bank (IDB, through FONTAGRO), and the US National Science Foundation (NSF, through the Plant Genome Research Program). Future funding is critical to help continue moving coffee to the forefront of Plant Genomics Research.

During our 16th ICGN Coffee Genomics workshop at PAG, we will highlight not only the major accomplishments of the ICGN network over the past two decades, but also ongoing efforts and progress towards generating the first true telomere-to-telomere (T2T) coffee genome assemblies. EviAnn (Zimin *et al.* 2025), a novel data driven annotation pipeline developed by our collaborators at Johns Hopkins University is being used to annotate our coffee genomes (https://github.com/alekseyzimin/EviAnn_release). We are working closely with our collaborators

at the Colombian National Coffee Growers Federation (FNC)/Colombian National Coffee Research Center (CENICAFE) to continue adapting state of the art multi-omics approaches to *Coffea arabica*, the most widely cultivated coffee species in the world, which is highly appreciated by consumers for its superior cup quality, to gain the most complete reconstruction of the genome of this complex allotetraploid species, while facilitating a deeper understanding of genome function and organization. Development of computational tools to help uncover evolutionary relationships, and identify and interpret genetic variants is being customized for our target genotypes.

We are using long read sequencing platforms such as Pacific BioSciences High Fidelity (HiFi) Revo system and Oxford Nanopore to generate our first T2T coffee genome assemblies, and Dovetail Genomics' proprietary proximity ligation technologies which allow capture of the 3D architecture of the genome alongside primary sequence information. Overall, our approach will help improve pipelines for variant detection in coffee, *de novo* chromosome assembly, haplotype phasing, and epigenetics analysis towards characterizing key genotypes to help us tackle the untapped diversity in *Coffea*. Integration of large multi-omics datasets will provide deeper insights for linking genotypic and phenotypic variation of interest and speed up adaptation of the crop to climate change.

Reference

Zimin, Aleksey, Mihaela Pertea, Daniela Puiu, and Steven Salzberg. 2025. Ultrafast high-quality data centric eukaryotic genome annotation with Eviann. Abstract PAG32, Computational Genome Decoding Workshop, Abstract Presentation at 9:28 am, January 12th, 2025, See full abstract at <https://plan.core-apps.com/pag32/abstract/a6d651bb-6aac-45a4-8e91-9b6e9744fc32>

Exploring the Soil-Associated Microbiome of *Coffea arabica* Small-Holder Farms in Colombia: A Metagenomics Bar-Coding Approach to Bacterial and Fungal Taxonomic and Functional Metagenomics Analysis

Marco Cristancho¹, Valeria Faggioli², Lorena Gomez³, Victor Hugo Ochoa⁴, Maximo Rivarola⁵, Herb Aldwinckle⁶, Marcela Yepes⁶ and Alvaro Gaitán¹, ⁽¹⁾FNC/CENICAFE Colombian National Coffee Research Center, Chinchiná, Colombia, ⁽²⁾Instituto Nacional de Tecnología Agropecuaria (INTA), Argentina, ⁽³⁾National Center for Genetic Resources of the National Institute for Agricultural and Livestock Forestry Research (CNRG-INIFAP), Jalisco, Mexico, ⁽⁴⁾Arizona State University, Tempe, AZ, ⁽⁵⁾Instituto Nacional de Tecnología Agropecuaria (INTA), Argentina, ⁽⁶⁾Cornell University/ School of Integrative Plant Sciences/ Plant Pathology and Plant Microbe Biology Section, Geneva, NY

Coffea arabica is a commercially very important perennial polyploid crop species, representing 60-70% of the world's coffee market, with underdeveloped genomic resources yet highly appreciated by coffee consumers for its superior cup quality and health benefits. Colombia is the second world producer of *C. arabica*, and produces only this species, renowned in the global specialty coffee markets for its superior cup quality. In Colombia, *C. arabica* is produced throughout the country, and harvested fresh throughout the year, with unique cup quality profiles depending on the region and 'terroir'. Colombian coffee quality profiles are characterized very

stringently using not only organoleptic characteristics but also through rigorous chemical profiling (NIRS, lipids, etc.).

Coffee growing is a vital source of income for more than 125 million people in many developing countries in Latin America, Asia, and Africa. It is estimated that 70% of coffee farms are managed by small-holders. A strong, climate-resilient coffee sector will contribute to poverty eradication around the globe, and to conservation efforts for the unique biodiversity that major coffee producing countries harbor. As an example, Colombia is a megadiverse country with the richest biodiversity in the world per square meter.

We performed a metagenomics study to describe soil bacteria associated with *Coffea arabica* cultivation in Colombia using 16S rRNA metabar-coding. Samples were analyzed from three Departments: Cauca and Risaralda, in the center of the Colombian coffee growing area, and Magdalena in the Northern part of the country. Key genera identified included: *Janthinobacterium*, *Bacillus*, *Actinomadura*, and *Actinoallomurus*, which have potential biotechnological applications such as promoting plant growth and producing antibiotics. This research highlights the importance of understanding soil microbiomes for coffee cultivation. Additionally, our study explored soil fungi from six selected coffee farms in the same Departments, revealing diverse fungal communities, particularly *Mortierella* and *Saitozyma*, with varying functional roles. These findings underscore the need for further research to examine how different management practices can enhance beneficial soil health microbiomes and support sustainable coffee production.

FNC/CENICAFE are actively engaged in the development and use of climate-resilient *Coffea arabica* varieties throughout the Colombian coffee growing area, as coffee small-holder farmers are able to improve yields while reducing the use of pesticides, making coffee production sustainable from the social, economic, and environmental perspectives. The use of more resilient, high cup quality coffee varieties also benefits consumers, providing more subtle aromas and new flavors for the two billion cups of coffee consumed daily.

References

- Lorena Jacqueline Gómez-Godínez, Victor Ochoa, Valeria Faggioli, and **Marco Cristancho**. 2024. Exploring the soil-associated bacterial microbiome of coffee plantations in different regions of Colombia: a metabarcoding approach. *Tropical and Subtropical Agroecosystems* 27, Art. No. 047. <http://doi.org/10.56369/tsaes.5196>.
- Victor Hugo Ochoa-Henriquez, Valeria Faggioli, Lorena Jacqueline Gómez-Godínez, Maximo Rivarola, and **Marco Cristancho**. 2024. Colombian coffee (*Coffea arabica* L.) plantations: a taxonomic and functional survey of soil fungi. *Frontiers in Sustainable Food Systems*. Open Access. <https://doi.org/10.3389/fsufs.2024.1345383>

Cracking Diploid and Polyploid Plant and Animal Genome Assemblies with Advanced Assembly Solutions from Dovetail Genomics

Sierra McWilson, Dovetail Genomics, a part of Cantata Bio, Scotts Valley, CA

The complexity of plant genomes characterized by polyploidy, high heterozygosity, and repetitive sequences, has long posed challenges for genome researchers. In this presentation, we demonstrate how Dovetail Genomics' advanced assembly technologies enable both animal and plant genome researchers to overcome these obstacles, paving the way for high-quality genome assemblies. Dovetail's genome assembly technologies address the unique demands of polyploid and diploid plant and animal species by resolving homologous chromosome sets and capturing structural variations. Leveraging cutting-edge linked-read (akin to Hi-C) and long-read sequencing technologies, the services facilitate variant phasing, accurate assembly of repetitive regions, and comprehensive structural resolution.

The Dovetail® LinkPrep™ technology simplifies and accelerates the process of generating the long-range scaffolds critical for robust assemblies, ensuring researchers can efficiently transition from raw sequence data to contiguous, high-quality genomes. Incorporating LinkPrep data into your assembly process makes it possible to tackle even the most complex plant and animal genomes with precision.

Dovetail's innovative approaches are playing a transformative role for many animal and plant species, facilitating genome research and beyond. Our technologies, available through services and kits, continue to play a transformative role in many crop species, including coffee, facilitating genome research and beyond. LinkPrep technology not only enhances the speed and accuracy of the genome assemblies but also empowers researchers to address increasingly complex genomes, driving scientific discovery and improving agricultural sustainability and biomedical applications.

Genomic-Assisted Breeding of Cacao and Coffee

Eduardo Beche, Nestlé Research, Tours, France

Global cocoa production is under significant threat from plant diseases, which impact farmers and disrupt raw material sourcing for the food industry. It is estimated that these diseases lead to an annual crop loss of 38%, with black pod rot (BPR) and frosty pod rot (FPR) being of particular concern. BPR is present in all cacao-growing regions, while FPR is primarily found in the Americas. A solid understanding of the genetic basis for disease resistance is essential for developing tolerant plant varieties, and several loci associated with resistance to both diseases have been previously described, though revealing limited predictive power. In our study performed at Centro Agronómico Tropical de Investigación y Enseñanza (CATIE) Turrialba, Costa Rica, we evaluated eight cacao crosses by artificially inoculating them with the pathogens responsible for BPR and FPR and assessed disease resistance in 508 F₁ hybrid plants across multiple families. Concurrently, we generated whole-genome sequencing data for these hybrids and performed variant calling. Using a genome-wide association study (GWAS), we identified 15 genomic regions associated with disease resistance and pod production. These regions explained 7% to 37% of the genetic variation when combining multiple loci, underscoring the polygenic nature of resistance to both BPR and FPR. Given the quantitative nature of these traits, genomic prediction is expected to outperform marker-assisted selection (MAS). Our genomic prediction models demonstrated improvements ranging from 170% to 470% over MAS. Overall, our findings highlight the potential of integrating quantitative genetics and genomic tools to improve cacao

breeding, not only for disease resistance but also for other key traits, paving the way for the development of superior cocoa varieties.

Similar studies for coffee on genomic assisted breeding have started and are ongoing.

Interactive Visualization of Genomic and Pangenomic Data with Persephone

MAXIM Troukhan, Persephone Software, LLC, AGOURA HILLS, CA

Introducing Persephone, a web application for the visualization and exploration of genomic data tailored for comparative genomic analysis. Designed to display multiple maps simultaneously, Persephone ensures rapid graphical rendering with seamless animation.

The browser supports the visualization of a wide range of data types, such as genetic maps, genomic sequences, gene models, synteny, QTLs, variants, RNA-seq, NGS read alignments, and other genome-mapped features. Users can easily upload their data and execute real-time comparative analyses, including entire chromosome alignments, through the web application available at <https://web.persephonesoft.com>. The recent improvements enhanced SNP/indel analysis, added statistic calculations, and introduced a compact multi-map view suitable for pangenome visualization.

Demonstrations will showcase the functionality of this powerful tool across various species, including coffee.

Thank you note to our ICGN workshop speakers and attendants, as well as to the PAG-32 organizers that made our workshop such a great success!

There were 2,901 attendees representing 66 countries at the PAG-32 meeting. Our **16th ICGN Coffee Genomics Workshop** was offered in conjunction with PAG-32 meeting, January 10-15, 2025. The PAG-32 meeting brought together leading genomics scientists and researchers in plant and animal genomics/multi-omics research, and over 131 exhibits, 213 scientific workshops, 17 industry workshops, 828 posters and over 2,261 abstracts were presented.