*antioxidants*

MDPI

*Article*

# Machine Learning-Assisted Discrimination of Organ-Specific Radish Pigmentation Through Targeted Phytochemical Profiling

**Sreeparna Chowdhury [1,†], Kingsley Ochar [1,2] and Seong-Hoon Kim [1,*,†]**

[1] National Agrobiodiversity Center, National Institute of Agricultural Sciences, Rural Development Administration, Jeonju 54874, Republic of Korea; sreeparna3@korea.kr (S.C.); king.ochar@csir.org.gh (K.O.)
[2] Council for Scientific and Industrial Research, Plant Genetic Resources Research Institute, Bunso P.O. Box 7, Ghana
* Correspondence: shkim0819@korea.kr
† These authors contributed equally to this work.

## Abstract

Phytochemical differentiation between radish leaves and roots has implications for crop utilization and quality-oriented breeding, yet its association with tissue pigmentation and genotype remains unresolved. We quantified 18 glucosinolate (GSL) compounds, total phenolic content (TPC), and antioxidant capacity, including ABTS and DPPH radical scavenging activities, in the leaves and roots of 23 early-maturing radish genotypes. Chemometrics, machine learning (ML), and Shapley additive explanations (SHAP) resolved pigmentation-associated phytochemical signatures. Total GSL concentrations in roots were 3.1 times higher than in leaves, while leaves showed 3.3 times higher TPC and greater antioxidant activity. Dark green leaves and white roots were GSL-enriched, while light green leaves and red or mixed-color roots exhibited higher TPC and ABTS activity. Within-group variation distinguished genotypes with high GSL or strong phenolic and antioxidant profiles. Under genotype-grouped cross-validation, elastic-net logistic regression classified leaf and root pigmentation with balanced accuracies of 0.97 and 0.95. SHAP prioritized gluconasturtiin, glucoerucin, and sinalbin in leaves, and TPC and glucoraphasatin as the leading contributors in roots. Collectively, these findings provide trait-specific germplasm resources and candidate phytochemical discriminators for radish quality improvement.

**Keywords:** Radish; organ-specific; pigmentation; glucosinolates; antioxidant; machine learning; SHAP analysis; phytochemical discriminators

Check for updates

Academic Editor: Mario Allegra

Received: 7 October 2026
Revised: 3 November 2026
Accepted: 4 November 2026
Published: 9 November 2026

**Copyright:** © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

## 1. Introduction

In recent years, increasing recognition of the relationship between diet and human health has accelerated consumer interest and demand for plant-based foods enriched with essential nutrients and bioactive compounds [1,2]. This trend has shifted global food systems toward the development of crop varieties that offer health-promoting functional properties [3]. Among these, Brassicaceae vegetables, such as broccoli, cauliflower, kale, cabbage, mustard, and radish, are valued for their diverse phytochemical constituents, including glucosinolates (GSLs), phenolic compounds, carotenoids, and antioxidant-associated metabolites [4,5].

Radish (*Raphanus sativus* L.; $2n = 18$) is an economically important Brassicaceae vegetable widely cultivated across Asia, Europe, and North America. Its roots are consumed in raw, cooked, and pickled forms and serve as a valuable source of vitamin C and dietary

fiber [6,7]. Furthermore, its nutritional significance is increasingly attributed to the presence of varied specialized metabolites, particularly GSLs and phenolic compounds [8]. GSLs are sulfur- and nitrogen-containing specialized metabolites that undergo myrosinase-mediated hydrolysis upon tissue disruption, producing biologically active compounds associated with characteristic pungency, plant defense responses, and potential chemopreventive and anti-inflammatory effects [9]. Depending on their amino acid precursors, GSLs are classified into aliphatic, indolic, and aromatic groups, with distinct distributions and biological functions [10,11]. Among these, glucoraphasatin and glucoraphenin are particularly prominent in radish and have been proposed as important chemotaxonomic markers with bioactive constituents [6].

Radish also accumulates phenolic constituents, including flavonoids and non-flavonoids, which enhance its alimentary and functional value [8]. These compounds are involved in defense and lignification processes and act as antioxidants by scavenging reactive oxygen species (ROS) and modulating redox homeostasis [12]. Their antioxidant potential, together with that of GSL-associated metabolites, contributes to the functional properties of radish and has been linked to antimicrobial and other health-related bioactivities [13,14]. TPC and antioxidant activity, often assessed through assays that measure free-radical scavenging such as 1,1-diphenyl-2-picrylhydrazyl (DPPH) and 2,2′-azino-bis-3-ethylbenzothiazoline-6-sulfonic acid (ABTS), are therefore widely used to characterize phytochemical and functional variation in plant germplasm. Importantly, the accumulation of GSLs and phenolics, along with antioxidant-associated traits, can vary significantly among genotypes in response to genetic, developmental, and environmental factors, highlighting the importance of characterizing phytochemical diversity within radish germplasm.

Extensive evolutionary diversification, domestication, and artificial selection have generated numerous radish landraces exhibiting variation in root morphology, color, and organoleptic properties [15]. Pigmented genotypes have exhibited enhanced antioxidant activity and differential levels of secondary metabolites [16], while darker-colored radish genotypes have shown comparatively higher GSL levels [17]. Red-pigmented germplasm has also exhibited greater DPPH scavenging than non-pigmented types [18]. Despite these observations, the relationships among pigmentation, plant organ, genotype, phytochemical composition, and antioxidant-associated traits remain insufficiently resolved. Prior studies have predominantly focused on whole-plant or root-level metabolite profiling, with comparatively limited attention to integrated leaf-root comparisons [17]. Although organ-specific profiling has provided valuable insights in other Brassicaceae species [19], the relationship between contrasting leaf and root pigmentation and organ-specific phytochemical profiles across genetically diverse radish genotypes has not been systematically investigated.

Mass spectrometry-based metabolite profiling enables sensitive and high-throughput characterization of chemically diverse metabolites [20,21]. In particular, ultra-performance liquid chromatography coupled with electrospray ionization tandem mass spectrometry (UPLC-ESI-MS/MS) and related UHPLC-MS/MS platforms have been successfully applied to GSL profiling in Brassicaceae germplasm [22,23]. Recent applications of triple-quadrupole mass spectrometry in radish have further demonstrated the utility of this approach for characterizing GSLs, phenolic acids, and flavonoids across contrasting cultivars [24].

Beyond compositional profiling, integrating mass-spectrometry-derived phytochemical datasets with supervised ML provides a predictive framework for biological classification and feature prioritization. In pigmented rice, high-resolution mass spectrometry metabolomics was integrated with phenolic, antioxidant, vitamin, and mineral traits to classify multinutritional ideotypes and identify superior germplasm using random forest modeling [25]. Similarly, multi-trait profiles acquired through UHPLC-Q-Orbitrap

HRMS and UHPLC-MS/MS were coupled with XGBoost and SHAP analysis to discriminate three *Zanthoxylum* species and identify constituents contributing to classification and anti-inflammatory activity [26]. Phenolic profiles have likewise supported accurate ML-based differentiation of olive cultivars, demonstrating that even relatively minor phenolic compounds can provide substantial discriminatory information [27]. These studies elucidate how ML can extend phytochemical profiling from descriptive characterization toward predictive classification and interpretable germplasm characterization. However, whether directly quantified GSL and antioxidant-associated profiles jointly discriminate organ-specific pigmentation groups in radish remains unexplored.

To address these gaps, the present study comprehensively evaluated 23 early-maturing radish genotypes from diverse geographic origins exhibiting contrasting pigmentation characteristics. GSL profiles in leaf and root tissues were analyzed using UPLC-ESI-MS/MS to quantify 18 GSLs, accompanied by assessments of TPC and antioxidant activity via DPPH and ABTS assays. The effects of plant organ, genotype, and pigmentation on GSL profiles and antioxidant-related traits were assessed to provide novel insights into phytochemical variation in radish. It further evaluated whether measured phytochemical and antioxidant profiles could discriminate the principal leaf and root pigmentation groups within the evaluated germplasm and employed SHAP analysis to identify features contributing strongly to classification. By integrating comparative phytochemical profiling, multivariate analysis, and interpretable ML, the study aimed to identify metabolically distinct genotypes and candidate organ-specific phytochemical discriminators, thereby providing a foundation for targeted germplasm utilization and nutritional quality improvement in radish.

## 2. Materials and Methods

### *2.1. Planting Materials and Growth Conditions*

Leaf and root tissues analyzed in this study were obtained from the previously described experiment [28]. The experimental panel comprised 23 radish genotypes from nine different origins, sourced from the Rural Development Administration (RDA) Gene Bank at the National Agrobiodiversity Center (Jeonju, Republic of Korea). Seeds were surface-sterilized using 70% ethanol (Sigma-Aldrich, St. Louis, MO, USA), followed by thorough rinsing with distilled water, and were subsequently sown in polyvinyl chloride (PVC) pipes (6 cm diameter × 40 cm height) filled with organic soil mix (Tobirang, Baekkwang Fertility, Andong, Republic of Korea). Plants were cultivated under natural daylight conditions in a greenhouse at the National Agrobiodiversity Center (35°49′18″ N, 127°08′56″ E), with 10 plants grown for each genotype. Greenhouse temperatures ranged from 18 ± 1 °C at night to a maximum of 32 ± 3 °C during the daytime, with relative humidity of 60–70%. Irrigation was carried out every 3–4 days, as needed to maintain adequate moisture in the growing medium throughout the cultivation period. Uniform and healthy plants were selected and harvested 20 days after sowing, representing the early-harvest stage of the Cherry Belle market class, which exhibits rapid growth with compact root development and is commonly utilized for fresh consumption. Further, leaf and root pigmentation were visually assessed at harvest following relevant radish descriptors provided in the guidelines of the International Union for the Protection of New Varieties of Plants (UPOV, 2023). The 23 genotypes were classified as dark green ($n = 7$), medium green ($n = 12$), and light green ($n = 4$) for leaf pigmentation; and as white ($n = 5$), mix (red-white, $n = 4$), and red ($n = 14$) for root pigmentation.

### *2.2. Chemicals and Reagents*

Eighteen GSL reference standards, including sinigrin (SIN), gluconapin (GNA), glucobrassicanapin (GBN), epiprogoitrin (EPI), progoitrin (PRO), glucoiberin (GIB), glu-

coalyssin (GAL), glucocheirolin (GCH), glucoraphanin (GRA), glucoraphenin (GRE), glucoraphasatin (GRH), glucoerucin (GER), sinalbin (SNB), glucotropaeolin (GTL), gluconasturtiin (GNS), neoglucobrassicin (NGBN), 4-hydroxyglucobrassicin (4HGBN), and 4-methoxyglucobrassicin (4MGBN) were purchased from Phytoplan Diehm, Neuberger GmbH (Neuenheimer, Heidelberg, Germany), with certified purity of ≥98%. Analytical-grade methanol, gallic acid, Folin–Ciocalteu's phenol reagent, sodium carbonate, DPPH, ABTS, potassium persulfate, and Trolox were procured from Sigma-Aldrich Co. (St. Louis, MO, USA). Trifluoroacetic acid and ethanol were obtained from Thermo Fisher Scientific Korea Ltd. (Seoul, Republic of Korea).

### *2.3. Sample Preparation*

For each genotype, three healthy plants were selected as biological replicates, and their leaf and root tissues were harvested separately and stored at −80 °C. Leaf and root samples from each plant were maintained as paired samples, with each tissue sample independently processed through extraction and analytical measurement. Subsequently, samples were freeze-dried for 48 h in an LP500 vacuum freeze dryer (Ilshinbiobase Co., Seoul, Republic of Korea), finely powdered, and preserved at −80 °C until extraction. Glucosinolates were extracted according to the previously reported protocol [29]. Briefly, 100 mg of ground tissue sample was extracted with 5 mL of 80% methanol, incubated at 25 °C for 30 min, and then agitated on an orbital shaker at 120 rpm for an additional 30 min. Following extraction, the mixture was centrifuged at 14,000 rpm for 10 min at 4 °C using a VS-180CFi centrifuge (Vision Scientific Co., Daejeon, Republic of Korea). The resulting supernatant was collected for GSL analysis.

### *2.4. UPLC–ESI–MS/MS Analysis and Glucosinolate Quantification*

Glucosinolate profiling was performed on an ACQUITY™ UPLC system (Waters Corp., Milford, MA, USA) interfaced with a Xevo™ TQ-S triple quadrupole mass spectrometer (Waters MS Technologies, Manchester, UK). Separation was achieved on an ACQUITY UPLC BEH C18 column (1.7 μm, 2.1 × 100 mm) maintained at 35 °C, with an injection volume of 5 μL and a flow rate of 0.5 mL $min^{-1}$. The mobile phase comprised solvent A (0.1% aqueous trifluoroacetic acid) and solvent B (0.1% trifluoroacetic acid in methanol), with the elution gradient as follows: initially set to 0–1 min, 100% A; followed by 1–7 min, 100–80% A; 7–10 min, 80–0% A; 10–11 min, 0–100% A; and 11–15 min, 100% A for column re-equilibration. Mass spectrometry was performed using both positive and negative electrospray ionization (ESI) interfaces with multiple-reaction monitoring (MRM) conditions. Ion source conditions were optimized using a capillary voltage of 3 kV and a cone voltage of 54 V. The source temperature was maintained at 150 °C and the desolvation temperature was set at 350 °C. The cone and desolvation gas flow rates were optimized at 150 and 650 L $h^{-1}$, respectively. GSL compounds were identified by comparing retention times as well as MS/MS fragmentation patterns with commercial standards. Quantification was performed using compound-specific external calibration curves constructed from five nominal standard concentrations (1, 10, 100, 1000, and 10,000 ppb) by applying 1/x-weighted linear regression. Calibration equations, coefficients of determination ($R^2$), retention times, and multiple reaction monitoring (MRM) transitions for the 18 targeted GSLs are presented in Table S1. The limits of detection (LOD) and quantification (LOQ) were estimated as three and ten times, respectively, the standard error of the regression intercept divided by the calibration slope, following previously reported analytical approaches for GSL determination [30,31]. The estimated LOD and LOQ values ranged from 1.08 to 19.83 ppb and from 3.60 to 66.09 ppb, respectively. Data acquisition and processing were conducted in

MassLynx 4.1 software (Waters Corp., Milford, MA, USA), and final GSL concentrations were reported as μmol $kg^{-1}$ dry weight (DW).

### *2.5. Determination of Total Phenolic Content and Antioxidant Activity*

TPC was assessed using the Folin–Ciocalteu colorimetric assay with modifications [32]. A 100 μL aliquot of the sample extract was combined with an equivalent volume of diluted Folin–Ciocalteu reagent (2 N, 1:1, *v*/*v* with distilled water) and thoroughly vortexed. Subsequently, the mixture was incubated for 30 min in the dark at room temperature; 100 μL of sodium carbonate solution (2%) was then added, and the mixture was further incubated for 30 min in the dark. Absorbance was measured at 750 nm using an Eon Microplate Spectrophotometer (BioTek Instruments, Winooski, VT, USA), with methanol serving as the blank sample. TPC values were quantified against a gallic acid calibration curve ($y = 0.0047x + 0.0231$, $R^2 = 0.9991$) and presented as milligrams of gallic acid equivalents per gram dry weight (mg GAE $g^{-1}$ DW).

Both DPPH and ABTS radical scavenging activities were quantified following a previously described protocol [18]. For DPPH radical scavenging activity, a DPPH stock solution of 0.2 mM was prepared using ethanol and standardized to an absorbance of 1.0 at 517 nm. Furthermore, 100 μL of each sample extract was reacted with 100 μL of 150 μM DPPH solution (freshly prepared) and kept for 30 min in the dark at room temperature. Subsequently, absorbance was recorded at 517 nm using an Eon Microplate Spectrophotometer (BioTek Instruments, Winooski, VT, USA), with a blank solution (100 μL each of sample extract with ethanol) and a control solution (100 μL of ethanol and DPPH solution each). The antioxidant activity was determined using an ascorbic acid calibration curve ($y = -0.0139x + 0.5503$; $R^2 = 0.9883$) and expressed as milligrams of ascorbic acid equivalent antioxidant capacity per gram dry weight (mg AAE $g^{-1}$ DW).

Moreover, the ABTS radical scavenging assay was performed using two stock solutions, 7.0 mM ABTS and 2.45 mM potassium persulfate. These were mixed in equal proportions to make the working solution and incubated at room temperature for 12–16 h under dark conditions to yield the ABTS radical cation. Prior to analysis, the ABTS working solution was diluted with ethanol (1:89) to attain an absorbance of 0.700 ± 0.02 at 734 nm. For the assay, 10 μL of each sample extract was mixed with 1 mL of ABTS solution and incubated in the dark for 30 min. Methanol was used as the blank and absorbance at 734 nm was recorded using the same spectrophotometer described above. ABTS radical scavenging activity was represented in milligrams of Trolox equivalent antioxidant capacity per gram dry weight (mg TE $g^{-1}$ DW), using the Trolox standard curve ($y = -0.0052x + 0.666$; $R^2 = 0.9876$).

For each biological replicate, aliquots of the same sample extract were used for TPC, DPPH, and ABTS measurements. Moreover, each measurement was performed in three technical replicates, and the resulting values were averaged before statistical analysis.

### *2.6. Statistical and Multivariate Analyses*

Statistical analyses were accomplished using R v4.4.2 in RStudio v2024.09.1 + 394 (Posit Software, Boston, MA, USA). Descriptive statistics were calculated, and univariate statistical analyses were performed in Microsoft Excel 2019 (Microsoft Corp., Redmond, WA, USA). Mean and standard deviation values were calculated across 23 genotype-level means for each organ, with each genotype mean derived from three biological replicates. Two-way repeated-measures ANOVA was conducted using the aov function in R to evaluate genotype, organ, and genotype-by-organ interaction effects based on biological replicate-level observations, with individual plants nested within genotypes to account for paired leaf–root measurements. Post-hoc comparisons among pigmentation groups of leaves and roots were conducted using Duncan's multiple range test implemented in the 'agricolae' package [33],

while differences between organs were examined using a paired Wilcoxon signed-rank test. Pearson correlation was estimated using the 'Hmisc' package [34]. $Log_{10}$-transformed data were subjected to multivariate analysis. Principal component analysis (PCA) was visualized using 'ggrepel' [35], while hierarchical cluster analysis (HCA) was performed employing the Bioconductor R package 'ComplexHeatmap' [36]. Orthogonal partial least squares discriminant analysis (OPLS-DA) and variable importance in projection (VIP) were performed using the Bioconductor R package 'ropls' [37]. OPLS-DA model performance was evaluated using three-fold cross-validation for both leaf and root datasets, together with 200 response-label permutations. Model fit and predictive ability were assessed using cumulative proportion of variation in the response variable explained by the model ($R^2Y$) and cumulative predictive ability estimated by cross-validation ($Q^2$), respectively.

### *2.7. Machine Learning Analysis and Model Interpretation*

Supervised ML classification was implemented using the 'caret' package [38]. Five ML models were evaluated: elastic-net logistic regression (ENLR) using 'glmnet' [39], radial support vector machine (RSVM) using 'kernlab' [40], random forest (RF) using 'ranger' [41], extreme gradient boosting (XGBoost) using 'xgboost' [42], and decision tree (DT) using 'rpart' [43]. Models were trained using replicate-level observations and 21 directly measured predictors. The consolidated pigmentation classes comprised 7 dark green and 16 medium/light green leaf genotypes (21 and 48 observations, respectively), and 5 white and 18 red/mixed root genotypes (15 and 54 observations, respectively). A fixed random seed (20260728) was used to ensure reproducibility. Model performance was assessed through 10 iterations of genotype-grouped three-fold cross-validation, while ensuring all replicate observations from the same genotype were retained within a single fold. ENLR hyperparameters were optimized by selecting the combination of $\alpha$ and $\lambda$ that maximized cross-validated balanced accuracy. The $\alpha$ parameter controls the balance between LASSO (L1) and ridge (L2) regularization, whereas $\lambda$ determines the overall strength of regularization. All 21 measured phytochemical variables were initially included without prior VIP or SHAP-based selection. Hyperparameter tuning was performed using automatically generated grids for ENLR and DT (tuneLength = 10) and RSVM (tuneLength = 8), whereas predefined parameter grids were used for RF and XGBoost. The evaluated tuning specifications and selected hyperparameter settings are provided in Table S2. Missing-value imputation, centering, scaling, and filtering of near-zero-variance or highly correlated predictors were performed within each cross-validation training fold. Class imbalance was addressed through majority-class downsampling within training folds.

Out-of-fold predictions from the three validation folds were pooled across 10 cross-validation repetitions. The pooled predictions were used to construct confusion matrices normalized within each observed class and to calculate precision, recall, specificity, balanced accuracy, F1 score, and Matthews correlation coefficient (MCC). These metrics were calculated from the numbers of true-positive (TP), true-negative (TN), false-positive (FP), and false-negative (FN) classifications as follows:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (1)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (2)$$

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (3)$$

$$\text{Balanced Accuracy} = \frac{\text{Recall} + \text{Specificity}}{2} \quad (4)$$

$$\text{F1 score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (5)$$

$$\text{MCC} = \frac{\text{TP}\times\text{TN}-\text{FP}\times\text{FN}}{\sqrt{(\text{TP}+\text{FP})(\text{TP}+\text{FN})(\text{TN}+\text{FP})(\text{TN}+\text{FN})}} \quad (6)$$

The area under the receiver operating characteristic curve (ROC-AUC) was calculated from pooled out-of-fold class probabilities using the 'pROC' package [44]. Contributions of the individual phytochemical predictors to the selected ENLR model were estimated using a custom Monte Carlo implementation of Shapley additive explanations [45,46]. SHAP values were calculated on the probability scale using 200 simulations per observation, with dark green leaves and white roots designated as the positive classes. SHAP importance and summary plots, correlation plot, and PCA plots were generated using the package 'ggplot2' [47], while multipanel figures were assembled using 'patchwork' [48]. SHAP feature-ranking stability was further assessed by refitting ENLR models across 30 cross-validation training subsets using the selected hyperparameters.

## 3. Results and Discussion

### *3.1. Phenotypic Diversity and Organ-Specific Glucosinolate Signatures in Radish Germplasm*

The evaluated radish genotypes displayed considerable phenotypic diversity, particularly in leaf and root pigmentation patterns (Figure 1A–F), representing a valuable resource for investigating natural variation in specialized metabolite accumulation. Comprehensive GSL profiling of 23 radish genotypes identified 18 glucosinolate compounds across leaf and root tissues, comprising 12 aliphatic, 3 aromatic, and 3 indolic GSLs. The detected aliphatic GSLs included SIN, GNA, GBN, EPI, PRO, GIB, GAL, GCH, GRA, GRE, GRH, and GER, while SNB, GTL, and GNS represented the aromatic GSLs, and NGBN, 4HGBN, and 4MGBN represented the indolic GSLs. Although most GSLs were widely distributed across the germplasm, several compounds were not detected in specific genotypes (Table S3), indicating substantial variation in GSL composition. Similar genotype-dependent differences in GSL profiles have previously been documented in radish and other Brassicaceae crops, suggesting strong genetic regulation of GSL biosynthesis and accumulation [17,49].

Significant variation in both individual and total GSL concentrations was observed across genotypes and between plant organs (Table 1; Figure 1), likely reflecting organ-specific differences in biosynthesis, transport, and storage associated with distinct physiological functions of leaf and root tissues [50]. Two-way repeated-measures ANOVA further confirmed significant genotype effects for all GSL compounds and classes except SIN, but organ effects were non-significant for SIN, SNB, GTL, and NGBN. Moreover, significant genotype × organ interactions for most compounds demonstrated that tissue-specific accumulation varied among genetic backgrounds rather than following a uniform organ response. Aliphatic glucosinolates (ALGSLs) constituted the predominant class across leaf and root tissues, making up the majority of total glucosinolate (TGSL) accumulation. Among the individual aliphatic compounds, GRE, GRH, and GER exhibited the maximum abundance, contributing approximately 40%, 27%, and 29% of the TGSL pool in leaves and 9%, 35%, and 52% in roots, respectively (Figure 1G). In contrast, aromatic (ARGSL) and indolic glucosinolates (INGSL) occurred at substantially lower concentrations and collectively contributed less than 4% of TGSLs. Overall, roots accumulated markedly higher levels of ALGSLs and TGSLs than leaves, with significantly higher accumulation of TGSLs in dark green leaves (34,372.79 μmol $kg^{-1}$ DW) and white roots (131,739.63 μmol $kg^{-1}$ DW) compared to other pigmentation groups (Figure 1H, Table S4). The predominance of these methionine-derived ALGSLs is consistent with previous studies reporting GRE, GRH, and GER as characteristic glucosinolates of radish [18,49,51]. Upon hydrolysis, these com-

pounds yield bioactive isothiocyanates, including sulforaphene, raphasatin, and erucin, which contribute to the pungency and health-associated properties of radish [52,53].

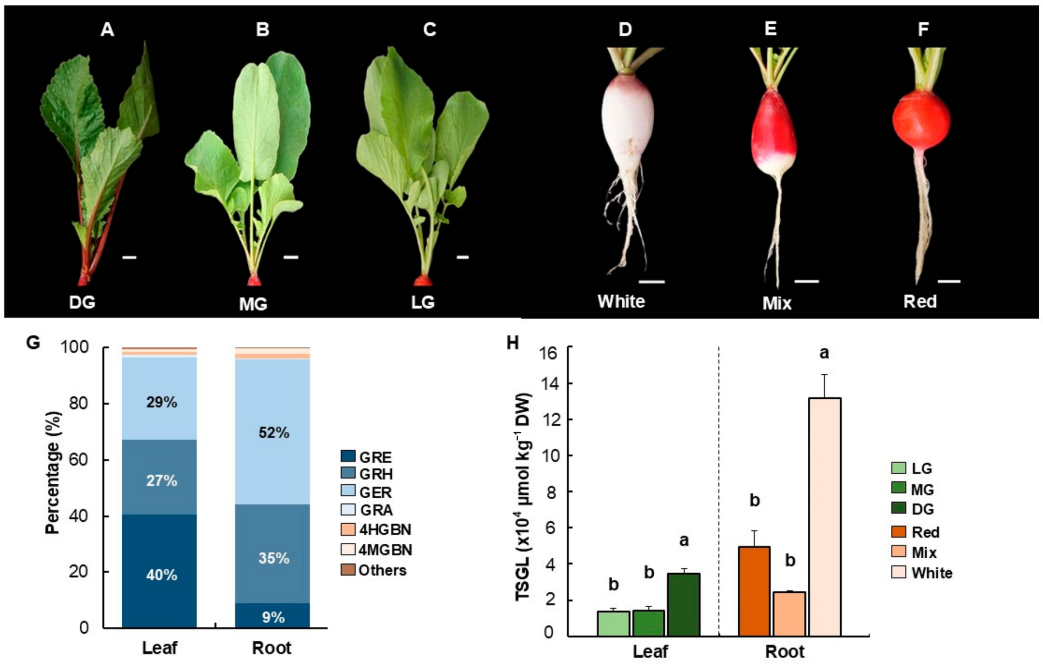



**Figure 1.** Phenotypic diversity and organ-specific glucosinolate profiles associated with pigmentation variation in radish genotypes. (**A**–**C**) Leaf pigmentation phenotypes, including dark green (DG), medium green (MG), and light green (LG) color. (**D**–**F**) Root pigmentation phenotypes, including white, mix (red-white), and red roots. Scale bars denote 1 cm. (**G**) Relative abundance (%) of the major glucosinolate compounds contributing to the total glucosinolate profiles of leaf and root tissues. GRE, Glucoraphenin; GRH, Glucoraphasatin; GER, Glucoerucin; GRA, Glucoraphanin; 4HGBN, 4-Hydroxyglucobrassicin; 4MGBN, 4-Methoxyglucobrassicin; and "Others" represents the combined contribution of the remaining individual GSLs. (**H**) Variation in average total glucosinolate (TGSL) accumulation among leaf and root pigmentation groups. Error bars represent standard error. Different letters indicate significant differences among pigmentation groups according to Duncan's multiple range test ($p < 0.05$).

In addition, SIN, GNA, GBN, and SNB were detected as minor constituents of the GSL pool, with average concentrations below 1 μmol $kg^{-1}$ DW. Comparable low-abundance patterns of specific glucosinolates have been reported among diverse *Brassica* germplasm [54,55]. Although these compounds occurred at relatively low levels, variation in their relative abundance may still contribute to genotype-specific GSL profiles, as the composition and abundance of individual GSLs are important determinants of quality traits in Brassicaceae crops [56,57]. Interestingly, among the evaluated genotypes, IT340147 (Turkey), IT297044, and IT344617 (Russia) exhibited the highest TGSL levels, while IT250789, IT308363, and IT278270 (China) showed the lowest TGSL levels (Table S3). The high-TGSL genotypes were characterized by dark green leaves and white roots, whereas low-TGSL genotypes generally exhibited medium to light green leaves with red roots. These findings indicate substantial organ-specific and pigmentation-associated variation in GSL signatures among the radish genotypes, providing a foundation for further evaluation of their relationship with antioxidant-related phytochemical traits.

**Table 1.** Comparative glucosinolate profiles and organ-specific variation among radish genotypes.

| GSL Compound (μmol kg$^{-1}$ DW) | Mean ± SD | | ANOVA (*p* Value) | | |
|---|---|---|---|---|---|
| | Leaf | Root | G | O | G × O |
| SIN | 0.14 ± 0.04 | 0.12 ± 0.08 | 0.668 $^{ns}$ | 0.352 $^{ns}$ | 0.745 $^{ns}$ |
| GNA | 0.08 ± 0.06 | 0.20 ± 0.14 | $3.73 \times 10^{-4}$ *** | $2.34 \times 10^{-7}$ *** | 0.443 $^{ns}$ |
| GBN | 0.14 ± 0.07 | 0.40 ± 0.37 | $2.85 \times 10^{-8}$ *** | $1.76 \times 10^{-11}$ *** | $1.19 \times 10^{-5}$ *** |
| EPI | 2.16 ± 1.97 | 12.97 ± 19.37 | $2.23 \times 10^{-9}$ *** | $1.57 \times 10^{-7}$ *** | $2.28 \times 10^{-3}$ ** |
| PRO | 2.31 ± 3.01 | 14.75 ± 18.93 | $<2 \times 10^{-16}$ *** | $<2 \times 10^{-16}$ *** | $2.63 \times 10^{-5}$ *** |
| GIB | 8.57 ± 6.47 | 5.11 ± 6.32 | $2.59 \times 10^{-15}$ *** | $7.72 \times 10^{-14}$ *** | $3.02 \times 10^{-7}$ *** |
| GAL | 15.51 ± 10.40 | 15.12 ± 14.18 | $<2 \times 10^{-16}$ *** | $2.68 \times 10^{-10}$ *** | $<2 \times 10^{-16}$ *** |
| GCH | 51.87 ± 27.10 | 29.73 ± 25.32 | $<2 \times 10^{-16}$ *** | $<2 \times 10^{-16}$ *** | $1.06 \times 10^{-15}$ *** |
| GRA | 210.60 ± 118.11 | 118.84 ± 99.89 | $9.57 \times 10^{-8}$ *** | $9.27 \times 10^{-10}$ *** | $2.47 \times 10^{-4}$ *** |
| GRE | 8196.77 ± 4220.40 | 5566.10 ± 4453.76 | $2.47 \times 10^{-14}$ *** | $5.15 \times 10^{-14}$ *** | $4.80 \times 10^{-15}$ *** |
| GRH | 5404.91 ± 5149.76 | 22,150.56 ± 15,314.22 | $<2 \times 10^{-16}$ *** | $<2 \times 10^{-16}$ *** | $1.57 \times 10^{-5}$ *** |
| GER | 5973.71 ± 6138.82 | 32,630.25 ± 28,368.10 | $<2 \times 10^{-16}$ *** | $<2 \times 10^{-16}$ *** | $<2 \times 10^{-16}$ *** |
| SNB | 0.05 ± 0.04 | 0.06 ± 0.06 | $7.08 \times 10^{-3}$ ** | 0.106 $^{ns}$ | 0.048 * |
| GTL | 1.36 ± 0.99 | 2.27 ± 3.50 | $<2 \times 10^{-16}$ *** | 0.083 $^{ns}$ | $1.76 \times 10^{-4}$ *** |
| GNS | 3.97 ± 3.88 | 22.21 ± 30.63 | $<2 \times 10^{-16}$ *** | $<2 \times 10^{-16}$ *** | $1.21 \times 10^{-15}$ *** |
| NGBN | 4.90 ± 4.06 | 8.14 ± 10.81 | $<2 \times 10^{-16}$ *** | 0.0616 $^{ns}$ | $<2 \times 10^{-16}$ *** |
| 4HGBN | 179.04 ± 126.21 | 1044.92 ± 376.02 | $6.95 \times 10^{-16}$ *** | $<2 \times 10^{-16}$ *** | $1.24 \times 10^{-15}$ *** |
| 4MGBN | 222.28 ± 151.71 | 1243.55 ± 1070.35 | $<2 \times 10^{-16}$ *** | $<2 \times 10^{-16}$ *** | $<2 \times 10^{-16}$ *** |
| ALGSL | 19,866.23 ± 11,408.04 | 60,544.12 ± 47,687.69 | $<2 \times 10^{-16}$ *** | $<2 \times 10^{-16}$ *** | $1.18 \times 10^{-5}$ *** |
| ARGSL | 5.22 ± 4.07 | 23.92 ± 32.61 | $<2 \times 10^{-16}$ *** | $<2 \times 10^{-16}$ *** | $<2 \times 10^{-16}$ *** |
| INGSL | 406.21 ± 248.94 | 2296.62 ± 1004.61 | $<2 \times 10^{-16}$ *** | $<2 \times 10^{-16}$ *** | $<2 \times 10^{-16}$ *** |
| TGSL | 20,277.67 ± 11,582.07 | 62,864.66 ± 48,182.85 | $<2 \times 10^{-16}$ *** | $<2 \times 10^{-16}$ *** | $1.65 \times 10^{-5}$ *** |

SIN, Sinigrin; GNA, Gluconapin; GBN, Glucobrassicanapin; EPI, Epiprogoitrin; PRO, Progoitrin; GIB, Glucoiberin; GAL, Glucoalyssin; GCH, Glucocheirolin; GRA, Glucoraphanin; GRE, Glucoraphenin; GRH, Glucoraphasatin; GER, Glucoerucin; SNB, Sinalbin; GTL, Glucotropaeolin; GNS, Gluconasturtiin; NGBN, Neoglucobrassicin; 4HGBN, 4-Hydroxyglucobrassicin; 4MGBN, 4-Methoxyglucobrassicin; ALGSL, Aliphatic glucosinolates; ARGSL, Aromatic glucosinolates; INGSL, Indolic glucosinolates; TGSLs, Total glucosinolates; SD, standard deviation; G, genotype effect; O, organ effect; G × O, genotype-by-organ interaction; ns, non-significant; *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$.

### *3.2. Functional Diversity in Phenolic Accumulation and Antioxidant Capacity*

Although GSLs are well-characterized specialized metabolites of radish, phenolic compounds also represent a major metabolite class contributing to both antioxidant capacity and nutritive quality of Brassicaceae vegetables [6,16]. Accordingly, TPC together with ABTS and DPPH radical scavenging activities were evaluated to characterize organ- and genotype-specific variation in antioxidant attributes among 23 radish genotypes. Significant effects of genotype, organ, and genotype-by-organ interaction were detected for all three traits ($p < 0.001$), indicating pronounced genetic and tissue-dependent variation across the evaluated germplasm collection (Figure 2A–C; Table S5).

Overall, leaves showed significantly greater ABTS and DPPH radical scavenging activity and TPC than roots. Mean ABTS, DPPH, and TPC values in leaves were 4.20 ± 0.29 mg TE g$^{-1}$ DW, 0.90 ± 0.08 mg AAE g$^{-1}$ DW, and 4.05 ± 0.56 mg GAE g$^{-1}$ DW, respectively, compared with 3.54 ± 0.66 mg TE g$^{-1}$ DW, 0.58 ± 0.07 mg AAE g$^{-1}$ DW, and 1.23 ± 0.36 mg GAE g$^{-1}$ DW in roots. The concurrent increase in TPC and radical scavenging activity indicates that photosynthetically active leaf tissues represent a major reservoir of antioxidant constituents compared with storage roots. This organ-specific pattern is consistent with previous findings in radish, where leaves exhibited higher phenolic contents and antioxidant capacity than roots [58–60]. These observations further emphasize the nutritional potential of radish foliage, given its comparatively high TPC and in vitro antioxidant activity.

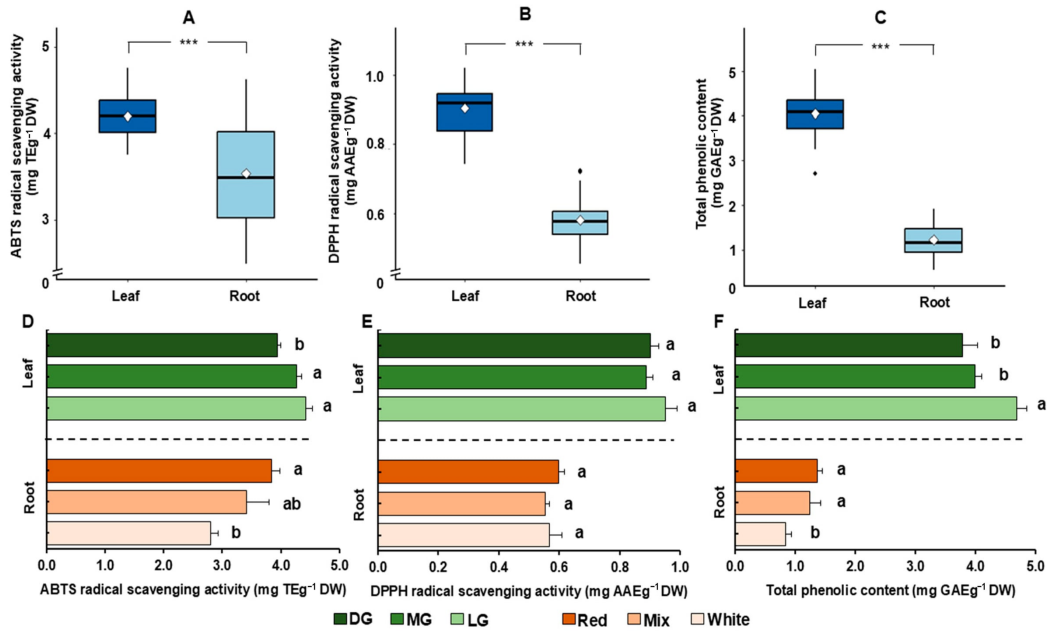



**Figure 2.** Organ- and pigmentation-associated differences in antioxidant activity and total phenolic content across radish genotypes. (**A**–**C**) Comparison of ABTS radical scavenging activity, DPPH radical scavenging activity, and total phenolic content between leaf and root tissues ($n = 23$ genotypes). Statistical differences between organs were determined using the paired Wilcoxon signed-rank test; *** indicates significance at $p < 0.001$. (**D**–**F**) Mean variations in ABTS, DPPH radical scavenging activity, and total phenolic content among leaf (DG, dark green, $n = 7$; MG, medium green, $n = 12$; LG, light green, $n = 4$) and root (red, $n = 14$; mixed, $n = 4$; and white, $n = 5$) pigmentation groups. Error bars represent the standard error across genotype-level means within each group, and the different lowercase letters indicate significant differences among pigmentation groups according to Duncan's multiple range test ($p < 0.05$).

Further, pigmentation-associated differences were also evident for ABTS activity and TPC, whereas DPPH radical scavenging activity did not differ significantly across pigmentation classes. Among leaf pigmentation groups, light green and medium green leaves exhibited significantly higher ABTS activity, and light green leaves showed the highest TPC than dark and medium green leaves (Figure 2D–F; Table S6). Alternatively, root pigmentation exhibited a distinct pattern. Red roots showed the greatest ABTS radical scavenging activity, followed by mixed-colour roots, while white roots showed the lowest ABTS activity. Similarly, TPC differed significantly among root color groups, with red and mixed-color roots exhibiting comparatively higher contents than white roots. This outcome aligns with prior studies reporting higher antioxidant and polyphenol activity in red-skinned radish [61]. The contrasting ABTS and DPPH responses across pigmentation groups may reflect differences in assay-specific reactivity toward the phytochemical constituents of radish extracts. Previous comparative studies have reported stronger associations between ABTS activity and phenolic content than DPPH activity [62], consistent with the positive TPC-ABTS relationships observed in the present study.

Considerable genotype-level variation in antioxidant activity and TPC was also observed within each organ (Table S5). In leaves, the light green IT308360 and medium green IT308361 (China) genotypes recorded comparatively higher ABTS and DPPH radical scavenging activities and TPC, while the dark green genotypes IT340147 (Turkey) and IT344617 (Russia) showed lower values. Correspondingly, for antioxidant activity and phenolic accumulation in roots, red genotypes IT250789 (China) and IT344615 (Russia), together with the mixed-colour genotype CHERISH 1 (Korea) exhibited higher levels, and

the white-colour genotype IT297044 (Russia) showed lower levels for all three traits. Notably, IT344617 (Russia), characterized by dark green leaves and white roots, also exhibited reduced ABTS activity and TPC in roots, with lower ABTS and DPPH radical scavenging activities and TPC in leaves, highlighting a consistently lower antioxidant profile across both tissues. The marked variation among individual genotypes therefore identifies contrasting radish genetic resources with distinct antioxidant profiles, which may be valuable for subsequent nutritional characterization and germplasm utilization. Several genotypes also showed contrasting TGSL and antioxidant profiles, underscoring the complexity of their interrelationships across the evaluated germplasm.

### *3.3. Multivariate Characterization of Glucosinolates and Antioxidant-Related Profiles*

Pairwise correlation analysis revealed organ-specific association patterns among GSL and antioxidant-related traits (Figure S1). In radish leaves, moderate-to-strong positive correlations were predominantly observed among individual aliphatic GSLs and within the three indolic GSLs ($r \geq 0.5$, $p < 0.05$), while aromatic GSLs exhibited more variable relationships. Roots displayed comparable positive associations among several aliphatic compounds, while the indolic and aromatic GSLs displayed compound-specific association patterns. In both organs, ABTS activity was negatively associated with several individual GSLs, whereas TPC was positively correlated with ABTS and DPPH activities in leaves, and with ABTS activity in roots ($r = 0.6$, $p < 0.05$). This aligns with a previous study in *Brassica juncea* reporting a positive association between total phenolics and antioxidant activity, along with a negative association of ABTS with several GSLs [63]. These associations suggest that greater accumulation of certain GSLs does not necessarily correspond to higher ABTS radical scavenging activity, while the positive TPC-ABTS relationships indicate a possible contribution of phenolic-associated constituents. To further characterize coordinated variation among the measured traits, PCA was conducted separately for leaf and root tissues using 18 individual GSLs, TPC, ABTS, and DPPH activity (Figure 3). GSL subclasses and TGSLs were excluded because they were derived from the individual compounds and would therefore introduce mathematical redundancy. In leaves, PC1 and PC2 explained 34.4% and 21.8% of the variance, accounting for 56.2% cumulatively (Figure 3A,C). Most GSLs were positioned opposite to TPC and the antioxidant traits along PC1, indicating a major disparity between GSL-dominated and antioxidant-associated profiles. Medium and light green genotypes overlapped considerably, whereas several dark green genotypes, particularly IT340147, IT344617, and IT297044, occurred toward the GSL-associated region. Subsequently, root PCA explained 70.6% of the variance across PC1 (57.1%) and PC2 (13.5%) (Figure 3B,D). PC1 primarily represented a compositional gradient between GSL-associated profiles and those characterized by higher ABTS activity and TPC; however, DPPH activity contributed mainly to variation along PC2. White root genotypes were more frequently distributed toward the GSL-associated regions, while red and mixed roots tended toward the opposite end of PC1; nevertheless, substantial overlap remained among pigmentation groups.

Similarly, HCA revealed that distinct phytochemical profiles did not correspond exclusively to pigmentation classes (Figure 3E,F). Comparable color-associated metabolic variation has been reported in radish. Park et al. [64] differentiated white and green root cultivars based on primary and secondary metabolites, with white roots containing higher levels of several GSLs and green roots showing greater phenolic accumulation. Likewise, Mei et al. [15] reported clustering of eight radish landraces into three PCA groups broadly corresponding to white-fleshed, green- and white-fleshed, and red-fleshed taproots. In contrast, Zhang et al. [65] found similar anthocyanin and flavonoid profiles between black and white radishes despite their contrasting coloration, indicating that visible pigmen-

tation does not invariably correspond to overall metabolic composition. Together, these findings highlight that pigmentation captures part of the phytochemical variation in radish, but does not fully account for genotype-specific variation in GSL composition, TPC, and antioxidant activities. Pigmentation thus provides a useful context for phytochemical comparison, but its value as a quality indicator depends on the trait examined, as illustrated by pigmentation-associated differences in ABTS activity but not in DPPH activity.

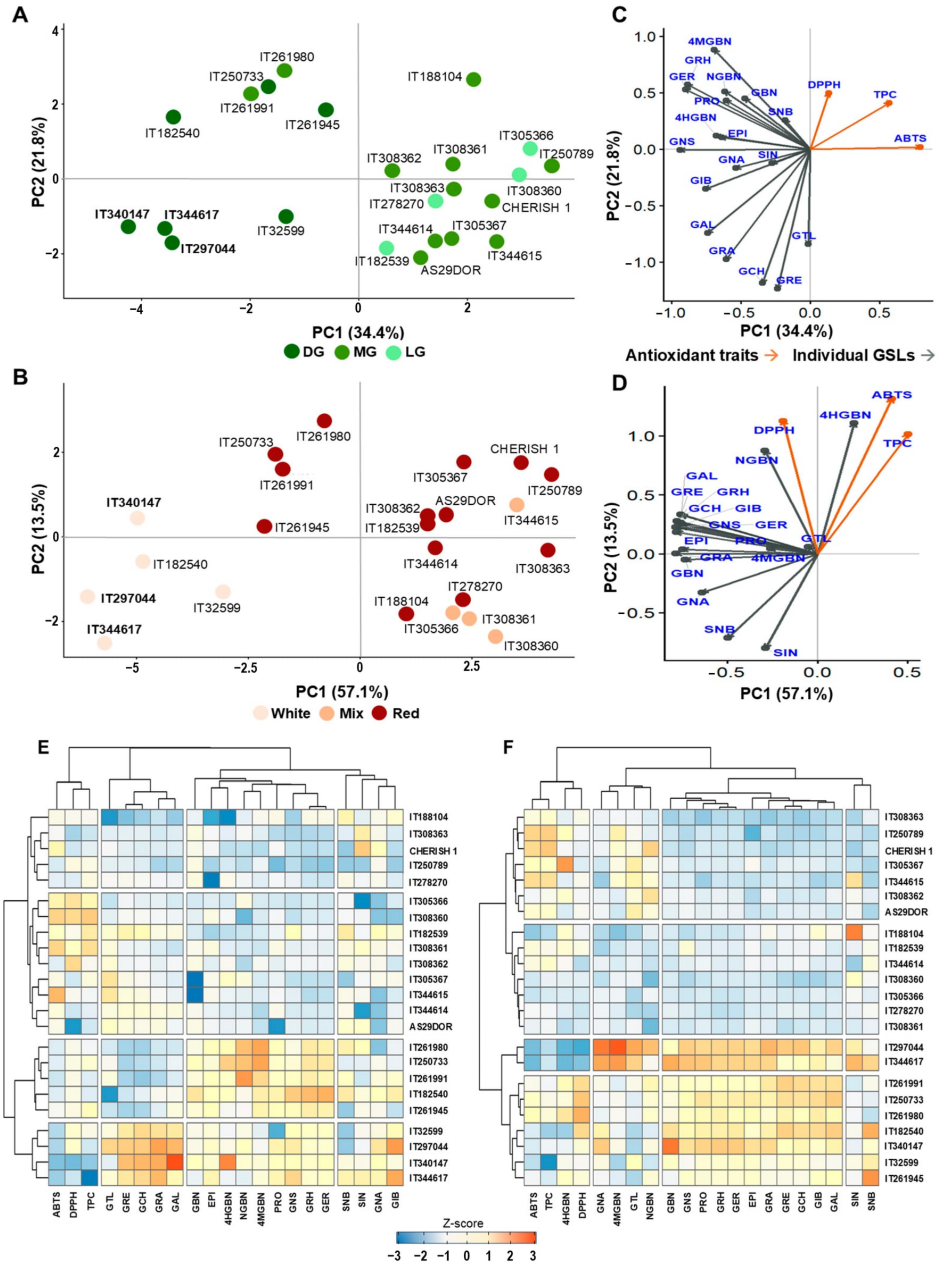



**Figure 3.** Principal component and hierarchical clustering of individual GSLs and antioxidant-associated profiles in radish genotypes. (**A**–**D**) are PCA score and loading plots of leaf and root tissues, respectively. In PCA score plots, the point colors represent leaf or root pigmentation groups; DG, dark green; MG, medium green; LG, light green. In the loading plots, orange vectors depict TPC and antioxidant-associated traits, while gray vectors indicate the 18 individual GSLs. (**E**,**F**) Hierarchical clustering maps of GSL compounds, antioxidant activity, and TPC across radish genotypes in leaf and root tissues, respectively. Warmer and cooler colors highlight relatively higher and lower values.

Since PCA and HCA identified only partial correspondence between measured phytochemical profiles and pigmentation classes, supervised analyses were subsequently performed to evaluate class differentiation (Figure 4). In radish metabolomics, OPLS-DA has previously provided clearer group differentiation than PCA and HCA [66]. Unlike unsupervised ordination, PLS-DA and OPLS-DA incorporate class information when extracting latent variables, thereby enabling targeted assessment of variation associated with predefined groups [67,68]. An initial multiclass PLS-DA model and exploratory pairwise OPLS-DA comparisons were conducted to assess the overall differentiation among three pigmentation groups of leaves and roots (Figures S2 and S3). Dark green leaves were more clearly distinguished from both medium and light green leaves, while medium and light green leaves exhibited poor predictive discrimination. Similarly, white root genotypes were well differentiated from both red and mixed genotypes, unlike red and mixed root genotypes (Table S7). These patterns, together with the overlap observed in the unsupervised analyses, supported the consolidation of medium and light green leaves, and red and mixed roots for subsequent binary OPLS-DA.

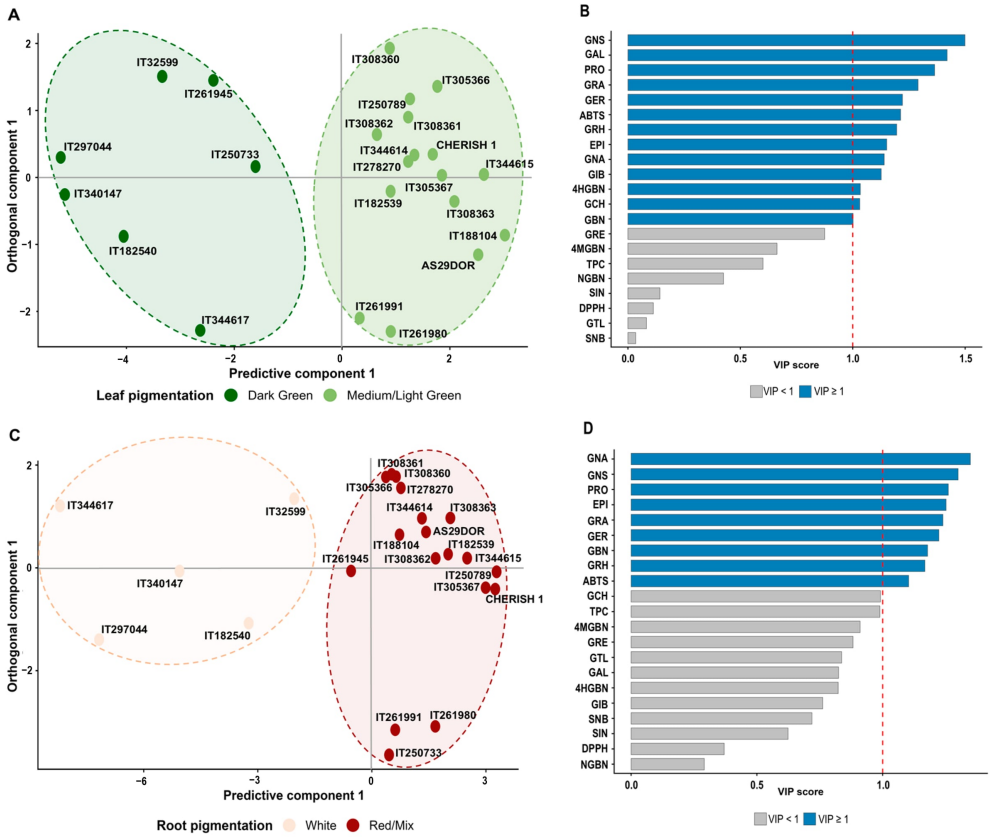



**Figure 4.** OPLS-DA of phytochemical profiles according to pigmentation groups in radish leaves and roots. (**A**,**C**) OPLS-DA score plot comparing dark green leaves with the combined medium/light group, and white roots with the combined red/mix group, respectively. (**B**,**D**) VIP score for the leaf and root OPLS-DA models, respectively. The red dashed line indicates the VIP threshold of 1; blue variables (exceeding VIP > 1) represent major contributors to class discrimination.

The binary OPLS-DA models demonstrated cross-validated discrimination between dark green and combined medium/light green leaves ($R^2Y = 0.86$, $Q^2 = 0.64$). Similarly, white roots showed distinct separation from the combined red/mix group ($R^2Y = 0.78$, $Q^2 = 0.57$), indicating that the major phytochemical differences were associated with these broader pigmentation classes rather than among the intermediate pigmentation groups (Figure 4A,C; Table S7). Further, VIP analysis identified the variables contributing most

strongly to the observed discrimination (Figure 4B,D). Remarkably, ABTS activity exceeded the VIP threshold in both leaf and root models, signifying that variation in antioxidant capacity also contributed to pigmentation differentiation alongside individual GSLs. Variables with VIP values > 1 were considered the principal contributors to pigmentation-associated discrimination and were compared with SHAP-based feature rankings from the subsequent ML analyses to identify candidate biochemical discriminators prioritized across both approaches. This is partially consistent with a prior study, which applied pairwise OPLS-DA to differently colored radish flesh and skin tissues and used VIP score criteria to prioritize metabolites contributing to their differentiation [61].

### *3.4. Machine Learning Identifies the Best Performing Predictive Model for Radish Pigmentation*

Five supervised ML algorithms were compared for phytochemical-associated classification of the consolidated pigmentation groups defined by OPLS-DA (Figure 5; Figure S4). The binary groups (Dark Green vs. Medium/Light Green leaves, and White vs. Red/Mix roots) were retained to include all genotypes and evaluate broader pigmentation-associated phytochemical contrasts rather than restricting classification to the most readily separable pairwise groups. All five models exhibited high predictive performance for both leaf and root datasets, indicating that consolidated pigmentation groups were associated with reproducible multivariate differences in phytochemical composition. The concordant performance across multiple algorithms reinforced the OPLS-DA results and indicated that class discrimination was not restricted to a particular modeling framework.

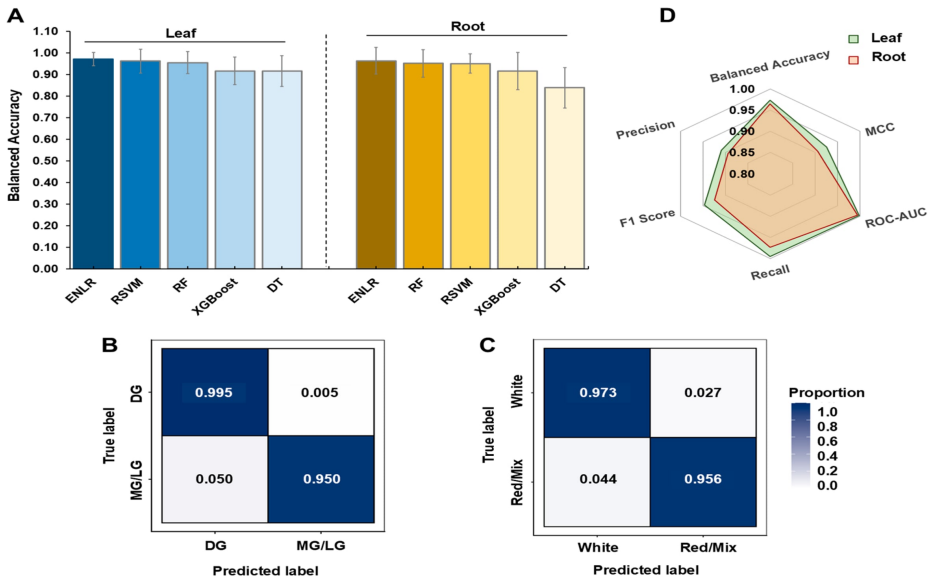



**Figure 5.** Comparative performance of supervised machine learning models for phytochemical-based pigmentation classification in radish. (**A**) Performance comparison of five machine learning algorithms for pigmentation classes in leaf and root datasets. Bars represent mean balanced accuracy, and error bars indicate the standard deviation calculated across 30 individual validation folds (10 repetitions × 3 folds). (**B**,**C**) Normalized confusion matrices of the selected elastic-net logistic regression model for leaf and root pigmentation classification, respectively. Values on the right indicate the proportion of samples classified into each observed and predicted class. (**D**) Radar plot summarizing the performance of the selected elastic-net logistic regression model across six evaluation metrics (balanced accuracy, precision, recall, F1 score, MCC, and ROC-AUC) for leaf and root datasets. Abbreviations are as follows: ENLR, Elastic-net logistic regression; RSVM, Radial support vector machine; RF, Random forest; XGBoost, extreme gradient boosting; DT, Decision tree; DG, Dark green; MG, Medium green; LG, Light green; MCC, Matthews correlation coefficient; ROC-AUC, Area under the receiver operating characteristic curve.

Among the evaluated models, ENLR consistently achieved the highest overall performance in both organs. ENLR achieved balanced accuracy of 0.97, MCC of 0.93, and ROC-AUC of 1.00 for leaf classification, with the corresponding values of 0.95, 0.91, and 0.99 for root classification, respectively. RSVM and RF also demonstrated excellent predictive performance, while XGBoost and DT showed comparatively lower classification performance. In addition, high balanced accuracy and MCC values further indicate that the classification performance was consistently maintained across both pigmentation classes rather than being driven predominantly by the larger class. Although using the same cross-validation procedure for hyperparameter optimization and model evaluation may introduce some optimism into the estimated performance, the results support discrimination between pigmentation groups within the evaluated germplasm collection. Given the number of predictors ($n = 21$) relative to the genotype panel ($n = 23$), the transferability of these findings to independent radish germplasm remains to be established.

The robust performance of ENLR may reflect its capacity to accommodate correlated predictors. ENLR combines lasso and ridge penalties; the lasso component reduces the contribution of less informative predictors and can eliminate them. Moreover, the ridge component stabilizes the coefficient estimates when multiple predictors contain overlapping information [69]. This property is relevant to the present dataset because several GSLs and antioxidant-associated traits were significantly correlated (Figure S1). Furthermore, the lower performance of XGBoost and DT models suggests that their non-linear partitioning provided no additional predictive benefit under the present sample size and validation design. This result, however, does not demonstrate that the underlying biological relationships are entirely linear.

Previous studies provide complementary evidence that ML can recover phytochemical and pigmentation-related information from multivariate analytical profiles. In *Brassica juncea*, hyperspectral models predicted anthocyanin, GSL, flavonoid, and chlorophyll contents, highlighting that spectral variation captured underlying differences in leaf phytochemistry [70]. Earlier, in radish SHAP was applied to interpret the contributions of image-derived color and lighting variables to ML-based prediction of leaf SPAD values [71]. In addition, a previous study in onion reported that volatile-metabolite profiles successfully discriminate red, yellow, and white onion bulbs, and subsequently applied SHAP analysis to determine features contributing most strongly to color-group prediction [72]. The present study applies this classification and interpretation framework to radish by examining whether directly quantified GSLs, TPC, and antioxidant activity can discriminate organ-specific pigmentation groups. Based on its superior cross-validated performance, ENLR was selected for subsequent SHAP analysis to facilitate biological interpretation of the classification model. Although logistic regression coefficients are intrinsically interpretable, SHAP enables quantification of both the magnitude and direction of each phytochemical feature's contribution to individual predictions relative to model baseline [45], thereby enabling identification of the metabolites most strongly associated with pigmentation classification.

### *3.5. SHAP Analysis Reveals Candidate Phytochemical Discriminators*

To elucidate the feature-level architecture of ENLR classification, SHAP analysis was conducted for individual phytochemical predictors in leaf and root models (Figure 6). Global feature importance was determined from the mean absolute SHAP value calculated across replicate observations, providing an aggregate ranking of predictors according to their overall influence on model predictions. SHAP summary plots further illustrated the sign of each attribution and its relationship with the standardized feature value. Positive SHAP values favored dark green leaf or white root classification, while negative values

favored the corresponding combined reference groups. Group-specific SHAP importance profiles are provided in Figure S5.

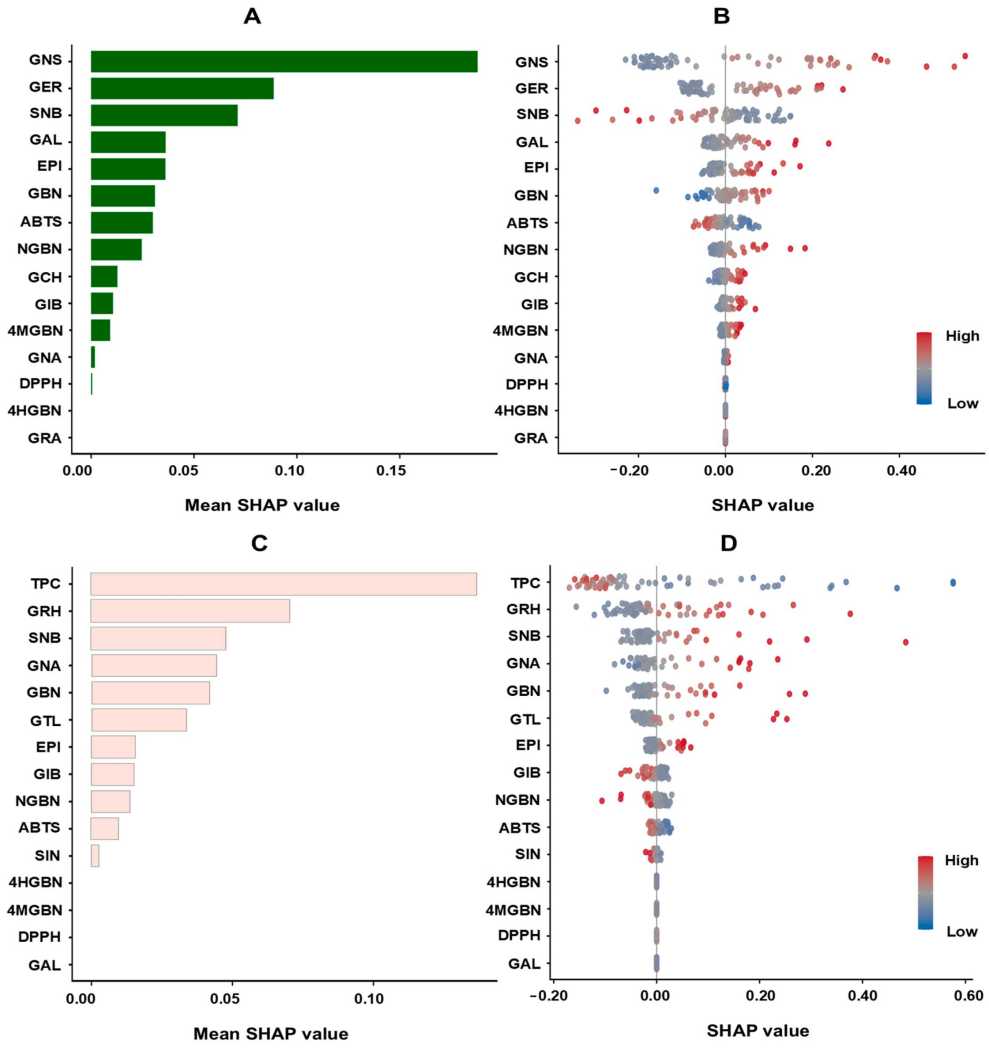



**Figure 6.** SHAP-based interpretation of feature importance for organ-specific pigmentation in radish. (**A**,**C**) Global SHAP feature importance for leaf and root pigmentation classification, respectively, based on the ENLR model. (**B**,**D**) SHAP summary plots illustrating the direction and magnitude of individual feature contributions to the corresponding classification models. Each point represents one observation, and point color indicates the standardized feature value (blue = low, red = high). Positive SHAP values favor prediction towards the positive class (dark green leaves or white roots), and negative SHAP values favor the reference class (medium/light green leaves or red/mixed roots).

In leaves, GNS contributed most strongly to pigmentation classification, followed by GER and SNB, whereas GAL, EPI, and GBN showed comparatively smaller but consistent contributions (Figure 6A,B). High GNS and GER predominantly generated positive SHAP values and shifted prediction towards the dark green class, while elevated SNB values contributed toward the combined medium and light green class. The contributions of GNS and GER were consistent with their higher mean concentrations in dark green leaves (Table S3), indicating that their predictive importance reflected observed differences in GSL composition among pigmentation groups. GER serves as a precursor of GRH through conversion mediated by glucoraphasatin synthase 1, which is highly expressed in leaves [73]. Furthermore, GNS has been documented as a major GSL in wild radish [74]. These findings provide biochemical and comparative context for their contributions to the leaf phytochem-

ical differences observed among pigmentation groups. Collectively, the concentration of model attribution in GNS, GER, and SNB delineated a compact GSL signature encompassing both aliphatic and aromatic compounds, suggesting that the ENLR discriminatory signal associated with leaf pigmentation was primarily attributable to a limited subset of GSLs.

In roots, TPC was the dominant predictor of pigmentation class, substantially exceeding the contribution of the remaining variables (Figure 6C,D). GRH, SNB, GNA, GBN, and GTL also contributed to model predictions, although with progressively lower importance. The SHAP summary plot indicated that reduced TPC values and higher GRH, SNB, GNA, and GBN concentrations favored white root classification. This pattern is consistent with the higher mean GRH and SNB concentrations and lower TPC in white roots than in red and mixed roots (Tables S3 and S6). The inverse correlations of TPC with GRH, SNB, and GNA further suggest coordinated variation in phenolic content and GSL composition among root genotypes (Figure S1). The association of lower TPC with white root classification aligns with previous findings from red and white radish hairy-root cultures, where white hairy roots exhibited lower phenolic content and no detectable anthocyanins compared with red hairy roots [75]. Because TPC represents the cumulative response of chemically diverse phenolics, its SHAP importance indicates a broad phenolic signature rather than the contribution of a specific pigment.

Comparison of SHAP feature rankings with the OPLS-DA VIP analysis highlighted considerable agreement for leaf pigmentation, with GNS, GER, GAL, EPI, GBN, ABTS, GCH, and GIB consistently prioritized by both approaches. Among these, GNS and GER exhibited the highest predictive importance in the ENLR model, while SNB emerged as an additional SHAP-prioritized feature despite receiving comparatively lower emphasis in the VIP analysis. In roots, agreement between the two approaches was more limited, with GNA, GBN, EPI, GRH, and ABTS consistently identified, whereas SHAP additionally highlighted TPC, SNB, and GTL as important predictive variables. This divergence reflects distinct measures of feature importance: VIP summarizes predictor contributions to OPLS-DA discrimination, while SHAP quantifies their contributions to ENLR predictions. When predictors are correlated, shared discriminatory information may be attributed differently by the two approaches, resulting in differences in feature rankings [76]. To further examine the stability of SHAP feature rankings, ENLR models were refitted across 30 cross-validation splits (Table S8). GNS, GER, and SNB were frequently ranked among the five leading leaf predictors (83.30–96.70%). Similarly, TPC, GBN, and GRH were frequently prioritized in roots (86.70–96.70%); however, SNB showed lower stability (26.70%).

Accordingly, variables supported by both analytical approaches (OPLS-DA and SHAP) were regarded as consensus candidate phytochemical discriminators, and SHAP-specific variables were considered predictive candidate phytochemical discriminators; both warrant further validation in independent germplasm collections. Validation in a larger, independent germplasm panel is required to establish external generalizability and assess the suitability of these features for breeding and quality assessment. In addition, analytical precision, extraction recovery, and matrix effects were not assessed as part of the present analytical validation, and their evaluation would further characterize the repeatability, extraction efficiency, and matrix-dependent response of GSL quantification. Collectively, integration of VIP and SHAP rankings provides a comprehensive basis for prioritizing organ-specific phytochemical signatures associated with radish pigmentation.

## 4. Conclusions

In summary, evaluation of 23 radish genotypes (20 days after sowing) revealed pronounced phytochemical differentiation among genotypes, between organs, and across pigmentation groups. Roots were characterized by greater GSL accumulation, and leaves

exhibited stronger phenolic and antioxidant profiles. Pigmentation reflected broad phytochemical tendencies, but variation within color classes suggests that these trends did not uniformly characterize individual genotypes. Positive associations between TPC and antioxidant activity, together with inverse relationships between ABTS activity and several individual GSLs, further defined contrasting organ-specific phytochemical profiles. As indirect, assay-dependent measurements, TPC, ABTS, and DPPH collectively offer complementary evidence of antioxidant-associated variation among radish genotypes, establishing a basis for further identification of individual contributing metabolites and assessment of their physiological relevance. Specifically, IT340147 (Turkey), along with IT297044 and IT344617 (Russia), were identified as high-GSL genotypes. IT308360 and IT308361 from China exhibited promising leaf phenolic and antioxidant profiles, whereas IT250789 from China, IT344615 from Russia, and CHERISH 1 from Korea were prominent for these traits in roots. These complementary genotypes may provide valuable parental resources for developing radish cultivars with enhanced organ-specific GSL or antioxidant-associated characteristics and also for combining contrasting phytochemical profiles through targeted breeding. Despite within-class variation, ML classification demonstrated that combined phytochemical profiles discriminated the consolidated pigmentation groups within the evaluated collection. Integration of OPLS-DA, ML, and SHAP analyses prioritized GNS, GER, and SNB as the principal contributors in leaves and identified TPC and GRH as the dominant contributors in roots, with GBN additionally showing consistent prioritization across cross-validation splits. Partial concordance between SHAP and OPLS-DA VIP rankings supported a focused set of candidate phytochemical discriminators. The identified phytochemical patterns were observed at a defined early-harvest stage under standardized greenhouse conditions. Further evaluation across developmental stages, cultivated environments, and independent germplasm collections will help establish the consistency and broader applicability of these patterns and assess potential genotype-by-environment interactions. Together, these trait-specific genotypes and candidate phytochemical features provide a foundation for strategic germplasm evaluation and targeted improvement of organ-specific GSL and antioxidant-associated profiles in radish.

**Supplementary Materials:** The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/antiox15111449/s1, Figure S1: Correlation plot for all individual GSLs, subclasses, TGSLs along with phenolic content and antioxidant activity in leaf and root; Figure S2: Pairwise OPLS-DA comparisons among dark green, medium green, and light green leaf pigmentation groups and multiclass PLS-DA of all three groups; Figure S3: Pairwise OPLS-DA comparisons among red, mix, and white root pigmentation groups and multiclass PLS-DA of all three groups; Figure S4: Cross-validated classification performance of five supervised machine learning models for leaf and root pigmentation groups in radish; Figure S5: Pigmentation group-specific SHAP importance of phytochemical predictors in radish; Table S1: UPLC-MS/MS parameters for the targeted quantification of glucosinolates in radish tissues, including retention times, calibration equations, coefficients of determination, and multiple-reaction-monitoring transitions; Table S2: Optimal hyperparameters for machine learning classification of radish pigmentation; Table S3: Variations in glucosinolate compounds for leaf and roots among 23 radish genotypes; Table S4: Pigmentation-associated differences in major glucosinolate classes across radish leaf and root tissues; Table S5: Total phenolic content and antioxidant capacities of leaf and root tissues in radish; Table S6: Mean variations in ABTS, DPPH radical scavenging activity, and total phenolic content among pigmentation groups; Table S7: Performance statistics of supervised multivariate models used to discriminate pigmentation groups in radish leaves and roots; Table S8: Stability of SHAP feature-importance rankings across cross-validation in leaf and root pigmentation classification of radish germplasm.

**Author Contributions:** S.C.: Writing—Original draft preparation, Visualization, Methodology, Formal analysis, Data curation, Validation, Conceptualization; K.O.: Investigation, Data curation, Analysis;

S.-H.K.: Writing-review, Supervision, Resources, Funding acquisition, Project administration, Data curation. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Research Program for Agricultural Science and Technology Development (Project NO. PJ01425501/RS-2019-RD007776) and by the 2026 RDA Fellowship Program, National Institute of Agricultural Sciences, Rural Development Administration (RDA), Republic of Korea.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The original contributions presented in this study are included in the article/Supplementary Materials. Further inquiries can be directed to the corresponding author.

**Acknowledgments:** The authors acknowledge the technical support and institutional facilities that contributed to this study.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

| | |
|---|---|
| GSL | Glucosinolate |
| TPC | Total phenolic content |
| ABTS | 2,2′-azino-bis-3-ethylbenzothiazoline-6-sulfonic acid |
| DPPH | 1,1-diphenyl-2-picrylhydrazyl |
| ML | Machine learning |
| SHAP | Shapley additive explanations |
| UPLC-ESI-MS/MS | Ultra-performance liquid chromatography-electrospray ionization-tandem mass spectrometry |
| SIN | Sinigrin |
| GNA | Gluconapin |
| GBN | Glucobrassicanapin |
| EPI | Epiprogoitrin |
| PRO | Progoitrin |
| GIB | Glucoiberin |
| GAL | Glucoalyssin |
| GCH | Glucocheirolin |
| GRA | Glucoraphanin |
| GRE | Glucoraphenin |
| GRH | Glucoraphasatin |
| GER | Glucoerucin |
| SNB | Sinalbin |
| GTL | Glucotropaeolin |
| GNS | Gluconasturtiin |
| NGBN | Neoglucobrassicin |
| 4HGBN | 4-hydroxyglucobrassicin |
| 4MGBN | 4-methoxyglucobrassicin |
| RF | Random forest |
| ENLR | Elastic-net logistic regression |
| MCC | Matthews correlation coefficient |
| ROC-AUC | Area under the receiver operating characteristic curve |

## References

1. Le, T.N.; Chiu, C.H.; Hsieh, P.C. Bioactive compounds and bioactivities of *Brassica oleracea* L. var. italica sprouts and microgreens: An updated overview from a nutraceutical perspective. *Plants* **2020**, *9*, 946. [CrossRef]
2. Partap, M.; Sharma, D.; Thakur, M.; Verma, V.; Bhargava, B. Microgreen: A tiny plant with superfood potential. *J. Funct. Foods* **2023**, *107*, 105697. [CrossRef]
3. Tachie, C.; Nwachukwu, I.D.; Aryee, A.N. Trends and innovations in the formulation of plant-based foods. *Food Prod. Process. Nutr.* **2023**, *5*, 16. [CrossRef]

4. Liu, Y.; Wang, C.; Chen, H.; Dai, G.; Cuimu, Q.; Shen, W.; Gao, L.; Zhu, B.; Gao, C.; Chen, L. Comparative transcriptome analysis reveals transcriptional regulation of anthocyanin biosynthesis in purple radish (*Raphanus sativus* L.). *BMC Genom.* **2024**, *25*, 624. [CrossRef]
5. Zhou, R.; Qin, X.; Hou, J.; Liu, Y. Research progress on Brassicaceae plants: A bibliometrics analysis. *Front. Plant Sci.* **2024**, *15*, 1285050. [CrossRef]
6. Gamba, M.; Asllanaj, E.; Raguindin, P.F.; Glisic, M.; Franco, O.H.; Minder, B.; Bussler, W.; Metzger, B.; Kern, H.; Muka, T. Nutritional and phytochemical characterization of radish (*Raphanus sativus*): A systematic review. *Trends Food Sci. Technol.* **2021**, *113*, 205–218. [CrossRef]
7. Kim, J.H.; Kim, H.; Park, M.; Park, B.; Park, B.; Park, H.Y.; Lee, G.; Yu, J.S. Comparative metabolic profiling of different Raphanus sativus cultivars and their antioxidant and anti-inflammatory activities. *Appl. Sci.* **2024**, *15*, 247. [CrossRef]
8. Chen, C.; Kim, R.H.; Hwang, K.T.; Kim, J. Chemical compounds and bioactivities of the extracts from radish (*Raphanus sativus*) sprouts exposed to red and blue light-emitting diodes during cultivation. *Eur. Food Res. Technol.* **2023**, *249*, 1551–1562. [CrossRef]
9. Maina, S.; Misinzo, G.; Bakari, G.; Kim, H.-Y. Human, Animal and Plant Health Benefits of Glucosinolates and Strategies for Enhanced Bioactivity: A Systematic Review. *Molecules* **2020**, *25*, 3682. [CrossRef]
10. Kitainda, V.; Jez, J.M. Structural studies of aliphatic glucosinolate chain-elongation enzymes. *Antioxidants* **2021**, *10*, 1500. [CrossRef]
11. Park, C.H.; Park, S.Y.; Park, Y.J.; Kim, J.K.; Park, S.U. Metabolite profiling and comparative analysis of secondary metabolites in Chinese cabbage, radish, and hybrid xBrassicoraphanus. *J. Agric. Food Chem.* **2020**, *68*, 13711–13719. [CrossRef]
12. Jiang, W.; Zhou, X. Hydrolysis of radish anthocyanins to enhance the antioxidant and antiproliferative capacities. *Food Chem.* **2019**, *294*, 477–485. [CrossRef]
13. Kasprzak-Drozd, K.; Oniszczuk, T.; Stasiak, M.; Oniszczuk, A. Beneficial effects of phenolic compounds on gut microbiota and metabolic syndrome. *Int. J. Mol. Sci.* **2021**, *22*, 3715. [CrossRef]
14. Višnjevec, A.M.; Barp, L.; Fanesi, B.; Iacumin, L.; Lucci, P.; Moret, S. Phenolic Compounds and Glucosinolates Composition of Cauliflower and Broccoli Byproducts Puree after the Lactic Acid Fermentation. *Food Chem.* **2025**, *492*, 145346. [CrossRef]
15. Mei, S.; He, Z.; Zhang, J. Identification and analysis of major flavor compounds in radish taproots by widely targeted metabolomics. *Front. Nutr.* **2022**, *9*, 889407. [CrossRef]
16. Manivannan, A.; Kim, J.H.; Kim, D.S.; Lee, E.S.; Lee, H.E. Deciphering the nutraceutical potential of *Raphanus sativus*—A comprehensive overview. *Nutrients* **2019**, *11*, 402. [CrossRef]
17. Yi, G.; Lim, S.; Chae, W.B.; Park, J.E.; Park, H.R.; Lee, E.J.; Huh, J.H. Root Glucosinolate Profiles for Screening of Radish (*Raphanus sativus* L.) Genetic Resources. *J. Agric. Food Chem.* **2016**, *64*, 61–70. [CrossRef] [PubMed]
18. Kim, B.; Hur, O.; Lee, J.E.; Assefa, A.D.; Ko, H.C.; Chung, Y.J.; Rhee, J.h.; Hahn, B.S. Characterization of phenotypic traits and evaluation of glucosinolate contents in radish germplasms (*Raphanus sativus* L.). *Korean J. Plant Res.* **2021**, *34*, 575–599. [CrossRef]
19. Li, Z.; Zheng, S.; Liu, Y.; Fang, Z.; Yang, L.; Zhuang, M.; Zhang, Y.; Lv, H.; Wang, Y.; Xu, D. Characterization of glucosinolates in 80 broccoli genotypes and different organs using UHPLC-Triple-TOF-MS method. *Food Chem.* **2021**, *334*, 127519. [CrossRef]
20. Dunn, W.B.; Broadhurst, D.; Begley, P.; Zelena, E.; Francis-McIntyre, S.; Anderson, N.; Brown, M.; Knowles, J.D.; Halsall, A.; Haselden, J.N.; et al. Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. *Nat. Protoc.* **2011**, *6*, 1060–1083. [CrossRef] [PubMed]
21. Wishart, D.S. Metabolomics for Investigating Physiological and Pathophysiological Processes. *Physiol. Rev.* **2019**, *99*, 1819–1875. [CrossRef]
22. Iwar, K.; Desta, K.T.; Ochar, K.; Kim, S.-H. Unveiling Glucosinolate Diversity in Brassica Germplasm and In Silico Analysis for Determining Optimal Antioxidant Potential. *Antioxidants* **2024**, *13*, 376. [CrossRef]
23. Yao, M.; Xing, S.; Yao, G.; Yong, W.; Ling, Y.; Chu, B. Simultaneous quantification of 14 glucosinolates in rapeseeds by ultra high performance liquid chromatography—Tandem mass spectrometry. *Food Chem.* **2025**, *467*, 142302. [CrossRef]
24. Zhong, Y.; Xie, Y.; Zhang, D.; Li, G.; Yu, J. Integrated metabolomic and transcriptomic analysis of metabolic diversity and biosynthesis of glucosinolates and flavonoids in various cultivars of radish microgreens. *Food Biosci.* **2024**, *59*, 104055. [CrossRef]
25. Tiozon, R.J.N.; Sreenivasulu, N.; Alseekh, S.; Sartagoda, K.J.D.; Usadel, B.; Fernie, A.R. Metabolomics and machine learning technique revealed that germination enhances the multi-nutritional properties of pigmented rice. *Commun. Biol.* **2023**, *6*, 1000. [CrossRef]
26. Yang, S.; Li, Z.; Li, W.; Li, C.; Yang, X.; Zhao, Y.; Liu, R. Unraveling the variations and relationships between phytochemical constituents and anti-inflammatory potentials of three Zanthoxylum species through metabolomics and explainable machine learning. *Ind. Crops Prod.* **2024**, *212*, 118281. [CrossRef]
27. Eugelio, F.; Mascini, M.; Marone, E.; Fanti, F.; Palmieri, S.; Sergi, M.; Del Carlo, M.; Compagnone, D. Phenolic Profile as a Powerful Machine Learning Tool for Identification, Traceability, and Quality Control of Olive Cultivars. *J. Agric. Food Chem.* **2025**, *73*, 23671–23683. [CrossRef]

28. Ochar, K.; Ki, D.W.; Moon, S.; Bissah, M.N.; Kim, S.H. Quantitative analysis of root system architecture and fresh weight biomass traits highlight phenotypic variation in radish (*Raphanus sativus* L.) germplasm. *Sci. Rep.* **2026**, *16*, 8001. [CrossRef]
29. Ochar, K.; Iwar, K.; Na, Y.W.; Kim, S.H. Phytochemical profiling, antioxidant activity and molecular docking analysis in Chinese cabbage (*Brassica rapa* ssp. pekinensis) germplasm. *Sci. Hortic.* **2025**, *343*, 114078. [CrossRef]
30. Rhee, J.-H.; Choi, S.; Lee, J.-E.; Hur, O.-S.; Ro, N.-Y.; Hwang, A.-J.; Ko, H.-C.; Chung, Y.-J.; Noh, J.-J.; Assefa, A.D. Glucosinolate Content in Brassica Genetic Resources and Their Distribution Pattern within and between Inner, Middle, and Outer Leaves. *Plants* **2020**, *9*, 1421. [CrossRef]
31. Kim, S.-H.; Ochar, K.; Iwar, K.; Lee, Y.-J.; Kang, H.J.; Na, Y.-W. Variations of Major Glucosinolates in Diverse Chinese Cabbage (*Brassica rapa* ssp. pekinensis) Germplasm as Analyzed by UPLC-ESI-MS/MS. *Int. J. Mol. Sci.* **2024**, *25*, 4829. [CrossRef]
32. Nithiyanantham, S.; Varadharajan, S.; Siddhuraju, P. Differential effects of processing methods on total phenolic content, antioxidant and antimicrobial activities of three species of Solanum. *J. Food Drug Anal.* **2012**, *20*, 21. [CrossRef]
33. De Mendiburu, F. *Agricolae: Statistical Procedures for Agricultural Research, R Package Version 1.2-2*; CRAN: Vienna, Austria, 2015.
34. Harrell, F.E., Jr.; Dupont, C. *Hmisc: Harrell Miscellaneous, R Package Version 4.4-0*; CRAN: Vienna, Austria, 2020.
35. Slowikowski, K.; Schep, A.; Hughes, S.; Lukauskas, S.; Irisson, J.O.; Kamvar, Z.N.; Ryan, T.; Christophe, D.; Hiroaki, Y.; Gramme, P. *Package ggrepel. Automatically Position Non-Overlapping Text Labels with ggplot2, R package version 0.8.0*; CRAN: Vienna, Austria, 2018.
36. Gu, Z.; Eils, R.; Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **2016**, *32*, 2847–2849. [CrossRef] [PubMed]
37. Thevenot, E.A. *Ropls: PCA, PLS (-DA) and OPLS (-DA) for Multivariate Analysis and Feature Selection of Omics Data, R Package Version 1.44.0*; Bioconductor: Cambridge, MA, USA, 2016.
38. Kuhn, M. Building Predictive Models in R Using the caret Package. *J. Stat. Softw.* **2008**, *28*, 1–26. [CrossRef]
39. Friedman, J.H.; Hastie, T.; Tibshirani, R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J. Stat. Softw.* **2010**, *33*, 1–22. [CrossRef]
40. Karatzoglou, A.; Smola, A.; Hornik, K.; Zeileis, A. kernlab—An S4 Package for Kernel Methods in R. *J. Stat. Softw.* **2004**, *11*, 1–20. [CrossRef]
41. Wright, M.N.; Ziegler, A. ranger: A fast implementation of random forests for high dimensional data in C++ and R. *J. Stat. Softw.* **2017**, *77*, 1–17. [CrossRef]
42. Chen, T.; Guestrin, C. Xgboost: A scalable tree boosting system. In Proceedings of the 22nd ACM Sigkdd International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794.
43. Therneau, T.M.; Atkinson, E.J. *An Introduction to Recursive Partitioning Using the RPART Routines*; Technical Report; Mayo Foundation: Rochester, MN, USA, 1997; p. 452.
44. Robin, X.; Turck, N.; Hainard, A.; Tiberti, N.; Lisacek, F.; Sanchez, J.-C.; Müller, M. pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinform.* **2011**, *12*, 77. [CrossRef] [PubMed]
45. Lundberg, S.M.; Lee, S.I. A unified approach to interpreting model predictions. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 1–3.
46. Strumbelj, E.; Kononenko, I. An efficient explanation of individual classifications using game theory. *J. Mach. Learn. Res.* **2010**, *11*, 1–18. Available online: https://www.jmlr.org/papers/v11/strumbelj10a.html (accessed on 3 November 2026).
47. Wickham, H. Getting Started with ggplot2. In *ggplot2: Elegant Graphics for Data Analysis*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 11–31.
48. Pedersen, T.L. *The Composer of Plots, R Package 1.1.3*; CRAN: Vienna, Austria, 2025.
49. Bhandari, S.R.; Jo, J.S.; Lee, J.G. Comparison of Glucosinolate Profiles in Different Tissues of Nine Brassica Crops. *Molecules* **2015**, *20*, 15827–15841. [CrossRef] [PubMed]
50. Brown, P.D.; Tokuhisa, J.G.; Reichelt, M.; Gershenzon, J. Variation of glucosinolate accumulation among different organs and developmental stages of Arabidopsis thaliana. *Phytochemistry* **2003**, *62*, 471–481. [CrossRef]
51. Hanlon, P.R.; Barnes, D.M. Phytochemical composition and biological activity of 8 varieties of radish (*Raphanus sativus* L.) sprouts and mature taproots. *J. Food Sci.* **2011**, *76*, C185–C192. [CrossRef]
52. Barillari, J.; Cervellati, R.; Costa, S.; Guerra, M.C.; Speroni, E.; Utan, A.; Iori, R. Antioxidant and Choleretic Properties of *Raphanus sativus* L. Sprout (*Kaiware Daikon*) Extract. *J. Agric. Food Chem.* **2006**, *54*, 9773–9778. [CrossRef]
53. Blažević, I.; Montaut, S.; Burčul, F.; Olsen, C.E.; Burow, M.; Rollin, P.; Agerbirk, N. Glucosinolate structural diversity, identification, chemical synthesis and metabolism in plants. *Phytochemistry* **2020**, *169*, 112100. [CrossRef] [PubMed]
54. Bhandari, S.R.; Rhee, J.; Choi, C.S.; Jo, J.S.; Shin, Y.K.; Lee, J.G. Profiling of individual desulfo-glucosinolate content in cabbage head (*Brassica oleracea* var. capitata) germplasm. *Molecules* **2020**, *25*, 1860. [CrossRef]
55. Chae, S.H.; Lee, O.N.; Park, H.Y.; Ku, K.M. Seasonal effects of glucosinolate and sugar content determine the pungency of small-type (Altari) radishes (*Raphanus sativus* L.). *Plants* **2022**, *11*, 312. [CrossRef]
56. Bell, L.; Oloyede, O.O.; Lignou, S.; Wagstaff, C.; Methven, L. Taste and Flavor Perceptions of Glucosinolates, Isothiocyanates, and Related Compounds. *Mol. Nutr. Food Res.* **2018**, *62*, e1700990. [CrossRef]

57. Cartea, M.E.; Velasco, P. Glucosinolates in Brassica foods: Bioavailability in food and significance for human health. *Phytochem. Rev.* **2008**, *7*, 213–229. [CrossRef]
58. Chae, S.H.; Lee, S.H.; Moon, J.H.; Cho, J.Y. Comparison of metabolites and antioxidative activity in leaves and roots of purple radish cultivars. *Food Sci. Preserv.* **2024**, *31*, 985–998. [CrossRef]
59. Goyeneche, R.; Roura, S.; Ponce, A.; Vega-Gálvez, A.; Quispe-Fuentes, I.; Uribe, E.; Di Scala, K. Chemical characterization and antioxidant capacity of red radish (*Raphanus sativus* L.) leaves and roots. *J. Funct. Foods* **2015**, *16*, 256–264. [CrossRef]
60. Koley, T.K.; Khan, Z.; Oulkar, D.; Singh, B.; Maurya, A.; Singh, B.; Banerjee, K. High resolution LC-MS characterization of phenolic compounds and the evaluation of antioxidant properties of a tropical purple radish genotype. *Arab. J. Chem.* **2020**, *13*, 1355–1366. [CrossRef]
61. Cai, D.; Dong, Y.; Wang, L.; Zhao, S. Integrated metabolomics and transcriptomics analysis provides insights into biosynthesis and accumulation of flavonoids and glucosinolates in different radish varieties. *Curr. Res. Food Sci.* **2025**, *10*, 100938. [CrossRef]
62. Floegel, A.; Kim, D.-O.; Chung, S.-J.; Koo, S.I.; Chun, O.K. Comparison of ABTS/DPPH assays to measure antioxidant capacity in popular antioxidant-rich US foods. *J. Food Compos. Anal.* **2011**, *24*, 1043–1048. [CrossRef]
63. Sun, B.; Tian, Y.-X.; Jiang, M.; Yuan, Q.; Chen, Q.; Zhang, Y.; Luo, Y.; Zhang, F.; Tang, H.-R. Variation in the main health-promoting compounds and antioxidant activity of whole and individual edible parts of baby mustard (*Brassica juncea* var. gemmifera). *RSC Adv.* **2018**, *8*, 33845–33854. [CrossRef]
64. Park, C.H.; Ki, W.; Kim, N.S.; Park, S.-Y.; Kim, J.K.; Park, S.U. Metabolic Profiling of White and Green Radish Cultivars (*Raphanus sativus*). *Horticulturae* **2022**, *8*, 310. [CrossRef]
65. Zhang, J.; Qiu, X.; Tan, Q.; Xiao, Q.; Mei, S. A comparative metabolomics study of flavonoids in radish with different skin and flesh colors (*Raphanus sativus* L.). *J. Agric. Food Chem.* **2020**, *68*, 14463–14470. [CrossRef]
66. Wang, J.; Wei, Q.; Wang, W.; Hu, H.; Yan, Y.; Wang, Y.; Li, Y.; Jiang, Y.; Wu, G.; Hu, T.; et al. Understanding the nutraceutical diversity through a comparative analysis of the taproot metabolomes of different edible radish types via UHPLC–Q–TOF–MS. *Food Chem.* **2023**, *403*, 134469. [CrossRef]
67. Barker, M.; Rayens, W. Partial least squares for discrimination. *J. Chemom.* **2003**, *17*, 166–173. [CrossRef]
68. Trygg, J.; Wold, S. Orthogonal projections to latent structures (O-PLS). *J. Chemom.* **2002**, *16*, 119–128. [CrossRef]
69. Zou, H.; Hastie, T. Regularization and Variable Selection Via the Elastic Net. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **2005**, *67*, 301–320. [CrossRef]
70. Yoon, H.I.; Lee, H.; Yang, J.-S.; Choi, J.-H.; Jung, D.-H.; Park, Y.J.; Park, J.-E.; Kim, S.M.; Park, S.H. Predicting Models for Plant Metabolites Based on PLSR, AdaBoost, XGBoost, and LightGBM Algorithms Using Hyperspectral Imaging of *Brassica juncea*. *Agriculture* **2023**, *13*, 1477. [CrossRef]
71. Kamiwaki, Y.; Fukuda, S. Effect of the Light Environment on Image-Based SPAD Value Prediction of Radish Leaves. *Algorithms* **2024**, *17*, 16. [CrossRef]
72. Cheng, K.; Xiao, J.; He, J.; Yang, R.; Pei, J.; Jin, W.; Abd El-Aty, A.M. Unraveling volatile metabolites in pigmented onion (*Allium cepa* L.) bulbs through HS-SPME/GC–MS-based metabolomics and machine learning. *Front. Nutr.* **2025**, *12*, 1582576. [CrossRef]
73. Kakizaki, T.; Kitashiba, H.; Zou, Z.; Li, F.; Fukino, N.; Ohara, T.; Nishio, T.; Ishida, M. A 2-Oxoglutarate-Dependent Dioxygenase Mediates the Biosynthesis of Glucoraphasatin in Radish. *Plant Physiol.* **2017**, *173*, 1583–1593. [CrossRef]
74. Malik, M.S.; Riley, M.B.; Norsworthy, J.K.; Bridges, W., Jr. Variation of Glucosinolates in Wild Radish (*Raphanus raphanistrum*) Accessions. *J. Agric. Food Chem.* **2010**, *58*, 11626–11632. [CrossRef]
75. Lim, J.; Kim, J.K.; Kim, K.; Kwon, H.; Seo, H.; Sathasivam, R.; Park, S.U.; Park, C. Comparative analysis of secondary metabolites and antioxidant activities in white and red radish (*Raphanus sativus*) hairy roots. *J. Plant Biochem. Biotechnol.* **2025**, *34*, 774–784. [CrossRef]
76. Aas, K.; Jullum, M.; Løland, A. Explaining individual predictions when features are dependent: More accurate approximations to Shapley values. *Artif. Intell.* **2021**, *298*, 103502. [CrossRef]


**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.