

Supplementary Information

Nutritional quality and genetic differences of five amaranth cultivars revealed by metabolome profiling and whole-genome sequencing

Khalid Sedeek, Andrea Zuccolo, Umair Ali Toor, Krishnaveni Sanikommu, Sunitha Kantharajappa, Luis Rivera-Serna, Rod A. Wing & Magdy M. Mahfouz

Correspondence: magdy.mahfouz@kaust.edu.sa

This file contains

Supplementary Figures S1–S7.

Supplementary Tables S1–S12 (descriptions; data files provided separately as XLSX/PDF).

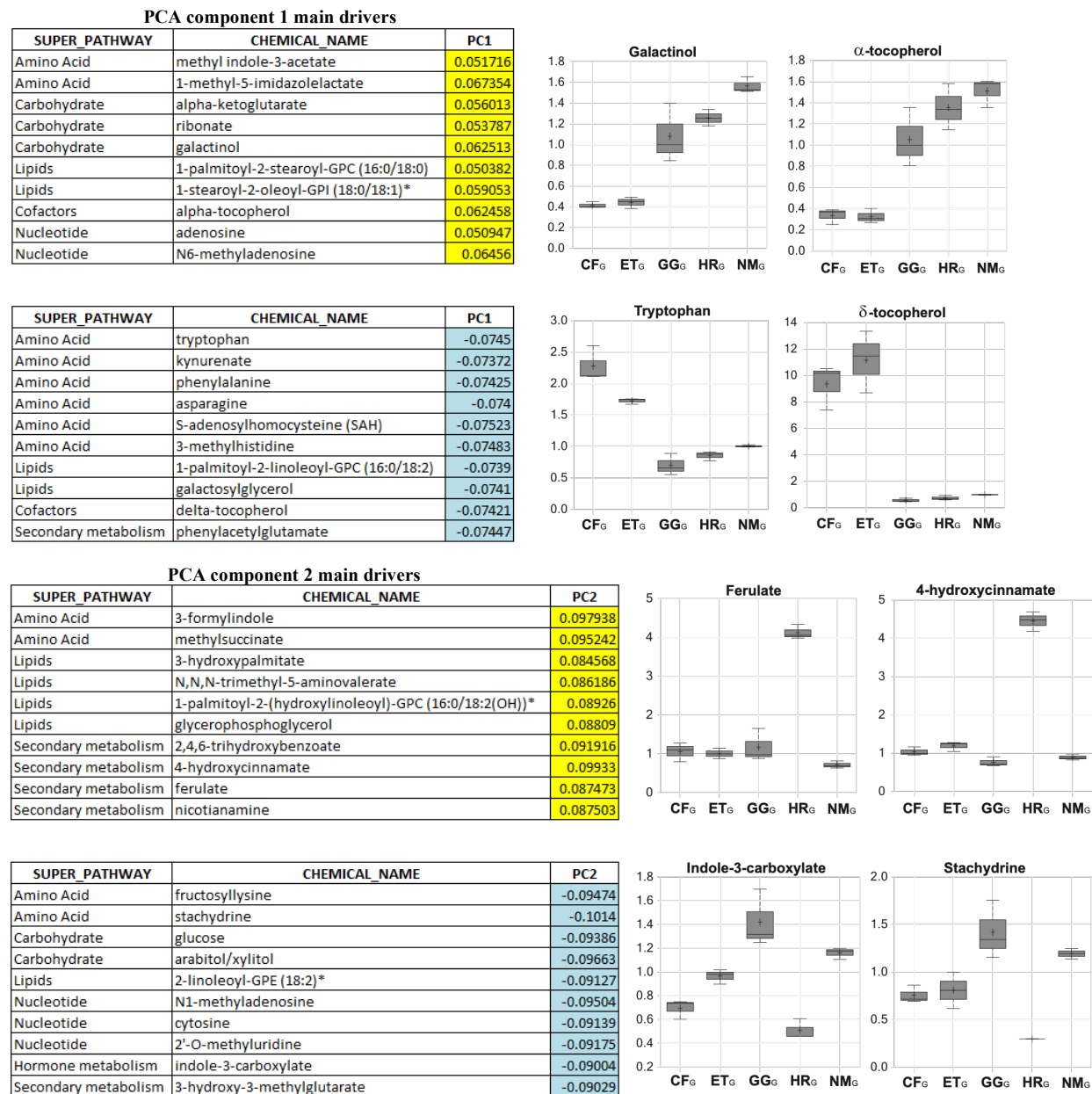
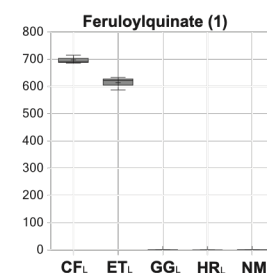
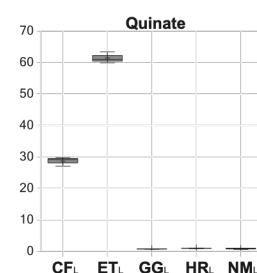


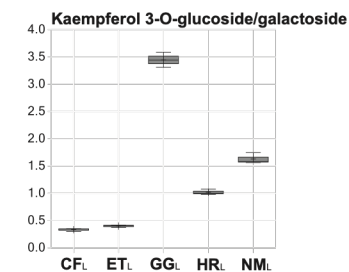
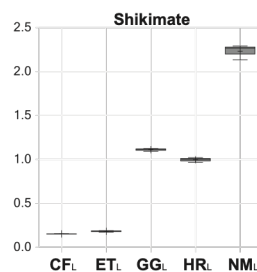
Fig. S1. Key grain metabolites separating cultivars along principal component axes. Boxplots showing the abundance of metabolites driving separation of cultivars along PC1 and PC2 of the grain metabolome PCA (Fig. 2C). Metabolites shown include galactinol, α -tocopherol, δ -tocopherol, tryptophan, ferulate, 4-hydroxycinnamate (*p*-coumarate), indole-3-carboxylate, and stachydrine. Data are from three biological replicates per cultivar. CF, Coral Fountain; ET, Emerald Tassels; GG, Golden Giant; HR, Hopi Red Dye; NM, New Mexico.

PCA component 1 main drivers

SUPER_PATHWAY	CHEMICAL_NAME	PC1
Amino Acid	quinate	0.06607
Amino Acid	5-methylthioribose**	0.06599
Cofactors	dehydroascorbate	0.06559
Cofactors	delta-tocopherol	0.06646
Secondary metabolism	coumaroylquininate (2)	0.06663
Secondary metabolism	coumaroylquininate (5)	0.0667
Secondary metabolism	coumaroylquininate (3)	0.06648
Secondary metabolism	feruloylquininate (1)	0.06578
Secondary metabolism	feruloylquininate (3)	0.06647
Secondary metabolism	feruloylquininate (4)	0.06588

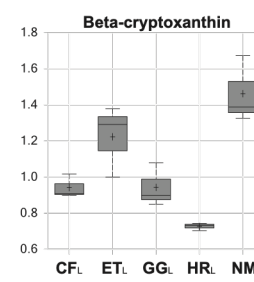
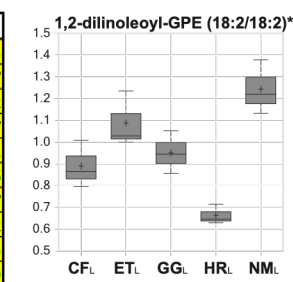


SUPER_PATHWAY	CHEMICAL_NAME	PC1
Amino Acid	glycine	-0.06121
Amino Acid	shikimate	-0.06262
Amino Acid	o-tyramine	-0.0642
Amino Acid	isoleucine	-0.05895
Carbohydrate	galactinol	-0.0595
Carbohydrate	glucarate (saccharate)	-0.0622
Nucleotide	uridine 5'-monophosphate (UMP)	-0.05954
Peptide	prolylalanine	-0.06094
Secondary metabolism	3-hydroxybenzoate	-0.06067
Secondary metabolism	kaempferol 3-O-glucoside/galactoside	-0.06136



PCA component 2 main drivers

SUPER_PATHWAY	CHEMICAL_NAME	PC2
Amino Acid	N-acetylglutamate	0.082733
Amino Acid	ornithine	0.081247
Lipids	1-stearoyl-2-oleoyl-GPC (18:0/18:1)	0.083412
Lipids	1-oleoyl-2-linoleoyl-GPC (18:1/18:2)*	0.081927
Lipids	1,2-dilinoeoyl-GPE (18:2/18:2)*	0.086101
Lipids	1-palmitoleoyl-2-linolenoyl-GPG (16:1/18:3)*	0.083106
Lipids	1,2-dilinoeoyl-galactosylglycerol (18:2/18:2)*	0.084997
Cofactors	nicotinamide ribonucleotide (NMN)	0.082082
Secondary metabolism	nicotianamine	0.082575
Secondary metabolism	beta-cryptoxanthin	0.084299



SUPER_PATHWAY	CHEMICAL_NAME	PC2
Amino Acid	tryptophan	-0.07069
Amino Acid	tyramine	-0.07057
Amino Acid	homoserine lactone	-0.06559
Amino Acid	saccharopine	-0.07528
Amino Acid	S-methylmethionine	-0.07138
Amino Acid	N6-acetyllysine	-0.07771
Amino Acid	histidine	-0.08707
Carbohydrate	stachyose	-0.07867
Lipids	1-palmitoyl-GPI (16:0)	-0.0658
Secondary metabolism	salidroside	-0.07235

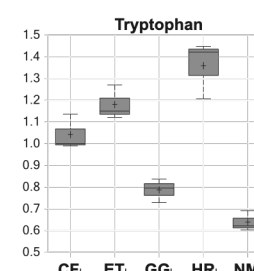
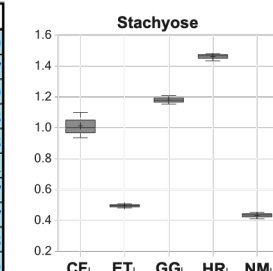


Fig. S2. Key leaf metabolites separating cultivars along principal component axes. Boxplots showing the abundance of metabolites driving separation of cultivars along PC1 and PC2 of the leaf metabolome PCA (Fig. 2D). Metabolites include quinate, feruloylquininate, shikimate, kaempferol-glycoside, 18:2/18:2 glycerophosphoethanolamine (GPE), β -cryptoxanthin, stachyose, and tryptophan. Data are from three biological replicates per cultivar.

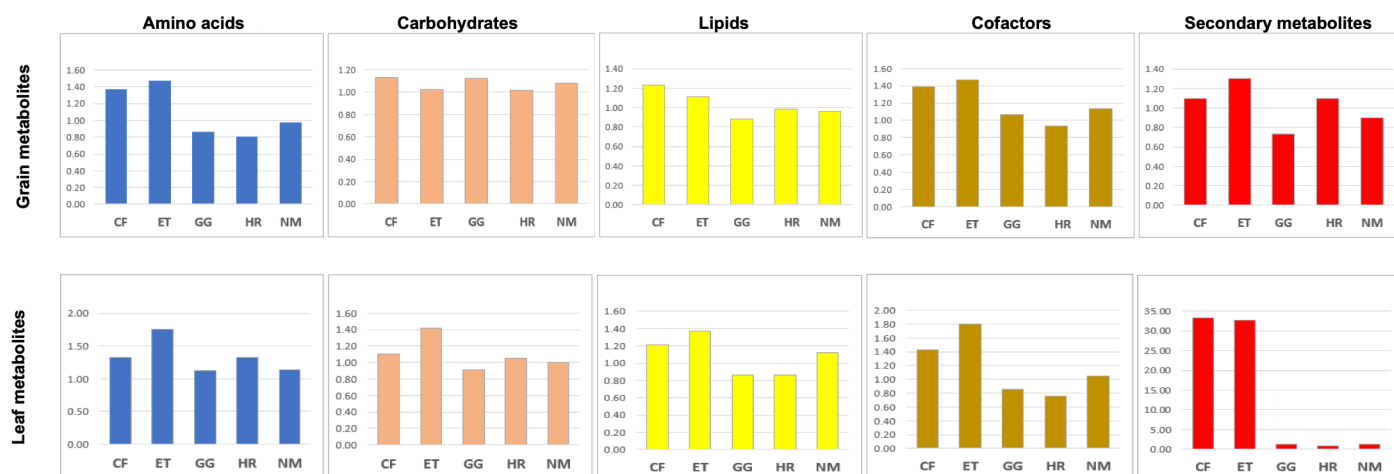


Fig. S3. Average abundance of named compounds within the five major compound classes in each amaranth cultivar. Stacked/ordered comparisons are shown for amino acids, carbohydrates, lipids, cofactors/prosthetic groups/electron carriers, and secondary metabolites in grain and leaf tissues of the five cultivars.

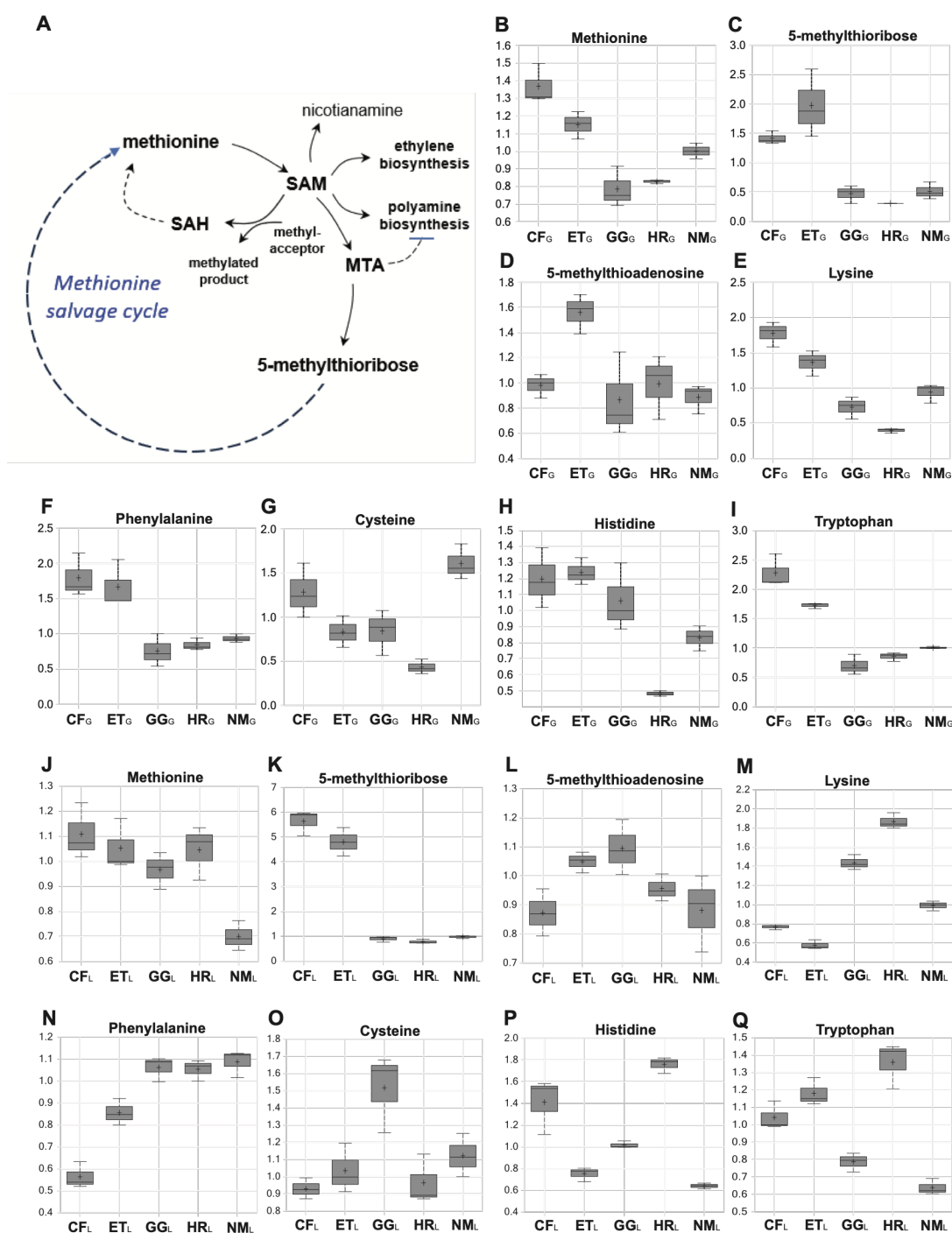


Fig. S4. Sulfur-containing compounds and the methionine-salvage pathway in grain and leaf tissues. (A) Schematic of the methionine-salvage pathway. (B) Methionine, (C) 5-methylthioribose, and related sulfur-containing metabolites in grains. (I–J) Equivalent compounds in leaves. Abbreviations as in Fig. 2.

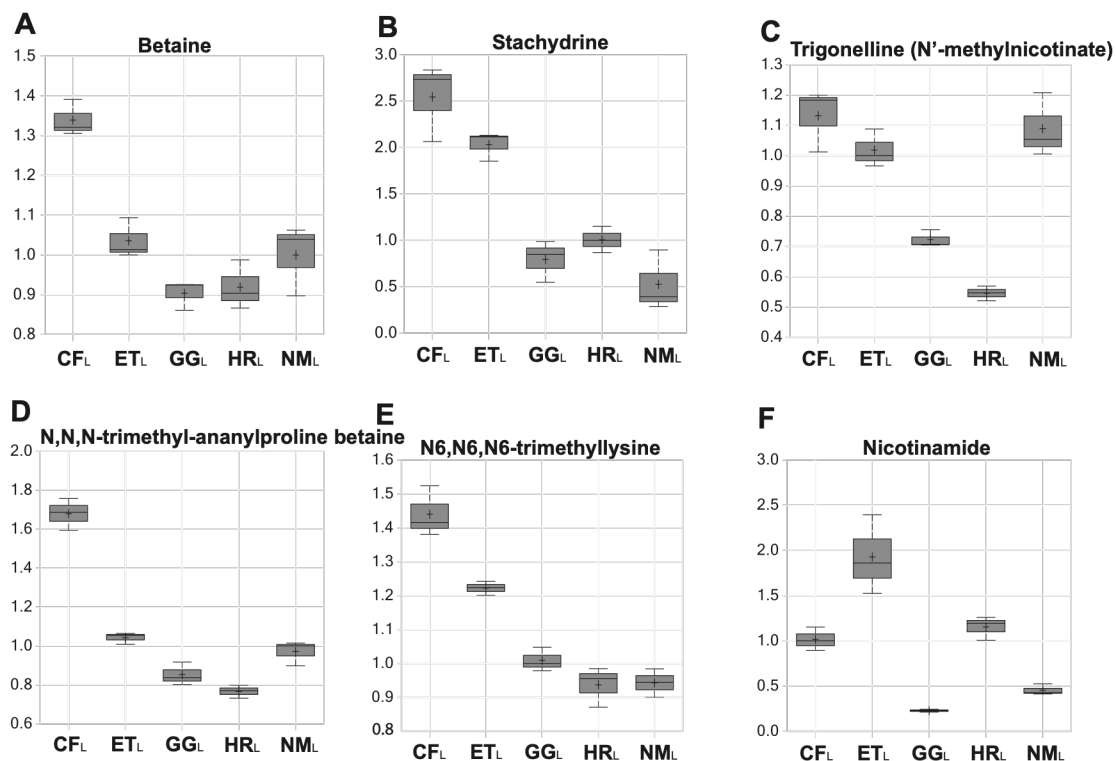


Fig. S5. Stress-related metabolites in leaf tissues. (A–F) Leaf-tissue abundance of (A) glutathione (reduced, GSH), (B) glutathione (oxidized, GSSG), (C) glycine betaine, (D) stachydrine, (E) trigonelline, and (F) nicotianamine. Abbreviations as in Fig. 2.

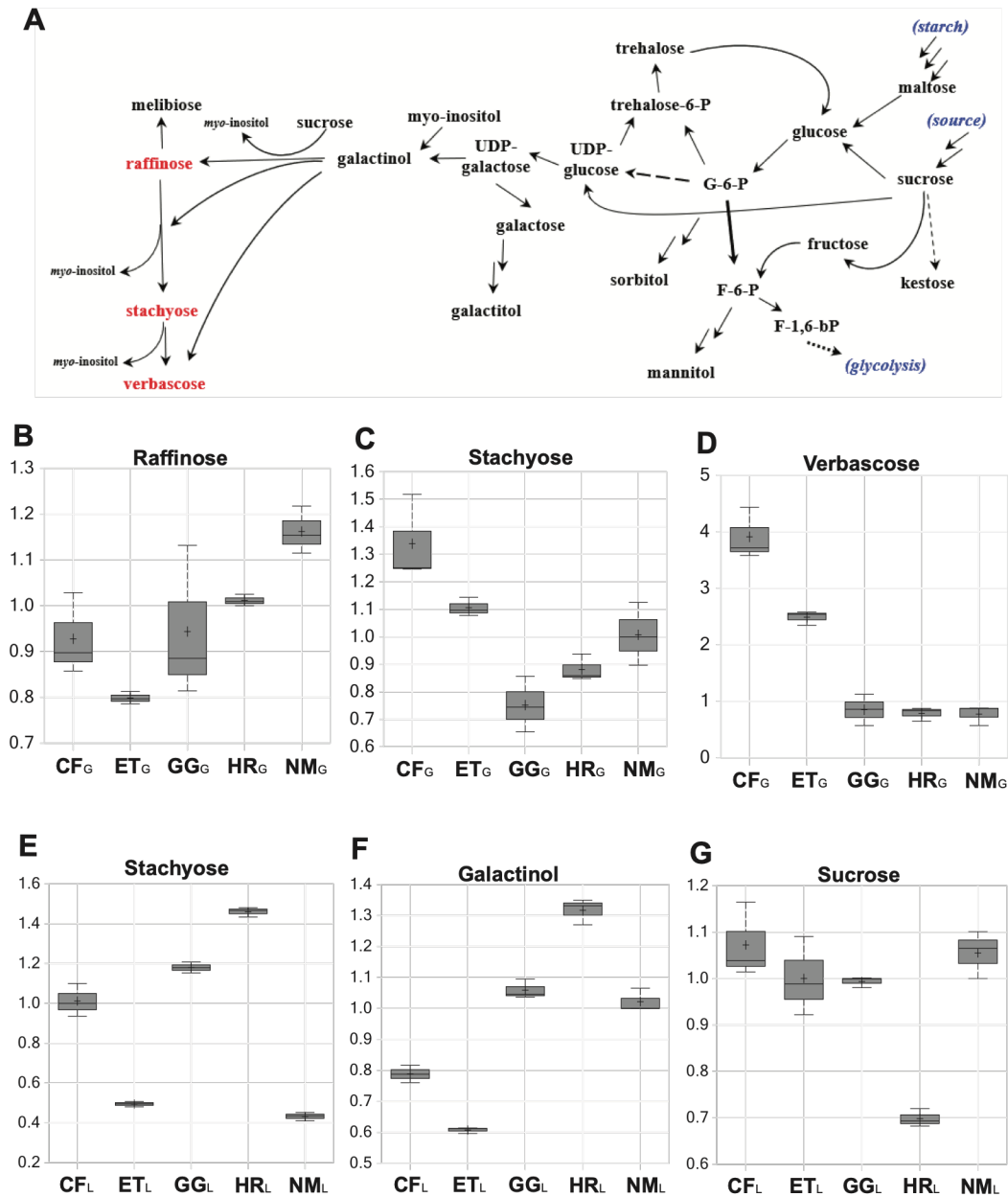


Fig. S6. Raffinose-family oligosaccharides (RFOs) in grain and leaf tissues. (A) RFO biosynthetic pathway schematic. (B–E) Raffinose, stachyose, verbascose, and galactinol in grains. (F–G) Stachyose and galactinol in leaves. Abbreviations as in Fig. 2.

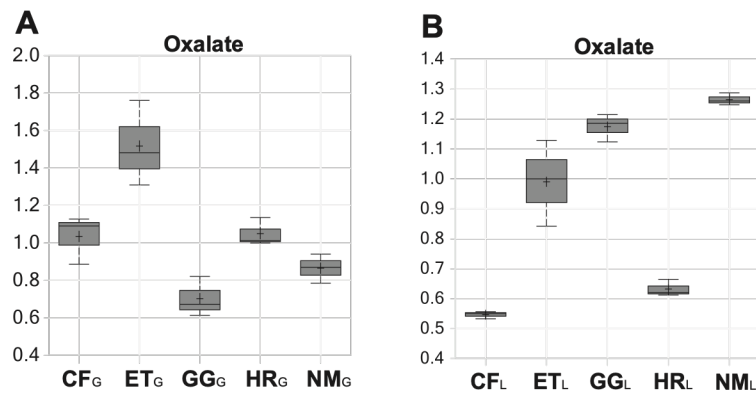


Fig. S7. Oxalate (ethanedioate) content in grain and leaf tissues. Boxplots of oxalate abundance in (A) grains and (B) leaves of the five cultivars. Abbreviations as in Fig. 2.

Supplementary Tables

Table S1. Grain metabolite dataset. Full list of 496 compounds identified in amaranth grain tissues by UPLC–MS/MS (426 with known identities and 70 unknowns), including super-pathway and sub-pathway assignments and raw abundance data per sample (three biological replicates per cultivar). *[File format: XLSX]*

Table S2. Leaf metabolite dataset. Full list of 559 compounds identified in amaranth leaf tissues by UPLC–MS/MS (420 with known identities and 139 unknowns), including super-pathway and sub-pathway assignments and raw abundance data per sample. *[File format: XLSX]*

Table S3. Grain metabolome — pairwise statistical comparisons. Welch's two-sample t-test results for all pairwise cultivar comparisons in grain samples, including fold-change and p-values for each compound. *[File format: XLSX]*

Table S4. Leaf metabolome — pairwise statistical comparisons. Welch's two-sample t-test results for all pairwise cultivar comparisons in leaf samples, including fold-change and p-values for each compound. *[File format: XLSX]*

Table S5. Boxplots of each grain metabolite across cultivars. Per-compound boxplots visualising abundance distributions across the five cultivars for all grain metabolites identified in Table S1. *[File format: XLSX]*

Table S6. Boxplots of each leaf metabolite across cultivars. Per-compound boxplots visualising abundance distributions across the five cultivars for all leaf metabolites identified in Table S2. *[File format: XLSX]*

Table S7. Vitamin-metabolite ranking. Per-cultivar ranks (1 = highest, 5 = lowest) for each vitamin-related leaf metabolite and the average rank across compounds, used to compare overall vitamin content among cultivars. *[File format: PDF]*

Table S8. BUSCO statistics for the raw genome assemblies of the five amaranth cultivars. *[File format: PDF]*

Table S9. Structural variants (SVs) detected relative to the New Mexico reference genome. Numbers of insertions, deletions and inversions in CF, ET, GG and HR genomes relative to NM, as well as the number of variants within 3 kb of the 5' end of supported gene models. *[File format: PDF]*

Table S10. Transposable-element (TE) content across the five amaranth genomes. Classified TE abundance (% of genome) for each TE family across CF, ET, GG, HR and NM genomes. *[File format: XLSX]*

Table S11. Summary statistics for Iso-seq libraries and reads. Per-tissue (leaf, flower, root) and per-cultivar Iso-seq read counts, high-quality non-chimeric transcript counts, and library metrics. *[File format: PDF]*

Table S12. Summary of predicted protein-coding genes and functional annotation.

Numbers of putative genes predicted in each genome, functionally annotated gene counts, and annotation sources (InterProScan and OmicsBox). *[File format: PDF]*

Table S13. Gene-ontology (GO) term assignments across the five cultivars.

Per-gene GO term assignments (biological process, molecular function, cellular component) derived from functional annotation of the predicted protein-coding genes. *[File format: XLSX]*