☕ Innovea Breeders Meeting Key Information

Colombia, November 17 to 21, 2025

Monday, Nov 17, 2025

I. General Network Overview & Objectives

The meeting's primary goal was to **strengthen collaboration** across the network, particularly between World Coffee Research (WCR) and its partners, as well as among the partners themselves.

- **Network Status (Arabica):** The Innovea Arabica network has progressed from an idea to a reality. Most participants have received and planted their material, and the network is moving into the critical phase of **phenotypic evaluation**.
- **New Network Launch:** The meeting served as the official launch for the **Global Robusta Network**, which will operate under the "Innovea" umbrella. This network includes representatives from Ghana, Vietnam, Indonesia, Uganda, and Rwanda.
- **Recognition:** The collective work of the network was recognized by **Time magazine** as a **"Best Invention of 2025,"** highlighting the global, partnership-driven innovation needed in the coffee sector.
- **WCR Role:** WCR functions as an industry-driven research organization, representing over 200 member companies invested in securing the future coffee supply chain.
- **Future Focus:** The agenda includes deep dives into **phenotyping/data collection protocols, administrative topics** (budgets, plant importation), **variety release strategies**, and specialized **Genomic Selection** training
- **Breeding Strategy:** The Innovea Arabica network utilizes a **Population Improvement Strategy**, focusing on creating new diversity and improving genetic materials through continuous cycling, rather than solely a product development strategy aiming for immediate variety release.

II. Technical Protocols & Data Management

The meeting sessions emphasized the need for standardized and reliable methodologies across all international sites.

A. Experimental Design & Data Collection

- **Design:** Experiments use an **Incomplete Block Design** to account for block, row, and column effects, making data comparable across sites.
- **Tools:** Data collection uses **digital tools** like the **FieldBook app** on tablets to ensure consistency, minimize typos, and facilitate immediate data sharing and quality checks
- **Data Quality:** Data quality is paramount, as results are only as good as the data

collected. WCR implements rigorous checks for missing data and outliers.

B. Yield Component Data Collection (Faster/Quick Process)

This process aims to quantify the value of the seed itself and requires processing individual trees separately.

- **Target Cherries:** Cherries should be harvested as they ripen, slightly **red to full red** (excluding green).
- **Key Measurements: Fresh cherry yield** (total weight), **green bean weight**, **screen size** (using a single benchmark screen), **defective bean percentage**, and **seed density (SD)** are collected.
- **Processing:** Beans are processed immediately after harvest, depulped, and subsequently weighed for various components.

C. Quality/Post-Harvest Protocol (Ideal/Detailed Process)

This separate, more rigorous process is designed to isolate the genetic contribution to quality by controlling external factors.

- **Harvesting:** Only **fully ripe (deep red) cherries** are collected. Unripe, overripe, and floating (low-density) cherries must be **removed**.
- **Processing Timing:** Depulping should occur quickly, ideally within **four hours of harvest**, to prevent negative fermentation that introduces defects.
- **Fermentation (Washed):** Samples are fermented anaerobically (covered but not sealed) with clean water (approx. **30% volume** of wet parchment) for about **12–15 hours** (e.g., overnight) until the mucilage is removed (pH ~ 3.9–4.2).
- **Drying:** The parchment is dried slowly in thin layers (e.g., on African beds) and must be **covered overnight** after reaching about 25% moisture to avoid rapid drying, which reduces shelf life. Drying temperature should not exceed **40°C**
- **Storage:** Fully dried beans (10–12% moisture) must be stored in **hermetic bags** for a minimum of **two months** before hulling and tasting to allow for proper aging/maturation.

D. Quality Assessment Tools

- WCR is transitioning from the older **SCA (Specialty Coffee Association) 2004 Protocol** to the newer, more scientifically accepted **Coffee Value Assessment (CVA)**.
- The CVA includes a **Descriptive Assessment** (objective sensory profile) and an **Effective Assessment** (hedonic quality rating), providing broader, more consistent results for differentiation across genotypes than the old SCA form.

III. Country-Specific Updates on Innovea Arabica

Country	Institution (Local/WCR Partners)	Establishment Status/Key Findings	Challenges & Next Steps
Kenya	Kenya Coffee Research Institute (KARLO)	Plants established April 2024. Observing high genetic variation in **Coffee Leaf Rust (CLR)** resistance; some families consistently susceptible/resistant. Some early flowering observed (9–12 months post-planting).	**CLR Management:** Need to balance data collection with disease control (fungicide application) to avoid crop loss. **Data Analysis:** Need sharing of R scripts from WCR for local analysis. **Future:** Multi-location trials (MLTs), multiplying promising lines via grafting/tissue culture.
Peru	Instituto Nacional de Innovación Agraria (INIA)	Established December 2024. Material was kept in a post-entry quarantine area overseen by SENASA (National Agrarian Health Service). Currently collecting morphological data. Plot established under shade (Pacae trees) at **600m altitude**.	**Pests:** Facing issues with leaf miners and leaf-cutting pests due to low altitude. **Future:** Plan to validate best hybrids across **three ecological levels** (low, medium, high altitude).

Costa Rica	ICAFE CICAFE	Collecting data on plant height, stem diameter, CLR, and conducting the **first harvest**. Site is in the Central Valley at **1,000m altitude**. ICAFE is also heavily involved in hybrid development, e.g., Centroamericano, Esperanza.	**Environmental Stress:** Harvesting occurs under combined effects of rain/drought, looking for fruit drop resilience. **Pests/Diseases:** Concerned about **Ojo de Gallo (Mycenea)** and **Anthracnosis** in addition to CLR. **Limitation:** Labor shortages limit capacity for full fruit collection, requiring subsampling.
Mexico	Universidad Autónoma de Chapingo (UACH)	Established November 2024. Seed material received in Feb 2024 showed **67% germination** rate. Plot established under **46% diverse shade** cover at **1,350m altitude**. Observed first flowering trials one year after planting.	**Climate Extremes:** Used thermal mesh to protect plants from potential frost due to high elevation, observing highly variable weather patterns. **Breeding Focus:** Needs varieties adapted to agroforestry systems typical for small-scale Mexican producers.

Rwanda	Rwanda Agriculture Board (RAB)	Established in 2024. Site at **1,700m altitude**. Initiated first harvest, with samples ready for quality assessment.	**Breeding Goals:** High demand for **dwarf/semi-dwarf varieties** and materials that mature early. **Climate:** Facing erratic rains and irregular temperatures. **Future:** Utilizing high-performing hybrids as a "quick win" solution while pursuing long-term pure line development via seed propagation.
India	Central Coffee Research Institute (CCRI)	Innova samples (seedlings) were received via in-vitro transfer in Sept 2024 and are currently established in the nursery. Field planting is targeted for December 2025/January 2026.	**Logistics:** Significant challenges in **seed importation** led to the use of in vitro plantlets. **Pests/Disease:** Primary threats are **CLR** and **White Stem Borer** (WSB), requiring shade management. **Future:** Active participant in the new Robusta network due to low existing robusta germplasm.

Uganda	NaCORI / National Coffee Research Institute	Plants established in April 2024. Site is a hotspot for **CLR** and **Coffea Mealybug** at **1,400m altitude**. Initiated harvesting, collecting yield data.	**Pests/Disease:** Requires intensive insect management (Mealybug) and efforts for CLR resistance. **Robusta:** Has a strong history of robusta breeding (Uganda type, Congolese genetic group), with a recent focus on **drought tolerance**.
Vietnam	Western Highlands Agriculture and Forestry Science Institute (WASI)	**Robusta Focus:** Primary focus on Robusta breeding since the 1990s, selecting high-yielding clones. High Robusta yields (avg. 2.8 tons/ha; selected clones >4 tons/ha).	**Drought:** Utilizing irrigation cycles (25–30 days) and urgently breeding for **drought-tolerant Robusta** to reduce water use. **Pests:** Successfully controlling nematodes using biological control agents and resistant rootstocks.

Indonesia	Indonesian Coffee and Cocoa Research Institute (ICCRI)	**Arabica:** Low germination rate in previous attempts, preparing for a new seed shipment. **Robusta:** Over 60% of production is robusta, using clonal selection and recurrent reciprocal selection.	**Logistics:** Importing Arabica seed is highly **complicated and time-consuming** due to government procedures. **Pests:** Major challenges are **Coffee Berry Borer (CBB)** and **Nematodes**. **Consumer Preference:** Indonesian farmers prefer large-bean Congolese Robusta types.
Ghana	Cocoa Research Institute of Ghana (CRIG)	Long history of robusta breeding since 1966, initially focusing on local collections and later introductions from Ivory Coast and Togo. **Breeding Approach:** Uses both **direct clonal selection** and **hybrid programs** (bi-clonal seed gardens) focusing on high-yielding and distinct parents.	**Pest Control:** Developed successful methods to ensure isolation of seed gardens to maintain hybrid purity, as contamination from nearby plots significantly reduces yield. **Materials:** Recently released five new clones (2019) and four new hybrids (2025).

Tuesday, Nov 18, 2025

I. Cenicafé: Overview, Research, and Production System

Cenicafé is the **scientific center of the National Federation of Coffee Growers of Colombia**, founded in 1938. Its primary objective is to develop **scientific and technological research** applied to coffee production, focusing on increasing **productivity, quality, and sustainability**.

A. Regional Research and Experimental Stations

- Cenicafé operates **nine experimental stations** across the country's coffee-growing regions.
- Research is developed **at a regional level** for the specific coffee zones of influence of each station.
- The Colombian coffee sector is divided into three zones based on latitude (South, Central, North), each with different harvest cycles (one main harvest in the South and North, two harvests in the Central zone).
- The experimental station **El Naranjal** (located in the Central Mountain Range at 1,400 meters above sea level) serves as the primary site for the initial, critical stages of the breeding program.

B. Cenicafé Breeding Program Strategy

- The breeding program is structured around two main pillars: the **Colombian Coffee Collection** (germplasm conservation) and **variety development**.
- Breeding uses a **conventional pedigree method** involving selection of parental lines, cross-breeding, field planting, and advancing generations up to F4/F5.
- The early stages of selection are conducted at **El Naranjal** to ensure:
  - **Expression of full productive potential** due to optimal agro-climatic conditions.
  - **High pressure of Coffee Leaf Rust (CLR)**, as fungicides are **never** used in breeding plots to evaluate resistance effectiveness.
  - **Close supervision** by Cenicafé researchers.
- **Key Goals of the Breeding Program:**
  - **Durable resistance** against diseases (primarily **CLR**) and potential diseases like **Coffee Berry Disease (CBD)**.
  - **High production** (must produce at least as much as the current commercial variety).
  - **Broad adaptation** to Colombia's diverse, changing environmental conditions.
  - **High quality** with the characteristic Colombian sensory profile (a non-negotiable selection parameter).
- Since 1983, Cenicafé has developed **14 improved varieties**.
- **Impact: 88%** of the Colombian coffee area is planted with Cenicafé's CLR-resistant

improved varieties. This has saved farmers an estimated **$250 million USD per year** (by avoiding fungicide application) and contributed to sustainability by preventing the use of 1 million liters of fungicide annually.

C. Cenicafé: Seed Production and Distribution System

Cenicafé's seed production system is considered **unique worldwide**. It involves a fully integrated system from breeding to delivery, managed entirely by the institutional framework to serve Colombian coffee growers exclusively.

- **Infrastructure:** Cenicafé maintains **85 hectares** of a Seed Multiplier Park across its nine experimental stations.
- **Process Control:** To guarantee **genetic identity** of its multiline varieties, each row in the multiplication fields corresponds to a specific progeny line. Harvest combines these rows to ensure the final blend is correct.
- **Post-Harvest for Seed:** Strict standards are applied to preserve the embryo's viability:
  - **Depulping:** Revolutions per minute (RPM) are kept low (150-160 RPM) compared to commercial processing (200-220 RPM).
  - **Drying:** Begins with cold air (18 hours) before increasing temperature, never exceeding **36°C**.
  - **Moisture Content:** Distributed at **11-13%** moisture content.
  - **Viability/Storage:** The seed is typically distributed **fresh** (within 20 days of harvest). For temporary storage (e.g., during year-end holidays), it is kept at **10°C** with 65% relative humidity for about one month. Under controlled storage, viability above 60-70% can be maintained for up to six months.
- **Distribution & Traceability:** Seed distribution is mandated through **Service Level Agreements** with Departmental Coffee Committees, ensuring no third parties profit. Each bag includes a **QR code** for traceability, linking it to its lot, harvest, and processing history for quality control and claims.
- **Legal/External Control:** The system is regulated by the **ICA** (Colombian Agricultural Institute), which performs inspections and mandates quality parameters for viability (≥80% germination), physical health (≤1.5% defects), appearance, and identity.

II. Cenicafé Quality and Post-Harvest Research

The Quality Discipline measures quality in four dimensions: **physical quality** (moisture, defects), **sensorial quality**, **safety** (micotoxin prevention), and **chemical composition** (volatile and non-volatile compounds).

A. Cenicafé:Sensory and Chemical Findings

- **Genomic Quality:** A study on **373 Ethiopian accessions** identified a **34%** subset with differentiated sensory notes (citrus, floral, spice), highly desirable for quality

breeding.

- **Impact of Over-Ripeness:** Leaving fruit on the plant longer (to a 'grape' color/maturity stage three) does **not** increase total sugar content in the bean after ripening, but it favors sucrose hydrolysis (breakdown).
- **Quality Impact of Delay:** Delaying the harvest to the latest mature stage, when combined with fermentation delays, creates a **high risk of ferment defects** in the cup (88% chance in one study) due to the loss of physiological quality of the seed.
- **Fruit Reservation (48h):** Storing ripe cherry in common sacks (the most common farm practice) for **48 hours** before depulping promotes the formation of the volatile compound **ethyl acetate**. This shifts the sensory profile from classic notes (chocolate, caramel) to desirable notes (cherry, wine) characteristic of natural coffees.

B. Cenicafé:Quality/Processing Recommendations

- **Quality Standard:** Cenicafé uses the in-house **Fer Maestro** tool to determine the **optimal point of mucilage washing** to secure a coffee free of sensory defects.
- **Natural/Honey Process:** Cenicafé successfully developed protocols for obtaining **Natural and Honey coffees** with good cup scores, but advises farmers that this requires **2.3 times** the area needed for conventional drying, increasing infrastructure costs.
- **Water Activity (aW):** Measuring water activity is redundant when moisture content is correctly managed (10-12%), as water activity simply measures the physical condition at one point in time and does not capture the drying history (e.g., moisture spikes/interruptions).

III. Presentation of International Breeding Program Summaries to Cenicafé

Country/Organization	Focus/Goals	Key Breeding Methods & Achievements	Challenges

World Coffee Research (WCR)	**Singular focus on varieties** as a tool for farmers. Promotes **population improvement** in a globally coordinated network (Innovea).	Coordinates **Arabica** (launched 2022) and **Robusta** (formally announced at this meeting) breeding networks across multiple environments.	Need for rapid development due to climate change.
Kenya (KARLO)	**Yield** (target 4 tons/ha for short varieties), **Quality** (80% Grade 16+; 60%+ specialty coffee ≥ 85 points), and **Resistance** to **CLR** and severe **CBD**.	Breeding since 1920s; released SL 28, SL 34, K7. Resistance breeding (since 1970) uses **gene stacking** (Sudan-R, Timor-H, Kent-R). Released **Ruiru 11** (hybrid, 1985) and **Batian** (pure line). Uses **hypocotyl inoculation** for early CBD resistance selection at 6 weeks.	**CBD** is the main challenge (80% crop loss risk). Difficulties with **hybrid production** (manual hybridization/vegetative) and failed **tissue culture** program. Root mealybug becoming a major problem.

Uganda (NaCORI)	Ambitious goal of **20 million bags/year by 2030**. Focus on **doubling/tripling productivity** in traditional areas and **expanding to new areas** with **drought/heat tolerance**.	Can produce both Arabica and Robusta. Developed 10 Robusta varieties resistant to **Coffee Wilt Disease** (Fusarium). Screened 300 Robusta accessions, finding 17 promising for drought tolerance.	**Seedling multiplication** is a key challenge (currently clonal/vegetative propagation). Need for capacity in multiplication techniques. Diseases: Coffee Wilt Disease (Robusta), Red Blister, Black Twig Borer, Coffee Berry Borer.
India (CCRI)	Develop **rust-tolerant Arabica** and **drought-tolerant Robusta** varieties. Must maintain shade-grown, agroforestry systems.	Focuses on stabilizing rust tolerance by F5/F6 generation. Exploring faster **F1 hybrid strategy** with large-scale clonal multiplication via **tissue culture** (PPA with Jain Irrigation). Released three Robusta varieties, including an **interspecific hybrid** (C. congensis x C. canephora). Uses **SCAR markers** for rust resistance screening in	Need for more trait-specific markers (e.g., root biomass/systems).

		breeding lines.	
Mexico (UACh)	Research covers the entire value chain, focusing on **docencia** (teaching), **investigación** (research), and **extension** (infancia, youth, gender). Manages the **National Coffee Germplasm Bank**.	Part of the **WCR Innovea project** (since 2022) to evaluate promising materials. Contributes to the national **agroecological transition** and the *Sembrando Vida* program.	Loss of a national coffee institute (IMECafe) in 1990 led to universities assuming research roles.
Costa Rica (ICAFE)	**Validating different cultivars** (many Brazilian) across seven regions/altitudes and microclimates. New focus on breeding for resistance to **CLR** and **CBD**.	Using **molecular markers** to screen for rust resistance/identity at the seedling stage and verify crosses. Innovea crosses are showing promising material.	Early breeding efforts are only a few years old.

Rwanda (RAB)	Goals are high yield, improved quality, disease resistance (CLR, CBD), and **climate resilient varieties**. Needs **"quick wins"**.	Three-track strategy: 1) **Introductions** (quick solution, ~ 5 years); 2) **Hybrids** (medium term, ~ 15 years), relying on **clonal propagation** (tissue culture/lab); 3) **Finished Varieties** (long term, ~ 25-30 years) by seed propagation.	Need for reliable, genetically pure clonal propagation of hybrids.
Vietnam (WASI)	Main focus is **Robusta** breeding. Current focus is on **cup quality, drought tolerance, and uniform ripening**.	Released **20 varieties** since 1995 (14 Arabusta, 6 Arabica). Varieties are known for **high yield, good bean quality, and CLR resistance**.	-
Ghana (CRIG)	Focus is on **Robusta**. Main goals are **productivity and cup quality**.	Develops both **clonal** and **hybrid** varieties. Seedlings are distributed **free** to farmers (minimum 1 million/year).	**CLR** is not a major problem currently. **Coffee Berry Borer** is a major problem.

CATIE	Regional center (Central/South America) for **conservation** and **breeding**. Holds one of the world's most important collections (4th largest, 2nd in LatAm), including ~40% wild species.	Pioneer in F1 hybridization (since 1990s). Released three F1 hybrids: **Centroamericano, Milenio, and Esperanza** (high yield, CLR tolerance, exceptional cup quality, **climate resilience**). Uses **somatic embryogenesis** and **cutting rooting** for clonal propagation (released 20M+ plants).	Primary challenges are **nematodes** (use of C. canephora rootstock) and **climate change/adverse conditions**.

Wednesday, Nov 19, 2025

Cenicafé: Near-Infrared Spectroscopy (NIR) for Coffee Quality

Cenicafé utilizes NIR technology to safeguard the **quality and reputation of Colombian coffee**.

- **Core Concept:** NIR detects the **unique spectral fingerprint** of coffee beans, enabling the identification of chemical compounds, origin, and quality. This technique is implemented at Colombian ports.

- **Adulteration Detection:** Models predict if coffee samples contain **mixtures** (adulteration) with coffee from other countries (e.g., Vietnam, Brazil).

  - NIR is crucial because sensory analysis alone cannot reliably detect adulteration below 2-3%.

- Models for detecting mixtures of national and foreign coffee have shown an **accuracy of 95%**.

- **Chemical Profiling:** Cenicafé has predictive models for **13 chemical compounds**.

  - **Caffeine** is a key indicator: *C. canephora* (Robusta) has >2% caffeine, while *C. arabica* (Arabica) has a maximum of 1.4%. An increased caffeine reading in a shipment of alleged Arabica coffee immediately flags it as suspicious.

- **Sensory Defect Screening:** Models can predict if parchment or green coffee has a **sensory defect** *before* roasting and cupping.

  - This is possible because defects primarily occur during the collection and drying/storage stages, affecting the parchment/bean first.

- **Portability:** Cenicafé is validating **portable NIR equipment** to move quality controls closer to collection centers, instead of waiting for shipment at the port.

- **Model Robustness:** Models are trained using the **full range of defects** (e.g., earthy, fermented, immature) and a wide range of quality scores to ensure high accuracy in predicting quality.

Molecular Breeding Tools

World Coffee Research (WCR) presented efforts to develop molecular tools to accelerate breeding through marker-assisted selection (MAS) and genomic prediction.

- **The Need:** Currently, there's no widely available **mid-density marker panel** for *C. arabica* and *C. canephora* (Robusta) that is accurate and cheap enough for routine use.

- **Arabica Mid-Density Panel (MDP):**
  - Developed by sequencing 94 accessions (the **Discovery Panel**) contributed by WCR and partners (CATIE, EMBRAPA, Rwanda RAB, USDA, others).
  - The goal was to filter ~252,000 initial markers down to ~ 5,000 markers that cover the genome effectively.
  - This MDP is intended for applications like **quality control, genomic prediction, and trait mapping** (e.g., rust resistance).

- **Robusta MDP Plan:** WCR is soliciting collaboration to develop a Robusta MDP to address the long **seed-to-seed cycle**.

  - **Value Proposition:** The panel can help **cut the breeding cycle time** and is needed

because most important robusta traits are **quantitative** (requiring predictive models) and sensitive to the environment.

- ○ WCR is asking for partners to contribute diverse wild, semi-wild, and cultivated robusta tissue samples for sequencing.

- **Collaborative Trait Mapping:** A specific project is underway (WCR, Cenicafé, USDA, Puerto Rico) to identify **Coffee Leaf Rust (CLR) resistance genes** and develop usable markers.

  - ○ This involves phenotyping three segregating F2 populations against three different rust isolates.

Emerging Issues and Next Steps

The discussion identified three key areas for breeding focus, reflecting current industry disruptions and opportunities:

Emerging Issue	Breeding/Research Opportunities
Disease Resistance	Develop a **unified screening system/protocols** for resistance evaluation across the network for data compatibility. Implement a **shared surveillance platform** to track disease incidence/evolution globally. Research **emerging diseases** (beyond CLR/CBB) to prepare for the future. Use **QTL mapping** and **interspecific crosses** to identify new resistance sources.
Climate Resilience & Abiotic Stress	Focus on **plasticity and broad adaptation** (handling variable stresses) rather than single specific stressors (e.g., just drought). Use **gene stacking** for multiple stressors. Optimize performance in **agroforestry systems**. Research **root performance** (e.g., grafting, root traits) for improved below-ground stress management.

Harvest Timing & Labor Efficiency	Breed for **uniform ripening/synchronized flowering** to concentrate harvest labor, or breed for **early and late varieties** to extend the harvest window and prevent labor bottlenecks. Select for plant architecture/physiology traits amenable to **mechanization**.

Robusta Breeding Network Design Decisions

The Robusta network will distribute the **same 1,000 elite genotypes** (clones) to every partner, as opposed to the Arabica network, where families are shared but individual plants vary.

Field Trial Design (Vote)

The Robusta partners voted on the most appropriate field design, balancing scientific rigor, land area, and labor costs.

- **Decision:** Partners voted for **Option B** as the design template.

- **Option B Summary:** This choice represents a **moderate-resource** approach with **partial replication** of test clones.
  - **Rationale:** It balances the scientific ideal (Option A) with the practical ability to manage the trial, collect data, and utilize available land/labor.

Priority Traits (Agreement)

The Robusta partners reviewed and agreed upon a list of core traits to be measured by all members of the network, recognizing the need to be realistic about the workload.

Trait Category	Core Priority Trait	Rationale/Comments
Yield & Production	**Cherry Yield per Tree**	Fundamental economic trait.

	Final Total Yield (kg/plot)	
Physical Quality	Bean Size	A key market requirement for Robusta.
	Bean Density	
	Defective Bean Percentage	Important for quality and market access.
Biotic Stress	Coffee Leaf Rust (CLR)	Remains a core disease to track, especially due to evolving variants.
Plant Architecture	Plant Height	Correlates to planting density and yield potential.
	Stem Diameter	Proxy for vigor and stability.

Thursday, Nov 20, 2025

Variety Development and Release

The morning session was dedicated to what happens **after breeding**, specifically focusing on variety development and release. World Coffee Research (WCR) views its entire research portfolio and effort as focused on **varieties** and a **pipeline** that flows from upstream breeding to downstream processes.

WCR's Variety Creation Strategy

WCR is focused on creating finished varieties out of the Innovea materials, particularly for countries not in the breeding network. The goal is to get new varieties into the hands of farmers. WCR's finished variety creation streams include:

- **Seed-propagated Arabica varieties:** Considered the most easily distributed and highly demanded by most producers.

- **Hybrid Arabica:** Production and distribution are planned for specific markets where farmers are ready to manage them appropriately.

WCR is operating with a target of **2030** for the first Arabica variety candidates to be ready for placement in trials.

Strategies for Speeding up Breeding

The traditional path for developing an inbred line (starting with F2 and evaluating for years, including two years of high yield for a good estimate) can take up to 24 years, plus the multi-environment trial time. WCR and partners are exploring strategies to reduce this time, such as:

Strategy	Timeframe to F5 (Approx.)	Pros	Cons
Traditional Path	24 years	Medium to high accuracy due to large dataset	Very long time
Short Evaluation	12 years	Faster	Lower accuracy due to less data
Speed Breeding	9 years	Much faster process (assuming seed-to-seed in 2 years)	Low accuracy

Mitigation Strategies for Low Accuracy (Shortcuts):

- More intensive use of **molecular markers** for specific disease resistances and **genomic prediction** for quantitative traits.

- Conducting **larger and/or intermediate multi-environment trials (METs)** earlier in the process (e.g., planting F4 in relevant environments for a proper, lengthy evaluation).

- **Early discarding** of inferior genotypes (e.g., about 30% of low yielders after the initial harvest) is feasible, as they tend to perform poorly early on.

- The optimal generation for variety release varies by country and institution, typically ranging from **F4 to F7**.

Variety Release and Registration Protocols

Variety registration is described as a **complicated and bureaucratic** process that differs in every country and can pose a significant **bottleneck** to farmers accessing new materials.

Country	Variety Release Custodian	Key Processes/Requirements	Protection/Registration Details
Kenya	National Variety Committee (NVC)	Distinct, Uniform, and Stable (**DUS**) testing is requested by the institution. DUS is only for the best-performing lines, while the International Performance Trial (**IPT**) is done by the institution itself. The NVC (involving researchers from universities, institutions, and farmers' orgs) discusses the DUS report and grants release.	Varieties are **protected** (requires an annual fee of $1,000 USD). Protection period is **20 years**, but discussions are ongoing to extend this to 30 years for tree crops due to the long development time (15–20 years).
Rwanda	Rwanda Inspectorate Competition and Consumer Protection Authority (**RICA**)	Requires filing an application and passing the following official variety testing: **DUS** (distinct, unique, and stable in terms of time and environment) and **VCU** (Value for Cultivation and Utilization). Testing requires at least **two seasons and three sites**.	Registration can be automatic if the variety is registered in a COMESA (Common Market of East and South Africa) member country. A registered variety can be **suspended or vacated** if requirements are not met.
Vietnam	Plant Protection Department (PPD) of the Ministry of Agriculture (MoA)	**Import/export** of varieties is legally difficult and lengthy, taking at least two months.	Mostly recognizes **clonally-propagated Robusta**. Only **one** variety (TR1) is

		Requires **VCU testing** (Value for Cultivation and Utilization) per national standards. **Controlled evaluation** focuses on rust resistance (must be high) and nematode resistance (both mandatory for Robusta). **Field evaluation** is on narrow areas, then multiple regions (3 for Arabica, 1 for Robusta, at least 3 sites per region). Requires **two successful harvest seasons** (after the first) with high standards for yield, bean quality (size, shape), and cup quality, taking at least **5 years**. **DUS testing** is only mandatory if protection is requested.	recognized as seed-propagated.
Mexico	Servicio Nacional de Inspección y Certificación de Semillas (**SNICS**)	Two procedures: 1) Obtaining a **Breeder's Title** (for new varieties) and 2) Registering a **Common Use Variety** (for materials past their protection period). Both require a review by an expert committee, an examination for **Distinction and Novelty**, and characterization/technical sheets. Registration legalizes the material and allows **seed certification**.	Mexico is a member of UPOV. Breeder's rights (ownership) are **18 years** for perennial species (like coffee) and 15 years for others. Cost for registering a new variety is approximately **$2,070 USD** plus an annual renewal fee of about $150 USD.

Administrative and Operational Topics

Funding and Budgeting

The establishment and operation of the Innovea Arabica and Robusta networks are **100% covered by World Coffee Research (WCR) members** (private funding). WCR establishes budgets based on **actual operational costs**, using its El Salvador farm (a dollarized, high-labor-cost economy) as a baseline, and adjusting for local restrictions and costs. Due to the private funding, there are **high standards for accountability and transparency**.

Capacity Building and Training

WCR actively incorporates **capacity building**, training, and skills development into the Innovea design.

- **Current Training:** Continuation of seminars on **experimental design and analysis** and a shift toward **Genomic Selection** and managing genetic data.

- **Requested Topics for Future Training:** Intellectual property (IP), targeted population environment (TPE), Coffee Value Assessment (CVA) training (quality assessment), seed production/quality assurance, and communication skills.
- **Expert Resources:** The technical team (Jorge, Santos, Robert, etc.) are key resources available to talk about processing, quality evaluation, experimental design, and genetic markers.

Publications and Intellectual Property (IP)

WCR's research integrity policy prioritizes **impact** first, so publications come second. The core principles for publication are:

- **Credit where credit is due:** All research participants, especially those with sites in collaborative trials like the International Multi-Location Variety Trial (IMLVT), must receive **co-authorship**.

- **Seek permission beforehand:** Partners must communicate with WCR before publishing results from joint work (courtesy and quality control). This applies to scientific literature, conference proceedings, and public media.

- **Authorship Order:** Authors will be listed alphabetically, except for the first and last authors, to avoid issues of seniority.

Intellectual Property (IP) of Germplasm:
- **Arabica Network:** The material planted by partners is owned by the partner institution; they are free to use it for pre-commercial trials or hybrids with no restrictions from WCR

(provided they continue collecting data).

- **Robusta Network:** The clones are owned by WCR. Partners can use the clones as parental lines for their own breeding programs without WCR involvement. However, if a partner wants to move a clone out to farmers or put it into official trials, it requires a separate negotiation with WCR, as WCR owns the clones and must maintain control.

Germplasm Shipment Logistics

The **movement of germplasm** is critical for breeding programs but also poses a **risk of pathogen transmission**. Shipment logistics often prove to be a **critical bottleneck**, delaying the establishment of the network.

- **Key Requirements:** The two most important documents required are the **Import Permit** and the **Phytosanitary Certificate**. These documents must be obtained **on time** for the shipment of live seeds or plantlets.

- **Robusta Shipments:** The robusta material, shipped as plantlets, is considered **more vulnerable** and expensive to replace if a shipment fails. Shipments will occur in **Quarter 4, 2026, into Quarter 1, 2027**.

- **Process:** The roadmap involves reviewing import/export requirements, aligning propagation facility procedures, shipment, and successful introduction/adaptation at the importing country's facilities.

Friday, Nov 21, 2025

🧬Workshop on Genomic Selection

Overview of **Genomic Selection (GS)**, its implementation within breeding programs, the required statistical and computational methods, and strategic considerations for its successful adoption.

1. Genomic Selection (GS) Fundamentals

Genomic Selection is a predictive modeling approach that uses **all available markers simultaneously** to predict the performance of an individual, primarily focusing on **additive genetic effects**.

- Core Concept: GS is essentially a multiple linear regression where the breeding value (the trait response, A_j) of an individual is predicted using thousands of marker effects (β_i) weighted by the individual's marker codes (X_{ji}, typically 0, 1, or 2 for SNPs).

$$A_j = \mu + \sum_{i=1}^{P} X_{ji}\beta_i + e_j \quad \text{or in matrix form} \quad \mathbf{A} = \mathbf{X}\beta + \mathbf{e}$$

- ○ **The Challenge:** The number of markers (P, up to 30,000 to 50,000 or even 3 million SNPs) is often **much larger** than the number of individuals (N, e.g., 1,000 to 2,000), which creates statistical issues like overfitting and collinearity
- ○ **Mitigation:** Techniques like **ridge regression** (used in GBLUP) are employed to manage collinearity and shrink the parameter estimates.
- **Purpose:** The primary goal is to create a **predictive model** (a "black box") of performance, not to pinpoint specific significant quantitative trait loci (QTLs) like in GWAS.

---

2. Benefits of Genomic Selection

Implementing GS can significantly improve a breeding program's efficiency.

- **Decreasing Generation Cycle:** GS can reduce the generation cycle, especially in perennial crops like coffee and trees, by allowing breeders to select the best parents much **earlier** (e.g., as seedlings in a greenhouse) instead of waiting years for field performance data.
- **Cost Reduction:** It decreases the cost of **testing and phenotyping** by allowing early discarding of inferior genotypes or predicting phenotypes for expensive, laborious, or destructive traits (e.g., meat quality in salmon or disease resistance).
- **Increased Intensity:** It enables **screening a large number of genotypes** (e.g., starting with 3,000 seeds and sending only the top 20% to the field), thereby increasing selection intensity.

---

3. Key Methodologies in Genomic Selection

GS models can be categorized into two statistically equivalent, but computationally different, types: **Marker-based Models** and **Relationship-based Models**.

Model Type	Primary Methodologies	Description & Key Features

Marker-based (SNP-based)	GSR/RR-BLUP (Ridge Regression BLUP), **BayesA, BayesB, BayesC**, **Bayesian LASSO**	Estimates a β parameter for **every marker**. Often computationally faster but requires careful handling of P > N. BayesB assumes a proportion of marker effects are exactly zero, which can lead to better performance for traits controlled by a few major QTLs.
Relationship-based	GBLUP (Genomic BLUP), **ssGBLUP** (Single-Step GBLUP)	Does not model individual markers but rather models the individual genetic effects using a **Genomic Relationship Matrix (GRM)**. This allows for the incorporation of complex mixed-model effects (e.g., multiple environments, blocks, spatial components) and raw data.

4. Data Preparation: The Two-Stage Process

Successful GS relies heavily on high-quality input data, which involves a multi-step preparation

A. Phenotypic Data Preparation

The first stage involves taking raw replicated/multi-environment trial data and converting it into a **single, adjusted phenotypic response** per genotype.

- **Adjustment:** Use a linear mixed model to correct the raw data for all non-genetic influences (e.g., year, block, lot, spatial components, GxE interaction) to get the "best guess" of the genotypic mean. These are often called **Adjusted Means** or **BLUES** (Best Linear Unbiased Estimates).
- **Heritability Check:** Before proceeding, test the genetic variance (heritability) of the raw data (using a random effect model) to ensure the trait is under sufficient **genetic control** and thus worth predicting.

B. Genomic Data Preparation

The marker matrix (M-matrix, N x P) must be meticulously cleaned.

- **Marker Coding:** Convert marker data (e.g., A/C/T/G) into numerical format (e.g., 0, 1,

2) based on a designated reference allele.

- **Filtering:** Eliminate markers with insufficient information, such as those with a low **Minor Allele Frequency (MAF)** (e.g., less than 5%), markers that are fixed, or individuals (subjects) with excessive **missing values** (e.g., over 10-15%).
- **Imputation:** For data with high missingness (e.g., 40% in GBS data), employ sophisticated imputation methods (like **Beagle** or **MaCH**) or use simple mean imputation for low missingness (<5%).
- **Genomic Relationship Matrix (GRM):** Calculate the G-matrix, which defines the observed genetic relationships between all individuals based on their markers. The G-matrix often needs **tuning** (e.g., blending with an Identity matrix or a Pedigree-based A-matrix) to ensure it is invertible and stable

---

5. Implementation Strategies

Integrating GS requires adjusting breeding practices to leverage its unique benefits.

- **Training Population:** The **training population** (the individuals used to build the predictive model) must be genetically representative of the target selection population for the predictions to remain accurate over future generations.
- **Reducing the Cycle:** In a typical plant breeding cycle, GS is used to select elite candidates early (e.g., F1/F2 seedlings) for advancement or to become parents, often allowing the elimination of a year or more of field trials.
- **Single-Step GBLUP (ssGBLUP):** For programs with valuable historical data (like in forestry or animal breeding), the Pedigree Matrix (A) and the Genomic Matrix ($\mathbf{G}$) can be combined into a **Hybrid Matrix (H)**. This model allows non-genotyped individuals (e.g., historical parents or non-selected siblings) to contribute to the breeding value estimation, maximizing the use of all accumulated data.
- **Genomics and AI:** While current Genomic Selection models use fundamental biological principles (additive inheritance) making them robust, **Artificial Intelligence (AI)/Machine Learning (ML)** tools are expected to emerge to improve:
  - **Phenotypic Data Quality:** Creating new, high-throughput proxies for complex traits (e.g., drone-based phenotyping for flower/fruit count).
  - **Predicting Non-Additive Effects:** Eventually improving the prediction of dominance or epistatic interactions, especially as data volume increases beyond the current limited scale.

---

6. Measures of Success and Model Validation

- **Predictability (P_A):** The correlation between the phenotyped value (Y) and the Genomic Estimated Breeding Value (G).

- **Accuracy (A_{cc}):** The measure of predictability divided by the square root of the trait's heritability ($\sqrt{h2}$). This represents the correlation between the **estimated** breeding value and the **true** (but unknown) breeding value. An accuracy value of over 1.0 is a strong indication of model **overfitting**.
- **Validation:** To obtain a reliable estimate of predictability, the model must be validated.
    - **True Validation:** Training the model on past data and predicting the performance of an independent, younger/future generation.
    - **Cross-Validation:** Splitting the existing data into multiple subsets (folds). The model is trained on most folds and predicts the remaining one, repeated until all subsets have been predicted. This provides a less inflated, though still imperfect, estimate of predictive ability.