Food Frontiers

WILEY

RESEARCH ARTICLE OPEN ACCESS

Integrative Metabolomic and Transcriptomic Profiling Reveals the Chemical and Molecular Basis for Aroma Divergence in *Isatis* Fruits

Rong Chen[1] | Yan Qing[1] | Xiaoshan Geng[2] | Qiaoyun Cai[1] | Lili Wang[1] | Fanglin Liu[1] | Huijun Cheng[1] | Yude Peng[3] | Chunfeng Tang[3] | Huaxi Huang[1] | Qin Liu[2]

[1]Key Laboratory for Resource Plants Protection and Utilization of Yili Valley in Xinjiang, College of Biological Sciences and Technology, Yili Normal University, Yining, China | [2]College of Agricultural Engineering, Guangxi Vocational University of Agriculture, Nanning, China | [3]National Center for TCM Inheritance and Innovation, Guangxi Botanical Garden of Medicinal Plants, Nanning, China

Correspondence: Huaxi Huang (18176608120@163.com) | Qin Liu (liuq@gxnzd.edu.cn)

Received: 14 January 2026 | **Revised:** 13 February 2026 | **Accepted:** 18 February 2026

Keywords: *Isatis* | key aroma compounds | odor activity value (OAV) | species differentiation | transcriptional regulation

ABSTRACT

The fruits of *Isatis* species hold significant but underexplored potential in flavor and phytochemical applications. This study comprehensively investigated the volatile organic compounds (VOCs) and their transcriptional regulation underlying aroma divergence in the fruits of three *Isatis* species. Using headspace solid-phase microextraction coupled with gas chromatography–mass spectrometry (HS-SPME–GC–MS), a total of 1042 VOCs were identified, with terpenoids, esters, and ketones being the most abundant classes. Principal component and differential analyses revealed distinct, species-specific VOC profiles. Quantitative analysis showed that *Isatis oblongata* had the highest total VOC content, characterized by esters and alcohols, whereas *Isatis tinctoria* and *Isatis indigotica* exhibited balanced aldehyde-terpenoid and ketone-enhanced profiles, respectively. The application of relative odor activity value (rOAV) modeling identified 44 key aroma-active compounds, defining the sensory profiles: *I. oblongata* as mushroom dominant (driven by the high rOAV of 1-octen-3-one), *I. tinctoria* as fatty-green-earthy (driven by C9 aldehydes), and *I. indigotica* as sweet-green with a spicy undertone (associated with benzene, (isothiocyanatomethyl)-). Integrated transcriptomic analysis and weighted gene co-expression network analysis (WGCNA) uncovered the molecular basis for this divergence. The fatty-green character of *I. tinctoria* was linked to a *NAC* transcription factor (Iin12486) that co-expressed with specific *lipoxygenase* (*LOX*) genes (Iin24854, Iin09829). Conversely, the mushroom aroma of *I. oblongata* was associated with a different *NAC* TF (Iin05720) co-expressed with a *hydroperoxide lyase* (*HPL*) gene (Iin21448). This multiomics study deciphers the coordinated transcriptional regulation underlying aroma divergence in *Isatis*, providing a scientific basis for chemotaxonomy, quality control, and the targeted utilization of aroma resources in these species.

Rong Chen, Yan Qing, and Xiaoshan Geng contributed equally to this paper.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2026 The Author(s). *Food Frontiers* published by John Wiley & Sons Australia, Ltd and Nanchang University, Northwest University, Jiangsu University, Zhejiang University, Fujian Agriculture and Forestry University.

1 | Introduction

The genus *Isatis* (Brassicaceae) comprises perennial herbs with notable historical and economic importance. Certain species, particularly *Isatis indigotica* (known as "Banlangen"), are valued in traditional medicine, with roots extensively studied for their phytochemical and pharmacological properties (Kang et al. 2020; Chang et al. 2024; M. Chen et al. 2012). In contrast, secondary metabolites from other plant parts, especially fruits, remain underexplored. This gap is particularly evident for volatile organic compounds (VOCs), which are crucial as aroma constituents and ecological mediators (Dadwal et al. 2025). A systematic analysis of fruit VOCs across *Isatis* species is therefore needed to understand their role in sensory characteristics and species-specific chemistry.

Plant VOCs are low-molecular-weight metabolites that contribute significantly to a plant's distinct aroma, a trait shaped by genetics, environment, and metabolism (Feng et al. 2024). Beyond their ecological roles (e.g., in pollinator attraction and defense), VOCs serve as chemical fingerprints for species identification and are of interest for flavor and fragrance applications (Fuentes-Lopez et al. 2025; Shukurova et al. 2020; Schwab et al. 2008). However, comparative VOC profiling within *Isatis* remains limited, with most studies focusing on single species or nonfruit tissues. Comprehensive characterization of VOC profiles, interspecific variation, and their formation mechanisms in *Isatis* fruits will provide a foundation for chemotaxonomy, quality assessment, and the utilization of aroma resources in this genus.

Despite the well-documented medicinal value of *Isatis*, research on its VOCs remains inadequate. Current studies are often limited to a few volatile components in a single species, lacking systematic interspecific comparisons within the genus. Most analyses remain qualitative, with limited integration of quantitative data and odor activity analysis (Condurso et al. 2006; Dupré et al. 2025). Moreover, the molecular regulatory mechanisms underlying VOC biosynthesis, particularly the pathways and regulatory networks of key aroma compounds, are still unclear. These knowledge gaps hinder the clarification of species-specific VOC patterns and the chemical and genetic basis of odor formation, thereby limiting their application in species identification, flavor marking, and high-value product development.

The formation of plant volatile odors results from the combined effects of genetic background and environmental factors (Loreto and Schnitzler 2010). In recent years, advances in gas chromatography–mass spectrometry (GC–MS) and efficient enrichment methods such as fully automated headspace solid-phase microextraction (HS-SPME) have enabled high-throughput and highly sensitive capture of the plant volatilome. Previous studies on medicinal plants such as *Phyllostachys edulis* and *Liquidambar* have established species-specific chemical profiles (L. Zhang et al. 2025; Zhao et al. 2025). However, research on *Isatis* has largely been confined to leaves or roots, focusing mainly on qualitative analysis of single species, with limited systematic comparison at the genus level. Furthermore, among the hundreds of VOCs detected instrumentally, only a subset significantly contributes to the overall scent. The relative odor activity value (rOAV) model, which relates compound concentration to human olfactory thresholds, has been successfully applied in flavor research on tea, fruits, and other products to identify key odor-active compounds (S. Zhang et al. 2023; Zhu et al. 2016). Thus, identifying these key components is essential for deciphering the sensory characteristics of *Isatis* plants.

At the molecular level, the biosynthesis of plant VOCs is tightly regulated at the transcriptional level, primarily through the terpenoid and fatty acid derivative pathways. In particular, C6 and C9 aldehydes and alcohols generated via the lipoxygenase pathway often contribute to key odor attributes such as "green" and "fatty" notes. The expression of structural genes in these pathways is precisely controlled by upstream transcription factors (TFs) (Picazo-Aragonés et al. 2020). In recent years, weighted gene co-expression network analysis (WGCNA) has emerged as a powerful bioinformatics tool for identifying gene modules highly correlated with target traits, such as metabolite accumulation, and for pinpointing hub TFs that play central regulatory roles (Eshaghi and Rashidi-Monfared 2025). This multiomics strategy, integrating metabolomics and transcriptomics, has been successfully applied to elucidate the formation of flavor in *Amomum tsao-ko* fruits and aroma in roses (Ma et al. 2024; Xiao et al. 2018). Nevertheless, the biosynthetic pathways and transcriptional regulation of key aroma compounds in *Isatis* have not yet been deciphered.

Based on this background, this study selects fruits from three representative *Isatis* species, including *I. indigotica*, *Isatis tinctoria*, and *Isatis oblongata*, as research materials. *Isatis indigotica*, the original species of "Banlangen," has been used for thousands of years in traditional medicine for its heat-clearing and detoxifying effects (Q. Chen et al. 2021). *Isatis tinctoria*, rich in indigo pigments, is widely used in the natural dye industry and shows promising potential for medicinal development (Rao et al. 2025). *Isatis oblongata*, a relative distributed in arid and semi-arid regions of northwestern China, remains underexplored in terms of its medicinal value and resource characteristics. This study aims to establish a comprehensive qualitative and quantitative profiling system for VOCs in these three species using headspace solid-phase microextraction coupled with gas chromatography–mass spectrometry (HS-SPME–GC–MS); identify key aroma-active compounds via rOAV; pinpoint characteristic odor contributors in each species; and reveal interspecific variation patterns and key genetic modules regulating VOC accumulation through differential metabolite analysis and WGCNA. Ultimately, this work will clarify the composition, odor variation mechanisms, and molecular regulatory basis of VOCs in these three *Isatis* species. The findings will fill a gap in comparative volatilome research within the genus, provide precise chemical markers for species identification, and offer theoretical and technical support for the targeted exploration and high-value utilization of VOC resources, such as in natural fragrance development and screening of medicinal scent-active components, while also establishing a reference methodology for systematic research on volatile metabolites in medicinal plants.

2 | Materials and Methods

2.1 | Plant Materials

To investigate the interspecific divergence in aroma profiles and transcriptional regulation, a common garden experiment was

established for three *Isatis* species: *I. oblongata* (Io), *I. tinctoria* (It), and *I. indigotica* (Ii). All plants were cultivated under identical edaphic and climatic conditions at the Experimental Base of Yili Normal University (YNU), located in Sulagong Village, Panjin Town, Yining City, Xinjiang, China (44°1′N, 81°23′E; altitude: 660 m). The region features a temperate continental climate. Standardized agronomic management, including irrigation and fertilization, was applied uniformly across all experimental plots to minimize environmental variance, ensuring that observed metabolic and transcriptomic differences were primarily driven by genetic factors. Voucher specimens for each species were authenticated and deposited at the Herbarium of YNU under accession numbers YNU-ISA-2025-001, YNU-ISA-2025-002, and YNU-ISA-2025-003, respectively.

Mature fruits were harvested in late May 2025, corresponding to the optimal ripening stage for flavor accumulation. A rigorous pooling strategy was employed to reduce individual biological noise: for each biological replicate ($n = 3$), mature fruits were collected from eight distinct, healthy individuals and combined into a single composite sample. Immediately upon harvest, the fresh fruit samples were flash-frozen in liquid nitrogen to arrest metabolic activities and prevent RNA degradation. Samples were subsequently stored at −80°C. For analysis, the frozen fresh tissues were ground into a fine powder in liquid nitrogen. This processed material was divided for parallel transcriptomic sequencing and volatile metabolome analysis (HS-SPME–GC–MS), ensuring direct correlation between gene expression and metabolite accumulation.

2.2 | Sample Extraction

The samples were ground into a fine powder using liquid nitrogen and homogenized via vortex mixing. Subsequently, approximately 500 mg of each sample was precisely weighed into a headspace vial. Then, 2 mL of saturated NaCl solution and 20 μL of an internal standard solution (10 μg/mL) were added to each vial. The VOCs were subsequently extracted using fully automated HS-SPME in preparation for gas chromatography–mass spectrometry (GC–MS) analysis.

2.3 | Chromatographic and Mass Spectrometric Acquisition Conditions

2.3.1 | HS-SPME Extraction

The extraction process was conducted under constant agitation at 60°C for 5 min. A 120-μm DVB/CWR/PDMS SPME Arrow was then inserted into the headspace of the vial for a 15-min extraction period. Following extraction, the fiber was desorbed in the GC inlet at 250°C for 5 min. Prior to initial use, new SPME Arrows were conditioned in the Fiber Conditioning Station for 2 h at 250°C; before each subsequent sampling, the fiber was reconditioned for 5 min at the same temperature. It is noteworthy that the SPME Arrow utilized offers a sensitivity approximately 10-fold greater than that of traditional SPME fibers.

2.3.2 | Gas Chromatography

Separation was performed on a DB-5MS capillary column (30 m × 0.25 mm × 0.25 μm; Agilent J&W Scientific, Folsom, CA, USA). High-purity helium (≥99.999%) served as the carrier gas at a constant flow rate of 1.2 mL/min. The injection port temperature was maintained at 250°C with a solvent delay of 3.5 min. The oven temperature program was as follows: an initial hold at 40°C for 3.5 min, ramp to 100°C at 10°C/min, then to 180°C at 7°C/min, and finally to 280°C at 25°C/min, followed by a final hold at 280°C for 5 min.

2.3.3 | Mass Spectrometry

The mass spectrometer was equipped with an electron ionization (EI) source operated at 70 eV. The temperatures of the ion source, quadrupole, and transfer line were set at 230°C, 150°C, and 280°C, respectively. Data acquisition was performed in selected ion monitoring (SIM) mode to ensure high sensitivity and selectivity for both qualitative and quantitative analysis.

2.4 | Qualitative and Quantitative Analysis of VOCs

The raw data files acquired from GC–MS were processed using MassHunter software. For accurate identification and quantification, a custom in-house database was established. This database incorporated information from multiple sources, including literature, authentic standards, and retention indices, and contained defined retention times (RTs) along with characteristic quantitative and qualitative ions for each compound. In the SIM method, one quantitative ion and two to three qualitative ions were selected for each metabolite. All target ions were monitored within specific time windows according to their elution order. A metabolite was positively identified if its detected RT matched the reference standard and all its designated ions were present in the sample mass spectrum after background subtraction (Yuan et al. 2022; Yuan et al. 2024). For quantification, the peak areas of the quantitative ions were integrated and calibrated, thereby enhancing the accuracy of the quantitative results.

2.5 | Analysis of Metabolite Relative Content

The internal standard semiquantitative method involves adding an appropriate chemical standard to the sample during quantitative analysis. The measured value of this internal standard serves as a reference for calculating the content of the target components. In quantitative analysis, to ensure that the procedure meets methodological requirements, a suitable compound is often selected as a reference and added to the sample at a known concentration. Both the target analytes and the reference compound are then detected simultaneously. According to the literature, isotopically labeled analogs are currently among the most commonly used internal standards. Therefore, an isotopic internal standard was selected in this study, and the relative content of VOCs in the samples was calculated based on the reference formula (D. Huang et al. 2024), as shown below:

$$X_i = (V_s \times C_s) \ / \ M \times PI_i \ / \ PI_s \times 10^{-3},$$

where X_i is the content of compound i in the sample (μg/g); V_s is the volume of the internal standard added (μL); C_s is the concentration of the internal standard (μg/mL); M is the mass of the sample (g); PI_s is the peak area of the internal standard; and PI_i is the peak area of compound i in the sample.

2.6 | Analysis of Metabolite rOAV

The rOAV is a method established based on the odor thresholds of compounds to identify key aroma contributors in food, illustrating the contribution of each volatile compound to the overall aroma profile. Generally, an rOAV of ≥1 indicates that the compound directly contributes to the sample's flavor. With reference to relevant literature (Fang et al. 2024; W. Huang et al. 2022; Xue et al. 2022), the rOAV was calculated as follows:

$$rOAV_i = C_I \ / \ T_I,$$

where $rOAV_i$ is the relative odor activity value of compound i, C_i is the relative content of the compound (μg/g), and T_i is its odor threshold (μg/g).

2.7 | Selection of Differential VOCs and KEGG Annotation and Enrichment Analysis

Differential VOCs were identified based on a variable importance in projection (VIP) value greater than 1 and an absolute $\log_2$ fold change ($|\text{Log}_2\text{FC}|$) of at least 1.0. The VIP values were derived from orthogonal projections to latent structures discriminant analysis (OPLS-DA), which was performed using the R package MetaboAnalystR. The OPLS-DA results also included score plots and permutation plots. Prior to OPLS-DA, the data were log-transformed (Log_2) and mean-centered. To avoid overfitting, a permutation test with 200 iterations was conducted.

The identified metabolites were annotated using the KEGG (Kyoto Encyclopedia of Genes and Genomes) compound database (http://www.kegg.jp/kegg/compound/). Subsequently, the annotated metabolites were mapped to the KEGG pathway database (http://www.kegg.jp/kegg/pathway.html) for pathway analysis.

2.8 | RNA-seq

Total RNA was extracted from samples using the RNAprep Pure Plant Kit (Tiangen, Beijing, China) according to the manufacturer's instructions. RNA concentration and purity were determined using a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, Wilmington, DE, USA). RNA integrity was subsequently assessed with the RNA Nano 6000 Assay Kit on an Agilent Bioanalyzer 2100 system (Agilent Technologies, CA, USA).

For library construction, 1 μg of total RNA per sample was used as input. Sequencing libraries were prepared following the manufacturer's guidelines using the Hieff NGS Ultima Dual-mode mRNA Library Prep Kit for Illumina (Yeasen Biotechnology (Shanghai) Co., Ltd.), with index codes incorporated to assign sequences to respective samples. Specifically, mRNA was isolated from total RNA using poly-T oligo-attached magnetic beads. First-strand cDNA synthesis was conducted, followed by second-strand synthesis. The resulting DNA fragments were blunt-ended through exonuclease/polymerase treatment. After adenylating the 3′ ends, NEBNext Adaptors with hairpin loop structures were ligated. The library fragments were then purified using the AMPure XP system (Beckman Coulter, Beverly, USA). Subsequently, size-selected and adaptor-ligated cDNA was treated with 3 μL of USER Enzyme (NEB, USA) at 37°C for 15 min, followed by incubation at 95°C for 5 min prior to PCR amplification. PCR was performed using Phusion High-Fidelity DNA polymerase, universal PCR primers, and an index (X) primer. Finally, the PCR products were purified (AMPure XP system), and library quality was verified using the Agilent Bioanalyzer 2100 system. The prepared libraries were sequenced on an Illumina NovaSeq platform to generate 150-bp paired-end reads, in accordance with the manufacturer's protocol.

Gene functions were annotated by aligning sequences against the following databases: Nr (NCBI nonredundant protein sequences), Pfam (Protein Family), KOG/COG (Clusters of Orthologous Groups of proteins), Swiss-Prot (a manually annotated and reviewed protein sequence database), KO (KEGG Ortholog database), and GO (Gene Ontology). Differential expression analysis was performed using DESeq2, which applies a negative binomial distribution-based model to statistically evaluate differences in gene expression levels. The resulting p-values were adjusted using the Benjamini–Hochberg procedure to control the false discovery rate. Genes satisfying the criteria of an adjusted p-value <0.01 and a fold change ≥2 were identified as differentially expressed. Additionally, TFs were predicted by performing domain searches against the TFDB database using hmmsearch from the HMMER package.

2.9 | Quantitative PCR Analysis

Total RNA was extracted using a cetyltrimethylammonium bromide (CTAB)-based method, and its concentration and purity were assessed with a Nanodrop spectrophotometer. Subsequently, first-strand cDNA was synthesized from the total RNA using the PrimeScript RT reagent kit, which incorporates a gDNA Eraser step to remove potential genomic DNA contamination. Quantitative real-time PCR (qPCR) was then conducted on a CFX96 Real-Time System using SYBR Green dye. The thermal cycling protocol comprised an initial pre-denaturation at 95°C for 30 s, followed by 40 cycles of denaturation at 95°C for 5 s and annealing/extension at 60°C for 30 s. Each target gene was analyzed in triplicate to ensure data reliability, and relative expression levels were calculated using the $2^{-\Delta\Delta Ct}$ method. All qPCR primers were designed with Primer Premier 5.0 and are provided in Table S1.

2.10 | Data Analysis

The experimental data were processed using Microsoft Office 2021 (Microsoft Corporation, Redmond, WA, USA). Histograms were generated using Origin 2021 (OriginLab Corporation, Northamp-

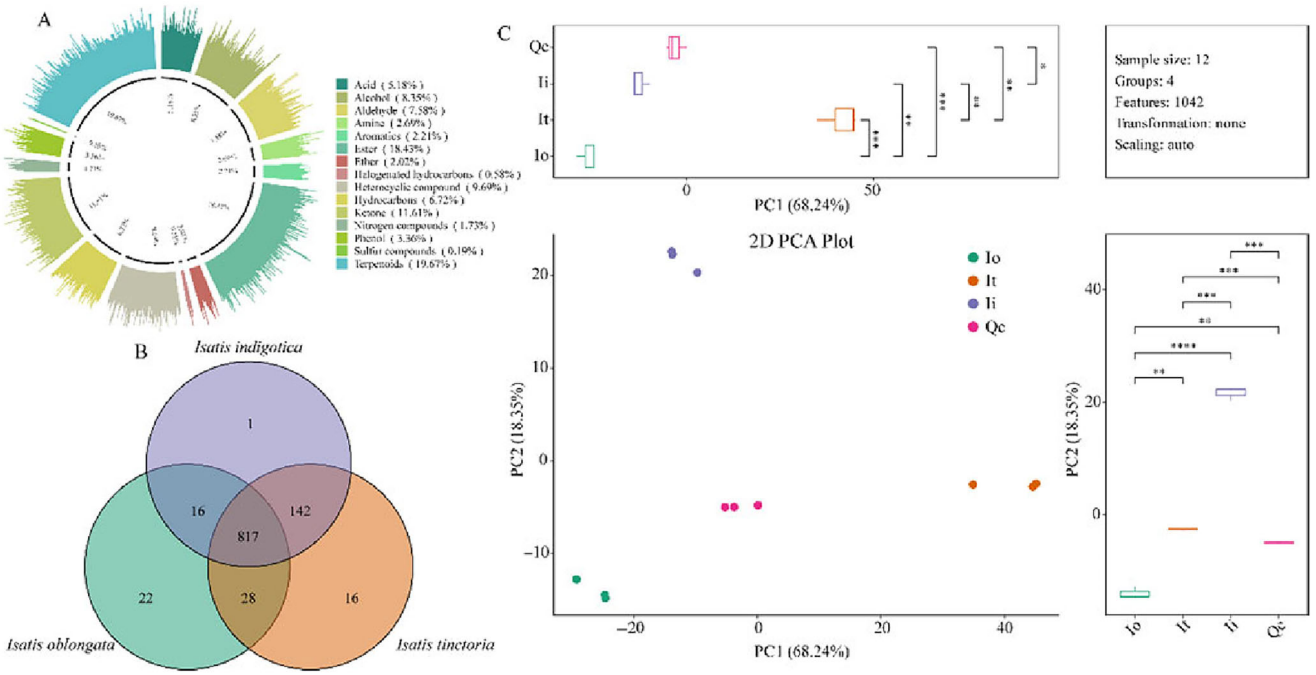


FIGURE 1 | Global profiling of volatile organic compounds (VOCs) in the fruit of three *Isatis* species. (A) Classification analysis of identified VOCs. (B) Venn analysis of VOCs. (C) Scatter plot of the principal component analysis (PCA) based on the identified VOCs. QC, quality control; Ii, *I. indigotica*; It, *I. tinctoria*; Io, *I. oblongata*. Asterisks represent statistical significance: $*p < 0.05$, $**p < 0.01$, $***p < 0.001$, and $****p < 0.0001$.

ton, MA, USA). One-way analysis of variance was conducted using Duncan's method in IBM SPSS Statistics 25. Classification pie charts, principal component analysis (PCA), and Venn diagrams were produced using R version 3.5.1. Advanced flavor radar charts were created with R version 4.1.2. Furthermore, heatmap visualization and *K*-means clustering analysis were performed using R version 4.2.0. Additionally, bubble plots were constructed with R version 4.2.2. Correlation networks among key aroma compounds, structural genes, and regulatory genes were analyzed based on the Pearson method and visualized using Cytoscape 3.9.1. Moreover, WGCNA was conducted with R version 4.2.2, where the minimum module size was set to 50 genes, the merge cut height was set to 0.25, and other parameters were kept as default.

3 | Results

3.1 | Comprehensive Profiling of VOCs

A total of 1042 VOCs were identified from the fruits of the three *Isatis* species (*I. oblongata*, *I. tinctoria*, and *I. indigotica*) using GC–MS analysis. These metabolites were categorized into 15 distinct classes (Figure 1A). Terpenoids constituted the most abundant class (19.67%), followed by esters (18.43%) and ketones (11.61%), indicating their significant role in the metabolic background of *Isatis* fruits.

A comparative analysis of the three species revealed both shared and distinct metabolic profiles. A total of 883, 1003, and 976 metabolites were identified in *I. oblongata*, *I. tinctoria*, and *I. indigotica*, respectively (Figure 1B). Venn diagram analysis highlighted a substantial common metabolic foundation, with 817 metabolites shared among all three species (Figure 1B), which likely represents the core VOCs of the *Isatis* genus. However, each species also exhibited unique characteristics: *I. oblongata* possessed the highest number of unique metabolites (22), including several terpenoids such as (+)-4-carene and (+)-3-carene, as well as compounds like creosol and 2-heptenoic acid. This suggests a distinct secondary metabolic pathway in this species. *Isatis tinctoria* contained 16 unique metabolites, characterized by compounds such as the 2,4-nonadienal, (*E*,*E*)- and 1-nonen-3-one, which may contribute to its specific aromatic profile. In stark contrast, *I. indigotica* had only a single unique metabolite, identified as panaxene. This indicates that while its overall metabolic profile is rich, its uniqueness lies not in the number of unique compounds but potentially in the relative abundance of shared metabolites. The pairwise shared metabolites further illuminate their phylogenetic relationships. The fact that *I. tinctoria* and *I. indigotica* shared a considerably higher number of metabolites (142) than either did with *I. oblongata* (28 and 16, respectively) suggests a closer metabolic relationship between *I. tinctoria* and *I. indigotica*.

PCA was employed to visualize the overall metabolic differences among the three species. The analysis showed a clear and distinct separation between the groups along the first two principal components (Figure 1C). PC1 and PC2 accounted for 68.24% and 18.35% of the total variance, respectively, cumulatively explaining 86.59% of the metabolic variation, which robustly represents the dataset's structure. Notably, the three biological replicates for each species clustered tightly together, indicating high reproducibility of the analytical results and the consistency of the metabolic profiles within each species.

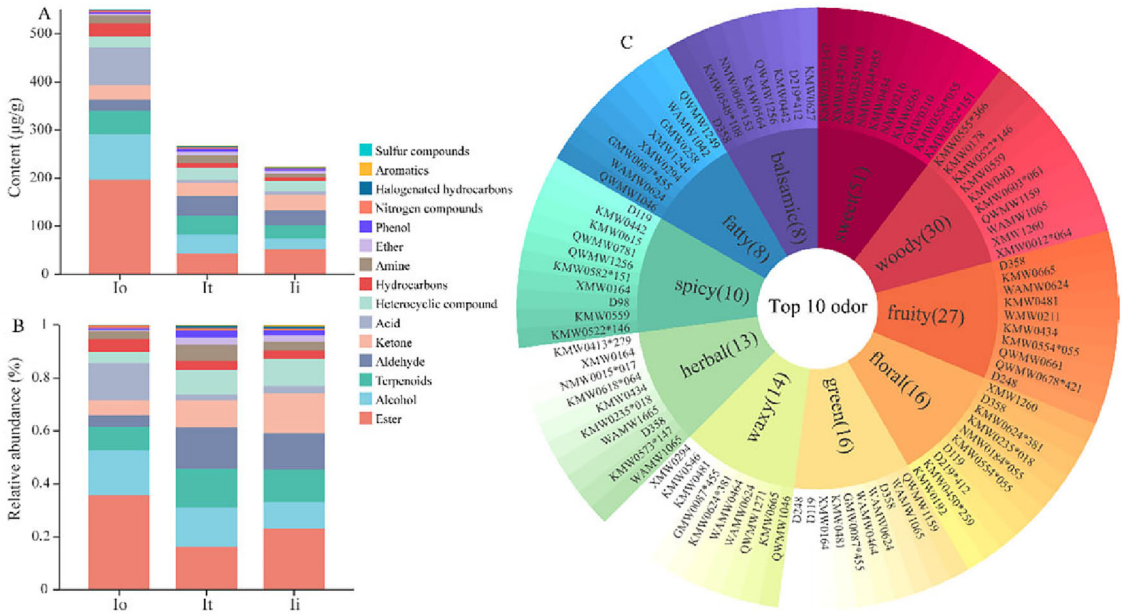


FIGURE 2 | Volatile organic compound (VOC) content and aroma profiles of the fruits from three *Isatis* species. (A) Quantitative VOC content. (B) Comparative relative abundance. (C) Flavor sunburst diagram of the predominant odor descriptors. Ii, *I. indigotica*; It, *I. tinctoria*; Io, *I. oblongata*.

The score plot (Figure 1C) demonstrates that *I. oblongata* (Io) samples were located in the negative quadrant of both PC1 and PC2, clearly separated from the others. *Isatis tinctoria* (It) samples clustered in the positive region of PC1 and the negative region of PC2. *Isatis indigotica* (Ii) samples were distinctly positioned in the positive region of PC2, separating them from both *I. oblongata* and *I. tinctoria*. The quality control (QC) samples formed a tight cluster, validating the stability and reliability of the analytical sequence. The statistical significance of the separations observed in the PCA score plot was confirmed by *T*-tests. For both PC1 and PC2, all interspecies comparisons and comparisons with QC samples were also statistically significant ($p < 0.05$). Notably, the separation between *I. oblongata* (Io) and *I. indigotica* (Ii) along PC2 was extremely significant ($p < 3.3 \times 10^{-6}$). These statistical results unequivocally demonstrate that the three *Isatis* species possess significantly distinct VOC profiles, corroborating the findings from the Venn analysis.

3.2 | Quantitative and Odor of VOCs

The relative concentrations of all 1042 identified VOCs were quantified to elucidate the distinct chemical landscapes of the three *Isatis* fruits. The total VOC content varied substantially among the species. *Isatis oblongata* (Io) exhibited the most abundant total VOC, with an average concentration of 551.03 μg/g, which was approximately 2.1 and 2.4 times higher than that of *I. tinctoria* (It, 266.57 μg/g) and *I. indigotica* (Ii, 223.67 μg/g), respectively (Figure 2A). This indicates a fundamentally higher level of volatile production in *I. oblongata*.

A more detailed analysis of the compositional profiles revealed pronounced species-specific patterns, which were driven by several key marker compounds. In *I. oblongata* (Io), esters were the most dominant class, constituting on average 35.77% of the total VOC profile (197.14 μg/g). This was largely attributed to exceptionally high levels of 2-propenoic acid, pentyl ester (up to 134.04 μg/g), and propanoic acid, 1,1-dimethylethyl ester (up to 18.89 μg/g), which contributed significantly to the overall ester abundance. This was followed by alcohols (14.15%, 93.65 μg/g) (Figure 2B), with 2-octanol, 2-octanol, (*S*)-, and 1-octen-3-ol being the major contributors, and acids (14.15%, 77.97 μg/g). The combination of high ester and acid content suggests a potentially strong fruity and acidic aroma note in Io fruits.

The VOC profile of *I. tinctoria* (It) was markedly different. While esters remained a significant class (16.22%, 43.34 μg/g), aldehydes (15.77%, 42.06 μg/g) and terpenoids (14.66%, 39.13 μg/g) contributed almost equally to the overall profile. Notably, the relative abundance of aldehydes in It was about 1.8 times higher than that in Io. This shift was characterized by substantial concentrations of 2,4-heptadienal, (*E*,*E*)- and benzaldehyde, compounds often associated with green, fatty, or almond-like odors. In the terpenoid fraction, Neral was a constituent of notable abundance. This unique blend likely defines a key aromatic characteristic differentiating It from Io.

Isatis indigotica (Ii) displayed a unique profile where ester was the most prominent class, representing 23.06% (51.64 μg/g) of its total VOCs, followed by ketone (15.05%, 33.62 μg/g), a proportion significantly higher than that in Io and It. This ketone-enhanced character was defined by high concentrations of 3(2H)-furanone, dihydro-2-methyl-, which can impart sweet, solvent, bread, buttery, and nutty-like notes.

In summary, the quantitative VOC analysis clearly demarcates the three *Isatis* species: *I. oblongata* is an "ester-alcohol"-rich type dominated by 2-propenoic acid, pentyl ester; *I. tinctoria* is a balanced aldehyde-terpenoid type characterized by (*E*,*E*)-2,4-heptadienal and benzaldehyde; and *I. indigotica* is a ketone-enhanced type marked by 3(2H)-furanone, dihydro-2-methyl-. These fundamental differences in both the total volatile

output and the specific marker compounds provide a quantitative basis for their unique fragrance profiles and set the stage for identifying key aroma-active compounds in subsequent analyses.

Based on the odor database annotations, we conducted a systematic odor analysis of the VOCs from three experimental materials to elucidate their overall aroma profiles. A total of 556 VOCs with definitive odor descriptions were detected, corresponding to 372 distinct odor descriptors. Through flavor sunburst diagram screening, the top 10 odor types identified were sweet (51 VOCs), woody (30 VOCs), fruity (27 VOCs), floral (16 VOCs), green (16 VOCs), waxy (14 VOCs), herbal (13 VOCs), spicy (10 VOCs), fatty (eight VOCs), and balsamic (8 VOCs). Together, these odor types constitute the complex and diverse aroma profiles of the three materials (Figure 2C).

Specifically, all three materials exhibited dominant aroma characteristics of sweet, woody, and fruity notes. The sweet odor was contributed by various compounds such as ionone (characterized by violet, sweet, floral, and woody notes), benzaldehyde, 3,4-dimethoxy- (sweet, woody, vanilla), and (*E*)-β-farnesene (woody, citrus, herbal, sweet), which collectively enhanced the overall sweet fragrance. The woody odor primarily originated from terpenoids, including trans-α-bergamotene (woody, warm, tea) and copaene (woody, spicy, honey), providing a solid woody foundation to the aroma. The fruity odor was represented by esters and terpenes such as 4-tert-butylcyclohexyl acetate (woody, cedary, floral, oily, herbal, balsamic, green, fruity) and α-ionone (sweet, woody, floral, violet, orris, tropical, fruity), imparting a fresh fruity sensation.

Furthermore, floral, green, and herbal odors added further complexity to the aroma profile. The floral odor was associated with compounds such as linalool (floral, green) and benzyl alcohol (floral, rose, phenol, balsamic), contributing floral nuances. The green odor was derived from compounds including 2-hexenal (sweet, almond, fruity, green, leafy, apple, plum, vegetable) and 3-hexen-1-ol (green, leafy), evoking fresh grassy and leafy impressions. The herbal odor was linked to compounds such as cubenol (spicy, herbal, green, tea) and thymol (herbal, thyme, phenol, medicinal, camphor), introducing herbal and spicy complexity. Additionally, waxy, spicy, fatty, and balsamic odors served as secondary components that further enhanced the depth and persistence of the aroma. For instance, waxy odor originated from *n*-butyl laurate (waxy, soapy, fruity, peanut), spicy odor from caryophyllene (sweet, woody, spice, clove, dry), fatty odor from 2-decenal (waxy, fatty, earthy, green, cilantro, mushroom, aldehydic, fried, chicken, tallow), and balsamic odor from 1H-cycloprop[e]azulene (woody, balsamic).

In summary, all three experimental materials (Ii, It, and Io) exhibited complex yet harmonious overall aroma profiles. This combination of odors may reflect conserved biosynthetic pathways in their metabolite composition and provides a theoretical basis for their potential applications, such as in perfumery or medicinal uses. Future studies may focus on quantifying differences in odor intensity among the materials to uncover more detailed species-specific characteristics.

3.3 | Identification and Analysis of Differential VOCs

Based on our established criteria for differential metabolites, a total of 838 VOCs were identified as significantly different across the three *Isatis* materials. To visualize their distinct accumulation patterns, a heatmap was constructed (Figure 3A), revealing clear metabolic divergence among *I. oblongata* (Io), *I. tinctoria* (It), and *I. indigotica* (Ii).

To further elucidate the accumulation dynamics of these differential VOCs, a *K*-means clustering analysis was performed, which classified the 838 compounds into six distinct subclasses (1–6) (Figure 3B). This clustering effectively delineated the species-specific and shared VOC profiles. Subclass 1 comprised VOCs that were nearly undetectable in Io but showed significant accumulation in both It and Ii. Notably, the concentration of these compounds was consistently and substantially higher in Ii compared to It. This group was dominated by terpenoids, such as di-epi-α-cedrene and γ-muurolene, and phenols, including 2-methoxy-4-vinylphenol. Key aldehydes like 2,5-furandicarboxaldehyde and esters were also prominent in this cluster. Subclass 2 included metabolites that were present at high levels in Io, reached their peak accumulation in It, and were found at intermediate levels in Ii. This group featured alcohols like (*E*)-2,6-dimethylocta-3,7-diene-2,6-diol, nitrogen-containing compounds such as hydrazinecarbothioamide, and ethers like (butoxymethyl)-benzene. Subclass 3 was the largest cluster and was characterized by VOCs with the most pronounced accumulation in It and the lowest levels in Ii and Io. This group encompassed a vast diversity of chemical classes, including a significant number of alcohols (e.g., phenylethyl alcohol, benzenemethanol), aldehydes (e.g., 2-propenal, 3-(2-furanyl)-, benzeneacetaldehyde), phenols, terpenoids (e.g., neral, carvone), ketones, heterocyclic compounds, and esters. Subclass 4 exhibited a unique pattern where VOCs accumulated to moderate and comparable levels in both It and Ii, while remaining low in Io. This group contained various aldehydes (e.g., benzaldehyde), acids, terpenoids (e.g., β-pinene), and ketones. Subclass 5 consisted of compounds that were highly abundant in Io and Ii but were largely absent in It. This cluster was rich in ketones such as 3(2H)-furanone, dihydro-2-methyl-, ethers, and esters. Subclass 6 represented VOCs that were predominantly and uniquely accumulated in Io, with very low or undetectable levels in both It and Ii. This group was characterized by high levels of specific alcohols (e.g., 1-octen-3-ol), acids (e.g., 2-heptenoic acid, hexanoic acid, 2-ethyl-), terpenoids (e.g., β-pinene), and amines.

In summary, the *K*-means analysis successfully categorized the complex VOC profiles into coherent classes, highlighting that It and Ii share a closer relationship in their volatile metabolism, particularly for terpenoids and phenols in Subclass 1. Conversely, Io possesses a more distinct VOC signature, enriched with a unique set of acids, alcohols, and terpenoids (Subclasses 5 and 6) that clearly differentiate it from the other two species.

To further dissect the metabolic disparities between the three *Isatis* materials, pairwise comparisons were conducted (Figure 3C). The Io_vs_It comparison yielded the highest

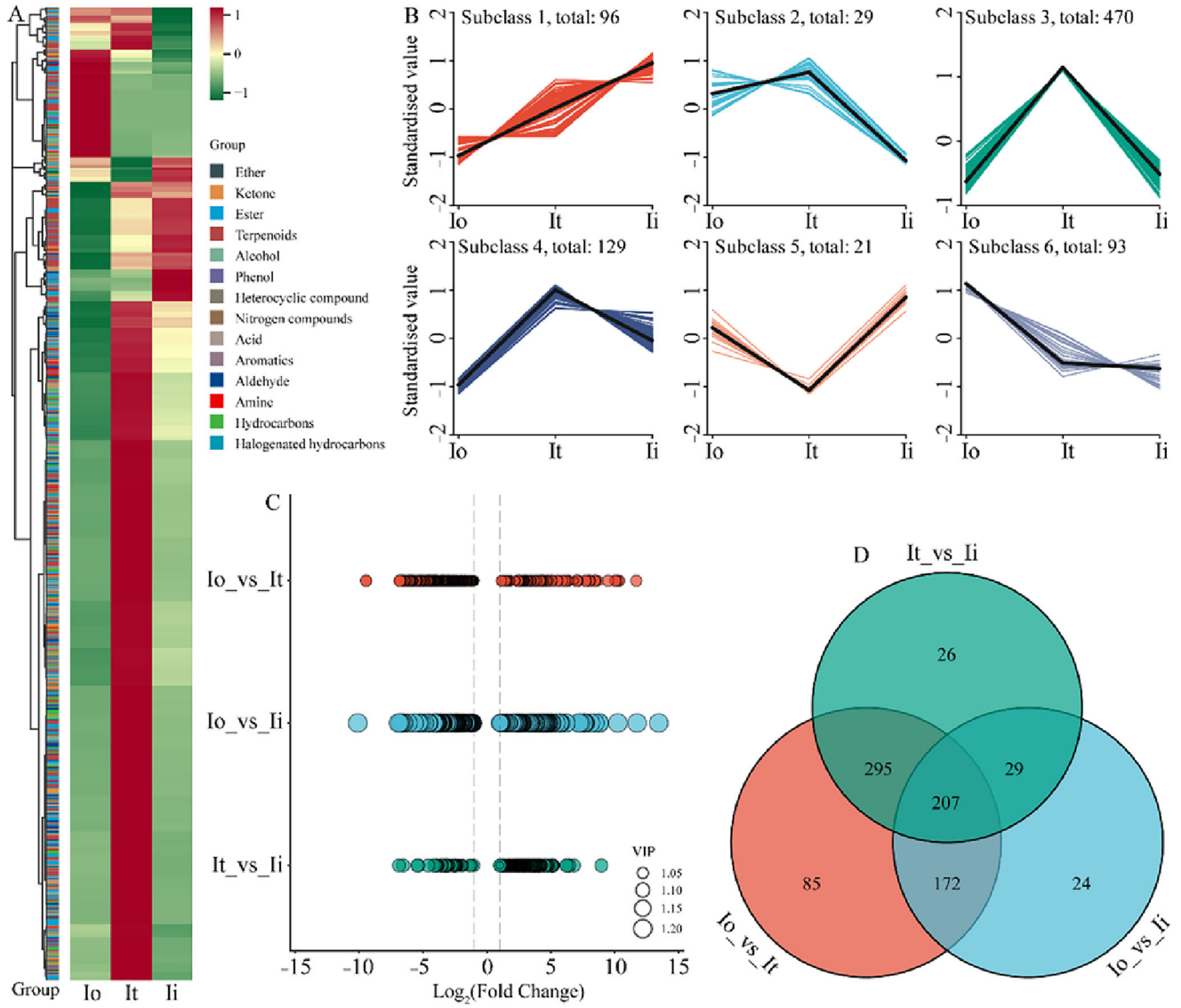


FIGURE 3 | Analysis of differential VOCs. (A) Heatmap of differential VOCs. (B) *K*-means clustering analysis of differential VOCs. (C) Bubble plot of differential VOCs across different comparison groups. (D) Venn diagram of differential VOCs from different comparison groups. Ii, *I. indigotica*; It, *I. tinctoria*; Io, *I. oblongata*.

number of differential VOCs (759), with 654 (86.17%) being down-accumulated in Io relative to It. Similarly, in the Io_vs_Ii comparison, 432 differential VOCs were identified, of which 302 (69.91%) were lower in Io. In stark contrast, the It_vs_Ii comparison revealed 557 differential VOCs, with an overwhelming majority (501 VOCs, 89.95%) being up-accumulated in It compared to Ii. This pattern strongly indicates that Io possesses a generally less abundant VOC profile compared to both It and Ii, while It and Ii, though closely related, exhibit significant quantitative differences, with It showing higher levels for most differential volatiles.

A Venn diagram was constructed to identify VOCs that are consistently differential across all pairwise comparisons (Figure 3D). This analysis revealed a core set of 207 VOCs that were universally discriminatory, underscoring their fundamental role in defining the metabolic identities of the three species. This common set was enriched in terpenoids (e.g., D-verbenone, β-pinene, α-ionone), aldehydes (e.g., various decadienals and nonenals), heterocyclic compounds (particularly pyrazines and pyrroles), esters, and ketones. Beyond the common core, each pairwise comparison featured a unique set of differential VOCs, highlighting specific metabolic differentiations. The Io_vs_It comparison possessed the largest number of unique differential VOCs (85). These were predominantly esters (e.g., formic acid, octyl ester; acetic acid, decyl ester) and terpenoids (e.g., trans-rose oxide, cubenol), suggesting a distinct divergence in flavor and fragrance-related pathways between these two species. The Io_vs_Ii comparison featured 24 unique VOCs. This group included several phenols and unique terpenoids, indicating specific metabolic pathways that differentiate Io from Ii, independent of It. The It_vs_Ii comparison contained 26 unique VOCs. This set was characterized by a mix of nitrogen-containing compounds (e.g., hydrazinecarbothioamide), alcohols, and specific terpenoids (e.g., D-limonene), pointing to metabolic specializations that distinguish the two more closely related species, It and Ii.

In summary, the pairwise analyses not only confirm the overall metabolic deficiency of Io but also delineate a hierarchy of metabolic relationships. The large core of 207 common VOCs forms the baseline for species discrimination. The unique VOC sets, particularly the 85 in Io_vs_It, further emphasize the pronounced metabolic distance between Io and the other two species, while the smaller unique sets in the Io_vs_Ii and It_vs_Ii comparisons refine the specific metabolic signatures that distinguish each species pair.

3.4 | Identification of Key Aroma-Active Compounds Based on rOAV

To identify volatile compounds that significantly contribute to the overall aroma profile of *Isatis* species, the rOAV was calculated for each of the 838 differential VOCs. The rOAV serves as a crucial index for evaluating the contribution of a compound to the overall odor, as it relates the concentration of the compound to its odor threshold. Among the analyzed VOCs, 262 were identified as aroma-active, meaning they possessed both a reported odor description and a calculable rOAV.

A preliminary screening based on the principle that compounds with rOAV >1 contribute to odor identified 104 such VOCs. To pinpoint the key aroma drivers, a more stringent criterion (rOAV > 10) was applied, which is widely accepted to indicate a substantial contribution. This refined screening revealed 44 key aroma-active compounds with rOAVs exceeding 10 across all three *Isatis* species (Table S2), confirming their pivotal role in shaping the characteristic scents.

The compounds belong to diverse chemical classes, including ketones (seven compounds), aldehydes (12 compounds), esters (six compounds), alcohols (six compounds), terpenoids (five compounds), phenols (three compounds), heterocyclic compounds (four compounds), and one nitrogen compound.

Notably, several compounds exhibited exceptionally high rOAVs (>10,000), identifying them as the predominant aroma contributors. These included 1-octen-3-one (rOAVs up to 88,673 in Io), imparting a potent mushroom note. 3(2H)-Furanone, dihydro-2-methyl- (rOAVs up to 2,355,345 in Ii) and 2(5H)-furanone, 5-ethyl-3-hydroxy-4-methyl- (rOAVs up to 525,005 in It), both were characterized by intense sweet, caramel, and nutty aromas.

A comparative analysis of the rOAV profiles across the three species revealed distinct and characteristic aroma patterns, allowing for their differentiation based on key odorants. *Isatis oblongata* was characterized as a mushroom-dominant type, primarily due to the overwhelmingly high rOAV of 1-octen-3-one (88673) in this species, which imparted a potent mushroom note that defined its overall aroma profile. In contrast, *I. tinctoria* presented a complex fatty-green-earthy character, distinguished by notably high levels of multiple C9 aldehydes and ketones such as (*Z*,*Z*)-3,6-nonadienal (rOAV = 2767), 2-nonenal, (*E*)- (rOAV = 2471), and 2,4-nonadienal (rOAV = 5228), which collectively contributed to intense fatty, cucumber-like, and earthy nuances. Meanwhile, *I. indigotica* exhibited a more balanced profile with prominent sweet and green notes, yet it was uniquely marked by a significantly higher rOAV of benzene, (isothiocyanatomethyl)- (rOAV = 1433), suggesting a potential spicy, medicinal undertone characteristic of this species.

In conclusion, the rOAV-directed screening successfully identified 44 VOCs as the key differential aroma-active compounds in the three *Isatis* species. The combination of these compounds, particularly those with exceedingly high rOAV values, forms the fundamental basis for their unique and complex aroma profiles. More importantly, the distinct combination and abundance of specific key odorants, such as mushroom alcohols/ketones in *I. oblongata*, fatty-green aldehydes in *I. tinctoria*, and a specific isothiocyanate in *I. indigotica*, provide a chemical basis for their sensory differentiation.

3.5 | WGCNA of Regulatory Genes and Key VOCs

Based on the expression data of regulatory genes (including TFs, transcriptional regulators [TRs], and protein kinases [PKs]) and the accumulation levels of seven key aroma compounds, we performed a WGCNA to decipher the transcriptional regulatory network underlying the distinct aroma profiles. A total of 2807 TFs, TRs, and PKs were predicted (Figure 4), with the top five most abundant families being *MYB* (146), *AP2/ERF-ERF* (101), *bHLH* (99), *RLK-Pelle_DLSV* (85), and *NAC* (74). Subsequently, WGCNA was carried out to identify TFs, TRs, and PKs strongly associated with the accumulation of key odor-related VOCs (Figure 4A). The 2807 regulatory genes were grouped into six distinct modules. Notably, the turquoise module contained the highest number of genes (919), followed by the blue (749), brown (376), and yellow (282) modules. In contrast, the green and gray modules each contained fewer than 100 regulatory genes.

Correlation analysis revealed that the expression levels of the 749 regulatory genes in the blue module exhibited a highly significant correlation ($p < 0.001$) with the accumulation of WAMW1895*260 (2,4-nonadienal), KMW0338*263 (2-nonenal, (*E*)-), KMW0287 ((*Z*,*Z*)-3,6-Nonadienal), and KMW0383 (2(5H)-furanone, 5-ethyl-3-hydroxy-4-methyl-). The 282 regulatory genes in the yellow module exhibited a highly significant correlation ($p < 0.01$) with the accumulation of KMW0170 (1-octen-3-one). The 919 regulatory genes in the turquoise module exhibited a highly significant correlation ($p < 0.01$) with the accumulation of w28 (3(2H)-furanone, dihydro-2-methyl-) and NMW0156 (benzene, (isothiocyanatomethyl)-).

Subsequently, five hub genes were identified within each module based on intramodular connectivity. A heatmap (Figure 4B) visualizing the expression patterns of these 15 hub genes revealed clear species-specific expression trends that aligned perfectly with the accumulation of their correlated aroma compounds.

3.5.1 | Hub Genes in the Blue Module

All five hub genes (e.g., Iin12486 [NAC TF], Iin17976 [PK], Iin28464 [PK]) displayed a conserved expression pattern characterized by exceptionally high expression in *I. tinctoria* and low expression in *I. oblongata* and *I. indigotica*. This pattern is consistent with the uniquely high abundance of fatty-green

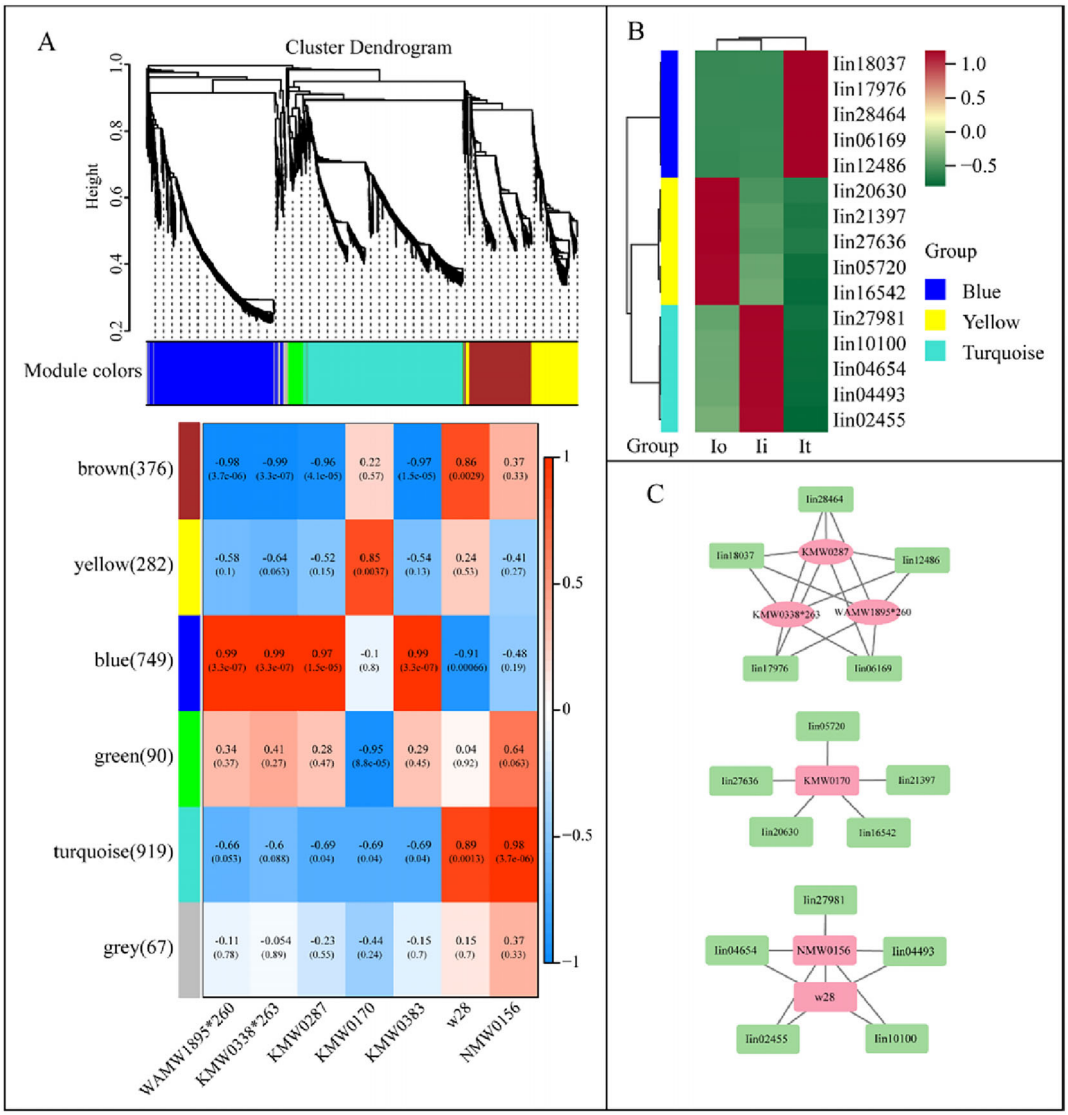


FIGURE 4 | Identification of co-expression modules and hub regulators associated with species-specific aroma profiles in *Isatis*. (A) Module–trait associations and key volatile organic compounds (VOCs). (B) Species-specific expression patterns of hub genes in key WGCNA modules. (C) Integrated regulatory network of hub genes and correlated aroma compounds. Green color represented regulatory genes, and pink color represented key aroma compounds. Ii, *I. indigotica*; It, *I. tinctoria*; Io, *I. oblongata*.

aldehydes (2,4-nonadienal, 2-nonenal, (*E*)-) that define the aroma of *I. tinctoria*.

3.5.2 | Hub Genes in the Yellow Module

The hub genes (including Iin05720 [NAC TF] and Iin21397 [FAR1 TF]) showed the highest expression in *I. oblongata*, moderate expression in *I. indigotica*, and negligible expression in *I. tinctoria*. This pattern directly corresponds with the dominant accumulation of 1-octen-3-one in *I. oblongata*, which is a key determinant of its mushroom-dominant aroma.

3.5.3 | Hub Genes in the Turquoise Module

The hub genes in this module (e.g., Iin04654 [EIL TF], Iin27981 [HSF TF], Iin02455 [C3H TF], Iin04493 [bHLH TF]) exhibited a contrasting profile, with the strongest expression in *I. indigotica*, very low expression in *I. tinctoria*, and intermediate levels in *I. oblongata*. This aligns with the high accumulation of both 3(2H)-furanone, dihydro-2-methyl- and, most notably, benzene, (isothiocyanatomethyl)- in *I. indigotica*, supporting its balanced sweet-green profile with a unique spicy undertone.

To further validate the regulatory relationships, a correlation network was constructed between the hub genes and their associated aroma compounds (Figure 4C). The network confirmed strong positive correlations (correlation coefficients ranging from 0.8 to 1.0) between each hub gene and its linked metabolites. Specifically, all five blue module hub genes showed near-perfect positive correlations (≥0.97) with the fatty-green aldehydes. Similarly, the turquoise module hub genes were strongly correlated with w28 (0.87–0.90) and NMW0156 (0.97–0.99), and the yellow module hub genes were robustly linked to KMW0170 (0.80–0.87).

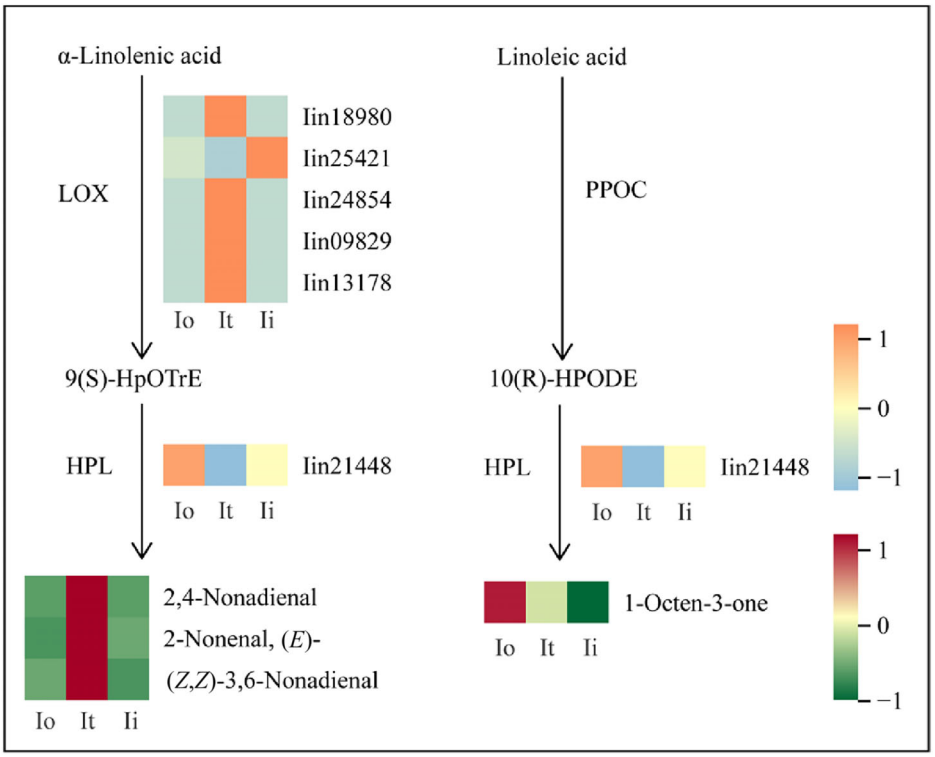


FIGURE 5 | Proposed biosynthetic pathway of key aroma compounds in *Isatis* species. Ii, *I. indigotica*; It, *I. tinctoria*; Io, *I. oblongata*.

In summary, the WGCNA successfully delineated three co-expression modules whose distinct spatiotemporal expression patterns strongly underpin the species-specific accumulation of the key aroma compounds. The identified hub genes, particularly the bHLH, NAC, EIL, and HSF TFs, represent high-priority candidates for being the core regulators responsible for the divergent aroma profiles of the three *Isatis* species.

3.6 | Biosynthetic Pathways of Key Aroma Compounds

To elucidate the biosynthetic origin of the key aroma compounds, we reconstructed the metabolic pathways leading to four critical volatiles: the C9 aldehydes (*Z*,*Z*)-3,6-nonadienal, 2-nonenal (*E*)-, and 2,4-nonadienal, and the C8 ketone 1-octen-3-one. Their proposed formation from α-linolenic and linoleic acid precursors via the lipoxygenase (LOX) pathway is depicted in Figure 5.

The accumulation patterns of these volatiles were highly species specific and aligned with the expression of key biosynthetic genes. *Isatis tinctoria* (It) accumulated the highest levels of the unsaturated C9 aldehydes, with 2,4-nonadienal and 2-nonenal (*E*)- being almost exclusively detected in this species. This remarkable abundance was strongly associated with the exceptionally high expression of two *LOX* genes, *Iin24854* and *Iin09829*, in *I. tinctoria*. This coordinated upregulation of *LOX* genes provides a clear genetic basis for the dominant fatty-green character of *I. tinctoria*.

In contrast, 1-octen-3-one, a potent mushroom-smelling ketone, was most abundant in *I. oblongata* (Io). This was consistent with the highest expression of the *hydroperoxide lyase* gene *Iin21448* (*HPL*) in this species, which is hypothesized to catalyze the formation of this volatile from 10(R)-HPODE. The strong correlation between the expression of *Iin21448* and the accumulation of 1-octen-3-one positions it as a key candidate gene responsible for the characteristic mushroom aroma of Io. The specific regulatory mechanisms driving the high co-expression of *Iin21448* and 1-Octen-3-one in Io, however, warrant further investigation.

The third species, *I. indigotica* (Ii), exhibited an intermediate profile for most compounds, consistent with its more balanced aroma signature. It showed substantial accumulation of (*Z*,*Z*)-3,6-nonadienal and moderate levels of 1-octen-3-one and 2-nonenal (*E*)-, which correlated with the intermediate expression levels of several *LOX* genes (e.g., *Iin25421*) and the *HPL* gene across the three species.

In conclusion, the divergent accumulation of key fatty acid–derived aroma compounds among the three *Isatis* species is underpinned by distinct expression patterns of genes in the LOX pathway. The specific upregulation of a subset of *LOX* genes in *I. tinctoria* is a key factor driving the abundant production of C9 aldehydes that define its unique aroma profile.

3.7 | An Integrated Model for the Transcriptional Regulation of Species-Specific Aroma Formation

Based on the comprehensive co-expression network analysis integrating key aroma compounds, structural genes, and TFs, we successfully delineated the distinct transcriptional regulatory modules underlying the biosynthesis of species-specific volatiles (Figure 6).

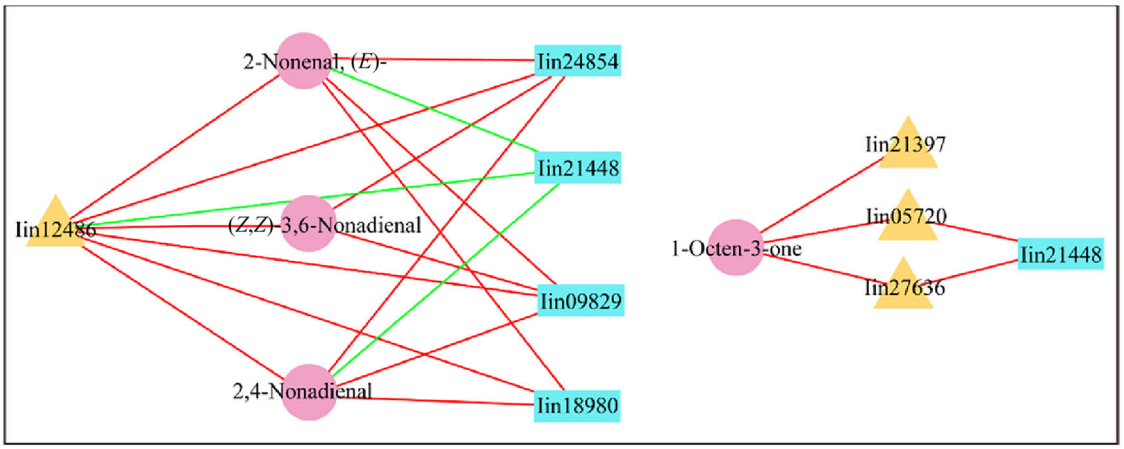


FIGURE 6 | Co-expression network of key odor compounds, structural genes, and transcription factors. Yellow triangles, pink circles, and blue squares represent transcription factors, key odor compounds, and structural genes, respectively. Red and green lines indicate positive and negative correlations, respectively.

A highly cohesive and species-specific regulatory module was identified for the biosynthesis of C9 aldehydes. The key fatty-green aldehydes, including 2,4-nonadienal, 2-nonenal, (*E*)-, and (*Z*,*Z*)-3,6-nonadienal, demonstrated exceptionally strong positive correlations ($r > 0.97$, $p < 0.001$) with the *NAC*-family TF *Iin12486* and two *LOX* genes, *Iin24854* and *Iin09829*. The correlation coefficients between Iin12486 and these structural genes were remarkably high ($r > 0.99$), suggesting a tightly co-expressed unit. This module exhibited a consistent and strong positive correlation among the TF, the *LOX* genes, and the C9 aldehydes, perfectly aligning with their collective high abundance in *I. tinctoria*. Intriguingly, the *HPL* gene *Iin21448* showed a significant negative correlation with these C9 aldehydes and the TF *Iin12486*, indicating a potential diversion of metabolic flux or a different substrate specificity in the pathway.

In contrast, a separate regulatory network governs the accumulation of 1-octen-3-one. The potent mushroom-smelling ketone was significantly correlated with a different set of TFs, namely *Iin05720* (*NAC*), *Iin21397* (*FAR1*), and *Iin27636* (*RWP-RK*). Notably, its formation was positively associated with the HPL gene Iin21448 ($r = 0.80$, $p < 0.01$), which was also co-expressed with the NAC TF Iin05720. This positions Iin21448 as a key candidate responsible for the high 1-octen-3-one production in *I. oblongata*, operating within a regulatory context distinct from that of the C9 aldehydes in *I. tinctoria*.

In summary, the co-expression network analysis provides a clear transcriptional blueprint for aroma divergence: the fatty-green character of *I. tinctoria* is orchestrated by an Iin12486-NAC-centered module activating specific *LOX* genes for aldehyde production, whereas the mushroom aroma of *I. oblongata* is associated with a separate set of TFs (*Iin05720*, *Iin21397*, *Iin27636*) co-expressed with the *HPL* gene *Iin21448*.

3.8 | qPCR Analysis

To validate the transcriptome sequencing (RNA-seq) results, the expression patterns of six candidate key genes and TFs (Iin21397, Iin27636, Iin21448, Iin24854, Iin09829, and Iin12486) were further examined using qPCR in the three materials (Io, It, and Ii). Overall, the relative expression levels determined by qPCR were generally consistent with the FPKM values obtained from RNA-seq, confirming the reliability of the transcriptomic data (Figure 7).

Notably, for genes Iin24854 and Iin09829, expression was markedly upregulated in It compared to Io and Ii, a trend robustly detected by both qPCR and RNA-seq. However, a notable discrepancy was observed for gene Iin12486 in It, where a dramatically high FPKM value contrasted with a moderate level of upregulation in qPCR. This suggests potential post-transcriptional regulation or technical distinctions between the two quantification methods for this specific gene. The strong correlation for the majority of the candidates underscores the accuracy of the transcriptome analysis and supports the selection of these genes for further functional investigation.

4 | Discussion

4.1 | VOC Profiling Unveils Species-Specific Chemotaxonomic Signatures

Our comprehensive VOC profiling identified a diverse volatilome in *Isatis* fruits, with terpenoids, esters, and ketones being the most abundant classes. This composition is consistent with reports in other medicinal plants, where terpenoids and esters often dominate the volatile spectrum and contribute to both ecological functions and aromatic properties (Li et al. 2025). The core set of 817 VOCs shared across all three species likely represents the conserved metabolic backbone of the *Isatis* genus. However, the distinct, species-specific VOC profiles revealed by PCA and the large number of differential VOCs (838) provide a powerful chemical basis for their taxonomic differentiation. Notably, the closer metabolic relationship between *I. tinctoria* and *I. indigotica*, as evidenced by their higher number of shared metabolites, aligns with traditional classifications and suggests a more recent evolutionary divergence compared to *I. oblongata*. The unique metabolite profile of *I. oblongata*, particularly its high abundance of esters and specific alcohols like 1-octen-3-ol, sets it apart. This

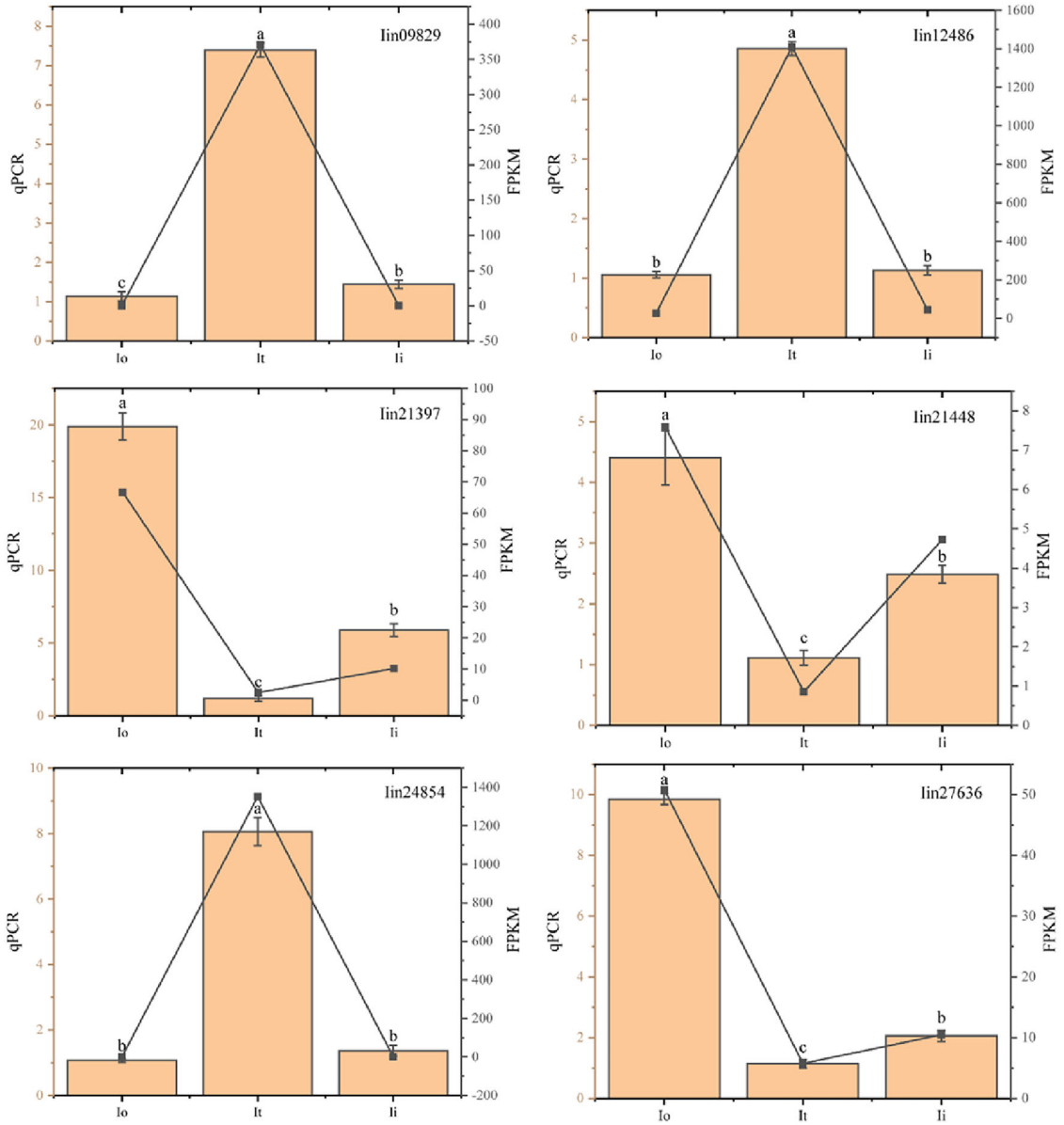


FIGURE 7 | Comparison of relative expression levels (qPCR) and FPKM values (RNA-seq) for six candidate genes in three *Isatis* species.

high ester content is unusual compared to many leafy or fruit-bearing plants but finds parallels in some aromatic fruits, where esters confer sweet and fruity notes (Bizzio et al. 2022). The stark contrast with *I. indigotica*, which possessed only one unique VOC (panaxene), indicates that the identity of this species may be defined more by quantitative differences in shared metabolites, a phenomenon observed in other closely related plant species where gene expression regulation, rather than gene presence/absence, drives phenotypic diversity (Shahzad et al. 2025).

4.2 | Odor-Activity Analysis Decodes the Divergent Sensory Impression of *Isatis* Species

Moving beyond mere presence/absence, the rOAV approach was crucial for identifying the compounds that truly shape the sensory perception of each species. Our finding that a small fraction of VOCs (44 out of 1042) are key aroma actors is a common principle in flavor science, often referred to as the component displacement theory (Guan et al. 2025). The species-specific aroma patterns we identified, mushroom-dominant for *I. oblongata*, fatty-green-earthy for *I. tinctoria*, and balanced sweet-green with a spicy undertone for *I. indigotica*, are not merely descriptive but are quantitatively grounded in their extreme rOAV values. The overwhelming contribution of 1-octen-3-one to *I. oblongata*'s aroma is reminiscent of its role in mushrooms and some truffles, where it is a key character-impact compound (Pons et al. 2011). The unique abundance of the spicy, medicinal benzene, (isothiocyanatomethyl)- in *I. indigotica* is of particular significance. Isothiocyanates are hydrolysis products of glucosinolates, hallmark compounds in the Brassicaceae family to which *Isatis* belongs. This finding bridges the gap between the traditional medicinal use of *Isatis* and its modern phytochemical profile, suggesting that this specific VOC could be a key marker for its authenticity and quality, a connection less emphasized in previous volatilomics studies on *Isatis*.

4.3 | Transcriptional Regulation Underpins Aroma Divergence Through Distinct Co-Expression Modules

Our WGCNA analysis provides a transcriptional narrative for the observed chemical divergence. The identification of three core modules, each strongly correlated with a specific aroma profile and driven by different TF families, offers a mechanistic hypothesis for the independent evolution of these scent signatures. The association of the *NAC* TF *Iin12486* with the fatty-green aldehydes in *I. tinctoria* is particularly compelling. *NAC* TFs are well-known master regulators of plant development and stress responses, but their role in controlling specific VOC biosynthesis pathways is increasingly being recognized (Y. Chen and Xia 2025; Liu et al. 2019). For instance, in tomato, certain *NAC* TFs regulate the synthesis of apocarotenoid volatiles (Wang et al. 2024). Our discovery extends this role to the LOX pathway-derived C9 aldehydes. Similarly, the link between *EIL* and *HSF* TFs (turquoise module) and the signature compounds of *I. indigotica* suggests a regulatory logic. *EIL* TFs are part of the ethylene signaling pathway, which is a key hormone influencing volatile emission in fruits, while *HSF*s can be induced by various stresses, potentially linking aroma production to environmental adaptation (Yin et al. 2010; B. Zhang et al. 2010). The fact that different *NAC* family members (*Iin12486* vs. *Iin05720*) are hub genes in different modules highlights the functional diversification within this large TF family and suggests that neofunctionalization after gene duplication could be a driver of aroma diversity in *Isatis*.

4.4 | Metabolic Pathway Analysis Reveals Shared Precursors but Divergent Fates Driven by Gene Expression

The reconstruction of the LOX pathway elegantly illustrates how genetic regulation can channel common precursors into distinct, species-characteristic end products. The high expression of specific *LOX* genes (*Iin24854*, *Iin09829*) in *I. tinctoria* directly explains its high yield of C9 aldehydes, which are classic products of the LOX-HPL cascade. This aligns with studies in cucumber and melon, where HPL activity is a major determinant of the green odor (Perpiñá et al. 2021). The most intriguing finding is the dual role of the *HPL* gene *Iin21448*. Its negative correlation with C9 aldehydes in *I. tinctoria* but positive correlation with 1-octen-3-one in *I. oblongata* suggests potential substrate specificity or competition between pathways. While *HPL*s typically cleave fatty acid hydroperoxides to aldehydes, the formation of 1-octen-3-one from 10-HPODE might involve alternative enzymatic steps or minor HPL activity followed by enzymatic oxidation. This complexity is seen in mushrooms, where the formation of 1-octen-3-ol from linoleic acid involves a *LOX* and a *hydroperoxide lyase*, and its subsequent oxidation leads to the ketone (Matsui et al. 2018). Our data suggest that *I. oblongata* may have evolved a specialized mechanism, potentially involving the specific TFs of the yellow module, to highly express *Iin21448* and possibly other downstream enzymes to shunt flux toward 1-octen-3-one, thereby creating its unique mushroom scent.

4.5 | Limitations and Future Perspectives

This study provides an integrated metabolomic and transcriptomic perspective on aroma formation in *Isatis*. However, some limitations should be acknowledged. First, the identification of key aroma compounds and the proposed biosynthetic pathways is primarily based on correlation analyses and database annotations. Functional validation of the candidate hub TFs (e.g., Iin12486, Iin05720) and structural genes (e.g., Iin21448) through techniques such as heterologous expression, gene silencing, or CRISPR-Cas9 is essential to confirm their precise regulatory and catalytic roles. Secondly, sensory evaluation by a trained human panel would be invaluable to directly correlate our chemical and molecular findings with the actual perceived aroma, thereby strengthening the link between instrumental analysis and sensory perception.

In conclusion, our integrated approach has delineated the chemical and transcriptional basis of aroma divergence among the three *Isatis* species. We have progressed from cataloguing volatile compounds to proposing a regulatory model in which species-specific transcriptional modules coordinate key metabolic pathways. This work not only provides a valuable resource for the chemotaxonomy and QC of *Isatis* but also lays a foundation for future research aimed at understanding and potentially modifying aroma profiles for flavor-related applications.

5 | Conclusion

This study provides an integrative volatilome and transcriptome landscape of three *Isatis* species, revealing the basis for aroma divergence in their fruits. We established distinct quantitative VOC profiles for *I. indigotica*, *I. tinctoria*, and *I. oblongata* and identified 44 key aroma-active compounds through rOAV analysis. This approach allowed for a quantitative definition of their unique sensory impressions: a mushroom-dominant character for *I. oblongata*, a fatty-green-earthy profile for *I. tinctoria*, and a balanced sweet-green aroma with a distinctive spicy note for *I. indigotica*. Crucially, transcriptomic integration revealed that this species-specific aroma divergence is orchestrated by distinct transcriptional regulatory modules. The fatty-green character of *I. tinctoria* is associated with a *NAC* TF (Iin12486)-centered module that co-expresses with and likely upregulates specific *LOX* genes (Iin24854, Iin09829), leading to abundant C9 aldehyde production. In contrast, the mushroom aroma of *I. oblongata* is linked to a separate regulatory module, where TFs (Iin05720, Iin21397) co-express with the *HPL* gene (Iin21448), promoting the production of 1-octen-3-one. This work elucidates the mechanistic link between TRs, structural genes, and aroma phenotypes. Our findings not only offer precise chemical markers for species identification and quality assessment but also reveal the core molecular regulatory network controlling aroma formation. This lays a solid foundation for the future high-value utilization of *Isatis* VOC resources and establishes a reference methodology for systematic research on volatile metabolites in medicinal plants.

Author Contributions

Rong Chen, Yan Qing, Xiaoshan Geng, Huaxi Huang, and Qin Liu designed the experiments. Rong Chen, Yan Qing, Qiaoyun Cai, Lili Wang, Fanglin Liu, Huijun Cheng, and Huaxi Huang performed the experiments. Yude Peng, Chunfeng Tang, and Yan Qing analyzed the data. Rong Chen, Yan Qing, Xiaoshan Geng, Huaxi Huang, and Qin Liu wrote and revised the manuscript. All authors have reviewed and approved the final version of the manuscript.

Funding

This study was supported by the Xinjiang Autonomous Region Science and Technology Aid Program (2024E02010), the Third Batch of the Tianchi Innovation Leading Talent Fund of the Xinjiang Autonomous Region (2025CXLJ004), and the Special Science and Technology Project of the Ili Autonomous Prefecture (YJC2025B03).

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Data that support the findings of this study are available from the corresponding author upon reasonable request.

References

Bizzio, L. N., D. Tieman, and P. R. Munoz. 2022. "Branched-Chain Volatiles in Fruit: A Molecular Perspective." *Frontiers in Plant Science* 12: 814138. https://doi.org/10.3389/fpls.2021.814138.

Chang, C. C., H. J. Su, H. L. You, C. W. Kao, I. L. Hung, and S. T. Huang. 2024. "*Isatis indigotica* Inhibits Influenza A Virus (H1N1) Virulent Protein Production and Autophagosome Accumulation." *Journal of Herbal Medicine* 43: 100827. https://doi.org/10.1016/j.hermed.2023.100827.

Chen, M., L. Gan, S. Lin, et al. 2012. "Alkaloids From the Root of *Isatis indigotica*." *Journal of Natural Products* 75, no. 6: 1167–1176. https://doi.org/10.1021/np3002833.

Chen, Q., H. Y. Lan, W. Peng, et al. 2021. "*Isatis indigotica*: A Review of Phytochemistry, Pharmacological Activities and Clinical Applications." *The Journal of Pharmacy and Pharmacology* 73, no. 9: 1137–1150. https://doi.org/10.1093/jpp/rgab014.

Chen, Y., and P. Xia. 2025. "NAC Transcription Factors as Biological Macromolecules Responded to Abiotic Stress: A Comprehensive Review." *International Journal of Biological Macromolecules* 308, no. Pt 1: 142400. https://doi.org/10.1016/j.ijbiomac.2025.142400.

Condurso, C., A. Verzera, V. Romeo, M. Ziino, A. Trozzi, and S. Ragusa. 2006. "The Leaf Volatile Constituents of *Isatis tinctoria* by Solid-Phase Microextraction and Gas Chromatography/Mass Spectrometry." *Planta Medica* 72, no. 10: 924–928. https://doi.org/10.1055/s-2006-946679.

Dadwal, V., D. K. Jha, M. Gaid, and R. Joshi. 2025. "Editorial: Exploring Volatile Organic Compounds in Fruits and Flowers: Aroma, Biosynthesis, and Ecological Impact." *Frontiers in Plant Science* 16: 1701678. https://doi.org/10.3389/fpls.2025.1701678.

Dupré, J., N. Joly, R. Vauquelin, et al. 2025. "*Isatis tinctoria* L.—From Botanical Description to Seed-Extracted Compounds and Their Applications: An Overview." *Plants* 14, no. 15: 2304. https://doi.org/10.3390/plants14152304.

Eshaghi, M., and S. Rashidi-Monfared. 2025. "Network-Based Identification of Hub Transcription Factors Associated With Benzylisoquinoline Alkaloid Biosynthesis in Papaver Somniferum." *Biochemistry and Biophysics Reports* 43: 102147. https://doi.org/10.1016/j.bbrep.2025.102147.

Fang, X., W. Xu, G. Jiang, et al. 2024. "Monitoring the Dynamic Changes in Aroma During the Whole Processing of Qingzhuan Tea at an Industrial Scale: From Fresh Leaves to Finished Tea." *Food Chemistry* 439: 137810. https://doi.org/10.1016/j.foodchem.2023.137810.

Feng, J., B. Zhang, H. Zhang, et al. 2024. "Combining With E-Nose, GC-MS, GC-IMS and Chemometrics to Explore Volatile Characteristics During the Different Stages of Zanthoxylum Bungeanum Maxim Fruits." *Food Research International* 195: 114964. https://doi.org/10.1016/j.foodres.2024.114964.

Fuentes-Lopez, K., J. Olivero-Verbel, and K. Caballero-Gallardo. 2025. "Aromatic Plant–Insect Dynamics: Molecular Insights for Advancing Pest Control Strategies." *Journal of Pest Science* 98, no. 6: 2501–2517. https://doi.org/10.1007/s10340-025-01953-8.

Guan, S., C. Liu, M. Ruan, et al. 2025. "Detection and Comparative Analysis of VOCs Between Tomato and Pepper Based on GC×GC-TOFMS." *Scientific Reports* 15, no. 1: 6140. https://doi.org/10.1038/s41598-025-90796-5.

Huang, D., X. Chen, R. Tan, et al. 2024. "A Comprehensive Metabolomics Analysis of Volatile and Non-Volatile Compounds in Matcha Processed From Different Tea Varieties." *Food Chemistry: X* 21: 101234. https://doi.org/10.1016/j.fochx.2024.101234.

Huang, W., S. Fang, J. Wang, et al. 2022. "Sensomics Analysis of the Effect of the Withering Method on the Aroma Components of Keemun Black Tea." *Food Chemistry* 395: 133549. https://doi.org/10.1016/j.foodchem.2022.133549.

Kang, M., H. Wu, Q. Yang, et al. 2020. "A Chromosome-Scale Genome Assembly of *Isatis indigotica*, an Important Medicinal Plant Used in Traditional Chinese Medicine: An *Isatis* Genome." *Horticulture Research* 7: 18. https://doi.org/10.1038/s41438-020-0240-5.

Li, D., J. Zhang, Y. Sun, et al. 2025. "The Utilization Value of Condensate Water From the Drying Process of *Lonicera japonica* via Metabolomics Analysis." *Metabolites* 15, no. 9: 569. https://doi.org/10.3390/metabo15090569.

Liu, M., Z. Ma, W. Sun, et al. 2019. "Genome-Wide Analysis of the NAC Transcription Factor Family in Tartary Buckwheat (*Fagopyrum tataricum*)." *BMC Genomics* 20, no. 1: 113. https://doi.org/10.1186/s12864-019-5500-0.

Loreto, F., and J. P. Schnitzler. 2010. "Abiotic Stresses and Induced BVOCs." *Trends in Plant Science* 15, no. 3: 154–166. https://doi.org/10.1016/j.tplants.2009.12.006.

Ma, M., H. Fu, T. Wang, L. Xiong, P. Feng, and B. Lu. 2024. "Widely Targeted Volatilomics and Transcriptome Analyses Reveal the Differences in Volatile Organic Components in Differently Shaped *Amomum tsao-ko* Fruits." *BMC Plant Biology* 24, no. 1: 915. https://doi.org/10.1186/s12870-024-05594-4.

Matsui, K., H. Takemoto, T. Koeduka, and T. Ohnishi. 2018. "1-Octen-3-Ol Is Formed From Its Glycoside During Processing of Soybean [*Glycine max* (L.) Merr.] Seeds." *Journal of Agricultural and Food Chemistry* 66, no. 28: 7409–7416. https://doi.org/10.1021/acs.jafc.8b01950.

Perpiñá, G., S. Roselló, C. Esteras, et al. 2021. "Analysis of Aroma-Related Volatile Compounds Affected by 'Ginsen Makuwa' Genomic Regions Introgressed in 'Vedrantais' Melon Background." *Scientia Horticulturae* 276: 109664. https://doi.org/10.1016/j.scienta.2020.109664.

Picazo-Aragonés, J., A. Terrab, and F. Balao. 2020. "Plant Volatile Organic Compounds Evolution: Transcriptional Regulation, Epigenetics and Polyploidy." *International Journal of Molecular Sciences* 21, no. 23: 8956. https://doi.org/10.3390/ijms21238956.

Pons, M., B. Dauphin, S. La Guerche, et al. 2011. "Identification of Impact Odorants Contributing to Fresh Mushroom Off-Flavor in Wines: Incidence of Their Reactivity With Nitrogen Compounds on the Decrease of the Olfactory Defect." *Journal of Agricultural and Food Chemistry* 59, no. 7: 3264–3272. https://doi.org/10.1021/jf104215a.

Rao, P. S., N. Fatima, M. H. Siddiqui, et al. 2025. "Indigofera Tinctoria: The Blue Gold of India's Sustainable Future." *Discover Sustainability* 6: 1135. https://doi.org/10.1007/s43621-025-01120-0.

Schwab, W., R. Davidovich-Rikanati, and E. Lewinsohn. 2008. "Biosynthesis of Plant-Derived Flavor Compounds." *The Plant Journal: For Cell*

and Molecular Biology 54, no. 4: 712–732. https://doi.org/10.1111/j.1365-313X.2008.03446.x.

Shahzad, Z., E. Hollwey, J. D. Moore, et al. 2025. "Gene Body Methylation Regulates Gene Expression and Mediates Phenotypic Diversity in Natural Arabidopsis Populations." *Nature Plants* 11, no. 10: 2084–2099. https://doi.org/10.1038/s41477-025-02108-4.

Shukurova, M. K., Y. Asikin, Y. Chen, M. Kusano, and K. N. Watanabe. 2020. "Profiling of Volatile Organic Compounds in Wild Indigenous Medicinal Ginger (*Zingiber barbatum* Wall.) From Myanmar." *Metabolites* 10, no. 6: 248. https://doi.org/10.3390/metabo10060248.

Wang, Q., L. Zhou, M. Yuan, F. Peng, X. Zhu, and Y. Wang. 2024. "Genome-Wide Identification of NAC Gene Family Members of Tree Peony (*Paeonia suffruticosa* Andrews) and Their Expression Under Heat and Waterlogging Stress." *International Journal of Molecular Sciences* 25, no. 17: 9312. https://doi.org/10.3390/ijms25179312.

Xiao, Z., J. Luo, Y. Niu, and M. Wu. 2018. "Characterization of Key Aroma Compounds From Different Rose Essential Oils Using Gas Chromatography-Mass Spectrometry, Gas Chromatography-Olfactometry and Partial Least Squares Regression." *Natural Product Research* 32, no. 13: 1567–1572. https://doi.org/10.1080/14786419.2017.1389933.

Xue, J., P. Liu, J. Yin, et al. 2022. "Dynamic Changes in Volatile Compounds of Shaken Black Tea During Its Manufacture by GC × GC-TOFMS and Multivariate Data Analysis." *Foods* 11, no. 9: 1228. https://doi.org/10.3390/foods11091228.

Yin, X. R., A. C. Allan, K. S. Chen, and I. B. Ferguson. 2010. "Kiwifruit EIL and ERF Genes Involved in Regulating Fruit Ripening." *Plant Physiology* 153, no. 3: 1280–1292. https://doi.org/10.1104/pp.110.157081.

Yuan, H., G. Cao, X. Hou, et al. 2022. "Development of a Widely Targeted Volatilomics Method for Profiling Volatilomes in Plants." *Molecular Plant* 15, no. 1: 189–202. https://doi.org/10.1016/j.molp.2021.09.003.

Yuan, H., Y. Jiangfang, Z. Liu, et al. 2024. "WTV2.0: A High-Coverage Plant Volatilomics Method With a Comprehensive Selective Ion Monitoring Acquisition Mode." *Molecular Plant* 17, no. 6: 972–985. https://doi.org/10.1016/j.molp.2024.04.012.

Zhang, B., J. Y. Shen, W. W. Wei, et al. 2010. "Expression of Genes Associated With Aroma Formation Derived From the Fatty Acid Pathway During Peach Fruit Ripening." *Journal of Agricultural and Food Chemistry* 58, no. 10: 6157–6165. https://doi.org/10.1021/jf100172e.

Zhang, L., X. Peng, X. Wu, et al. 2025. "Characterization and Decoding of Volatile Organic Compounds in Moso Bamboo Shoots (*Phyllostachys edulis* L.) at Different Growth Stages: A Combined Analysis by HS-GC-IMS, HS-SPME-GC-MS, E-Nose, and Molecular Docking." *Food Research International* 221, no. Pt 4: 117559. https://doi.org/10.1016/j.foodres.2025.117559.

Zhang, S., L. Sun, S. Wen, et al. 2023. "Analysis of Aroma Quality Changes of Large-Leaf Black Tea in Different Storage Years Based on HS-SPME and GC-MS." *Food Chemistry: X* 20: 100991. https://doi.org/10.1016/j.fochx.2023.100991.

Zhao, X., J. Zhao, S. Qi, Y. Fan, and J. Zhang. 2025. "Comprehensive Analysis of Volatile Compounds in Hybrid Sweetgum Population and Other Liquidambar Species by HS-SPME-GC-MS." *BMC Plant Biology* 25, no. 1: 381. https://doi.org/10.1186/s12870-025-06409-w.

Zhu, J., F. Chen, L. Wang, et al. 2016. "Characterization of the Key Aroma Volatile Compounds in Cranberry (*Vaccinium macrocarpon* Ait.) Using Gas Chromatography-Olfactometry (GC-O) and Odor Activity Value (OAV)." *Journal of Agricultural and Food Chemistry* 64, no. 24: 4990–4999. https://doi.org/10.1021/acs.jafc.6b01150.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.

Supplementary Table S1. Primers for quantitative real-time PCR. **Supplementary Table S2**. Volatile compounds, odor properties, and relative odor activity values.