Title: Decoding pathogen ecological memory from chemical-stress phenotypes

Running title: Decoding fungal ecological memory

Insuck Baek[1], Seunghyun Lim[2], Amelia Lovelace[2], Sookyung Oh[1], Masoud Kazem-Rostami[3], Helen Ngo[3], Moon S. Kim[1], Lyndel W. Meinhardt[2], Lalit Kandpal[1], Minhyeok Cha[1], Chansong Hwang[1], Richard D. Ashby[3], Ezekiel Ahn[2,*]

[1] Environmental Microbial and Food Safety Laboratory, Agricultural Research Service, United States Department of Agriculture, Beltsville, MD, 20705, USA

[2] Sustainable Perennial Crops Laboratory, Agricultural Research Service, United States Department of Agriculture, Beltsville, MD, 20705, USA

[3] Sustainable Biofuels and Co-Products Research Unit, Agricultural Research Service, United States Department of Agriculture, Eastern Regional Research Center, Wyndmoor, PA, 19038, USA

* Author to whom correspondence should be addressed (email: ezekiel.ahn@usda.gov; ORCID: https://orcid.org/0000-0003-4447-410).

Abstract

Can standardized chemical stress reveal an ecological "memory" of host origin in plant pathogens? We profiled six coffee-associated *Colletotrichum* isolates using four phenolic-branched compounds and compared them with a previously characterized cacao panel. Quantitative morphology, hyperspectral imaging (HSI), and supervised machine learning (ML) yielded pathosystem-specific fingerprints under uniform, isotropic *in vitro* conditions. Circularity, a measure of edge symmetry, was the most informative morphological feature, and ML classified host origin with 87% accuracy in cross-validation. HSI detected dose-dependent physiological shifts, including changes near 1930 nm in the short-wave infrared consistent with water-associated bands. Multi-locus phylogeny showed the coffee isolates are polyphyletic, indicating that these fingerprints likely reflect convergent adaptation rather than shared ancestry. We propose a "chemical priors" framework in which long-term environmental exposure imprints stress-response

pathways that become legible under simple, standardized probes. This integrative workflow supports scalable screening of eco-friendly antifungals and precision disease management.

Keywords: Ecological Memory, Machine Learning, Fungal Adaptation, Pathosystem-specificity, Hyperspectral Imaging, Morphological Fingerprint, *Colletotrichum*, Phenolic Antifungals

1. Introduction

Fungal pathogens like *Colletotrichum* threaten global food security, creating an urgent need for novel, sustainable control strategies that require a deeper understanding of pathogen adaptation [1–5]. A fundamental challenge is decoding an organism's adaptive history from its observable traits, a problem we address here using an integrated spectroscopy and machine learning approach. To this end, we turned to the anthracnose pathogen *Colletotrichum*. While notorious for its broad host range, we selected its interaction with two globally significant perennial crops, cacao (*Theobroma cacao*) and coffee (*Coffea arabica*). Their distinct chemistries and ecologies make them ideal natural experiments for testing the generality of our core hypotheses. For decades, the primary defense against such pathogens has been the application of synthetic chemical fungicides. However, this strategy is increasingly challenged by the evolution of fungicide-resistant pathogen strains and growing concerns about the environmental and toxicological effects of chemical residues [6–8]. This creates an urgent need for novel antifungal strategies that are not only effective but also ecologically sustainable and are often sought from natural sources [9, 10]. This led us to a central hypothesis: that a pathogen's observable response to chemical stress is an echo of its ecological past. To test this, we developed a novel, high-throughput workflow combining quantitative morphology, hyperspectral imaging (HSI), and machine learning (ML). We contend that this integrative platform provides a powerful new lens for studying microbial adaptation and has significant practical applications in phytopathology.

The distinct ecologies of coffee and cacao pathogens provide an ideal natural experiment to deploy this workflow. Here, we applied our integrative approach to decode what we call 'chemical priors'—stress-response phenotypes that reveal a pathogen’s ecological memory. To elicit these phenotypes, we utilized a 2x2 matrix of novel, stabilized phenolic-branched antifundals derived from sustainable feedstocks [11–15]. These compounds, with systematic variations in their

phenolic headgroups and backbones, allow us to probe the structural basis of antifungal activity, building upon our initial findings in the cacao pathosystem where specific derivatives induced reproducible morphological fingerprints [16].

This observation prompted a central question: do pathogens from different agroecosystems carry a 'memory' of their origin, one that is legible in their response to chemical stress? From this question arise two competing hypotheses. This memory could be imprinted by either the immediate host-pathogen interaction or by larger-scale biogeographical factors. The distinction is critical, as it determines whether a morphological response indicates a universal chemical mechanism or a pathosystem-specific adaptation. Given the contrasting biochemical environments of cacao and coffee plantations, each defined by distinct host metabolites, microbial communities, and abiotic pressures, it is plausible that fungal pathogens from these systems have evolved divergent physiological strategies for coping with xenobiotic stress (i.e., exposure to biologically foreign compounds, including our phenolic derivatives). For instance, cacao agroecosystems are typically characterized by the humid, shaded conditions of their native tropical understory [17], while arabica coffee is frequently grown in more open, higher-altitude environments, often above 1,000 meters [18, 19]. Furthermore, the host plants themselves produce distinct classes of defensive secondary metabolites that create unique chemical landscapes. Cacao relies on a defense chemistry rich in methylxanthines, primarily theobromine and caffeine, as well as compounds from the phenylpropanoid pathway [20]. In contrast, coffee leaves maintain high concentrations of caffeine and chlorogenic acid to suppress the growth of microbial pathogens [21]. To test this hypothesis, we utilized the identical four phenolic-branched compounds previously tested in cacao: a phenol-soy oil fatty acid (PhSOFA) and a beechwood creosote-branched grease fatty acid (BCBGFA), both consisting of a phenolic headgroup covalently linked to a fatty acid backbone, and their respective amide analogs, PhSOAM and a beechwood creosote-branched grease amide (BCBGAM), which replace the terminal carboxylic acid with an ethylenediamine-derived amide group. PhSOFA contains a simple, unsubstituted phenol, while BCBGFA incorporates a complex mixture of substituted phenols derived from beechwood creosote. Their amide counterparts (PhSOAM and BCBGAM) preserve the phenolic identity but differ in their increased polarity and nitrogen content. These systematic structural variations allow us to probe how differences in phenolic complexity and headgroup chemistry influence antifungal activity and the resulting

morphological response. Thus, the four compounds represent a 2 × 2 matrix of structural variation: simple vs. complex phenol and acid vs. amide headgroup, allowing direct inference of both phenolic and backbone contributions to antifungal activity.

Our primary goal was to establish and validate an integrated workflow that can decode these signatures. By applying this platform to the contrasting coffee and cacao pathosystems, we aimed to: (1) quantify compound-induced morphological fingerprints; (2) use ML to test their predictive power for pathogen origin; and (3) use HSI to explore the underlying physiological basis of these predictive phenotypes.

2. Materials and methods

2.1 Biobased compounds tested as antifungal agents

Four previously described phenolic-branched compounds (PhSOFA, PhSOAM, BCBGFA, and BCBGAM), synthesized from biobased feedstocks, were used in this study [13, 15]. Stock solutions (10 mg/mL) were prepared in pure ethanol for all assays.

Fungal isolates and culture conditions

Six coffee-pathogenic *Colletotrichum* spp. isolates were used, with origins and host details provided in Table S4. For comparative analyses, data from three previously characterized cacao-pathogenic isolates were also included. All isolates were maintained on Potato Dextrose Agar (PDA) at 24°C in the dark.

2.2 Phylogenetic analysis

To determine the evolutionary relationships among isolates, we performed a multi-locus phylogenetic analysis using the internal transcribed spacer (ITS), glyceraldehyde-3-phosphate dehydrogenase (GAPDH), and actin (ACT) gene regions. Phylogenetic trees were constructed in MEGA 11 using the Neighbor-Joining method with nodal support assessed by 1,000 bootstrap replications. Detailed procedures for DNA extraction, PCR amplification, sequencing, and alignment are provided in the Supplementary Methods.

2.3 Antifungal activity assay and morphological Analysis

Antifungal activity was assessed using a modified agar plate assay. Each compound (150 µg) was spread on PDA plates, with ethanol-treated plates serving as controls. A 5-mm mycelial plug of each isolate was placed at the center and incubated for 96 hours. Digital images of the colonies were captured under consistent lighting conditions. Colony morphology was quantified using SmartGrain software to measure seven traits: area, perimeter, length, width, length-to-width-ratio (LWR), circularity, and an asymmetry distance ((the distance between the intersection of length and width and the center of gravity (IS & CG)) [22].

2.4 Statistical and machine learning analysis

All statistical analyses were performed using JMP Pro 17 [23]. For the comparative analysis, the coffee pathogen dataset was merged with the previously published cacao pathogen dataset [16]. One-way ANOVA followed by Tukey's HSD test was used for pairwise comparisons. For multivariate analyses, all morphological data were log-transformed. We used Principal Component Analysis (PCA) to visualize overall relationships and Permutational Multivariate Analysis of Variance (PERMANOVA) to test for group differences. To predict a pathogen's host of origin, we evaluated a suite of machine learning classification models using a repeated K-Fold cross-validation strategy (k=5) on a class-balanced dataset. Full details on data transformation and model parameters are available in the Supplementary Methods.

2.5 HSI and data analysis

HSI was conducted on select isolates cultured on PDA with varying compound concentrations (0, 150, 300, and 450 µg per plate). Three imaging modalities were employed using custom-built line-scan systems: fluorescence (450–720 nm), visible and near-infrared (VNIR) reflectance (430–1000 nm), and short-wave infrared (SWIR) reflectance (1000–2400 nm). Spectral data were extracted from standardized regions of interest (ROIs), calibrated, and analyzed using one-way ANOVA F-statistics to identify wavelengths with significant treatment-induced variation [24]. Detailed system specifications are provided in the Supplementary Methods.

3 Results

3.1 Potent but distinct effects of amide derivatives PhSOAM and BCBGAM

157 The two ethylenediamine-carrying amide derivatives, PhSOAM and BCBGAM, were among the
158 most potent growth inhibitors tested. PhSOAM exhibited potent but selective antifungal activity,
159 reducing colony area in the most susceptible isolates, P24-192 and P24-85, by 19.4% ($p < 0.001$)
160 and 20.7% ($p < 0.001$), respectively (Fig. S1A). This reduction is visually illustrated for P24-192
161 (Fig. S1B, C). In contrast, BCBGAM induced slightly stronger inhibition in certain isolates (e.g.,
162 20.1% in P24-192) but was less effective against others (Fig. 1A). While both amides reduced
163 size metrics, their shape metrics (circularity, LWR, IS&CG) diverged, underscoring a size–shape
164 decoupling central to this study. PhSOAM triggered diverse morphological responses, in some
165 cases inducing irregular expansion (e.g., a decrease in circularity in isolate P24-193) (Fig. S1A).
166 Conversely, BCBGAM promoted a more compact, rounded morphology in three of the six
167 isolates (Fig. 1A). The most striking response was observed in isolate P24-192, where BCBGAM
168 treatment induced a significant increase in colony elongation (LWR), as illustrated in the
169 representative images (Fig. 1).
170

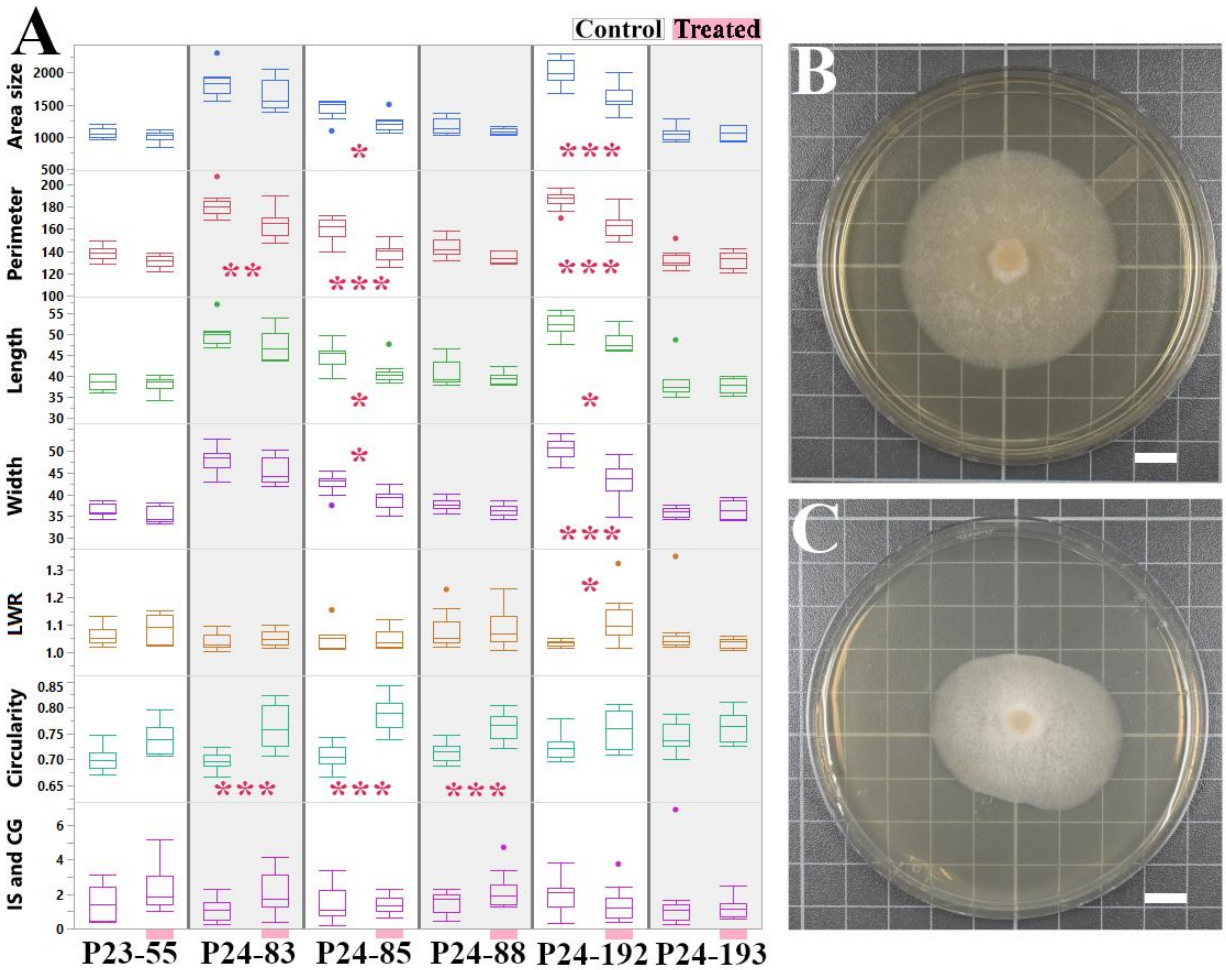

Fig. 1. Representative morphological disruption by phenolic antifungal compounds, illustrated using BCBGAM-treated isolates, including P24-192. (A) Box plots quantifying the morphological response of six *Colletotrichum* sp. isolates to one representative phenolic antifungal (BCBGAM), which exhibited isolate-specific effects. BCBGAM is featured because it elicited the most visually distinct morphological changes in isolate P24-192. While PhSOAM showed stronger overall inhibition (Fig. S1), the response to BCBGAM better illustrates the complex 'morphological fingerprint' central to this study, highlighting our focus on decoding characteristic stress responses rather than solely quantifying growth inhibition. Measured traits include colony area (mm²), perimeter (mm), length (mm), width (mm), LWR, circularity, and an asymmetry distance (IS & CG, mm). Significant differences between a treated group and its corresponding control (Tukey's HSD; $*p < 0.05$; $**p < 0.01$; $***p < 0.001$) are marked with asterisks. (B-C) Representative images of isolate P24-192 after 96 hours on (B) a control plate and (C) a plate treated with BCBGAM. Sample size, n = 10 for all groups except P24-192 and P24-193 controls (n = 11). Scale bars = 1 cm.

3.2 Variable and opposing effects of fatty acid analogs

The two phenol-branched fatty acids, PhSOFA and BCBGFA, exhibited overall weaker antifungal activity than their corresponding amide derivatives, yet each triggered highly distinctive and isolate-specific morphological responses (Fig. 2; Fig. S2–S3).

PhSOFA demonstrated selective growth inhibition, significantly reducing colony size in two of the six isolates: P24-83 and P24-192 (Fig. S2A). For example, PhSOFA treatment reduced the colony area of isolate P24-83 by 18.8% ($p < 0.01$), a change illustrated in the representative images (Fig. S2B–C). In contrast, BCBGFA's effects were highly variable. It had no significant inhibitory effect on any isolate but paradoxically promoted the growth (assessed by area size measurement) of P24-85 by 16.5% ($p < 0.05$), as detailed in Fig. S3. The remaining isolates were largely resistant to either compound. Morphological shape metrics highlighted further, though more sporadic, divergence between the two treatments. The effects on shape metrics were more sporadic: significant increases in colony circularity were rare, observed only for BCBGFA in P24-83 and PhSOFA in P24-193, while PhSOFA was the only compound to cause significant increases in both colony elongation (LWR) and asymmetry (IS & CG), effects that were limited to isolate P24-192 (Figs. S2 & S3).

When morphological data were pooled across all isolates (Fig. 2), a clear hierarchy of antifungal potency, based on the reduction in colony area, was observed. PhSOAM remained the most potent growth inhibitor. It was followed by PhSOFA. The other amide, BCBGAM, showed no significant difference from the control, while BCBGFA was the least effective, showing no significant inhibition. Despite variable effects on colony size, the compounds induced reproducible shifts in shape. PhSOFA uniquely elevated the LWR. A significant increase in colony circularity was observed for three compounds: PhSOFA, BCBGAM, and BCBGFA. Furthermore, only PhSOFA caused a significant increase in the asymmetry distance (IS & CG), pointing to non-lethal but geometry-altering cellular effects. Together, these data show isolate- and structure-dependent effects across compounds (Fig. 2; Figs. S2–S3).

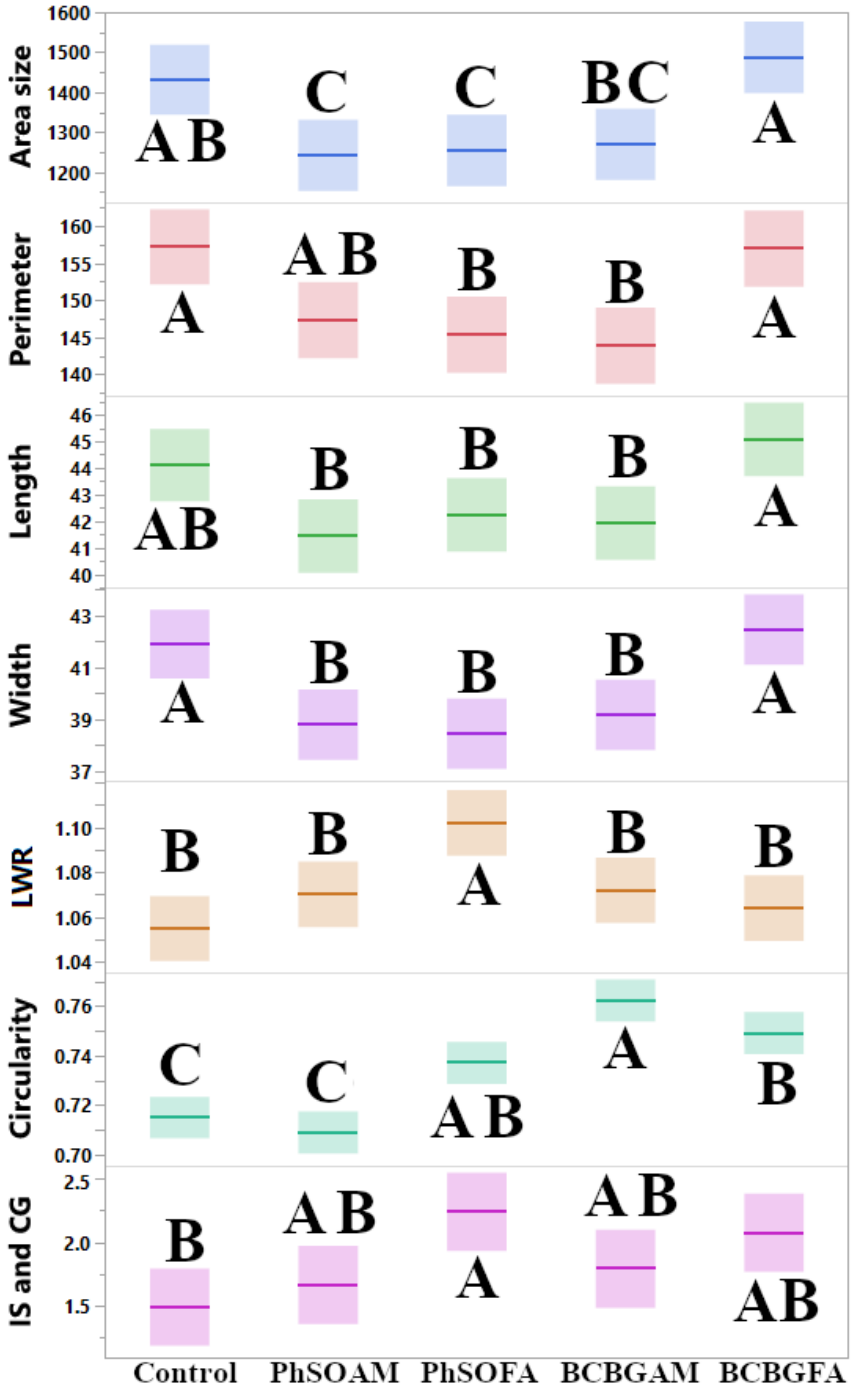

Fig. 2. Comparative analysis of the antifungal effects of four phenolic-branched compounds on coffee-pathogenic *Colletotrichum* isolates. Box plots show the pooled morphological data from six *Colletotrichum* isolates under four treatment conditions (PhSOAM, BCBGAM, PhSOFA, and BCBGFA) and an untreated control. Measured traits include colony area (mm^2), perimeter (mm), length (mm), width (mm), LWR, circularity, and an asymmetry distance (IS & CG, mm). Different letters denote statistically significant groupings ($p < 0.05$) determined by Tukey's HSD.

222 **3.3 Multivariate analysis reveals a dominant treatment effect on fungal morphology**

223 PCA of the morphological traits showed that the first two PCs captured 76.2% of the total variance
224 (Fig. S4). However, both chemical treatments and fungal isolates showed considerable overlap in
225 PCA space, indicating that linear ordination provided limited separation between groups. This
226 necessitated the use of supervised machine learning to test for nonlinear, diagnostic patterns.
227 Permutational multivariate analysis of variance (PERMANOVA) within the coffee dataset
228 indicated a small but significant treatment effect and a stronger isolate effect (Compound: pseudo-
229 F = 4.06, R^2 = 0.0518, p = 0.001; Strain: pseudo-F = 13.00, R^2 = 0.180, p < 0.0001). This partial
230 grouping in PCA, combined with substantial overlap, suggests that linear ordination provides
231 limited separation. Because PCA did not fully resolve isolates or pathosystem identity, we next
232 used supervised ML to test for nonlinear, diagnostic patterns.

233 Further analyses were conducted to explore the relationships between the seven morphological
234 traits (Fig. S5). The PCA loading plot (Fig. S5A) and partial contributions plot (Fig. S5B)
235 confirmed that the primary axis of variation, PC1, was driven by the four main size metrics (area,
236 perimeter, length, and width). Shape-related traits had greater impacts on subsequent components:
237 LWR and IS & CG contributed most to PC2, while circularity was a major driver of PC3. This
238 clear distinction between size and shape was corroborated by hierarchical clustering, which
239 grouped the four size metrics into one tight cluster and the three shape metrics into another (Fig.
240 S5C). Pearson's correlation analysis further quantified these relationships, revealing strong
241 positive correlations among all size-related traits (r > 0.9 for most pairs, detailed in Supplementary
242 Data 1).

243

244 **3.4 Phylogenetic analysis reveals coffee-pathogenic isolates are polyphyletic and intermixed**
245 **with cacao-associated isolates**

246 To determine the evolutionary relationships among the isolates, we first conducted a multi-locus
247 phylogenetic analysis using a concatenated dataset of the internal transcribed spacer (ITS),
248 glyceraldehyde-3-phosphate dehydrogenase (GAPDH), and actin (ACT) gene regions (Fig. 3).
249 The analysis revealed the six coffee-pathogenic isolates were polyphyletic, distributing across
250 several distinct clades rather than forming a monophyletic group.

Specifically, isolate P24-88 formed a sister group with *C. karstii* (79% bootstrap support), and this clade in turn was a sister group to isolate P24-193. Isolate P23-55 was placed in a separate, distant clade with *C. pseudoboninense*. The remaining three coffee isolates were located in different parts of the tree: P24-192 clustered with *C. siamense* (99% bootstrap), P24-83 was a sister taxon to *C. rhizophorae* and *C. fructicola*, and P24-85 was placed near *C. gloeosporioides* and *C. kahawae*. This polyphyletic distribution indicates that the coffee pathogens used in this study represent multiple distinct lineages within the *Colletotrichum* genus.

An analysis based on the standard fungal barcode, the ITS region, was also performed to directly compare the coffee isolates with cacao-associated isolates (Fig. S6). This analysis further highlighted the complex relationships between isolates from the two pathosystems. Notably, a single, strongly supported clade (99% bootstrap) contained an admixture of isolates from different hosts and origins. This inclusive clade grouped coffee pathogens P24-192, P24-85, and P24-83 together with cacao isolates (CGH17, CGH53) and shade tree isolates (CGH34, CGH38). Overall, based on the ITS region, some coffee pathogens are more closely related to cacao-associated isolates than to other coffee pathogens, such as P24-88 and P24-193, which formed a separate, well-supported clade (96% bootstrap).

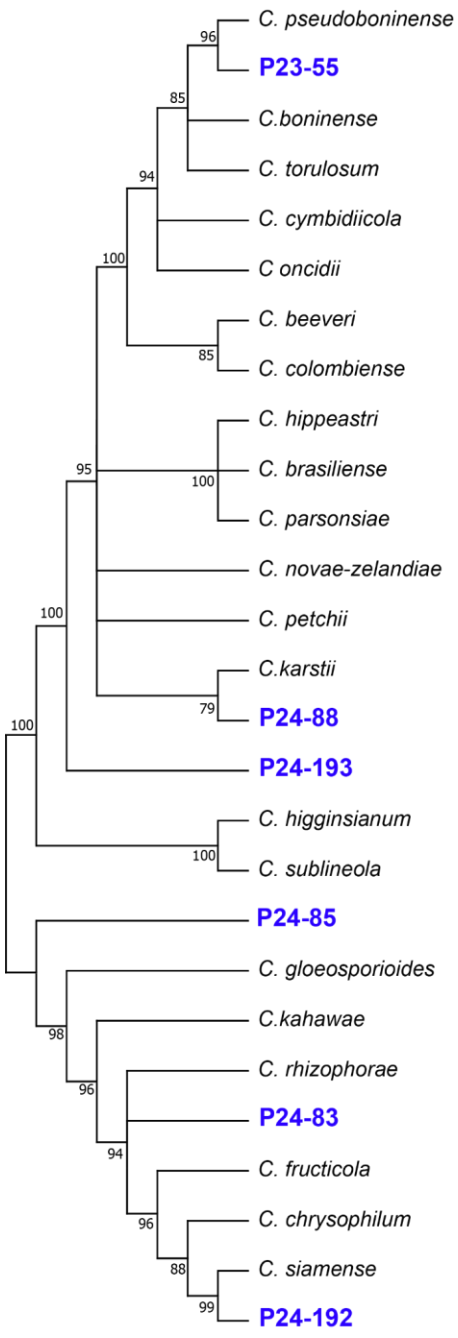

Fig. 3. Multi-locus phylogeny reveals the polyphyletic nature of coffee-pathogenic *Colletotrichum* isolates

The phylogenetic tree was inferred by the Neighbor-Joining method using a concatenated alignment of the ITS, GAPDH, and ACT sequences. The evolutionary distances were computed using the Tamura-Nei (TN+G) model. Numbers at the nodes represent bootstrap support values based on 1,000 replications; only values of 75% or greater are shown. The six coffee-pathogenic isolates (highlighted in blue) are shown to be polyphyletic, distributing across multiple distinct clades and indicating affiliations with different known *Colletotrichum* species.

3.5 Comparative multivariate analysis of coffee and cacao pathosystems

To investigate whether the chemical compounds induced environment-specific morphological responses, the coffee pathogen dataset was combined with a previously generated dataset from cacao pathogens[16], and a comparative PCA was performed on the combined, log-transformed data.

The first two principal components accounted for 78.6% of the total variance (PC1: 56.8%, PC2: 21.8%) (Fig. 4). When the data were colored by the pathogen's system of origin (Fig. 4A), no clear separation was evident. The coffee (black) and cacao (red) samples were heavily intermixed. PERMANOVA detected a weak host effect (Host: pseudo-F = 8.49, $R^2 = 0.0174$, $p = 0.0017$) and a small but significant treatment effect (Compounds: pseudo-F = 5.57, $R^2 = 0.0447$, $p < 0.001$), consistent with the minimal variance explained by these factors.

When visualized by treatment type, the analysis showed significant overlaps between the groups (Fig. 4B). The untreated control samples clustered closely with those treated with BCBGFA. A separate cluster was formed by samples treated with PhSOAM and PhSOFA, while the BCBGAM-treated samples were positioned between these two main groupings. This strong overlap indicates that the morphological signatures distinguishing the two pathosystems, if they exist, are not linearly separable and thus are too subtle for PCA to resolve.

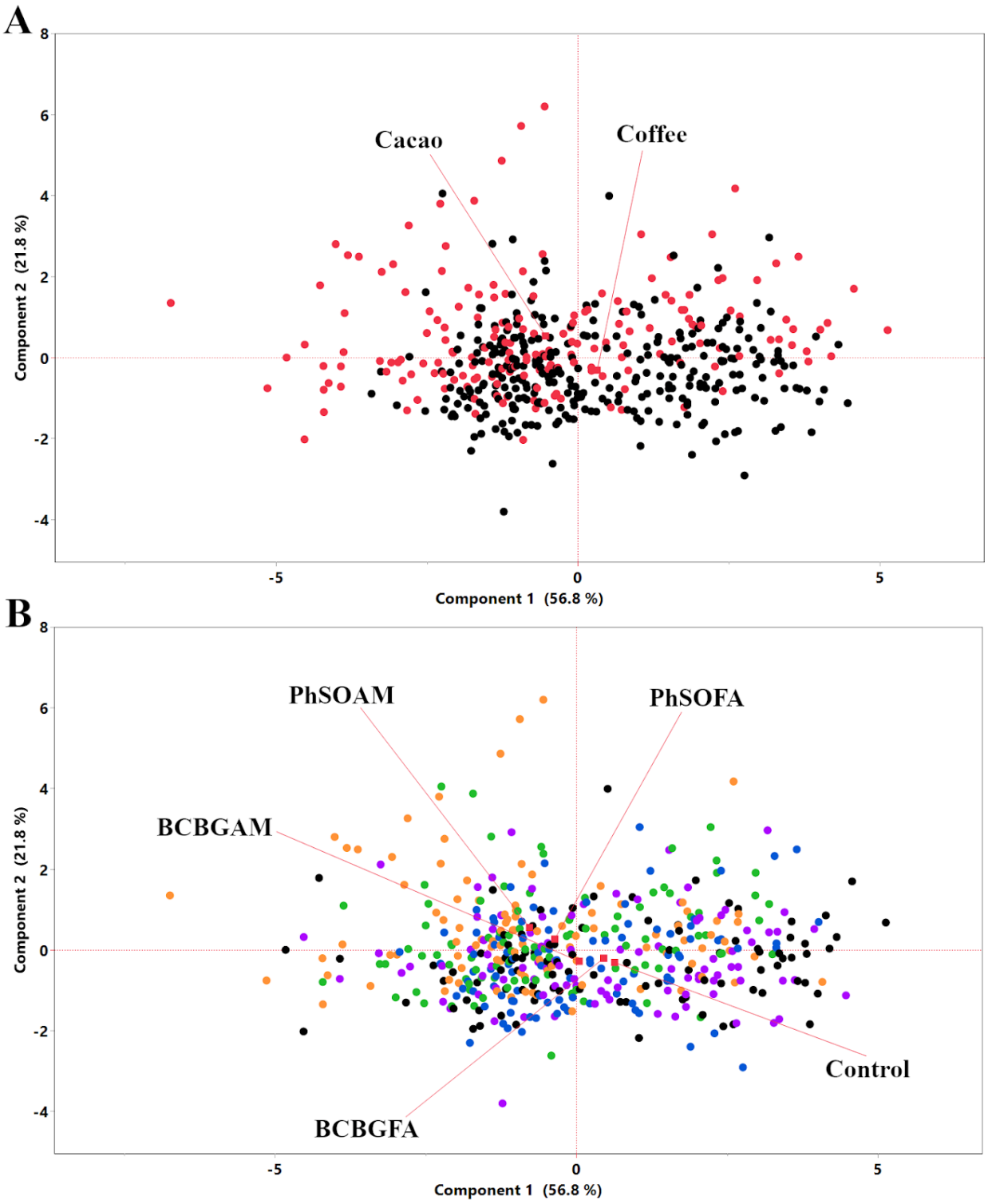

Fig. 4. Comparative PCA of coffee and cacao pathosystems reveals treatment-driven separation but pathosystem overlap. PCA of morphological traits from the combined coffee and cacao pathogen datasets (log-transformed). (A) PCA scores plot with samples color-coded by pathosystem: Cacao pathogens (red) and Coffee pathogens (black). Note the significant overlap between the two pathosystems, indicating the first two principal components cannot separate them. (B) The identical PCA scores plot with samples color-

coded by chemical treatment: Control (black), PhSOAM (orange), PhSOFA (green), BCBGAM (blue), and BCBGFA (purple). The red squares indicate the mean of each pathosystem group in (A) and each chemical treatment group in (B). PERMANOVA indicated a weak host effect (Host: pseudo-F = 8.49, $R^2 = 0.0174$, $p = 0.0017$) and a small but significant treatment effect (Compound: pseudo-F = 5.57, $R^2 = 0.0447$, $p < 0.001$), both explaining only a small proportion of total variance.

3.6 Machine learning models successfully classify pathogen host origin

Following the PCA, which could not clearly separate the two pathosystems, a suite of ML models was employed to quantitatively test if a predictive morphological signature of host origin existed within the combined dataset. The analysis revealed that several models could classify a pathogen's host of origin with high accuracy, confirming the existence of a robust 'fingerprint' based on the response to the chemical compounds (Fig. 5). A comprehensive screening of nine different algorithms identified the Neural Boosted model (an F1-score of 0.87, and a Matthews Correlation Coefficient (MCC) of 0.73) as the top performer, achieving a validation accuracy of 87%, which was followed closely by Nominal Logistic and Generalized Regression (Lasso) models (both 85%), as well as Support Vector Machines, Bootstrap Forest, and Fit Stepwise (all 84%) (Table S1).

To understand which traits were driving this successful classification, feature importance was analyzed using the top-performing Neural Boosted model (Table S2). The analysis revealed that circularity was the single most dominant predictor of host origin, with a Total Effect score of 0.592, which was followed by the primary size-related traits of area (Total Effect = 0.405) and perimeter (Total Effect = 0.168). This indicates that the model primarily learned to distinguish the pathosystems by identifying a unique 'shape signature', driven by how circular a colony becomes in response to chemical stress, complemented by information about the colony's overall size.

To further test the robustness of the classification, the analysis was repeated on the log-transformed dataset after removing circularity, the top predictor (full details in Supplementary Data 1). Even without this primary feature, the models maintained high predictive power. The Neural Boosted model was again the top performer with a validation accuracy of 86.1%, followed by the Nominal Logistic (85%), Lasso (85%), and Fit Stepwise (84%) models. Feature importance analysis for this new model revealed a clear shift in predictive strategy; in the absence of circularity, the model pivoted to a 'size signature', with area size (Total Effect = 0.825) and perimeter (Total Effect = 0.467) emerging as the most dominant predictors (Supplementary Data 1), demonstrating that the

morphological 'fingerprint' distinguishing the pathosystems is robust and encoded in multiple features.

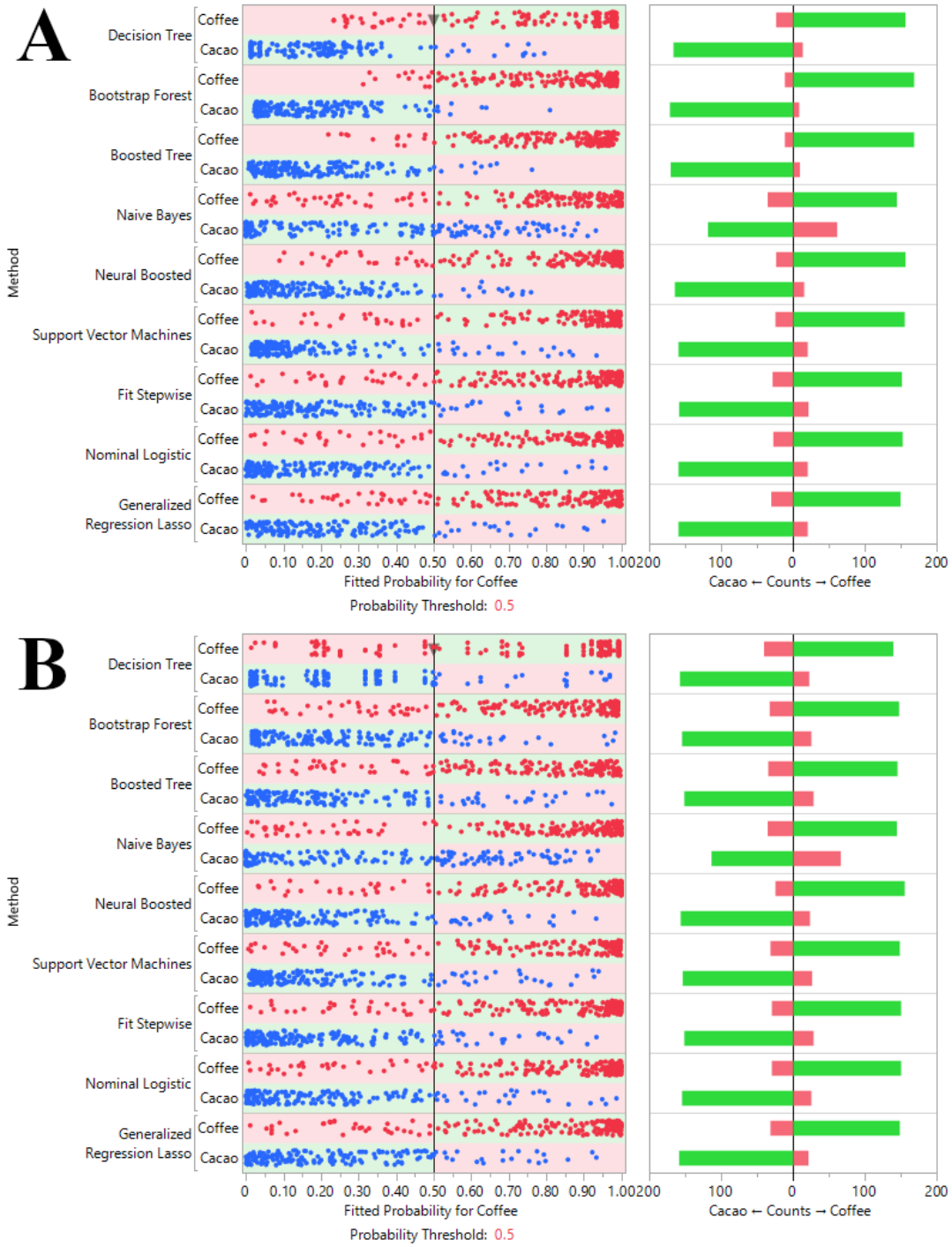

Fig. 5. Machine learning classification of pathogen host origin based on chemical-induced morphology. Decision thresholds plot showing the performance of nine ML models in classifying a

pathogen's host of origin (cacao vs. coffee) based on its morphological fingerprint from chemical-induced morphology (A: Training set, B: Validation set). Blue dots represent isolates from the cacao pathosystem, and red dots represent isolates from the coffee pathosystem. The plot shows the aggregated results from the k-fold cross-validation, with the confusion matrix on the right summarizing classification performance at a probability threshold of 0.5. 'Neural Boosted' denotes the JMP Pro 17 boosted neural network; default base learner with boosting enabled; repeated 5-fold CV with fixed random seed.

3.7 Dose-dependent hyperspectral fingerprints in CGH5 under PhSOAM treatment

To probe the physiological basis of these morphological fingerprints, we first employed three modalities of HSI: fluorescence HSI, visible and near-infrared (VNIR) reflectance imaging, and short-wave infrared (SWIR) reflectance imaging. We began by testing the response of CGH5, a cacao-derived isolate known to exhibit a dramatic 48.8% reduction in fungal growth upon PhSOAM treatment [16]. HSI revealed distinct, dose-dependent responses of *Pestalotiopsis* sp. CGH5 to PhSOAM treatment across all three spectral modalities (Fig. 6).

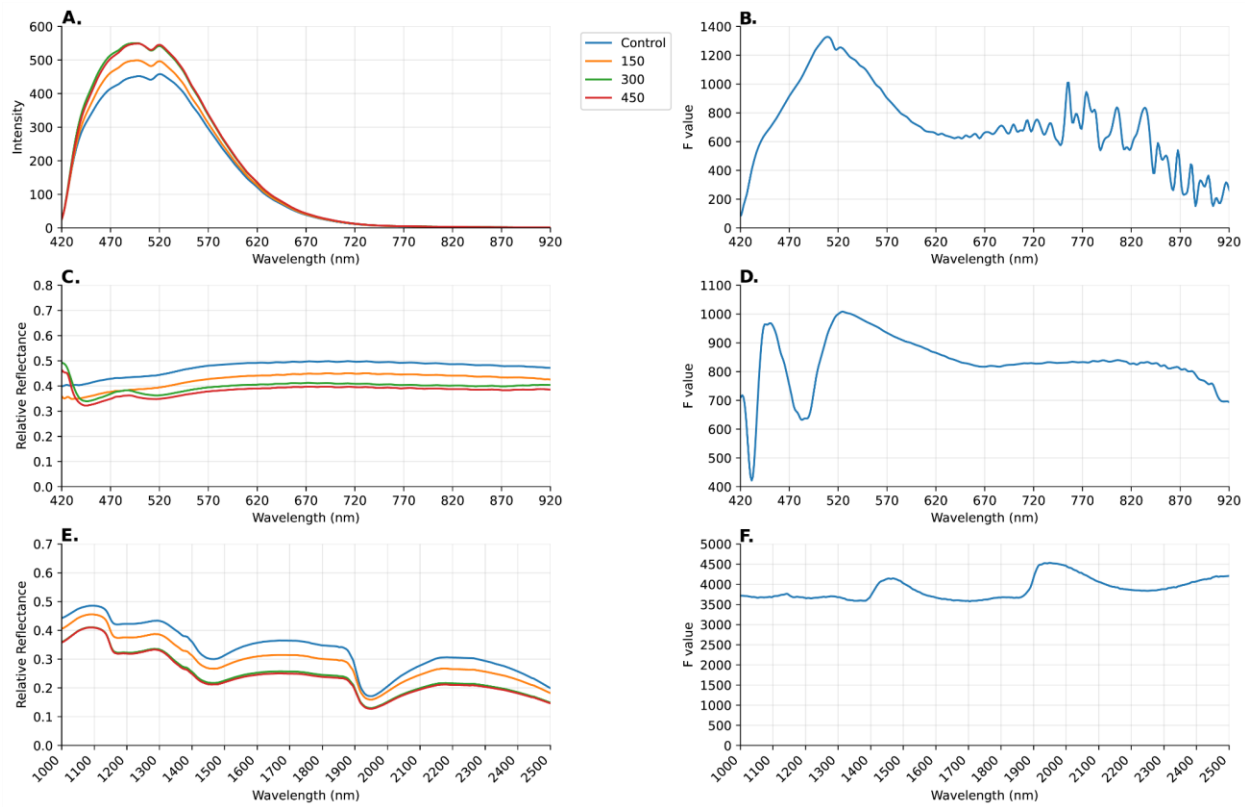

Fig. 6. Hyperspectral imaging of *Pestalotiopsis* sp. CGH5 treated with the biobased fungicide PhSOAM across three spectral modalities. (A) Fluorescence intensity profiles of CGH5 colonies

treated with 0, 150, 300, and 450 µg of PhSOAM; color indicates dose (control=blue, 150 µg = orange, 300 µg = green, 450 µg= red). (B) F-value curve corresponding to the fluorescence data. (C) Reflectance intensity curves under broadband illumination. (D) F-value profile corresponding to the reflectance data. (E) SWIR reflectance curves showing spectral shifts. (F) F-value curve corresponding to the SWIR data. All data represent the mean of four biological replicates per treatment (n = 4).

3.8 Attenuated hyperspectral response in a tolerant coffee isolate

In contrast to the strong response of CGH5, the more tolerant coffee isolate P24-85 exhibited an attenuated and nonlinear hyperspectral response to PhSOAM (Fig. 7). This included a non-monotonic fluorescence signal and muted shifts in the SWIR water-associated bands. These attenuated spectral changes are consistent with a more resilient physiological state capable of active damage management.

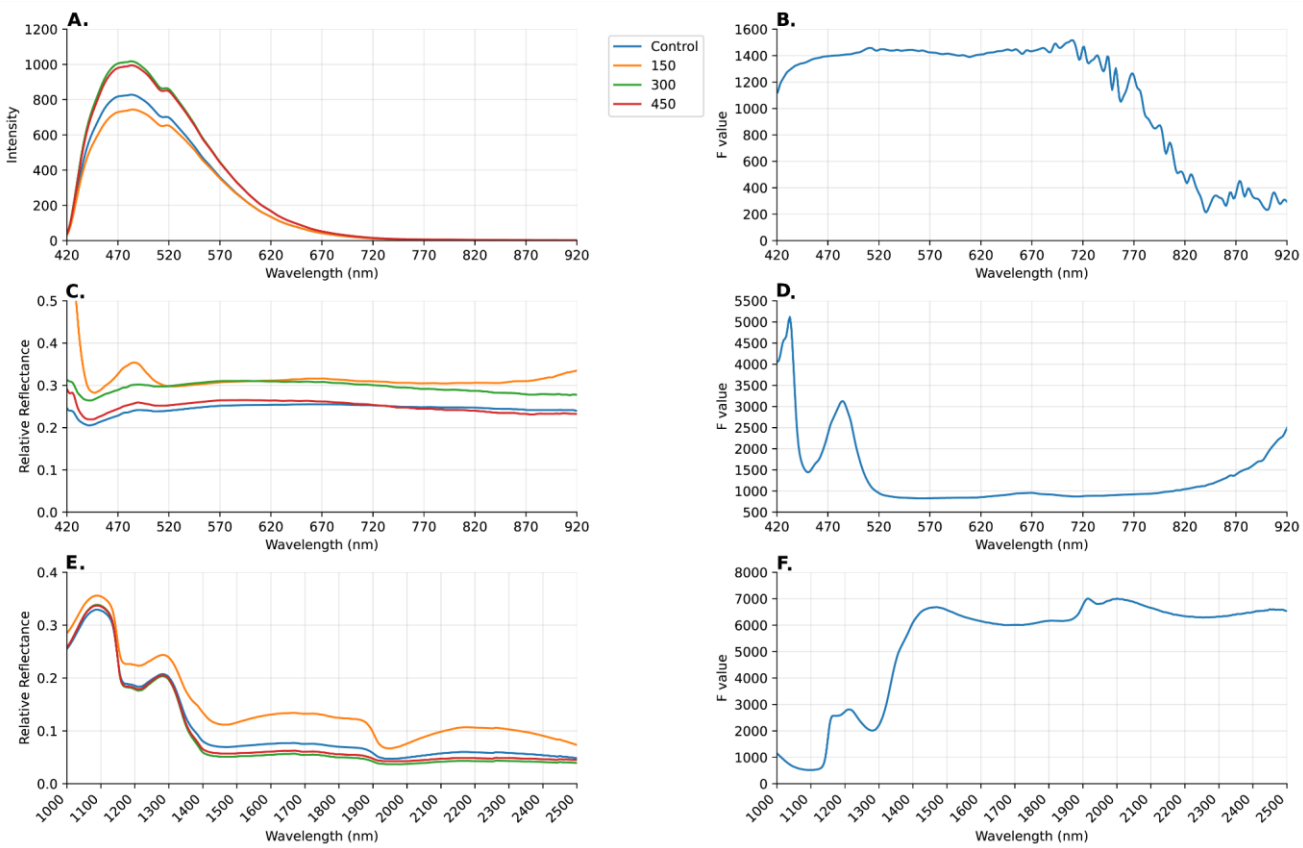

Fig. 7. Hyperspectral imaging of *Colletotrichum* sp. P24-85 treated with the biobased fungicide PhSOAM across three spectral modalities. (A) Fluorescence intensity profiles of P24-85 colonies treated with 0 µg (control, blue), 150 µg (orange), 300 µg (green), and 450 µg (red) of PhSOAM. (B) F-value curve corresponding to the fluorescence data. (C) Reflectance intensity curves under broadband

illumination. (D) F-value profile corresponding to the reflectance data. (E) SWIR reflectance curves. (F) F-value curve corresponding to the SWIR data. All data represent the mean of four biological replicates per treatment (n = 4).

3.9 Compound-specific hyperspectral fingerprints of *Colletotrichum* sp. P24-192

Different compounds also induced unique spectral fingerprints in the same isolate, *Colletotrichum* sp. P24-192, suggesting divergent modes of action (Fig. S7). For example, in the SWIR modality, BCBGAM produced a focused diagnostic band near ≈1950 nm, whereas PhSOAM yielded a broader separation across multiple bands. These distinct spectral signatures are consistent with compound-specific biochemical interactions with the fungus.

4 Discussion

Here, we established that a microbe's response to a uniform stressor can reveal its ecological history. More importantly, we developed and validated a powerful workflow to decode this "ecological memory." By integrating quantitative morphology, HSI, and ML, we showed that pathosystem-specific fingerprints could predict an isolate's host of origin with high accuracy. Our work demonstrates that this scalable, phenotype-driven platform is a powerful tool for studying how ecological history shapes microbial adaptation, with direct applications in screening new antifungals and decoding complex morphological signatures.

The varied fungal responses underscored a clear structure-activity relationship dependent on both the headgroup and backbone of the four phenolic compounds. Consistent with previous reports, amide derivatives generally showed stronger antifungal activity than their fatty acid analogs. Specifically, PhSOAM, with its simple phenol and amide moiety, was the most potent growth inhibitor. This suggests that both the increased polarity from the amide group and the steric accessibility of the unsubstituted phenol contribute to efficacy [12, 25, 26].

An unexpected observation was that BCBGFA appeared to promote growth in isolate P24-85. While this may reflect experimental variability, an alternative explanation is that certain fungal isolates possess enzymatic adaptations that allow them to tolerate or even metabolize specific phenolic compounds. This possibility is supported by earlier studies demonstrating that some fungi can grow on phenolic acids and related aromatic compounds [27, 28].

This divergence in morphological response suggests a deeper biological basis rooted in ecological adaptation. We propose that long-term exposure to a native environment, whether a specific host, soil habitat, aquatic niche, or built environment, imprints an ecological memory in stress-response pathways. When challenged with a novel stressor, this legacy is expressed as what we term a ‘chemical prior’: a reproducible, system-specific phenotypic fingerprint. In this study, the coffee–cacao comparison serves as one example, but the same principle could apply to other fungi, bacteria, protists, or mixed microbial communities [29, 30]. While our previous work showed that PhSOAM induces a dramatic increase in colony elongation (LWR) in susceptible cacao pathogens [16], this response was attenuated and inconsistent among coffee isolates. Notably, P24-88 exhibited moderate elongation, but this phenotype was not broadly observed across the coffee panel. Instead, circularity emerged as the more robust and pathosystem-distinctive morphological trait. Intriguingly, this response is not limited to chemical insults; it is further reinforced by our earlier findings where physical (UVC) stress also induced increased circularity in multiple cacao isolates [24]. Circularity may therefore be a generalizable indicator of stress, as disparate pressures from chemical and UVC treatments appear to converge on a similar morphological outcome.

While linear methods like PCA failed to separate the pathosystems, supervised machine learning successfully classified host origin with 87% accuracy, validating the existence of a predictive 'interaction phenotype' [31, 32]. Feature-importance analysis consistently revealed colony circularity, a measure of isotropic growth, as the most robust biomarker of this latent ecological memory. The model’s reliance on this shape metric, even after its removal from the feature set, indicates the fingerprint is robustly encoded across multiple morphological traits [33, 34]. The importance of circularity as a stress indicator is also consistent with recent work on biotic antagonism [35–37].

We interpret this circularity shift as a direct readout of the pathogen's pre-programmed growth machinery responding to a uniform, isotropic stress field. An alternative hypothesis, that these shifts reflect rapid selection of resistant mutants, is unlikely. The 96-hour incubation is insufficient for the multi-generational selection typically required for such changes[38, 39], and the consistent, pathosystem-wide nature of the fingerprint is inconsistent with the stochastic nature of mutation

[40]. Therefore, we conclude these phenotypes represent plastic responses reflecting an evolutionarily imprinted memory [41].

HSI provided physiological evidence for this ecological memory, revealing stressor- and isolate-specific spectral fingerprints. Previous studies show that near- and shortwave-infrared reflectance and fluorescence spectroscopy can distinguish fungal physiological responses, including hydration level, pigmentation, and cell-wall composition [42–45]. In our study, strong isolate- and dose-dependent signals emerged across all three imaging modalities. HSI provided a physiological basis for these fingerprints, revealing two distinct stress-response archetypes. The susceptible cacao isolate CGH5 exhibited strong spectral shifts, including increased fluorescence (470–520 nm) and prominent absorption near water-associated SWIR bands (≈1930 nm), potentially indicating a loss of homeostasis or 'metabolic collapse' [24, 42, 46, 47]. In contrast, the more tolerant coffee isolate P24-85 showed a muted or attenuated spectral response at these same bands, suggesting an active 'damage management' strategy that maintains physiological stability under stress. This attenuated spectral response, particularly at water-associated bands, may reflect a distinct and potentially more resilient physiological state under chemical stress. Consistent with the SWIR water band and the circularity shift, this resilience may arise from rapid water-balance control (osmolyte accumulation, aquaporin gating) coupled with cell-wall remodeling at the colony boundary. These mechanisms are hypotheses consistent with our readouts and will require direct molecular validation.

The spectral contrast observed between PhSOAM and BCBGAM treatments further suggests divergent biochemical interactions (Fig. S7, Table S3). These compound-specific signatures, likely reflecting distinct impacts on cell hydration, membrane lipids, and phenolic compounds, require future molecular validation to elucidate precise mechanisms [24, 48]. Because the ≈1930 nm band is highly moisture-sensitive, future work should explicitly model or filter water effects (e.g., variable band weighting or EMSC) to isolate biochemical signals [47]. Interestingly, some of these morphological and spectral stress markers, particularly increased circularity and shifts in SWIR bands, converge with those we previously observed under physical (UVC) stress, suggesting a potentially generalizable stress-response program in these pathogens [24].

469 These parallels between xenobiotic chemical exposure and UVC irradiation suggest that HSI can
470 capture both compound-specific and isolate-specific physiological responses. The recurrence of
471 CGH53's specific spectral features under both conditions raises the possibility that its native
472 agroecosystem has shaped a stress-response pathway that is preferentially activated when
473 homeostasis is disrupted. This pathway, once triggered by a sufficiently strong stressor, may lead
474 to a cascade of metabolic and physiological changes, though the precise molecular mechanisms
475 remain to be determined.
476
477 Crucially, the convergence of findings from macroscopic colony morphology and physiological
478 hyperspectral signatures suggests the observed "ecological memory" is a robust biological signal.
479 This multi-modal evidence allows us to propose a conceptual model of two contrasting
480 physiological archetypes: a 'metabolic collapse' in susceptible isolates like CGH5, characterized
481 by spectral shifts indicative of lost homeostasis, and a buffered 'damage management' in tolerant
482 isolates like P24-85. This framework, while requiring future molecular validation, provides a
483 powerful new model for understanding the diverse strategies pathogens employ under chemical
484 stress.
485
486 The polyphyletic nature of the coffee isolates provides critical evolutionary context for our
487 findings (Fig. 3, Fig. S6). The fact that these isolates do not form a single clade, and are sometimes
488 more closely related to cacao pathogens, argues that the 'coffee fingerprint' is not a trait inherited
489 from a single ancestor. Instead, this suggests convergent adaptation, where distinct lineages
490 independently evolved similar stress-response strategies when faced with the unique chemical
491 landscape of the coffee agroecosystem [49]. This interpretation is supported by the ability of our
492 ML models to classify isolates by their host of origin, detecting a true, host-imprinted ecological
493 signal that transcends the complex phylogenetic backdrop and the limited resolution of ITS-based
494 phylogenies.
495
496 While our data suggest host adaptation as the primary driver, we do not rule out the influence of
497 biogeography and lineage-specific traits, although direct biogeographical effects were not
498 explicitly tested in this study. The coffee isolates' distribution across multiple clades, some of
499 which are shared with isolates from a geographically distinct cacao agroecosystem, underscores

the complex interplay between evolutionary history and environmental adaptation [50, 51]. Disentangling these factors presents a critical and now-testable question for the field. Ultimately, our work demonstrates that a pathogen's observable response to chemical stress is a rich, multi-layered phenotype, reflecting a powerful ecological memory forged at the intersection of its lineage and its environment.

To synthesize these findings, we propose a generalized conceptual model that links a microbe's long-term ecological history to its observable stress-response phenotype (Fig. 8). We term this the "chemical priors" framework, which posits that persistent exposure to a native environment, whether a specific host, soil, or aquatic niche, imprints a latent "ecological memory" in stress-response pathways. When challenged with a novel stressor, such as the phenolic compounds used in this study, this embedded memory is expressed as a reproducible, system-specific phenotypic fingerprint, such as the shifts in colony circularity observed here. This fingerprint can then be decoded by supervised machine learning to accurately infer the pathogen's origin, providing a quantitative readout of its ecological history. While demonstrated with the coffee–cacao pathosystem, this framework offers an adaptable approach for studying microbial adaptation across diverse biological contexts. This "chemical priors" framework suggests a form of heritable, environmentally-induced memory, potentially rooted in epigenetic mechanisms that warrant future investigation.

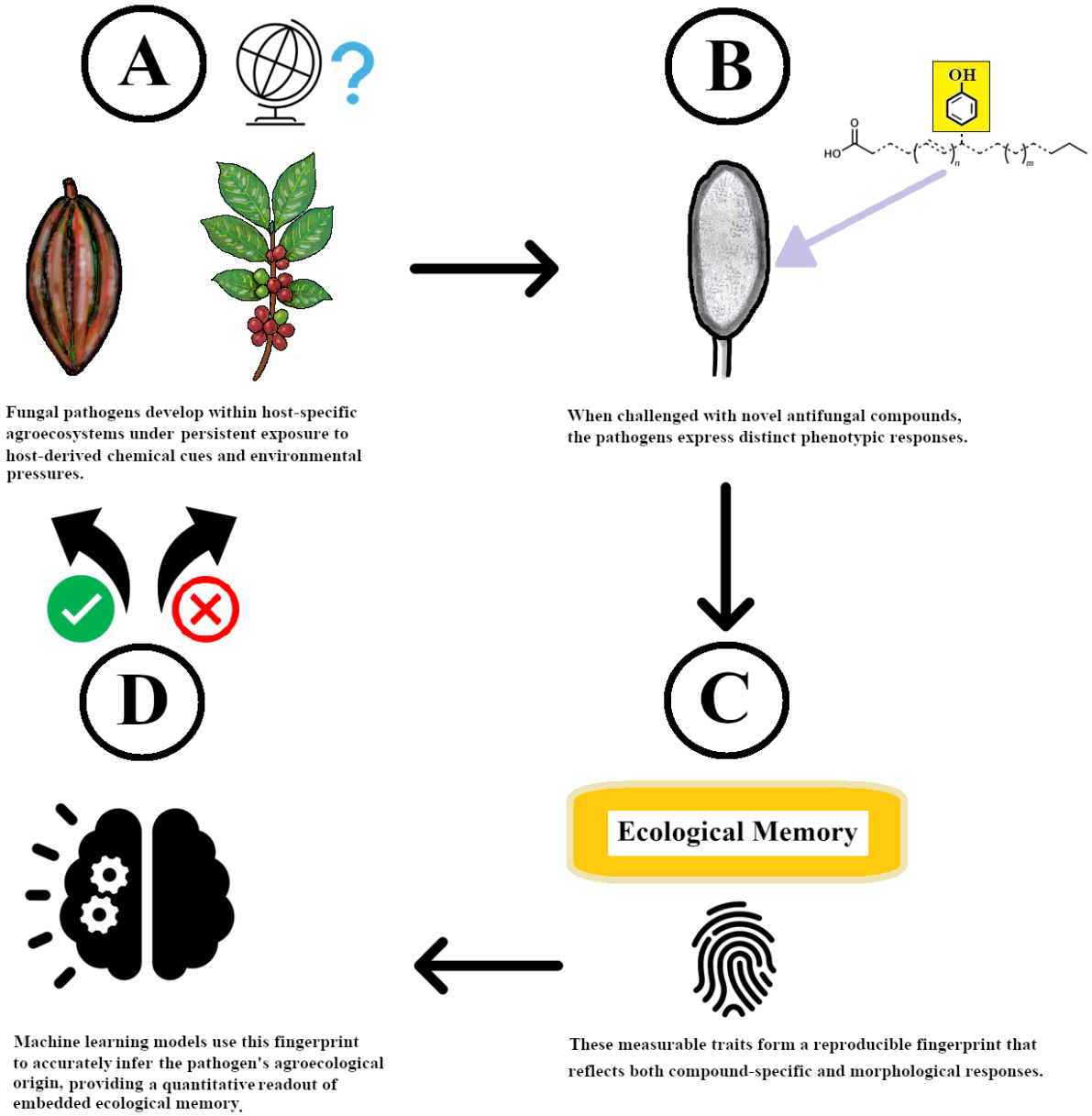

Fig. 8. Conceptual model of ecological memory and chemical priors in microbial response to environmental stress. (A) Long-term exposure of a microorganism to a specific native environment, whether a host–pathogen agroecosystem (e.g., coffee or cacao), soil, aquatic, or built environment, imprints an "ecological memory" shaped by environmental chemistry and associated microbial communities. (B) When challenged *in vitro* with a novel environmental stressor (chemical, physical, or biological), the fungus (spore; center in the panel for representative image) expresses a distinct, reproducible phenotypic fingerprint (e.g., circularity shift, hyperspectral change) shaped by this memory. (C) Supervised machine learning classifies these fingerprints to

infer the organism's ecological origin or adaptation history with high accuracy, quantitatively linking past ecological exposure to present-day phenotypic stress responses. (D) This framework, termed the "chemical priors" model, provides a testable hypothesis for how historical selective pressures are embedded in microbial stress physiology and can be leveraged to predict responses in novel environments. While illustrated here for coffee–cacao *Colletotrichum* under phenolic antifungal stress, the same conceptual workflow could be adapted to other microorganisms and environmental contexts, including bacterial, algal, and mixed microbial systems exposed to chemical, physical, or biological challenges.

Our study establishes a scalable, phenotype-driven platform for decoding microbial ecological memory. The immediate value of this methodological advance is its application in high-throughput screening of eco-friendly fungicides and precision disease management. While future work should aim to validate these findings in planta and resolve their molecular bases, the "chemical priors" framework and the integrative workflow presented here provide a transformative new tool for the phytopathology and microbial ecology communities.

Our study establishes a scalable, phenotype-driven platform for decoding microbial ecological memory. The immediate value of this methodological advance is its application in high-throughput screening of eco-friendly fungicides and precision disease management. Building on this foundation, future work will prioritize three key directions: (1) validating these phenotypic fingerprints *in planta* to connect them to real-world disease outcomes; (2) resolving the underlying molecular and genomic bases of these stress responses to move from prediction to mechanism; and (3) testing the generality of the "chemical priors" model across diverse pathogen systems to establish it as a universal principle.

Acknowledgements

Mention of any trade names or commercial products in this article is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U. S. Department of Agriculture. USDA is an equal opportunity provider and employer, and all agency services are available without discrimination. This work was supported by the U.S. Department of Agriculture, Agricultural Research Service, In-House Projects Nos. 8042-21220-258-000-D and

8072-41000-113-000-D, and by the National Institute of Food and Agriculture (NIFA) under Grant No. 2021-67021-34502. We thank Lisa Keith (Daniel K. Inouye U.S. Pacific Basin Agricultural Research Center, USDA-ARS, Hilo, Hawaii) for kindly providing the *Colletotrichum* isolates used in this study.

Competing interests

The authors declare no competing interests.

Author contributions

E.A. conceptualized and supervised the project. A.L., C.H., and M.C. developed the methodology. L.K., C.H., S.O., and E.A. conducted the experiments. I.B., S.L., and E.A. performed the formal analysis. M.K-R., H.N., R.D.A., and A.L. provided resources. E.A., L.W.M., and M.S.K. acquired funding. E.A. wrote the original manuscript draft, and all authors contributed to reviewing and editing the final version.

Data and materials availability

All raw and processed data, including colony morphology measurements, per-wavelength mean spectra, and F-statistics for HSI, are provided in Supplementary Data 1. Multi-locus sequence data (ITS, GAPDH, ACT) generated for this study are openly available at figshare: https://figshare.com/articles/dataset/29924960.

Code availability

Not applicable.

Supplementary Materials

This PDF file includes: Figs. S1 to S7
Table S1 to S4
Supplementary methods

References

592 1. Ferrucho RL, Marín-Ramírez GA, Gaitan A. Integrated Disease Management for the Sustainable
593 Production of Colombian Coffee. *Agronomy* 2024;**14**. https://doi.org/10.3390/agronomy14061286
594 2. Ploetz R. The impact of diseases on cacao production: a global overview. *Cacao Dis Hist Old*
595 *Enemies New Encount* 2016;33–59.
596 3. Kifle Belachew KB, Demelash Teferi DT. Climatic variables and impact of coffee berry diseases
597 (Colletotrichum kahawae) in Ethiopian coffee production. *J Biol Agric Healthc* 2015;**5**:55–64.
598 4. Mohali Castillo SR, Miller S, Stewart J. Colletotrichum spp. and other fungi associated with
599 anthracnose on Coffea arabica L. in Mérida State, Venezuela. *Summa Phytopathol* 2022;**48**:99–111.
600 https://doi.org/10.1590/0100-5405/245876
601 5. Ploetz RC. Cacao diseases: important threats to chocolate production worldwide. *Phytopathology*
602 2007;**97**:1634–1639.
603 6. Cortaga CQ et al. Mutations associated with fungicide resistance in Colletotrichum species: A
604 review. *Phytoparasitica* 2023;**51**:569–592.
605 7. Schnabel G et al. Inherent tolerance of Colletotrichum gloeosporioides to fludioxonil. *Pestic*
606 *Biochem Physiol* 2021;**172**:104767.
607 8. Tsalidis GA. Human Health and Ecosystem Quality Benefits with Life Cycle Assessment Due to
608 Fungicides Elimination in Agriculture. *Sustainability* 2022;**14**. https://doi.org/10.3390/su14020846
609 9. Cantrell CL, Dayan FE, Duke SO. Natural Products As Sources for New Pesticides. *J Nat Prod*
610 2012;**75**:1231–1242. https://doi.org/10.1021/np300024u
611 10. Thind TS. Role of Fungicides in Crop Health Management: Prospects and Challenges. In:
612 Satyanarayana T, Deshmukh SK, Johri BN (eds.), Developments in Fungal Biology and Applied
613 Mycology. Singapore: Springer Singapore, 2017, 433–447.
614 11. Borges DF et al. Formulation of botanicals for the control of plant-pathogens: A review. *Crop Prot*
615 2018;**110**:135–140. https://doi.org/10.1016/j.cropro.2018.04.003
616 12. Du S et al. Synthesis, antifungal activity and QSAR of some novel carboxylic acid amides.
617 *Molecules* 2015;**20**:4071–4087.

618 13. Huang K et al. Structure-activity relationship of antibacterial bio-based epoxy polymers made from
619 phenolic branched fatty acids. *Prog Org Coat* 2021;**155**:106228.
620 14. Kazem-Rostami M et al. Antibacterial agents from waste grease: arylation of Brown grease fatty
621 acids with beechwood creosote and derivatization. *ACS Sustain Chem Eng* 2023;**11**:17760–17768.
622 15. Lew HN et al. Methods for preparing phenolic branched chain alkyl fatty acids or esters thereof and
623 methods for killing microorganisms. 2018.
624 16. Ahn E et al. Structure-Activity Relationship of Novel Phenolic-Branched Amides Reveals Potent,
625 Isolate-specific Antifungal Activity and Morphological Disruption of Cacao Fungal Pathogens.
626 2025. ChemRxiv, 2025.
627 17. Lahive F, Hadley P, Daymond AJ. The physiological responses of cacao to the environment and the
628 implications for climate change resilience. A review. *Agron Sustain Dev* 2019;**39**:5.
629 18. Nugroho D, Basunanda P, Mw S. Physical bean quality of Arabica coffee (Coffea arabica)
630 cultivated at high and medium altitude. *Pelita Perkeb* 2016;**32**:151–161.
631 19. Cassamo CT et al. Shade and altitude implications on the physical and chemical attributes of green
632 coffee beans from Gorongosa Mountain, Mozambique. *Agronomy* 2022;**12**:2540.
633 20. Winters NP et al. A combination of conserved and diverged responses underlies Theobroma cacao's
634 defense response to Phytophthora palmivora. *BMC Biol* 2024;**22**:38.
635 https://doi.org/10.1186/s12915-024-01831-2
636 21. de Sousa LP, Guerreiro Filho O, Costa Mondego JM. Differences between the leaf mycobiome of
637 Coffea arabica and wild coffee species and their modulation by caffeine/chlorogenic acid content.
638 *Microorganisms* 2021;**9**:2296.
639 22. Tanabata T et al. SmartGrain: high-throughput phenotyping software for measuring seed shape
640 through image analysis. *Plant Physiol* 2012;**160**:1871–1880.
641 23. Klimberg R. Fundamentals of predictive analytics with JMP. Sas institute, 2023.
642 24. Baek I et al. Unveiling differential responses to UVB (305 nm) and UVC (275 nm) in cacao-
643 infecting Colletotrichum gloeosporioides and Pestalotiopsis sp. *bioRxiv* 2025;2025.05.21.655292.

644 https://doi.org/10.1101/2025.05.21.655292
645 25. Dembitsky VM. Microbiological aspects of unique, rare, and unusual fatty acids derived from
646 natural amides and their pharmacological profile. *Microbiol Res* 2022;**13**:377–417.
647 26. Ōmura S et al. Relationship between the structures of fatty acid amide derivatives and their
648 antimicrobial activities. *Antimicrob Agents Chemother* 1974;**6**:207–215.
649 27. Middelhoven WJ. Catabolism of benzene compounds by ascomycetous and basidiomycetous yeasts
650 and yeastlike fungi: A literature review and an experimental approach. *Antonie Van Leeuwenhoek*
651 1993;**63**:125–144.
652 28. Middelhoven WJ et al. Degradation of some phenols and hydroxybenzoates by the imperfect
653 ascomycetous yeasts Candida parapsilosis and Arxula adeninivorans: evidence for an operative
654 gentisate pathway. *Antonie Van Leeuwenhoek* 1992;**62**:181–187.
655 29. Brown AJ et al. Stress adaptation in a pathogenic fungus. *J Exp Biol* 2014;**217**:144–155.
656 30. Gasch AP. Comparative genomics of the environmental stress response in ascomycete fungi. *Yeast*
657 2007;**24**:961–976.
658 31. Carobene A et al. The multicenter European Biological Variation Study (EuBIVAS): a new glance
659 provided by the Principal Component Analysis (PCA), a machine learning unsupervised algorithms,
660 based on the basic metabolic panel linked measurands. *Clin Chem Lab Med CCLM* 2022;**60**:556–
661 568.
662 32. Tarca AL et al. Machine learning and its applications to biology. *PLoS Comput Biol* 2007;**3**:e116.
663 33. Lyou ES et al. Single-cell phenotypes revealed as a key biomarker in bacterial–fungal interactions: a
664 case study of Staphylococcus and Malassezia. *Microbiol Spectr* 2023;**11**:e00437-23.
665 34. Talento AF et al. A prospective study of fungal biomarkers to improve management of invasive
666 fungal diseases in a mixed specialty critical care unit. *J Crit Care* 2017;**40**:119–127.
667 35. Baek I et al. Dissecting Trichoderma antagonism: Role of strain identity, volatiles, biomass, and
668 morphology in suppressing cacao pathogens. *Biol Control* 2025;**207**:105807.
669 https://doi.org/10.1016/j.biocontrol.2025.105807

670 36. Zhang F et al. The stress response regulator AflSkn7 influences morphological development, stress
671 response, and pathogenicity in the fungus Aspergillus flavus. *Toxins* 2016;**8**:202.
672 37. Sterflinger K, Krumbein W. Multiple stress factors affecting growth of rock-inhabiting black fungi.
673 *Bot Acta* 1995;**108**:490–496.
674 38. Fisher KJ, Lang GI. Experimental evolution in fungi: An untapped resource. *Fungal Genet Biol*
675 2016;**94**:88–94.
676 39. Schoustra SE et al. Reducing the cost of resistance; experimental evolution in the filamentous
677 fungus Aspergillus nidulans. *J Evol Biol* 2006;**19**:1115–1127.
678 40. Mani G, Clarke BC. Mutational order: a major stochastic process in evolution. *Proc R Soc Lond B*
679 *Biol Sci* 1990;**240**:29–37.
680 41. Alster CJ et al. Phenotypic plasticity of fungal traits in response to moisture and temperature. *ISME*
681 *Commun* 2021;**1**:43.
682 42. Kong W et al. Estimating leaf water content at the leaf scale in soybean inoculated with arbuscular
683 mycorrhizal fungi from in situ spectral measurements. *Int J Agric Biol Eng* 2019;**12**:149–155.
684 43. Neurauter M et al. Exploring near-infrared spectroscopy and hyperspectral imaging as novel
685 characterization methods for anaerobic gut fungi. *FEMS Microbes* 2024;**5**:xtae025.
686 44. Teena M et al. Near infrared (NIR) hyperspectral imaging to classify fungal infected date fruits. *J*
687 *Stored Prod Res* 2014;**59**:306–313.
688 45. Valenzuela-Gloria M et al. Molecular characterization of fungal pigments. J Fungi (Basel) 7. 2021.
689 46. Melo D et al. Characterization of fungal melanins from black stains on paper artefacts. *Heritage*
690 2022;**5**:3049–3065.
691 47. Jochemsen A, Alfredsen G, Burud I. Hyperspectral imaging as a tool for profiling basidiomycete
692 decay of Pinus sylvestris L. *Int Biodeterior Biodegrad* 2022;**174**:105464.
693 48. Park E et al. Hyperspectral reflectance imaging for nondestructive evaluation of root rot in Korean
694 ginseng (Panax ginseng Meyer). *Front Plant Sci* 2023;**14**:1109060.
695 49. Rojas EI et al. Colletotrichum gloeosporioides sl associated with Theobroma cacao and other plants

696 in Panama: multilocus phylogenies distinguish host-associated pathogens from asymptomatic
697 endophytes. *Mycologia* 2010;**102**:1318–1338.
698 50. Fumagalli M et al. Signatures of environmental genetic adaptation pinpoint pathogens as the main
699 selective pressure through human evolution. *PLoS Genet* 2011;**7**:e1002355.
700 51. Kroymann J. Natural diversity and adaptation in plant secondary metabolism. *Curr Opin Plant Biol*
701 2011;**14**:246–251.
702