

Supplementary Methods

S1 | CT Scanning Protocol

CT scanning was performed using a SOMATOM Definition AS+ (Siemens Healthineers, Germany) dual-source clinical CT scanner at the Haikou Affiliated Hospital of Central South University Xiangya School of Medicine. Scanning parameters: tube voltage 120 kV, tube current 200 mAs, slice thickness 0.6 mm, reconstruction increment 0.4 mm, matrix size 512 Å— 512, field of view 250 Å— 250 mm, B30f medium-smooth reconstruction kernel. Each scan produced approximately 150–200 axial DICOM slices per fruit, yielding a total of 7,686 DICOM slices across 76 time-series scans (five varieties Å— five developmental stages Å— 3–5 biological replicates).

S2 | CT Image Processing Pipeline

All DICOM processing was implemented in Python 3.11 using pydicom (v2.4.3), OpenCV (v4.8.1), scikit-image (v0.22.0) and NumPy (v1.24.3).

Automated equatorial slice selection: For each fruit, the equatorial slice was automatically identified by maximizing the foreground pixel count within a circular Hounsfield window (–1,000 to +500 HU). This approach exploits the anatomical principle that the equatorial plane contains the maximum cross-sectional area. Validation against 100 manually selected slices confirmed 100% agreement.

Tissue-level density profiling: Four tissue compartments (fibre layer, shell, endosperm, water cavity) were segmented using Hounsfield-based thresholding: fibre layer (–950 to –50 HU), shell (+100 to +500 HU), endosperm (–50 to +100 HU), water cavity (0 to +50 HU). Tissue boundaries were refined using morphological operations (erosion/dilation with 3 Å— 3 kernel) and connected component analysis. For each compartment, mean CT value (HU), area (mm²) and thickness (mm) were computed.

Morphological feature extraction: Fifteen DICOM-derived features were computed: fruit weight (from CT volume estimation calibrated against balance measurements), equatorial diameter, polar diameter, shell thickness (mean of 8 radial measurements), fibre layer thickness, endosperm thickness, shell-to-fruit ratio, endosperm-to-fruit ratio, water-to-fruit ratio, gas cavity area, and the four tissue-specific CT values.

S3 | Metabolomics Data Processing

Freeze-dried endosperm samples (100 mg) were extracted with 1 mL of 70% methanol/water containing internal standards, ultrasonicated for 30 min, centrifuged at 12,000 rpm for 15 min at 4°C, and filtered through 0.22 µm PTFE membranes. UHPLC-MS/MS analysis was performed on a Q-

Exactive Orbitrap (Thermo Fisher Scientific) equipped with a HESI-II electrospray ionisation source. Chromatographic separation used an ACQUITY UPLC HSS T3 column (2.1 Å— 100 mm, 1.8 Å¼m) with a 0.4 mL/min gradient of water (0.1% formic acid) and acetonitrile (0.1% formic acid) over 25 min. Mass spectrometry was operated in positive and negative ion switching mode with full scan (m/z 70–1,050) at 70,000 resolution (FWHM at m/z 200), followed by data-dependent MS/MS at 17,500 resolution. Metabolites were identified against an in-house standard library (1,200+ authentic standards), mzCloud and HMDB databases, yielding 624 identified metabolites across 12 major classes.

S4 | Statistical Analysis Details

Permutation test (10,000 shuffles): To assess whether the observed number of near-perfect CT–metabolite correlations in W5 (55 pairs with $|r| > 0.99$) could arise by chance, time-point labels were randomly permuted 10,000 times while preserving the correlation structure of the data. For each permutation, Pearson correlations between all 18 CT parameters and 624 metabolites were recomputed. The median number of $|r| > 0.99$ pairs under the null distribution was 0 (95% CI 0–1), versus 55 observed ($p < 0.0001$).

Robustness checks:

1. **Spearman rank correlation:** Non-parametric correlation was computed for the same 18 Å— 624 = 11,232 pairs. In W5, 48 pairs retained $|\rho| > 0.95$; in CK, 0 pairs exceeded this threshold.
2. **First-differencing Pearson correlation:** To remove shared temporal trends, each time-series was first-differenced ($x_{t+1} - x_t$) before Pearson correlation. In W5, 41 pairs retained $|r| > 0.95$; in CK, 2 pairs exceeded this threshold.

Coefficient of variation analysis: To test whether the asymmetry in correlation strength could be explained by narrower metabolic variation in CK, coefficients of variation ($CV = SD/mean$) were computed across 92 differentially abundant metabolites ($|\log_2 FC| > 1$, FDR $q < 0.05$). Median CV was higher in CK (1.13) than in W5 (0.98) ($p = 0.017$, two-tailed Wilcoxon rank-sum test), ruling out the hypothesis that CK’s lack of strong correlations is due to insufficient metabolic variation.