Genotype-specific non-additive heat–drought responses and recovery hysteresis in coffee

Insuck Baek
Beltsville Agricultural Research Center
Sunchung Park
Beltsville Agricultural Research Center
Rakesh K. Upadhyay
Bowie State University
Moon S. Kim
Beltsville Agricultural Research Center
Lyndel W. Meinhardt
Beltsville Agricultural Research Center
Ezekiel Ahn
ezekiel.ahn@usda.gov

Beltsville Agricultural Research Center

Article

Keywords: Coffee, Coffea arabica, Coffea canephora, heat stress, drought stress, combined stress, non-additivity, RNA-seq, recovery, transcriptomics

Posted Date: June 12th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9849440/v1

Additional Declarations: No competing interests reported.

Abstract

Coffee frequently experiences overlapping heat and water-deficit episodes, yet combined-stress responses are not always predictable from single-stress treatments. We reconstructed public RNA-seq data for *Coffea arabica* cv. Icatu and *Coffea canephora* cv. Conilon Clone 153 (CL153) into a balanced genotype-specific design comprising two water regimes, well-watered and severe water deficit, across T25, T37, T42, and 14-day recovery stages, with three biological replicates per condition. Transcript abundance was quantified against genotype-specific references and modeled with DESeq2 using water, stage, and water-by-stage interaction terms. Interaction statistics were combined with non-additivity scores to classify genes as synergistic positive, antagonistic or buffered, or emergent only. Icatu showed a large early interaction burst at T37, with 2,680 non-additive genes dominated by synergistic-positive responses. CL153 showed only 165 non-additive genes at T37, but delayed responses at T42 and recovery, with 650 and 668 non-additive genes, respectively. Functional summaries identified a recovery-stage photosynthesis signal in CL153, whereas Icatu candidates pointed to sugar transport, thylakoid assembly, proteostasis, autophagy, vacuolar trafficking, ubiquitin/ERAD remodeling, and hormone-related regulation. These results show that coffee heat–drought responses are genotype-specific, temporally structured, and non-additive, and recovery-stage expression retained signatures of stress history.

Introduction

Coffee is a perennial tropical crop whose growth, yield, and quality depend strongly on water availability, temperature, irradiance, and their interactions. The two major cultivated species, *Coffea arabica* and *Coffea canephora*, differ in ecological origin and stress tolerance: *C. arabica* is generally associated with milder environments, whereas *C. canephora* is better adapted to warmer lowland conditions. Still, both species can be constrained when water limitation and high temperature coincide, especially through altered stomatal conductance, reduced carbon assimilation, oxidative stress, and damage to photosynthetic and reproductive processes[1,2]. Controlled studies have shown that coffee photosynthesis can tolerate moderate heat in some genotypes, but severe heat and severe water deficit can exceed the buffering capacity of photosynthetic enzymes, electron transport, and protective systems[3]. Recent integrative work also indicates that coffee acclimation under harsh drought and/or heat depends on multiple molecular layers, including aquaporins, dehydrins, heat shock proteins, antioxidant enzymes, chloroplast membrane lipids, and recovery-associated responses[4].

This complexity is not surprising, because heat and drought impose partly conflicting physiological demands. Heat can increase evaporative demand and favor transpirational cooling, whereas drought promotes stomatal closure and limits both water loss and CO_2 diffusion. Their combination therefore does not necessarily behave like a stronger version of either single stress. Studies in several plant systems have shown that combined or sequential stresses can trigger molecular responses that differ from those observed under individual treatments, including interaction-specific transcriptomic and metabolic signatures[5,6]. Multi-omics analyses in crop systems support the view that combined stress responses can involve distinct changes in photosynthesis, hormones, amino acid metabolism, and stress-response networks rather than simple extrapolations from single-stress data[7]. For coffee, this means that stress markers discovered under drought alone or heat alone may not capture the biology of simultaneous heat–drought exposure.

The coffee heat–drought RNA-seq dataset generated by Marques et al.[8] is well suited to this question. That study compared *C. arabica* cv. Icatu and *C. canephora* cv. Conilon Clone 153 (CL153) under drought, heat, combined heat–drought, and recovery conditions, and reported that combined-stress responses were complex, genotype-dependent, and not readily explained by additive effects of single stressors[8]. The dataset includes both acute stress stages and a recovery time point, making it possible to ask not only which genes respond to each treatment, but also which genes depart from additive expectations and whether these departures persist after stress relief.

That distinction is central to the present study. Under an additive expectation, the transcriptional response to combined heat and drought should be predictable from the heat-associated and drought-associated components. Under a non-additive model, the combined condition may amplify, buffer, reverse, or generate responses that are not apparent from the individual treatments. This is more than a statistical distinction: it changes how candidate stress markers are interpreted. A gene that responds strongly to heat alone may be a poor marker of combined stress, whereas a gene with a modest single-stress response may become important if it marks the interaction itself.

Recovery adds another layer to this problem. In perennial crops, stress response cannot be judged only at the point of maximum exposure; the plant must also restore, reorganize, or maintain altered physiological states after stress release. Coffee studies on

repeated drought have shown that previous stress exposure can modify later physiological and metabolic responses, suggesting that stress history can shape subsequent acclimation[9]. More generally, plant stress-memory research has shown that recovery involves active transcriptome resetting as well as memory-like expression patterns, rather than a passive return to the pre-stress state[10,11]. In this study, we use "recovery hysteresis" in this practical transcriptomic sense: persistent recovery-stage expression differences that retain information about the combined-stress history. We do not infer epigenetic memory directly; rather, we use recovery-stage non-additivity as a measurable expression signature of stress-history-dependent transcription.

Here, we reconstructed the public coffee heat–drought RNA-seq data into a balanced genotype-specific design and modeled non-additive responses separately in *C. arabica* Icatu and *C. canephora* CL153. For each genotype, the retained design included two water regimes, well-watered and severe water deficit, across T25, T37, T42, and 14-day recovery stages, with three biological replicates per water-by-stage combination. We quantified transcript abundance using genotype-specific references and fitted a factorial model with water, stage, and water-by-stage interaction terms. We then combined interaction statistics with non-additivity scores to classify genes as synergistic positive, antagonistic or buffered, or emergent only, and to prioritize recovery-stage candidate markers.

This interaction-centered reanalysis revealed a clear difference in timing between the two genotypes. Icatu showed a large early non-additive burst at T37, whereas CL153 showed a smaller T37 response but larger delayed responses at T42 and recovery. Functional summaries further indicated that CL153 recovery-stage non-additivity was associated with a photosynthesis-linked signal, while Icatu candidates pointed to broader early rewiring involving sugar transport, thylakoid assembly, proteostasis, autophagy, vacuolar trafficking, ubiquitin/ERAD-related remodeling, and hormone-related regulation. Together, these results show that combined heat and drought in coffee are better understood as genotype-specific interaction trajectories than as the sum of single-stress responses.

Results

A balanced factorial design isolates heat–drought interaction in two coffee genotypes

To resolve interaction-specific transcriptional behavior under combined heat and drought, we reconstructed a balanced public RNA-seq design for two coffee genotypes, *Coffea arabica* cv. Icatu and *Coffea canephora* cv. CL153. The retained dataset comprised 48 RNA-seq libraries: two water regimes, well-watered (WW) and severe water deficit (SWD), across four stress or recovery stages (T25, T37, T42, and 14-day recovery [REC14]), with three biological replicates for each water-by-stage combination within each genotype (Table 1; Supplementary Data 1). This structure enabled genotype-specific factorial modeling of water regime, stage, and their interaction.

Read mapping differed substantially between the two genotype-specific analyses. Icatu libraries mapped consistently to the *C. arabica* reference, with a median Salmon mapping rate of 89.2% and a range of 84.5–93.1%. CL153 libraries showed greater mapping-rate variability, with a median of 65.7% and a range of 32.1–87.1% (Table 1). Five CL153 libraries, all from T42 or REC14 conditions, mapped below 50% and were flagged for QC-aware interpretation (Supplementary Table S1; Supplementary Data 1). All retained libraries were kept in the primary analysis to preserve the balanced factorial design, while the lower-mapping CL153 libraries were considered explicitly when interpreting severe-heat and recovery-stage results.

PCA of variance-stabilized expression profiles showed genotype-specific stress trajectories (Fig. 1). In Icatu, PC1 explained 33.3% of the variance and separated samples along a temperature/recovery trajectory, while PC2 explained 16.7% and contributed to water-regime separation. In CL153, PC1 explained 29.1% and PC2 explained 21.9%, with greater dispersion among severe-heat and recovery-stage libraries. Thus, both genotypes showed structured transcriptional responses across the experimental series, but CL153 required more cautious interpretation because mapping-rate variability was concentrated in the same severe and recovery stages that drive its strongest downstream signals.

(a) Salmon mapping-rate comparison between Icatu and CL153 retained RNA-seq libraries. Points represent individual libraries, boxes summarize genotype-level distributions, and the bracket indicates a Wilcoxon rank-sum test ($p < 0.001$). (b) PCA of variance-stabilized expression profiles for *C. arabica* Icatu. (c) PCA of variance-stabilized expression profiles for *C. canephora* CL153. PCA panels are colored by stress/recovery stage and shaped by water regime. Low-mapping samples are not labeled in the main PCA panels and are reported in Supplementary Table S1.

Table 1
Retained design and genotype-level QC summary.

Genotype	Water regimes	Stages	Retained libraries	Biological replication	Median mapping rate (%)	Mapping rate range (%)	Libraries < 50% mapped
C. arabica Icatu	WW, SWD	T25, T37, T42, REC14	24	3 per water-stage condition	89.200	84.5–93.1	0
C. canephora CL153	WW, SWD	T25, T37, T42, REC14	24	3 per water-stage condition	65.700	32.1–87.1	5

Concise genotype-level summary of retained libraries, water regimes, stages, biological replication, median mapping rate, mapping-rate range, and number of libraries below 50% mapping. Full sample-level QC is provided in Supplementary Data 1.

Non-additive heat–drought arithmetic differs in timing and class composition between genotypes

We next asked whether combined heat and drought produced transcriptional responses predictable from the component effects or whether the combined condition deviated from additive expectations. For each genotype, expression was modeled using a factorial design with water regime, stage, and water-by-stage interaction terms. Interaction FDRs were combined with non-additivity scores to classify genes as Synergistic Positive, Antagonistic or Buffered, or Emergent Only (Methods). The two genotypes differed strongly in the timing and composition of non-additive responses in the timing and composition of non-additive transcriptional responses (Table 2; Fig. 2; Supplementary Data 1).

Icatu showed a large early interaction burst at T37. At this stage, 2,680 genes were classified as non-additive, including 2,224 Synergistic Positive genes (83.0%), 433 Antagonistic or Buffered genes (16.2%), and 23 Emergent Only genes (0.9%). The Icatu non-additive set contracted at T42 to 947 genes and at REC14 to 719 genes. Although Synergistic Positive genes remained the largest class at both stages, the Antagonistic or Buffered fraction increased to 39.1% at T42 and 36.9% at REC14. Icatu therefore displayed a large, predominantly amplifying T37 response followed by a smaller and more mixed severe-heat/recovery response.

CL153 showed the opposite temporal profile. Only 165 genes were non-additive at T37, split almost evenly between Synergistic Positive genes (79; 47.9%) and Antagonistic or Buffered genes (84; 50.9%), with two Emergent Only genes (1.2%). The non-additive response then expanded at T42 to 650 genes, of which 458 (70.5%) were Synergistic Positive, and remained large at REC14 with 668 non-additive genes, including 493 Synergistic Positive genes (73.8%). Icatu was therefore characterized by an early T37 interaction burst, whereas CL153 showed delayed expansion of non-additive responses at severe heat and recovery.

The two genotypes also differed in the trajectory of non-additivity, not just its magnitude. Icatu and CL153 had similar numbers of REC14 non-additive genes (719 and 668, respectively), but these comparable recovery-stage totals followed different paths: contraction from a large T37 burst in Icatu and expansion from a small T37 response in CL153. This pattern supports a model in which heat–drought interaction timing is genotype-specific.

(a) Total number of non-additive genes by genotype and target stage. (b) Non-additive class counts by genotype and stage. (c) Within-stage class composition for *C. arabica* Icatu. (d) Within-stage class composition for *C. canephora* CL153. Icatu showed a dominant T37 Synergistic Positive burst, whereas CL153 showed delayed T42 and REC14 non-additive expansion.

Table 2
Non-additive gene totals and class composition.

Genotype	Target stage	Total non-additive genes	Synergistic Positive n (%)	Antagonistic or Buffered n (%)	Emergent Only n (%)
C. arabica Icatu	T37	2680	2224 (83.0%)	433 (16.2%)	23 (0.9%)
C. arabica Icatu	T42	947	563 (59.5%)	370 (39.1%)	14 (1.5%)
C. arabica Icatu	REC14	719	450 (62.6%)	265 (36.9%)	4 (0.6%)
C. canephora CL153	T37	165	79 (47.9%)	84 (50.9%)	2 (1.2%)
C. canephora CL153	T42	650	458 (70.5%)	192 (29.5%)	0 (0.0%)
C. canephora CL153	REC14	668	493 (73.8%)	172 (25.7%)	3 (0.4%)

Merged summary of total non-additive gene counts and class composition for each genotype and target stage. Full class-level data are provided in Supplementary Data 1.

Functional module enrichment reveals a recovery-stage photosynthesis signal in CL153

To place non-additive genes into functional context, we tested curated functional modules for enrichment within each genotype-stage non-additive set. Only two module-stage associations met the reporting threshold (Table 3; Fig. 3; Supplementary Table S2). The strongest result was a CL153 REC14 Photosynthesis enrichment: 7 of 668 REC14 non-additive genes belonged to the Photosynthesis module, compared with 35 photosynthesis-associated genes among 20,285 background genes, yielding FDR = 0.0033. This result identifies a discrete recovery-stage photosynthesis signal within the delayed CL153 response.

A second, suggestive association was observed for the CL153 T37 Water Transport / Aquaporin module, with 2 of 165 T37 non-additive genes in the module compared with 16 of 20,285 background genes (FDR = 0.0989). Although this enrichment did not meet FDR < 0.05, it supports the presence of a water-transport-related signal in the early CL153 interaction response.

In contrast, Icatu did not show a module enrichment passing FDR < 0.10 despite having a much larger T37 non-additive set. This absence of a dominant enriched module should not be interpreted as absence of biological structure; rather, it suggests that the Icatu T37 burst is distributed across several regulatory and cellular processes rather than concentrated in one curated functional module. Candidate-level analyses below therefore provide a more appropriate approach for interpreting the broad Icatu response.

(a) Functional module composition of non-additive genes in *C. arabica* Icatu. (b) Functional module composition of non-additive genes in *C. canephora* CL153. (c) FDR-supported or suggestive module enrichment, showing significant CL153 REC14 Photosynthesis enrichment and suggestive CL153 T37 Water Transport / Aquaporin enrichment. The dashed line marks FDR = 0.05. (d) High-confidence candidate counts by functional module and genotype.

Table 3
Significant and suggestive functional module enrichments.

Genotype	Stage	Functional module	Non-additive genes in module	Total non-additive genes in stage	Background module genes	Background genes	FDR
C. canephora CL153	REC14	Photosynthesis	7	668	35	20285	3.30E-03
C. canephora CL153	T37	Water Transport / Aquaporin	2	165	16	20285	9.89E-02

Compact summary of module enrichments passing FDR < 0.10. Full module enrichment statistics are provided in Supplementary Table S2.

Candidate marker axes identify genotype-specific responses to combined heat and drought

Because module enrichment can miss broad but biologically interpretable responses, we prioritized high-confidence candidates using interaction FDR, non-additive score, product interpretability, and candidate-axis relevance (Table 4; Figs. 4 and 5; Supplementary Tables S3-S4; Supplementary Data 1). This candidate-axis approach highlighted distinct biological themes in Icatu and CL153.

In Icatu, the T37 burst included candidates linked to carbon transport, photosynthetic assembly, vacuolar trafficking, autophagy, and chaperone-related proteostasis. Representative candidates included sucrose transport protein SUC4-like (*XM_027238646.1*; Synergistic Positive; non-additive score = 1.56; FDR = 4.68×10^{-19}), protein THYLAKOID ASSEMBLY 8 (*XM_027233171.1*; Emergent Only; non-additive score = 2.22; FDR = 2.94×10^{-7}), vacuolar protein sorting-associated protein 60.2-like (*XM_027255469.1*; Synergistic Positive; non-additive score = 2.00; FDR = 2.38×10^{-17}), autophagy-related protein 8C-like (*XM_027233866.1*; Synergistic Positive; non-additive score = 1.61; FDR = 2.72×10^{-17}), and chaperone protein DnaJ 50-like (*XR_003458410.1*; Emergent Only; non-additive score = 2.23; FDR = 4.82×10^{-5}). These candidates are consistent with a rapid Icatu response involving carbon redistribution, photosynthetic organization, and protein/organelle quality-control pathways.

At T42 and REC14, Icatu candidates shifted toward regulatory and proteostasis-associated responses. A bZIP transcription factor 16-like transcript (*XM_027225538.1*) was Synergistic Positive at T42 with a large non-additive score (9.61; FDR = 1.37×10^{-9}), while an ERAD-associated E3 ubiquitin-protein ligase component HRD3A transcript (*XM_027270143.1*) was Antagonistic or Buffered (non-additive score = -13.54; FDR = 2.28×10^{-8}). During REC14, probable E3 ubiquitin-protein ligase ARI1 (*XM_027234234.1*) remained Synergistic Positive (non-additive score = 4.74; FDR = 4.04×10^{-15}), indicating that ubiquitin-associated proteostasis remained part of the recovery-stage interaction signature.

In CL153, candidate axes were more delayed and were concentrated in transport, membrane trafficking, remodeling, and recovery-associated functions. At T37, Expansin (*GSCOC_T00039583001*) was Antagonistic or Buffered (non-additive score = -1.24; FDR = 6.30×10^{-7}), while probable carotenoid cleavage dioxygenase 4 (*GSCOC_T00024342001*) was Synergistic Positive (non-additive score = 5.82; FDR = 4.30×10^{-3}). At T42, transport-associated buffered candidates included NRT1/PTR FAMILY 6.2 (*GSCOC_T00031111001*; non-additive score = -5.04; FDR = 7.14×10^{-4}) and cation/H(+) antiporter 17 (*GSCOC_T00035493001*; non-additive score = -4.71; FDR = 3.03×10^{-4}). At REC14, CL153 candidates included Syntaxin-21 (*GSCOC_T00019798001*; Synergistic Positive; non-additive score = 7.23; FDR = 7.18×10^{-6}), Metalloendoproteinase 1-MMP (*GSCOC_T00030498001*; Antagonistic or Buffered; non-additive score = -5.86; FDR = 2.40×10^{-3}), and Early nodulin-like protein 2 (*GSCOC_T00021256001*; Antagonistic or Buffered; non-additive score = -3.49; FDR = 3.42×10^{-6}). Together with the REC14 Photosynthesis enrichment, these candidates support a delayed CL153 response involving membrane trafficking, water/ion transport, structural remodeling, and recovery-associated photosynthetic regulation.

(a) High-confidence candidate effect space showing non-additive score versus -log10(FDR). Vertical dotted lines indicate non-additive score thresholds of ± 1, and the horizontal dashed line indicates FDR = 0.05. (b) REC14 recovery-memory candidates. (c) Top candidate priority scores by genotype and stage. (d) Working model summarizing genotype-specific stress arithmetic: Icatu shows an early T37 interaction burst involving sugar transport and proteostasis-related axes, whereas CL153 shows delayed T42/REC14 rewiring associated with photosynthesis-linked recovery and transport-related axes.

Table 4
Representative high-confidence candidate axes.

Genotype	Stage	Candidate axis	Gene/transcript ID	Product annotation	Class	Non-additive score	FDR
C. arabica Icatu	T37	Carbon / sugar transport	XM_027238646.1	sucrose transport protein SUC4-like, transcript variant X1	Synergistic Positive	1.560	4.68E-19
C. arabica Icatu	T37	Photosynthetic assembly	XM_027233171.1	protein THYLAKOID ASSEMBLY 8, chloroplastic-like, transcript variant X1	Emergent Only	2.220	2.94E-07
C. arabica Icatu	T37	Vacuolar trafficking	XM_027255469.1	vacuolar protein sorting-associated protein 60.2-like, transcript variant X1	Synergistic Positive	2.000	2.38E-17
C. arabica Icatu	T37	Autophagy / proteostasis	XM_027233866.1	autophagy-related protein 8C-like, transcript variant X1	Synergistic Positive	1.610	2.72E-17
C. arabica Icatu	T37	Chaperone / proteostasis	XR_003458410.1	chaperone protein dnaJ 50-like, transcript variant X3	Emergent Only	2.230	4.82E-05
C. arabica Icatu	T42	Hormone / transcriptional regulation	XM_027225538.1	bZIP transcription factor 16-like, transcript variant X2	Synergistic Positive	9.610	1.37E-09
C. arabica Icatu	T42	ERAD / ubiquitin proteostasis	XM_027270143.1	ERAD-associated E3 ubiquitin-protein ligase component HRD3A, transcript variant X3	Antagonistic or Buffered	-13.540	2.28E-08
C. arabica Icatu	REC14	Recovery proteostasis	XM_027234234.1	probable E3 ubiquitin-protein ligase ARI1, transcript variant X2	Synergistic Positive	4.740	4.04E-15
C. canephora CL153	REC14	Membrane trafficking	GSCOC_T00019798001	Syntaxin-21	Synergistic Positive	7.230	7.18E-06
C. canephora CL153	REC14	Protease / remodeling	GSCOC_T00030498001	Metalloendoproteinase 1-MMP	Antagonistic or Buffered	-5.860	2.40E-03
C. canephora CL153	REC14	Water/transport-associated response	GSCOC_T00021256001	Early nodulin-like protein 2	Antagonistic or Buffered	-3.490	3.42E-06
C. canephora CL153	T37	Cell-wall remodeling	GSCOC_T00039583001	Expansin	Antagonistic or Buffered	-1.240	6.30E-07
C. canephora CL153	T37	Carotenoid / stress signaling	GSCOC_T00024342001	Probable carotenoid cleavage dioxygenase 4, chloroplastic	Synergistic Positive	5.820	4.30E-03
C. canephora CL153	T42	Nitrate/peptide transporter family	GSCOC_T00031111001	Protein NRT1/ PTR FAMILY 6.2	Antagonistic or Buffered	-5.040	7.14E-04
C. canephora CL153	T42	Ion transport / recovery response	GSCOC_T00035493001	Cation/H(+) antiporter 17	Antagonistic or Buffered	-4.710	3.03E-04

Representative candidate axes selected for interpretability and downstream validation. Full high-confidence candidate data and top-25 candidates by stage are provided in Supplementary Data 1.

Top interaction heatmaps support genotype-specific candidate structure

Top interaction-candidate heatmaps further supported genotype-specific response structures (Fig. 5). In Icatu, the top interaction candidates showed expression patterns consistent with the large T37 interaction burst and with candidate axes including carbon transport, photosynthetic assembly, vacuolar trafficking, autophagy, chaperone/proteostasis, and hormone/transcriptional regulation. In CL153, top candidates showed a distinct structure centered on delayed severe-heat and recovery-stage responses, with top-candidate module composition dominated by ABA-related, transport-associated, and recovery-linked categories. These patterns indicate that Icatu and CL153 do not share a single uniform coffee heat–drought response; rather, they encode distinct non-additive response trajectories.

(a) Heatmap of top interaction candidates in *C. arabica* Icatu. (b) Heatmap of top interaction candidates in *C. canephora* CL153. Rows represent top interaction candidates and columns represent samples ordered by stage, water regime, and replicate. Colors show row-scaled expression Z-scores. (c) Functional module composition of Icatu top-25 candidates by stage. (d) Functional module composition of CL153 top-25 candidates by stage.

Recovery is an active non-additive state in both genotypes

REC14 retained substantial non-additive gene sets in both genotypes: 719 genes in Icatu and 668 genes in CL153 (Table 2). In both cases, Synergistic Positive genes were the largest recovery-stage class, representing 62.6% of the Icatu REC14 set and 73.8% of the CL153 REC14 set. Thus, recovery was not a simple return to baseline expression after combined heat and drought. Instead, combined-stress history continued to shape transcriptional states 14 days after recovery.

The biological content of these recovery-stage signatures differed by genotype. Icatu recovery was anchored by candidate axes involving ubiquitin and proteostasis, including probable E3 ubiquitin-protein ligase ARI1. CL153 recovery was distinguished by significant Photosynthesis module enrichment and candidates related to membrane trafficking, water/ion transport, and remodeling. The results therefore support a recovery-hysteresis model in which both genotypes retain non-additive post-stress signatures, but the dominant recovery axes differ between Icatu and CL153.

In sum, these results show that coffee heat–drought responses are genotype-specific, temporally structured, and non-additive. Icatu displays an early T37 interaction burst dominated by Synergistic Positive responses and candidate axes related to carbon transport, photosynthetic assembly, autophagy, vacuolar trafficking, and proteostasis. CL153 displays delayed T42/REC14 non-additivity, including a significant recovery-stage Photosynthesis module and transport/remodeling-associated candidate axes. These findings provide a framework for interpreting coffee multistress responses through genotype-specific stress arithmetic and recovery-stage hysteresis.

Discussion

Coffee heat–drought responses are better described by interaction logic than by single-stress contrasts

This reanalysis shows that combined heat and drought in coffee cannot be treated as a simple extension of either heat stress or water deficit alone. By reconstructing a balanced public RNA-seq design and explicitly modelling the water-by-stage interaction, we identified genotype- and stage-specific sets of genes whose expression under combined stress deviated from additive expectations. This interaction-centred framing extends the original analysis of the same dataset, which showed that *C. arabica* Icatu and *C. canephora* CL153 exhibit complex responses to combined heat and drought, but did not focus primarily on classifying the arithmetic of the combined-stress response[8].

This result is consistent with a broader body of plant stress literature showing that stress combinations can generate molecular states that are not predictable from individual stress treatments. In *Sorghum bicolor*, combined heat and drought altered a substantial set of transcripts that were not changed under either single stress alone, including stress-specific transcription factors, chaperones, and metabolic pathways[12]. Similar conclusions have emerged from combined heat–drought transcriptomic studies in grasses and cereals, where combined stress produced unique or expanded transcriptional responses relative to individual stresses[13,14]. More broadly, plant omics studies have emphasized that heat and drought can interact through additive, synergistic, antagonistic, dominant, or equalizing

effects across transcripts, proteins, and metabolites[15]. These results place coffee within this model, but add a perennial tropical crop perspective and a recovery-stage dimension.

The important point is not simply that many genes respond to combined stress. Rather, the interaction model identifies genes for which the combined response is different from the response expected from heat and drought components. This distinction matters for marker discovery. A gene that responds strongly to heat or drought alone may not necessarily be informative for combined stress, while a gene with a modest single-stress response may become important if it marks the interaction itself. The non-additive classes used here—synergistic positive, antagonistic or buffered, and emergent only—therefore provide a more explicit vocabulary for describing coffee multistress transcriptional behaviour.

Icatu and CL153 differ in the timing of non-additive response

Icatu and CL153 differed most clearly in the timing of their non-additive responses. Icatu showed a large early non-additive burst at T37, with 2,680 non-additive genes, whereas its T42 and REC14 responses were smaller. CL153 showed the opposite profile: only 165 non-additive genes at T37, followed by larger delayed responses at T42 and REC14. Thus, the two genotypes differ not only in the number of interaction-responsive genes but also in when those genes are deployed.

This temporal contrast has a plausible biological basis: Combined-stress transcriptomes are known to vary with genotype, tissue, stress intensity, and developmental or sampling stage. Barley varieties with contrasting stress response phenotypes differed in the scale and composition of their heat, drought, and combined heat–drought transcriptomic responses, including enrichment of Hsp70 chaperone-associated genes under combined stress[14]. Stage-specific transcriptome profiling in wheat also showed that combined drought and heat can produce stage-dependent regulatory patterns rather than a uniform response across development[16]. Although those systems differ from coffee, they support the broader idea that stress combinations are temporally structured.

For coffee, this means that Icatu and CL153 should not be interpreted simply as “responsive” or “less responsive” genotypes. Icatu appears to activate a large early interaction program at moderate heat, whereas CL153 shows a delayed expansion of non-additivity at severe heat and recovery. This interpretation is also consistent with the integrated coffee stress-resilience study showing genotype- and stress-dependent responses across transcript, protein, and lipid layers in Icatu and CL153 under severe water deficit, heat, combined stress, and recovery[4]. The current reanalysis sharpens that conclusion by showing that the timing of deviation from additivity differs between the two genotypes.

Icatu’s early burst points to broad candidate-level remodeling rather than one dominant enriched pathway

Icatu's T37 response was the largest in the study. The large number of non-additive genes, dominated by synergistic-positive responses, suggests that Icatu rapidly amplifies a broad transcriptional program when heat and drought first converge. Functional enrichment did not resolve a single FDR-significant module for Icatu, but the high-confidence candidate set points to a biologically coherent set of axes: sugar transport, thylakoid assembly, DnaJ/chaperone-related activity, autophagy, vacuolar trafficking, and ubiquitin/ERAD-associated proteostasis.

The absence of a single enriched module should not be overinterpreted as absence of biological structure. Large transcriptional bursts often distribute across several functional categories and may be poorly captured by manually curated enrichment bins. In this case, the candidate-level pattern is consistent with known cellular consequences of heat and drought. Heat can destabilize proteins and membranes, while drought alters water status, carbon assimilation, and oxidative balance. Under these conditions, chaperones, ubiquitin-dependent protein quality control, and autophagy can help remove damaged or misfolded proteins and recycle cellular material.

Several lines of plant stress literature support this interpretation. NBR1-mediated selective autophagy targets ubiquitinated protein aggregates generated under heat and other abiotic stresses, and *nbr1* mutants accumulate insoluble ubiquitinated proteins under stress[17]. A related study showed that the chaperone-associated E3 ubiquitin ligase CHIP and NBR1-mediated selective autophagy act as complementary anti-proteotoxic pathways under plant stress, including heat stress[18]. Drought-focused work further shows that autophagy can regulate water-stress responses by removing damaged proteins, excess heme, and even aquaporins, and by contributing to cellular resetting during drought recovery[19]. The presence of Icatu candidates such as ATG8C-like, VPS60.2-like, DnaJ-

like, E3 ubiquitin ligase, and ERAD-related genes is therefore compatible with an early proteostasis and recycling response, although it remains a transcriptomic hypothesis rather than a validated mechanism.

This Icatu candidate set also links proteostasis with carbon and photosynthetic organization. SUC4-like sugar transport and thylakoid assembly candidates suggest that the early interaction burst may involve source-sink and chloroplast-associated adjustments in addition to stress-protein turnover. Coffee physiological studies have shown that photosynthesis and protective systems can remain relatively resilient up to moderate heat, but more severe heat and water deficit can impair photosynthetic enzymes, chloroplast components, and recovery dynamics[3]. Icatu's T37 burst may therefore represent an early attempt to preserve or reorganize carbon transport, chloroplast function, and protein quality control before more severe heat is reached.

CL153 shows delayed non-additivity and a recovery-stage photosynthesis signal

CL153 showed a smaller T37 response but larger non-additive sets at T42 and REC14. The most interpretable module-level result in the study was the CL153 REC14 photosynthesis enrichment, while the T37 water transport/aquaporin module was suggestive. This pattern supports a delayed-response interpretation: CL153 non-additivity is most visible during severe heat and recovery, rather than during the earlier T37 stage.

Photosynthesis is a credible axis for delayed recovery. Heat stress directly affects photosystem function, electron transport, Calvin-cycle enzymes, chlorophyll biosynthesis, and chloroplast protein stability[20]. In coffee, photosynthetic machinery can show resilience under moderate heat, but severe heat and superimposed water deficit can strongly affect photosynthetic performance and recovery behaviour[3]. The integrated Icatu/CL153 study also reported lasting transcript, protein, and lipid effects at Rec14, including altered chloroplast membrane lipid profiles and stress-responsive proteins such as aquaporins, chaperonins, dehydrins, HSP70, APXs, and CAT[4]. Against that background, the CL153 REC14 photosynthesis enrichment in our analysis is best interpreted as a recovery-stage photosynthetic reorganization signal rather than as evidence of full photosynthetic recovery.

The CL153 candidate genes are also consistent with delayed transport and membrane-related adjustment. Syntaxin-21, cation/H^+ antiporter 17, NRT1/PTR-family candidates, Early nodulin-like protein 2, and metalloendoproteinase candidates point toward membrane trafficking, ion transport, cell-wall or membrane remodeling, and signalling. These candidate axes do not prove mechanism, but they fit the idea that severe heat and recovery involve re-establishing membrane transport, cellular compartmentation, and photosynthetic function. Similar recovery transcriptome studies in tea and tomato have shown that rewatering or recovery can involve large-scale changes in hormone signalling, transcription factors, carbon metabolism, photosynthesis-related genes, and protein kinases rather than a simple reversal of the stress response[21,22].

The CL153 result should nevertheless be interpreted with explicit caution. Several CL153 samples from severe heat or recovery conditions had lower Salmon mapping rates and were retained to preserve the balanced factorial design. This does not invalidate the genotype-level pattern, but it does mean that CL153 T42 and REC14 results should be treated as QC-aware candidate modules. The photosynthesis enrichment is the strongest statistical module signal in the study, but independent validation is needed before assigning it a causal role in CL153 recovery physiology.

Recovery-stage non-additivity indicates persistent transcriptional effects of stress history

The recovery stage is one of the most important parts of this reanalysis. Both genotypes retained large non-additive sets at REC14: 719 genes in Icatu and 668 in CL153. These numbers are not trivial residuals after stress release. They indicate that recovery remains transcriptionally structured by the previous combined-stress history.

This observation fits with plant stress-memory literature, but it should be framed carefully. We do not infer epigenetic memory, transgenerational memory, or priming mechanism directly from these RNA-seq data. Rather, we use "recovery hysteresis" in a practical transcriptomic sense: expression at recovery does not simply return to the WW baseline, and part of the recovery state remains dependent on the prior combined-stress condition. This is consistent with broader plant stress-memory frameworks, which emphasize that recovery can involve active resetting, persistent expression states, RNA turnover, and memory-like response classes[10,11].

Coffee is relevant to this idea because it is perennial. In annual model systems, recovery is often treated as a short post-stress phase. In a perennial crop, however, recovery may influence the next cycle of growth, flowering, carbon allocation, and stress responsiveness. Coffee plants exposed to repeated drought cycles can show photosynthetic and metabolic acclimation, with altered antioxidant metabolism, carbon metabolism, and gene expression that help maintain an “alert” state for later stress exposure[9]. The REC14 non-additivity reported here is not the same as repeated-drought memory, but they point in a related direction: stress history remains visible after the acute combined treatment has ended.

The two genotypes appear to populate this recovery state differently. Icatu REC14 candidates retain ubiquitin/proteostasis and regulatory signatures, whereas CL153 REC14 is associated with photosynthesis enrichment and membrane/transport candidates. This difference suggests that recovery hysteresis may itself be genotype-specific. Icatu may emphasize regulatory and proteostasis-linked adjustment, while CL153 may emphasize delayed photosynthetic and transport-associated reorganization. These interpretations should be validated physiologically, but they provide testable hypotheses for future recovery experiments.

Candidate markers should be treated as genotype- and stage-specific axes

The high-confidence candidate panel is a practical output of the study, but it should not be presented as a universal marker list for coffee stress resilience. The candidates are best interpreted as genotype- and stage-specific axes that prioritize future validation. In Icatu, the most compelling axis is the T37 early burst, including sugar transport, thylakoid assembly, autophagy, DnaJ/chaperone activity, vacuolar trafficking, and ubiquitin/ERAD-related candidates. In CL153, the strongest axis is delayed T42/REC14 response, especially photosynthesis-linked recovery and transport or membrane-associated candidates.

This distinction is important for experimental follow-up. If validation resources are limited, the first validation target should not necessarily be the single gene with the smallest FDR. A stronger validation design would combine interaction FDR, non-additivity score, annotation interpretability, recovery-memory score where relevant, and consistency with the genotype-stage pattern. For example, Icatu T37 SUC4-like, THYLAKOID ASSEMBLY 8, ATG8C-like, VPS60.2-like, DnaJ-like, and E3/ERAD candidates would test whether the early burst reflects carbon transport, chloroplast organization, and proteostasis. CL153 REC14 photosynthesis and transport candidates would test whether delayed non-additivity marks photosynthetic repair, incomplete recovery, or persistent stress injury.

The marker panel also highlights a limitation of single-stress screening. Candidate genes selected from drought-only or heat-only contrasts may miss genes that matter specifically because they encode interaction behaviour. Combined-stress breeding or screening should therefore include interaction-aware candidate selection when factorial designs are available.

Strengths, caveats, and interpretation boundaries

This study reanalyzes a public dataset to extract interaction-specific information rather than to generate new sequencing data. By reconstructing a balanced factorial design and modeling water-by-stage interactions separately for each genotype, the analysis distinguishes main stress effects from interaction-specific responses. The workflow also provides non-additivity scores, class composition, module summaries, candidate prioritization, and QC-aware tables that can be reused or extended.

Several limitations should be stated clearly. First, this is a public-data transcriptomic reanalysis, not a new physiological experiment. The proposed candidate genes and modules should therefore be interpreted as prioritized hypotheses rather than validated causal determinants of coffee resilience. Second, the two genotypes were quantified against genotype-specific references. This is appropriate for within-genotype modeling, but it limits direct ortholog-by-ortholog comparison across species. The strongest cross-genotype claims in this study concern timing, class composition, module trends, and candidate categories, not one-to-one ortholog equivalence.

Third, CL153 had greater mapping-rate variability than Icatu, including five samples below 50% mapping concentrated in severe heat and recovery conditions. These samples were retained to preserve the balanced design, but CL153 severe/recovery interpretation should remain QC-aware. Fourth, functional modules were assigned by annotation keyword matching. This provides interpretable summaries, but it is not equivalent to curated pathway reconstruction. For this reason, the CL153 REC14 photosynthesis module can be discussed as the strongest enrichment-level result, whereas most Icatu biology should be framed as candidate-level evidence.

Finally, some of the broader literature on multifactorial stress combinations, multi-omics integration, epigenetics, and interpretable machine learning is relevant to the long-term field, but it is not central to the current claims. The present study does not model epigenetic mechanisms, microbiomes, machine learning prediction, or field performance. Keeping those boundaries clear makes the main claim stronger: coffee heat–drought transcriptional response is genotype-specific, temporally structured, and non-additive.

Future directions

The next step is targeted validation. For Icatu, validation should focus on the T37 interaction burst and test whether sugar transport, thylakoid assembly, autophagy, chaperone, vacuolar trafficking, and ubiquitin/ERAD candidates are reproducible under independent combined heat–drought experiments. For CL153, validation should focus on T42 and REC14, especially whether photosynthesis-linked and transport-associated candidates reflect repair, delayed injury, or acclimation.

A second direction is physiological integration. Candidate expression should be linked to photosynthetic parameters, stomatal conductance, leaf water potential, chlorophyll fluorescence, ROS markers, carbohydrate status, membrane stability, and recovery performance. This is particularly important for CL153, where the REC14 photosynthesis signal is statistically strong but QC-aware interpretation is needed.

A third direction is orthology-aware comparative analysis. Because this study used genotype-specific transcriptome references, future work could map candidate genes into orthogroups and ask whether Icatu and CL153 deploy conserved genes at different times, or whether their non-additive responses are largely species- or genotype-specific. Such analysis would be especially useful across broader coffee germplasm panels.

Finally, future experiments should treat recovery as part of the stress phenotype. For perennial tropical crops, the most useful marker may not be the gene that responds most strongly at peak stress, but the gene that predicts whether the plant can reorganize and resume function after stress release. The REC14 results in this study suggest that recovery-stage non-additivity is a rich source of such candidates.

Conclusion

In summary, this reanalysis shows that coffee heat–drought responses are genotype-specific, temporally structured, and non-additive. Icatu displays a large early T37 interaction burst dominated by synergistic-positive responses and candidate axes linked to sugar transport, thylakoid assembly, and proteostasis-related remodeling. CL153 shows delayed T42 and REC14 non-additivity, with the strongest enrichment-level signal in recovery-stage photosynthesis. Together, these patterns support a model in which coffee genotypes differ not only in the magnitude of combined-stress response, but also in the timing and biological composition of their interaction programs. Recovery is not a passive return to baseline; it remains an active transcriptional state shaped by previous combined-stress exposure.

Materials and Methods

Public RNA-seq data and study design

Public RNA-seq data were reanalysed from the coffee heat–drought transcriptomic study of Marques et al.[8], which first reported that *Coffea arabica* and *Coffea canephora* respond more complexly to combined heat and drought than to either stress in isolation. Run-level metadata for the five contributing BioProjects (PRJNA787748, PRJNA1087442, PRJNA1087679, PRJNA1088119, and PRJNA1135679) were retrieved from the European Nucleotide Archive (ENA) read-run file-report endpoint[23], with NCBI SRA RunInfo tables retained for inspection alongside the ENA fields.

From this combined metadata set, we curated a balanced factorial design covering both genotypes used in the original study: *C. arabica* cv. Icatu and *C. canephora* cv. Conilon Clone 153 (CL153). For each genotype, runs corresponding to the ambient-CO_2 design were retained under two water regimes (well-watered, WW; severe water deficit, SWD) and four stress/recovery stages (T25, T37, T42, and a 14-day recovery time point, REC14), each represented by three biological replicates. Runs outside this factorial scope were excluded, yielding 24 libraries per genotype and 48 libraries in total. Final genotype, water, stage, and replicate assignments were resolved from ENA sample aliases, sample titles, and experiment titles using study-specific curation logic provided in the accompanying repository.

Reference transcriptomes and transcript-to-gene mapping

Transcript abundance was quantified against genotype-specific references. For *C. arabica*, we used the NCBI RefSeq Cara_1.0 assembly (GCF_003713225.1), including the RefSeq RNA FASTA and corresponding GFF annotation[24]. For *C. canephora*, we used

Ensembl Plants release 62 files for the AUK_PRJEB4211_v1 assembly, including the cDNA FASTA, GFF3 annotation, and protein FASTA[25]. Each reference was retained in its native namespace so that transcript and gene identifiers remained traceable to their primary source.

Transcript-to-gene mapping tables were built directly from the transcript FASTA headers. Where headers exposed a gene-level identifier (RefSeq $[\geq \ \neq \ =$ or $Ge \neq ID:$ tags for *C. arabica*; Ensembl $\geq \ \neq \ :$ tags for *C. canephora*), the transcript was assigned to that gene; where a header carried only a transcript identifier, the transcript was retained as its own transcript-group identifier. This preserves a uniform feature definition across the two references and allows transcript-level features without a parent gene to flow through the same downstream model. For *C. arabica*, transcript version suffixes were stripped before joining annotation tables to the Salmon output, so that quantification identifiers matched the first column of the $tx2 \geq \ \neq$ table.

Read quality control and transcript quantification

Raw FASTQ files were assessed with FastQC[26] and the resulting per-sample reports were aggregated with MultiQC[27]. Genotype-specific Salmon indices were then constructed from the transcript FASTA files at *k*-mer length 31, and paired-end libraries were quantified with Salmon 1.10.3 in selective-alignment mode with automatic library-type detection, sequence-bias and GC-bias correction enabled ($-lA - \ -valteMap\pi ngs - \ -seqBias - \ -gcBias$)[28]. To make the workflow restartable on a single local workstation, both FastQC and Salmon calls were wrapped in a resume-aware bash script that skipped any sample for which a non-empty FastQC archive or a Salmon $quant.$ file was already present.

Sample-level Salmon statistics were extracted from the per-sample $aux_{\in} f\frac{o}{m}\eta_{\in} f \odot json$ files to obtain the number of processed and mapped fragments and the percent mapped, which were combined with the curated sample metadata into a single QC table. Libraries with a mapping rate below 50% were flagged but retained in the primary analysis so that the balanced factorial design of the interaction model was preserved. Their identities and mapping statistics were retained in the QC output tables and were considered during QC-aware interpretation of CL153 severe-heat and recovery-stage results.

Import and differential expression modelling

Salmon transcript-level quantifications were imported into R using $t\xi mp \ \text{or} \ t$ and summarised to the gene or transcript-group level using the mapping tables described above, with $countsFromAbundance = \text{lengthScaledTPM}$ to propagate effective transcript-length information while providing count matrices compatible with DESeq2[29,30]. Inferential replicates were not used. Features with low support — defined as fewer than five counts in fewer than 10% of samples, with a minimum of two samples — were removed before model fitting.

Differential expression and interaction modelling were carried out separately for each genotype in DESeq2[30]. For each genotype-specific dataset, we fitted the factorial design

$$\text{expression} \sim \ \text{water} + \text{stage} + \text{water} : \text{stage}$$

with $water$ (WW, SWD) and $sta \geq$ (T25, T37, T42, REC14) treated as fixed factors, and with WW and T25 as the respective baseline levels. The water-by-stage interaction coefficients for T37, T42, and REC14 were extracted as the primary test of departure from an additive heat–drought response. Variance-stabilised expression matrices were generated from the fitted models and saved as serialised R objects, providing the substrate for downstream principal component analysis on the 500 most variable features per genotype and for candidate-level inspection.

Non-additivity scoring and response classification

To complement the interaction *p*-values with an interpretable effect-size measure, we calculated a non-additivity score from the normalized $\log_2$-transformed expression. For each post-baseline stage *s* ∈ {T37, T42, REC14} we defined

$$\text{nonadditivity}\,(s) = (\text{SWD}_{\text{s}} - \text{WW}_{\text{T25}}) - (\text{WW}_{\text{s}} - \text{WW}_{\text{T25}}) - (\text{SWD}_{\text{T25}} - \text{WW}_{\text{T25}})$$

i.e. the deviation of the observed combined response from the sum of the component heat and drought effects measured against the WW T25 baseline. Positive scores indicate amplification beyond additive expectations, and negative scores indicate damping or reversal. A gene or transcript-group was classified as non-additive when the DESeq2 interaction FDR was below 0.05 and the absolute non-additivity score was at least 1 (corresponding to roughly a twofold deviation on the linear scale). Within the non-additive set, genes

with positive scores were assigned to the synergistic positive class, genes with negative scores to the antagonistic or buffered class, and genes that combined a substantial combined-stress response (|combo| ≥ 1) with weak individual-stress effects (|heat| < 0.5 and |drought| < 0.5) to the emergent only class. For REC14, we additionally computed a recovery-memory score as

$$\text{recovery-memory} = \text{SWD}_{\text{REC14}} - \text{WW}_{\text{REC14}}$$

in normalised $\log_2$ expression, providing a stage-specific magnitude of residual response 14 days after the temperature treatment had been released.

Functional annotation and module enrichment

Functional product annotations were generated from the reference GFF files and transcript or protein FASTA headers and were joined to the non-additivity score tables via genotype-specific transcript identifiers. Each annotated feature was assigned to one of nine manually curated functional modules by keyword matching of the product description: photosynthesis (e.g. photosystem, chlorophyll, RuBisCO, light-harvesting, thylakoid); proteostasis / HSP / chaperone; dehydration / LEA / osmoprotection; water transport / aquaporin; ABA signalling; ROS / redox; lipid / membrane; protease / vacuole / ubiquitin; and TF / hormone crosstalk. Features whose annotations did not match any module were collected into an "other or unannotated" category, which also served as the explicit complement set for enrichment testing.

Over-representation of each module within the non-additive set was tested against the genotype-specific feature background using the hypergeometric distribution. *p*-values were adjusted for multiple testing across all module × stage tests within each genotype using the Benjamini–Hochberg false discovery rate procedure[31]. Module enrichments are reported as significant at FDR < 0.05 and suggestive at FDR < 0.10.

Candidate prioritisation and high-confidence shortlist

Within each genotype × stage combination, non-additive features were prioritised jointly by interaction FDR, absolute non-additivity score, the REC14 recovery-memory score where applicable, and annotation interpretability. To support manuscript-level reporting we additionally filtered for a high-confidence shortlist by removing features whose product annotations matched any of $uncharacterized$, $hypothetical$, $unknown$, $ncRNA$, $rRNA$, or related low-information tags. The shortlist was used to build the representative candidate marker axes reported in the main tables; the full ranked candidate list is retained as supplementary data so that less interpretable features remain available for downstream re-evaluation.

Reproducibility, code, and data availability

The complete computational workflow was executed on a single Windows workstation with Ubuntu under the Windows Subsystem for Linux, using a project-local $micromamba$ environment named $coffee - rnaseq$ containing Python 3.11, Salmon 1.10.3, FastQC, MultiQC, R, DESeq2, $t\xi mp$ or t, and the $openxlsx$ R package used for workbook consolidation. The public analysis repository accompanying this study contains the scripts required for metadata retrieval, reference preparation, FASTQ download, read quality control, Salmon quantification, DESeq2 interaction modelling, non-additivity scoring, functional annotation, module enrichment, candidate prioritisation, and tabular output consolidation. Final analysis outputs are exported as TSV/CSV files and as a single consolidated Excel workbook. Manuscript figures were produced from these exported tables and from the saved variance-stabilised expression matrices outside the scope of the public repository; the figure-assembly step is not required to reproduce any of the computational results reported here.

Raw sequencing reads were obtained from the public BioProjects listed above and are not redistributed by this repository. The workflow downloads run-level metadata and FASTQ locations from ENA/NCBI using BioProject accessions PRJNA787748, PRJNA1087442, PRJNA1087679, PRJNA1088119, and PRJNA1135679. Genotype-specific reference files are downloaded programmatically from their primary providers: the NCBI RefSeq *Coffea arabica* Cara_1.0 assembly GCF_003713225.1, including the RefSeq RNA FASTA and corresponding genomic GFF annotation, and Ensembl Plants release 62 for *Coffea canephora* assembly AUK_PRJEB4211_v1, including the cDNA FASTA, GFF3 annotation, and protein FASTA.

Data availability

The public RNA-seq data reanalyzed in this study were originally generated and reported by Marques et al. in International Journal of Molecular Sciences 25, 7995 (2024), "Transcriptomic Analyses Reveal That *Coffea arabica* and *Coffea canephora* Have More Complex

Responses under Combined Heat and Drought than under Individual Stressors," DOI: https://doi.org/10.3390/ijms25147995. Raw sequencing reads and run-level metadata are publicly available through NCBI SRA/ENA under BioProject accessions PRJNA787748, PRJNA1087442, PRJNA1087679, PRJNA1088119, and PRJNA1135679.

Reference files used for transcript quantification were obtained from their primary providers. For *Coffea arabica*, the workflow uses the NCBI RefSeq Cara_1.0 assembly GCF_003713225.1, including the RefSeq RNA FASTA and corresponding genomic GFF annotation. For *Coffea canephora*, the workflow uses Ensembl Plants release 62 for assembly AUK_PRJEB4211_v1, including the cDNA FASTA, GFF3 annotation, and protein FASTA. These large public source files are not redistributed in the repository but are downloaded programmatically by the workflow.

Curated sample metadata, Salmon mapping-rate summaries, low-mapping sample flags, DESeq2 interaction results, non-additivity score tables, response-class summaries, module enrichment outputs, recovery-memory scores, high-confidence candidate marker summaries, and top interaction candidate tables supporting the findings of this manuscript are provided in Supplementary Data 1 associated with this article. The custom code used for metadata retrieval, reference preparation, FASTQ download, transcript quantification, DESeq2 interaction modeling, non-additivity scoring, module enrichment, candidate prioritization, and tabular output consolidation is openly available at: https://github.com/EJSAHN/coffee-heat-drought-nonadditivity. All other data and materials supporting this study are available from the corresponding author upon reasonable request.

Declarations

Competing interests

The authors declare no competing financial or non-financial interests.

Author Contribution

E.A. conceptualized and supervised the project. I.B., S.P. and E.A. performed the formal analysis. I.B. developed the software and validated the data. I.B., S.P. and E.A. were responsible for data curation. S.P., R.K.U., M.S.K., L.W.M, and E.A. provided resources and methodology. L.W.M. acquired the funding. E.A. wrote the original draft. All authors reviewed and edited the manuscript.

Acknowledgement

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 is supported by the U.S. Department of Agriculture, Agricultural Research Service, In-House Projects No. 8042-21220-258-000-D and 8042-21000-303-000-D. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Data Availability

The public RNA-seq data reanalyzed in this study were originally generated and reported by Marques et al. in International Journal of Molecular Sciences 25, 7995 (2024), "Transcriptomic Analyses Reveal That Coffea arabica and Coffea canephora Have More Complex Responses under Combined Heat and Drought than under Individual Stressors," DOI: https://doi.org/10.3390/ijms25147995. Raw sequencing reads and run-level metadata are publicly available through NCBI SRA/ENA under BioProject accessions PRJNA787748, PRJNA1087442, PRJNA1087679, PRJNA1088119, and PRJNA1135679.Reference files used for transcript quantification were obtained from their primary providers. For Coffea arabica, the workflow uses the NCBI RefSeq Cara_1.0 assembly GCF_003713225.1, including the RefSeq RNA FASTA and corresponding genomic GFF annotation. For Coffea canephora, the workflow uses Ensembl Plants release 62 for assembly AUK_PRJEB4211_v1, including the cDNA FASTA, GFF3 annotation, and protein FASTA. These large public source files are not redistributed in the repository but are downloaded programmatically by the workflow.Curated sample

metadata, Salmon mapping-rate summaries, low-mapping sample flags, DESeq2 interaction results, non-additivity score tables, response-class summaries, module enrichment outputs, recovery-memory scores, high-confidence candidate marker summaries, and top interaction candidate tables supporting the findings of this manuscript are provided in Supplementary Data 1 associated with this article. The custom code used for metadata retrieval, reference preparation, FASTQ download, transcript quantification, DESeq2 interaction modeling, non-additivity scoring, module enrichment, candidate prioritization, and tabular output consolidation is openly available at: https://github.com/EJSAHN/coffee-heat-drought-nonadditivity. All other data and materials supporting this study are available from the corresponding author upon reasonable request.

References

1. DaMatta, F. M., Avila, R. T., Cardoso, A. A., Martins, S. C. & Ramalho, J. C. Physiological and agronomic performance of the coffee crop in the context of climate change and global warming: a review. *J. Agric. Food Chem.* 66, 5264–5274 (2018).
2. Venancio, L. P. *et al.* Impact of drought associated with high temperatures on Coffea canephora plantations: a case study in Espírito Santo State, Brazil. *Sci. Rep.* 10, 19719 (2020).
3. Dubberstein, D. *et al.* Resilient and sensitive key points of the photosynthetic machinery of Coffea spp. to the single and superimposed exposure to severe drought and heat stresses. *Front. Plant Sci.* 11, 1049 (2020).
4. Ramalho, J. C. *et al.* Stress resilience in Coffea arabica and Coffea canephora under harsh drought and/or heat conditions: selected genes, proteins, and lipid integrated responses. *Front. Plant Sci.* 16, 1623156 (2025).
5. Georgii, E. *et al.* Relationships between drought, heat and air humidity responses revealed by transcriptome-metabolome co-analysis. *BMC Plant Biol.* 17, 120 (2017).
6. Alptekin, B. & Kunkowska, A. B. How plants adapt to combined and sequential abiotic stresses: A transcriptomics approach. (2025).
7. Zagorščak, M. *et al.* Integration of multi-omics data and deep phenotyping provides insights into responses to single and combined abiotic stress in potato. *Plant Physiol.* 197, kiaf126 (2025).
8. Marques, I. *et al.* Transcriptomic analyses reveal that Coffea arabica and Coffea canephora have more complex responses under combined heat and drought than under individual stressors. *Int. J. Mol. Sci.* 25, 7995 (2024).
9. Menezes-Silva, P. E. *et al.* Photosynthetic and metabolic acclimation to repeated drought events play key roles in drought tolerance in coffee. *J. Exp. Bot.* 68, 4309–4322 (2017).
10. Crisp, P. A., Ganguly, D., Eichten, S. R., Borevitz, J. O. & Pogson, B. J. Reconsidering plant memory: Intersections between stress recovery, RNA turnover, and epigenetics. *Sci. Adv.* 2, e1501340 (2016).
11. Alves de Freitas Guedes, F., Menezes-Silva, P. E., DaMatta, F. M. & Alves-Ferreira, M. Using transcriptomics to assess plant stress memory. *Theor. Exp. Plant Physiol.* 31, 47–58 (2019).
12. Johnson, S. M. *et al.* Transcriptomic analysis of Sorghum bicolor responding to combined heat and drought stress. *BMC Genomics* 15, 456 (2014).
13. Martin, R. C., Kronmiller, B. A. & Dombrowski, J. E. Transcriptome analysis of Lolium temulentum exposed to a combination of drought and heat stress. *Plants* 10, 2247 (2021).
14. Mahalingam, R. *et al.* Heat and drought induced transcriptomic changes in barley varieties with contrasting stress response phenotypes. *Front. Plant Sci.* Volume 13-2022, (2022).
15. Xu, X., Fonseca de Lima, C. F., Vu, L. D. & De Smet, I. When drought meets heat – a plant omics perspective. *Front. Plant Sci.* Volume 14-2023, (2023).
16. Han, J. *et al.* Stage-specific transcriptomic responses of wheat to heat, drought, and combined stress: AP2/ERF regulatory networks and functional characterization of TaEREBP1-L. *BMC Plant Biol.* 25, 1594 (2025).
17. Zhou, J. *et al.* NBR1-Mediated Selective Autophagy Targets Insoluble Ubiquitinated Protein Aggregates in Plant Stress Responses. *PLOS Genet.* 9, e1003196 (2013).
18. Zhou, J. *et al.* E3 Ubiquitin Ligase CHIP and NBR1-Mediated Selective Autophagy Protect Additively against Proteotoxicity in Plant Stress Responses. *PLOS Genet.* 10, e1004116 (2014).
19. Tang, J. & Bassham, D. C. Autophagy during drought: function, regulation, and potential application. *Plant J.* 109, 390–401 (2022).
20. Wang, X., Xu, C., Cai, X., Wang, Q. & Dai, S. Heat-Responsive Photosynthetic and Signaling Pathways in Plants: Insight from Proteomics. *Int. J. Mol. Sci.* 18, 2191 (2017).

21. Liu, S.-C. *et al.* Transcriptomic Analysis of Tea Plant Responding to Drought Stress and Recovery. *PLOS ONE* 11, e0147306 (2016).
22. Iovieno, P. *et al.* Transcriptomic Changes Drive Physiological Responses to Progressive Drought Stress and Rehydration in Tomato. *Front. Plant Sci.* Volume 7-2016, (2016).
23. Cummins, C. *et al.* The European nucleotide archive in 2021. *Nucleic Acids Res.* 50, D106–D110 (2022).
24. O'Leary, N. A. *et al.* Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.* 44, D733–D745 (2016).
25. Yates, A. D. *et al.* Ensembl Genomes 2022: an expanding genome resource for non-vertebrates. *Nucleic Acids Res.* 50, D996–D1003 (2022).
26. Andrews, S. FastQC: a quality control tool for high throughput sequence data. 2010. (2017).
27. Ewels, P., Magnusson, M., Lundin, S. & Käller, M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* 32, 3047–3048 (2016).
28. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* 14, 417–419 (2017).
29. Soneson, C., Love, M. I. & Robinson, M. D. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Research* 4, 1521 (2016).
30. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550 (2014).
31. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B Methodol.* 57, 289–300 (1995).

Figures

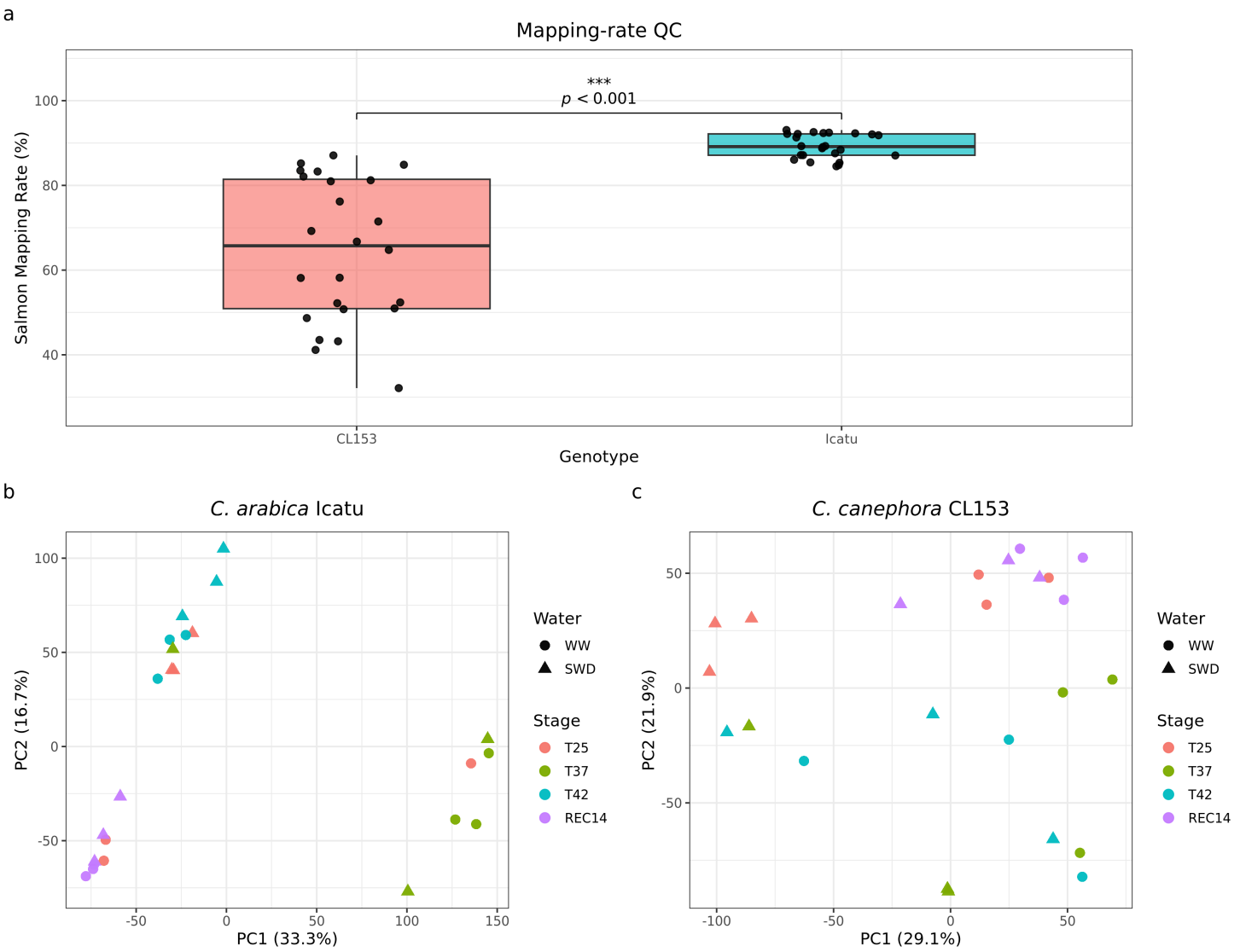

Figure 1

QC-aware reconstruction of coffee heat–drought transcriptomic trajectories.

(a) Salmon mapping-rate comparison between Icatu and CL153 retained RNA-seq libraries. Points represent individual libraries, boxes summarize genotype-level distributions, and the bracket indicates a Wilcoxon rank-sum test ($p < 0.001$). (b) PCA of variance-stabilized expression profiles for *C. arabica* Icatu. (c) PCA of variance-stabilized expression profiles for *C. canephora* CL153. PCA panels are colored by stress/recovery stage and shaped by water regime. Low-mapping samples are not labeled in the main PCA panels and are reported in Supplementary Table S1.

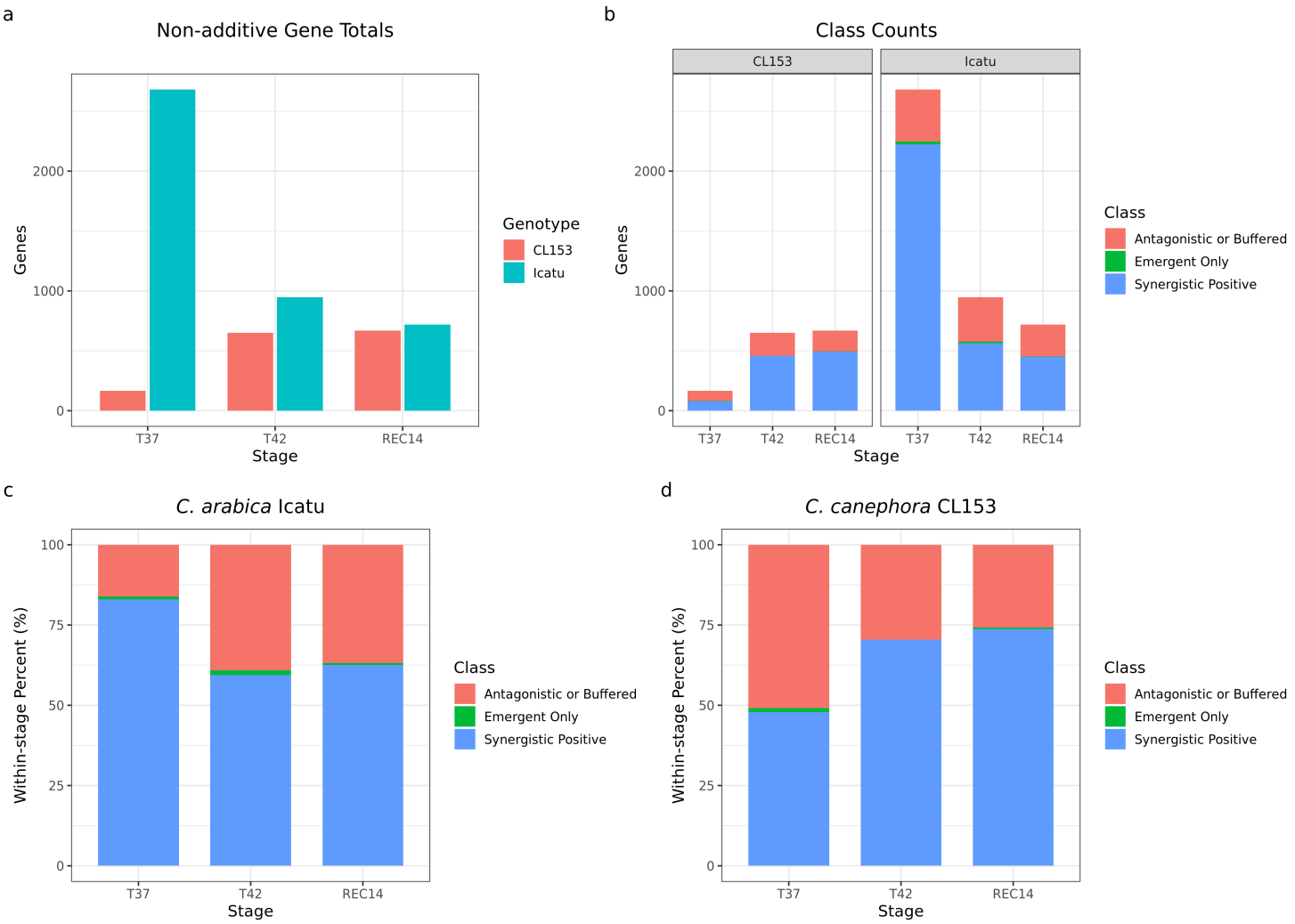

Figure 2

Genotype-specific non-additive heat–drought stress arithmetic.

(a) Total number of non-additive genes by genotype and target stage. (b) Non-additive class counts by genotype and stage. (c) Within-stage class composition for *C. arabica* Icatu. (d) Within-stage class composition for *C. canephora* CL153. Icatu showed a dominant T37 Synergistic Positive burst, whereas CL153 showed delayed T42 and REC14 non-additive expansion.

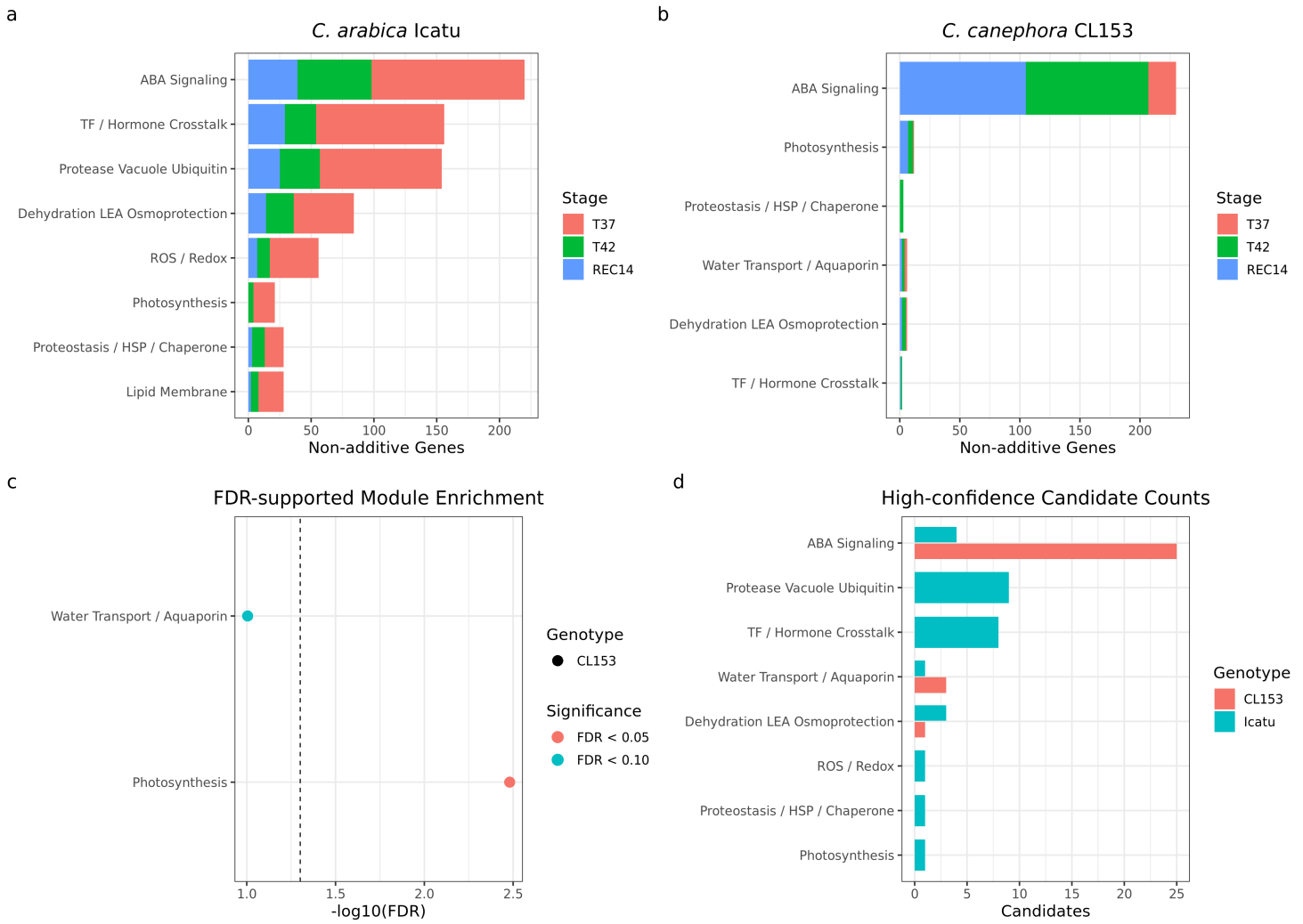

Figure 3

Functional module summaries of non-additive responses.

(a) Functional module composition of non-additive genes in *C. arabica* Icatu. (b) Functional module composition of non-additive genes in *C. canephora* CL153. (c) FDR-supported or suggestive module enrichment, showing significant CL153 REC14 Photosynthesis enrichment and suggestive CL153 T37 Water Transport / Aquaporin enrichment. The dashed line marks FDR = 0.05. (d) High-confidence candidate counts by functional module and genotype.

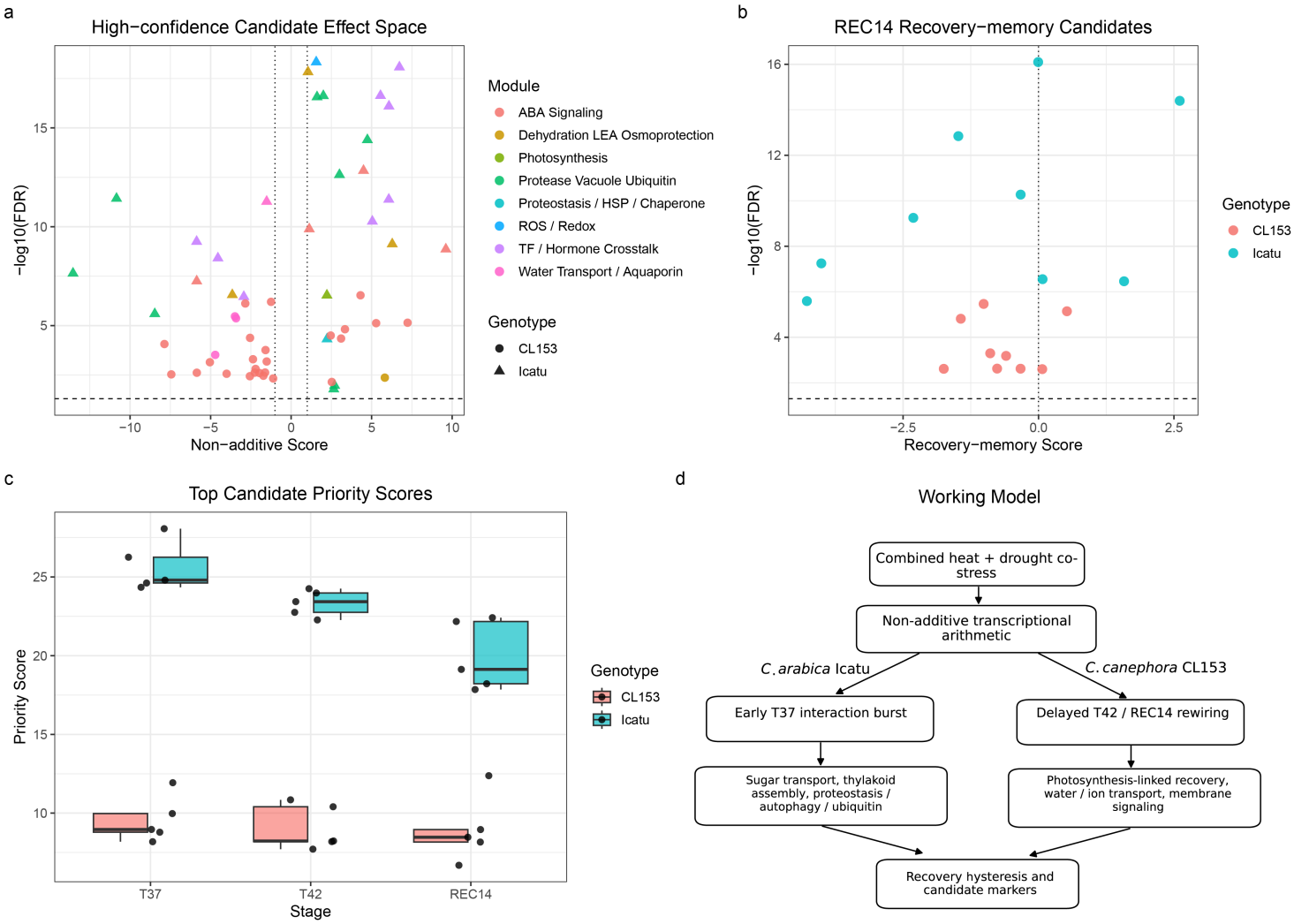

Figure 4

Candidate marker space and working biological model.

(a) High-confidence candidate effect space showing non-additive score versus -log10(FDR). Vertical dotted lines indicate non-additive score thresholds of ±1, and the horizontal dashed line indicates FDR = 0.05. (b) REC14 recovery-memory candidates. (c) Top candidate priority scores by genotype and stage. (d) Working model summarizing genotype-specific stress arithmetic: Icatu shows an early T37 interaction burst involving sugar transport and proteostasis-related axes, whereas CL153 shows delayed T42/REC14 rewiring associated with photosynthesis-linked recovery and transport-related axes.

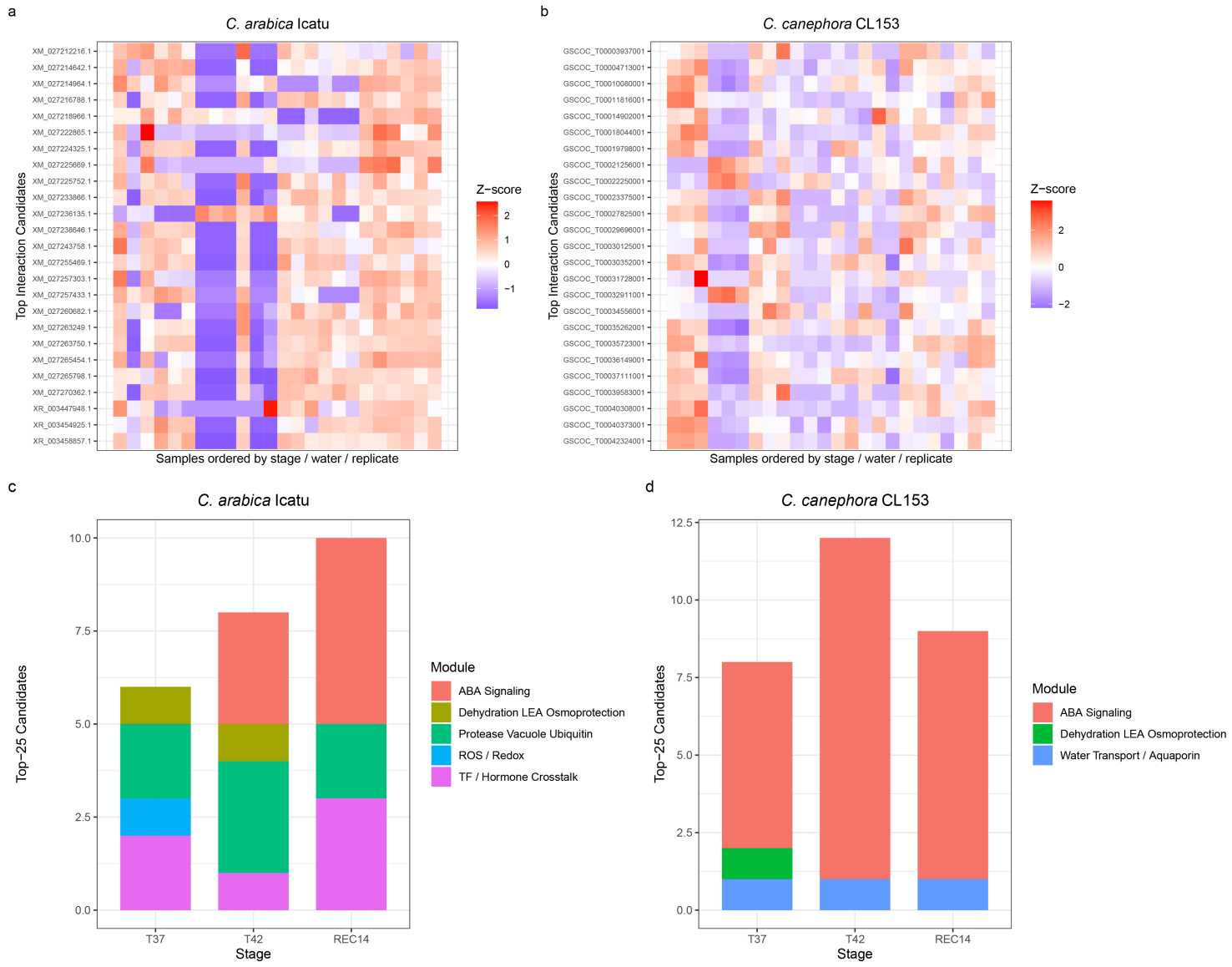

Figure 5

Top interaction candidate heatmaps and candidate module composition.

(a) Heatmap of top interaction candidates in *C. arabica* Icatu. (b) Heatmap of top interaction candidates in *C. canephora*CL153. Rows represent top interaction candidates and columns represent samples ordered by stage, water regime, and replicate. Colors show row-scaled expression Z-scores. (c) Functional module composition of Icatu top-25 candidates by stage. (d) Functional module composition of CL153 top-25 candidates by stage.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryData1Final.xlsx
- SupplementaryInformationTablesS1S4Final.pdf