

Article

Not peer-reviewed version

Integrating Metabolomics and Machine Learning to Analyze Chemical Markers and Ecological Regulatory Mechanisms of Geographical Differentiation in *Thesium chinense* Turcz

[Cong Wang](#), [Ke Che](#), [Guanglei Zhang](#), [Hao Yu](#)^{*}, [Junsong Wang](#)^{*}

Posted Date: 20 May 2025

doi: 10.20944/preprints202505.1441.v1

Keywords: *Thesium chinense* Turcz; metabolomics; geo-authentic herbs; machine learning; antioxidant activity; environmental factors; UHPLC-Q-TOF/MS



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

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.

Article

Integrating Metabolomics and Machine Learning to Analyze Chemical Markers and Ecological Regulatory Mechanisms of Geographical Differentiation in *Thesium chinense* Turcz.

Cong Wang ¹, Ke Che ², Guanglei Zhang ^{2,3}, Hao Yu ^{3,*} and Junsong Wang ^{1,*}

¹ Center for Molecular Metabolism, School of Environmental and Biological Engineering, Nanjing University of Science and Technology, 200 Xiao Ling Wei Street, Nanjing 210094, People's Republic of China

² College of Food Science and Engineering, Anhui Science and Technology University, Chuzhou 233100, China

³ School of Chinese Medicine, Bozhou University, Bozhou 236800, China

* Correspondence: yhz_1230@163.com (H.Y.); wangjunsong@njust.edu.cn (J.W.)

Abstract: Background: The relationship between medicinal efficacy and geographical environment in *Thesium chinense* Turcz., a traditional Chinese herb, remains systematically unexplored. This study integrates metabolomics, machine learning, and ecological factor analysis to elucidate the geographical variation patterns and regulatory mechanisms of secondary metabolites in *Thesium chinense* from Anhui, Henan, and Shanxi provinces. **Methods:** Metabolomic profiling was conducted on *Thesium chinense* samples collected from three geographical origins across Anhui, Henan, and Shanxi provinces. Machine learning algorithms (Random Forest, LASSO regression) identified region-specific biomarkers through intersection analysis. Metabolic pathway enrichment employed MetaboAnalyst 5.0 with target prediction. Antioxidant activity (DPPH/hydroxyl radical scavenging) was quantified spectrophotometrically. Environmental correlation analysis incorporated 19 WorldClim variables using redundancy analysis, Mantel tests, and Pearson correlations. **Results:** we identified 43 geographical marker compounds (primarily flavonoids and alkaloids). Random forest and LASSO regression algorithms determined core markers for each production area: Anhui (4 markers), Henan (6 markers), and Shanxi (3 markers). Metabolic pathway enrichment analysis revealed these markers exert pharmacological effects through neuroactive ligand-receptor interaction and PI3K-Akt signaling pathways. Redundancy analysis demonstrated Anhui samples exhibited significantly higher antioxidant activity (DPPH and hydroxyl radical scavenging rates) than other regions, strongly correlating with stable low-temperature environments (annual mean temperature) and precipitation patterns. **Conclusions:** This study established the first geo-specific molecular marker system for *Thesium chinense*, demonstrating that geographical environment critically influences metabolic profiles and bioactivity. Findings provide a scientific basis for quality control standards of geo-authentic herbs and offer insights into plant-environment interactions for sustainable cultivation practices.

Keywords: *Thesium chinense* Turcz.; metabolomics; geo-authentic herbs; machine learning; antioxidant activity; environmental factors; UHPLC-Q-TOF/MS

1. Introduction

Thesium chinense Turcz., a medicinal plant with a rich legacy in traditional Chinese medicine (TCM), has long been revered for its multifaceted therapeutic properties, including “heat-clearing, detoxifying, kidney-nourishing, and essence-securing” effects [1]. Its documented use spans centuries, appearing prominently in seminal works such as the Compendium of Materia Medica

(Bencao Gangmu) and the Illustrated Classic of Materia Medica (Bencao Tujing). These ancient texts not only highlight its efficacy but also establish its irreplaceable status in the treatment of inflammation, infections, and various systemic ailments. The meticulous documentation by herbal medicine pioneers, particularly Li Shizhen, underscores the plant's enduring value in addressing febrile diseases, rheumatism, and skin conditions like sores [2]. These ancient manuscripts illuminate not just the herb's effectiveness, but also cement its indispensable role in treating inflammatory conditions, infections, and diverse systemic disorders. The detailed records left by pioneers of herbal medicine, notably Li Shizhen, emphasize the plant's lasting significance in managing fever-related illnesses, rheumatic conditions, and dermatological ailments including ulcerations [2].

In contemporary clinical practice, *Thesium chinense* continues to play a pivotal role. Modern formulations, such as Bairui tablets and granules, are widely employed to treat pharyngitis and respiratory infections. These preparations leverage the plant's bioactive compounds to alleviate oxidative stress, inhibit ferroptosis, and modulate inflammatory factors [3]. This continuity from historical pharmacopeias to modern therapeutics reflects the plant's adaptability and relevance across different eras. However, despite its widespread use, there remains a need to delve deeper into the mechanisms underlying its efficacy, especially in light of advancements in molecular biology that have begun to unravel the complexities of its chemical composition.

Advances in molecular science have significantly enhanced our understanding of the pharmacological properties of *Thesium chinense*. The plant is rich in bioactive compounds, including flavonoids, alkaloids, coumarins, and polysaccharides, which collectively contribute to its anti-inflammatory, antimicrobial, antioxidant, and immunomodulatory effects [4]. Flavonoids, for instance, are known for their potent antioxidant activity, scavenging free radicals and mitigating oxidative damage. Alkaloids, on the other hand, exhibit antimicrobial properties, making them effective against a range of pathogens. Polysaccharides play a crucial role in immune modulation, enhancing the body's natural defense mechanisms. In modern clinical settings, *Thesium chinense* maintains its crucial therapeutic significance. Contemporary pharmaceutical preparations, including Bairui tablets and granules, are extensively utilized in treating pharyngitis and respiratory infections. These formulations harness the plant's bioactive constituents to mitigate oxidative stress, suppress ferroptosis, and regulate inflammatory mediators [3]. This seamless transition from historical pharmacopeias to modern therapeutic applications demonstrates the plant's versatility and enduring relevance across medical epochs. Nevertheless, despite its extensive clinical application, there remains a compelling need to further elucidate the underlying mechanisms of its therapeutic efficacy, particularly in light of recent molecular advances that have begun to decode its complex chemical architecture.

Recent breakthroughs in molecular science have substantially deepened our comprehension of *Thesium chinense*'s pharmacological profile. The plant harbors a diverse array of bioactive compounds, encompassing flavonoids, alkaloids, coumarins, and polysaccharides, which synergistically contribute to its anti-inflammatory, antimicrobial, antioxidant, and immunomodulatory properties [4]. Notably, flavonoids demonstrate robust antioxidant activity, neutralizing free radicals and attenuating oxidative injury. Alkaloids exhibit potent antimicrobial effects, effectively combating various pathogenic organisms. Additionally, polysaccharides play a vital role in immune regulation, strengthening the body's innate defense systems.

The transition of *Thesium chinense* from a traditional remedy to a cornerstone of modern molecular medicine has been facilitated by these discoveries. Contemporary metabolomics platforms, such as ultra-performance liquid chromatography-quadrupole time-of-flight mass spectrometry (UPLC-QTOF-MS) and liquid chromatography-mass spectrometry (LC-MS), have enabled comprehensive chemical profiling of the plant [5]. These high-throughput techniques allow for the simultaneous detection of hundreds of metabolites, providing unprecedented insights into the plant's chemical diversity. Network analysis further elucidates how environmental factors, ecological variables, and genetic background collectively shape the metabolome of *Thesium chinense*. This progress has redefined the concept of "geo-authenticity," transforming it from an empirical

observation rooted in traditional knowledge to a scientific evaluation system based on molecular markers, component networks, pathway regulation, and environmental adaptation [6]. These breakthrough findings have catalyzed the evolution of *Thesium chinense* from a traditional therapeutic agent into a cornerstone of modern molecular medicine. State-of-the-art metabolomics platforms, particularly ultra-performance liquid chromatography-quadrupole time-of-flight mass spectrometry (UPLC-QTOF-MS) and liquid chromatography-mass spectrometry (LC-MS), have enabled exhaustive chemical characterization of the plant [5]. These sophisticated high-throughput methodologies facilitate the concurrent detection of hundreds of metabolites, offering unprecedented insights into the plant's chemical complexity. Network analysis further unveils how environmental conditions, ecological parameters, and genetic foundations collectively influence the metabolomic profile of *Thesium chinense*. Such advances have fundamentally reconstructed the concept of "geo-authenticity," transforming it from an empirical observation grounded in traditional wisdom into a rigorous scientific framework encompassing molecular markers, component networks, pathway regulation, and environmental adaptation mechanisms [6].

Despite these advances, the geographical variation in *Thesium chinense* and its impact on efficacy and quality remain understudied. Current literature predominantly focuses on geo-authentic regions such as Anhui and Henan, often relying on single-component analyses, such as high-performance liquid chromatography (HPLC) for rutin quantification, without comprehensive multi-regional comparisons [7]. While studies have shown that *Thesium chinense* sourced from Anhui exhibits higher levels of saccharides, saponins, and flavonoids due to its unique climatic and edaphic conditions, the scientific comparison of geo-authenticity mechanisms and cross-regional compositional stability remains lacking [8]. This gap poses significant challenges for quality control, clinical application, and market authentication.

Traditional linear chemometric methods, such as principal component analysis (PCA), have proven inadequate for handling the complex nonlinear metabolomic data generated by modern analytical techniques. These methods fail to accurately identify geo-specific core markers or quantify environment-metabolite-efficacy relationships [9]. Furthermore, systematic validation of how climatic factors (temperature and precipitation variability), altitude, and soil properties influence the biosynthesis of alkaloids and flavonoids is notably absent. Current comparative studies on geographically diverse *Thesium chinense* populations are limited, hindering efforts to standardize and conserve this valuable medicinal resource [10].

To address these gaps, this study aims to evaluate *Thesium chinense* from different regions through a comprehensive approach that integrates metabolomics, climate response analysis, and bioactivity testing. Samples were collected from geo-authentic regions, including Anhui, Henan, and Shanxi, along with corresponding environmental data. Advanced analytical techniques, such as UPLC-QTOF-MS, were employed to identify region-discriminating metabolites and determine area-specific markers. Machine learning algorithms, including random forest (RF) and least absolute shrinkage and selection operator (LASSO), were used to screen core markers, while redundancy analysis (RDA) explored the relationships between environmental factors, antioxidant activity, and metabolite profiles. This study represents the systematic effort to establish a traceable geo-specific molecular marker system for *Thesium chinense* across three major production areas. By filling theoretical and data gaps in geographical adaptation and molecular tracing, this research contributes to a deeper understanding of Chinese geo-authentic herbs and paves the way for improved quality control, standardization, and conservation strategies.

Despite significant scientific progress, the geographical diversity of *Thesium chinense* and its implications for therapeutic efficacy and quality standards remain inadequately explored. Current research predominantly centers on established geo-authentic regions like Anhui and Henan, typically employing single-component analyses such as high-performance liquid chromatography (HPLC) for rutin quantification, while lacking comprehensive cross-regional comparative studies [7]. Although evidence indicates that *Thesium chinense* from Anhui demonstrates elevated levels of saccharides, saponins, and flavonoids owing to distinct climatic and soil conditions, rigorous scientific

investigation of geo-authenticity mechanisms and inter-regional compositional consistency remains insufficient [8]. This knowledge gap presents substantial challenges for quality assurance, clinical implementation, and market authentication protocols.

Conventional linear chemometric approaches, including principal component analysis (PCA), have proven insufficient for processing the complex nonlinear metabolomic data generated by contemporary analytical platforms. These methodologies fail to effectively identify region-specific core markers or accurately quantify the intricate relationships between environmental factors, metabolite profiles, and therapeutic efficacy [9]. Moreover, there is a notable absence of systematic research examining how climatic variables (temperature and precipitation patterns), elevation, and soil characteristics influence the biosynthesis of alkaloids and flavonoids. The current scarcity of comparative studies across geographically diverse *Thesium chinense* populations impedes efforts to standardize and preserve this invaluable medicinal resource [10].

This study addresses these research gaps through a comprehensive investigation that integrates metabolomics, climate response analysis, and bioactivity assessment. Samples were systematically collected from established geo-authentic regions, including Anhui, Henan, and Shanxi, alongside corresponding environmental data. Advanced analytical methodologies, particularly UPLC-QTOF-MS, were employed to identify region-specific metabolites and establish area-distinctive markers. Sophisticated machine learning algorithms, including random forest (RF) and least absolute shrinkage and selection operator (LASSO), were utilized to screen core markers, while redundancy analysis (RDA) investigated the correlations between environmental parameters, antioxidant activity, and metabolite profiles. This research represents a pioneering systematic effort to establish a traceable geo-specific molecular marker system for *Thesium chinense* across three major production regions. By addressing theoretical and empirical gaps in geographical adaptation and molecular tracing, this study advances our understanding of Chinese geo-authentic herbs while establishing foundations for enhanced quality control, standardization, and conservation strategies.

2. Materials and Methods

2.1. Sample Collection

Thesium chinense samples were collected from nine localities across three provinces in China (Anhui, Henan, and Shanxi) during September and October 2022 (Figure 1). Three samples were collected from each locality, selecting two-year-old plants with a length of 25–30 cm. The samples were identified by Professor Liu Hanzhen from Anhui University of Science and Technology and stored under the accession numbers 2022071901 (Anhui), 2022071902 (Henan), and 2022071903 (Shanxi) at the Center for Molecular Metabolism, Nanjing University of Science and Technology.

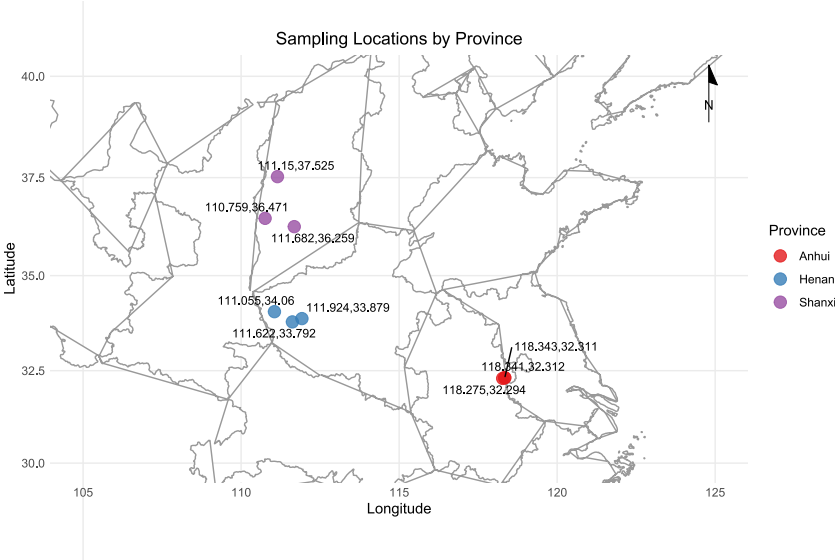


Figure 1. Detailed sample information of the geographical origins.

2.2. Metabolite Extraction

Fresh plant tissues were ground into fine powder in a liquid nitrogen-cooled mortar [5]. Approximately 15 mg of the tissue powder was accurately weighed and transferred into a 2 mL EP tube. A pre-cooled chloroform/water/methanol (20:20:60, v/v/v) mixture and tungsten carbide beads were added, and the mixture was homogenized at 70 Hz for three intermittent cycles (60 s each with a 5 s interval) using a low-temperature homogenizer. The homogenate was centrifuged at 12,000 rpm for 10 min at 4 °C. The supernatant was collected, dried under nitrogen, and re-dissolved in 500 µL of 50% methanol solution. The re-dissolved sample was vortexed for 30 s, followed by sonication in an ice bath for 1 min and centrifugation at 12,000 rpm for 15 min at 4 °C. The clear supernatant was transferred to a vial for UHPLC-MS injection. Quality control (QC) samples were also prepared to ensure the stability and reproducibility of the mass spectrometry.

2.3. LC-MS Detection Method

The LC-MS analysis was performed using a Triple TOF 5600+ system with an electrospray ionization source. Gradient elution was employed with formic acid and methanol as the mobile phase. The analysis was conducted in negative ion mode to capture full-scan mass spectra, with specific settings for voltage, declustering potential, collision energy, and gas pressure. Automatic MS/MS scans were performed on the most intense ions to ensure accuracy and calibration.

Data processing involved converting the raw files to mzML format using ProteoWizard software, followed by ion feature extraction using XCMS software. Metabolite identities were assigned by matching the spectra with the Human Metabolome Database (HMDB). Data normalization was performed using R, and further analysis was carried out using the mixOmics R package [11,12]. Potential biomarkers were identified based on VIP scores and t-tests.

2.4. Machine Learning-Based Biomarker Screening

Random Forest (RF) [13] and Lasso regression [14] algorithms were employed to select region-specific biomarkers from the chemical compositions of the three regions. Intersection analysis was conducted to identify region-specific biomarkers, and statistical analysis was used to verify the stability of the screening results.

2.5. Pathway Analysis of Metabolites and Related Genes and Proteins

Metabolite set enrichment analysis was performed using MetaboAnalyst 5.0 (<https://www.metaboanalyst.ca/>) [15]. The KEGG database (<http://www.kegg.jp/>) [16] was used for pathway mapping. The components of Thesium chinense were subjected to batch prediction of targets using TargetNet (<http://targetnet.scbdd.com/>), followed by GO and KEGG enrichment analysis.

2.6. Antioxidant Activity Assays

2.6.1. DPPH Radical Scavenging Activity Assay

Thesium chinense sample powder (50 mg) was mixed with DPPH solution and nitrogen radical extraction solution, and the absorbance was measured at 517 nm [17]. A control sample without DPPH solution was used for comparison. The DPPH scavenging rate was calculated.

2.6.2. Hydroxyl Radical Scavenging Activity Assay

Thesium chinense sample powder (100 mg) was extracted with 4 mL of methanol by ultrasonication. The supernatant was diluted and mixed with double-distilled water and substrate application solution, followed by incubation at 37°C for 1 min. The reaction was terminated by

adding a chromogenic agent [18], and the hydroxyl radical scavenging ability of the sample was determined.

2.7. Correlation Between Thesium Chinense Metabolites and Environmental Factors

Nineteen environmental variables (11 temperature variables and 8 precipitation variables) were recorded for each sample area from the WorldClim Version 2 dataset (<http://www.worldclim.org/>) [19,20], as shown in Table 1.

Table 1. Environmental variables recorded for each sample area.

Field	Description	Units
bio1	annual mean temperature	°C
bio2	mean diurnal range (mean of monthly (max temp - min temp))	°C
bio3	isothermality (bio2/bio7) (*100)	°C
bio4	temperature seasonality (standard deviation *100)	%
bio5	max temperature of warmest month	°C
bio6	min temperature of coldest month	°C
bio7	temperature annual range(bio5-bio6)	°C
bio8	mean temperature of wettest quarter	°C
bio9	mean temperature of driest quarter	°C
bio10	mean temperature of warmest quarter	°C
bio11	mean temperature of coldest quarter	°C
bio12	annual precipitation	mm
bio13	precipitation of wettest month	mm
bio14	precipitation of driest month	mm
bio15	precipitation seasonality (coefficient of variation)	%
bio16	precipitation of wettest quarter	mm
bio17	precipitation of driest quarter	mm
bio18	precipitation of warmest quarter	mm
bio19	precipitation of coldest quarter	mm

Redundancy analysis (RDA) and Mantel tests were performed in R to explore the relationships between environmental factors, antioxidant activity, and Thesium chinense metabolites. Pearson correlation analysis was used to associate the relative metabolite content with environmental factors and antioxidant activity.

3. Results

3.1. Differences in Chemical Composition and Metabolic Pathway Characteristics of Thesium Chinense from Different Regions

Liquid chromatography-mass spectrometry (LC-MS) was used to systematically analyze the chemical composition of Thesium chinense samples collected from Anhui, Henan, and Shanxi. A chemical fingerprint database was constructed, and differential biomarkers and secondary metabolic pathway characteristics were identified through multivariate statistics (PCA, PLS-DA, Heatmap) and metabolite set enrichment analysis.

The principal component analysis (PCA) model (Figure 2A, B) showed spatial overlap between Henan samples and the other two regions, with an overlap area of 63.5% with Shanxi samples, indicating high similarity in metabolite composition between these two regions. In contrast, the partial least squares discriminant analysis (PLS-DA) model (Figure 2C, D) exhibited stronger inter-group discrimination in positive ion mode, with Anhui samples forming a distinct cluster on the PC1 axis, suggesting significant specificity in metabolite composition for this region.

The heatmap analysis (Figure 2E) visually presented the relative abundance gradient distribution of 289 compounds in positive ion mode and 217 compounds in negative ion mode. Screening criteria of VIP ≥ 1 and p < 0.05 identified a total of 43 statistically significant differential

metabolites. Metabolite set enrichment analysis revealed ten metabolic pathways (Figure 2F), with the top three being flavone and flavonol biosynthesis, flavonoid biosynthesis, and indole alkaloid biosynthesis. These pathways indicate significant metabolic activity in the synthesis of flavonoids and indole alkaloids in *Thesium chinense* samples, which are closely related to their known pharmacological effects.

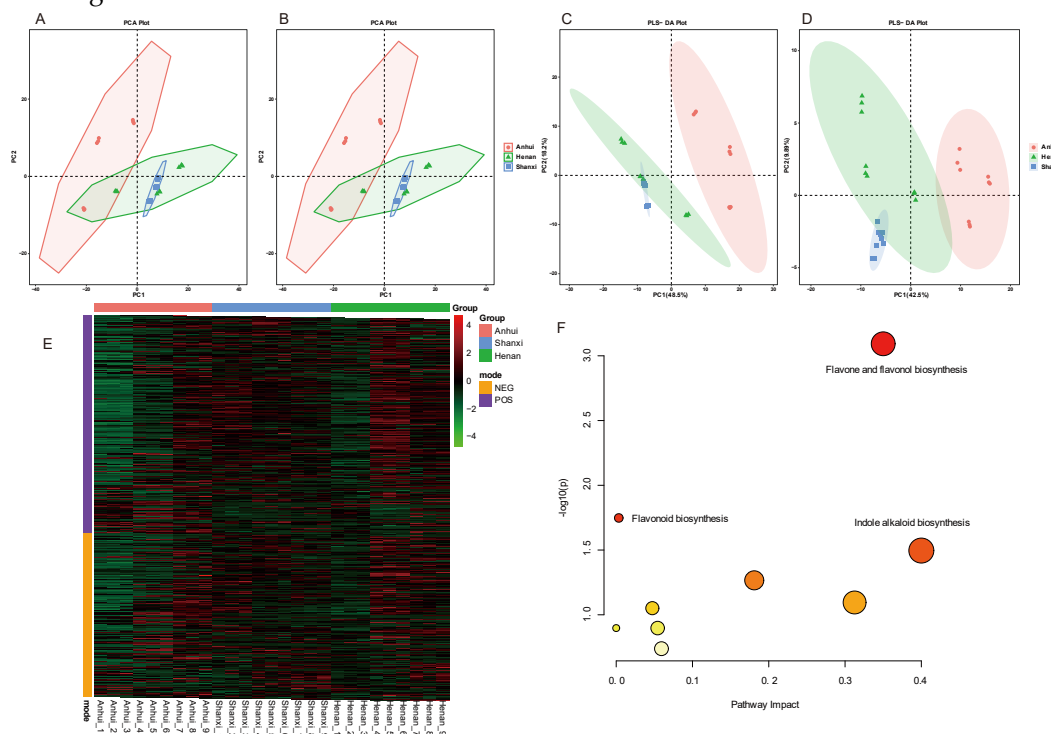


Figure 2. Multivariate analysis and metabolic pathway characteristics of Bairuicao samples from Anhui, Henan, and Shanxi. PCA score plots in positive (POS) (A) and negative (NEG) (B) ion modes; PLS-DA score plots in POS (C) and NEG (D) modes; (E) Heatmap of differential metabolites; (F) Metabolic pathway enrichment analysis.

3.2. Geographical Impact on the Chemical Composition of *Thesium chinense*

To further investigate the impact of geographical factors on the chemical composition of *Thesium chinense*, three comparative models were constructed (Anhui vs. others, Shanxi vs. others, Henan vs. others). PLS-DA analysis (Figure 3) showed that Anhui samples had no overlap with other regions, while Shanxi and Henan samples exhibited significant overlap with other regions, indicating that Anhui samples had the most unique chemical characteristics, whereas Shanxi and Henan samples had lower distinguishability from the other two regions.

At the level of differential metabolites, in the “Anhui vs. others” group, Glechomafuran and Cyclomethyltryptophan were significantly downregulated ($p < 0.001$), while Myristicin showed specific enrichment, but the expression was reversed in the Shanxi group. This cross-group expression heterogeneity may reflect the differential regulation of secondary metabolite synthesis by geographical environments. Kaempferol flavonoid metabolites exhibited bidirectional regulation, with Kaempferol 7-neohesperidoside abundance decreasing by 82%, while its isomer Kaempferol 3-O-beta-sophoroside increased by 3.2-fold. This reverse regulation of homologous metabolites may be due to tissue-specific modifications in homologous metabolites. In the “Shanxi vs. others” group, Myristoylcarnitine and Sanggenone H were significantly upregulated. Ferulic acid reversed the downward trend in the Anhui group in the Shanxi group. Notably, 6-Pentyl-2H-pyran-2-one entered the top 10 differential metabolites in both Anhui and Shanxi groups, but the expression direction was completely opposite. In the “Henan vs. others” group, 7-Hydroxy-3-(2-hydroxy-propyl)-5-methylisochromen-1-one and cis-Aconitic acid were significantly downregulated, and the expression of Glechomafuran in the Henan group completely reversed the downward trend in the Anhui group.

In all groups, 13-alpha-(21)-Epoxyeurycomanone showed highly significant differences, with its stable geographical response characteristics making it a potential key biomarker.

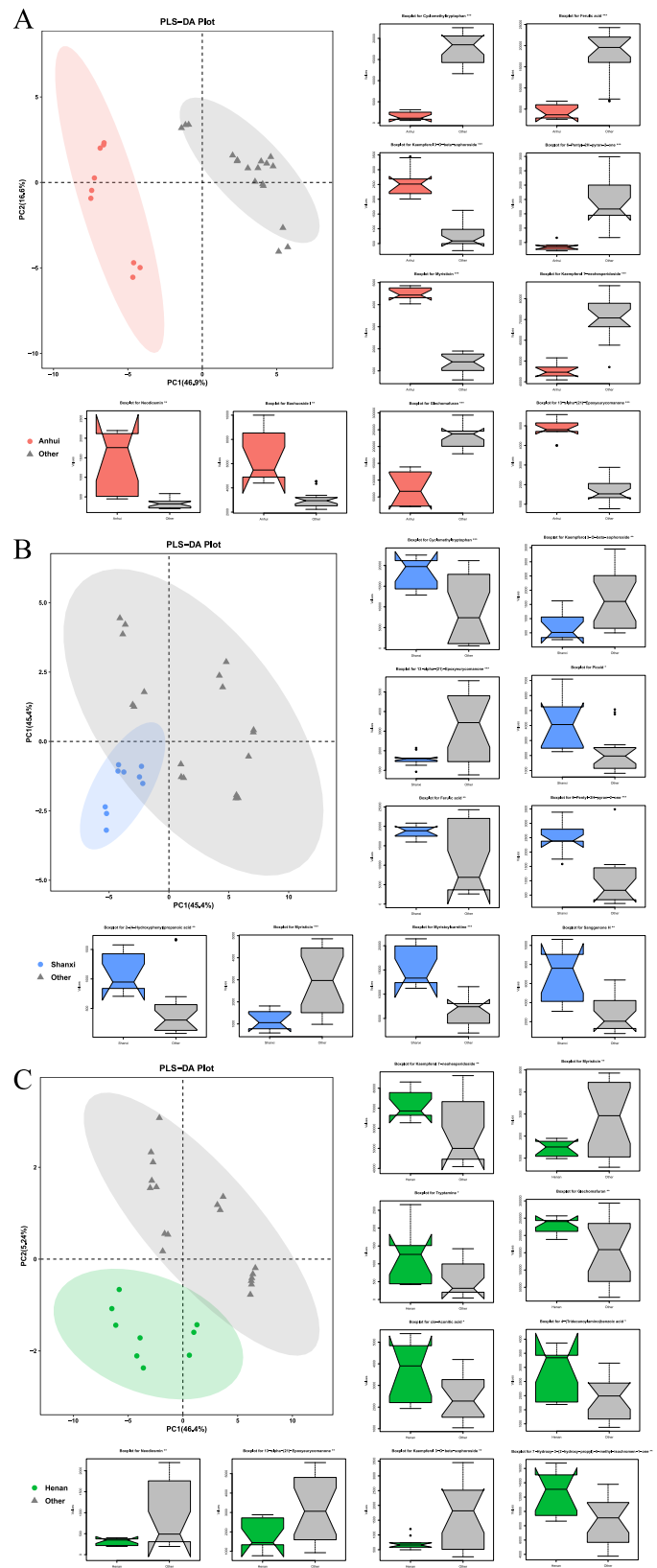


Figure 3. PLS-DA analysis and top 10 characteristic compounds of Baihuicao from Anhui, Shanxi, and Henan. (A) Anhui vs. other regions; (B) Shanxi vs. other regions; (C) Henan vs. other regions.

3.3. Machine Learning-Driven Screening System for Regional Biomarkers

By integrating Random Forest (RF) and Lasso regression algorithms (Figure 4), a dual screening system for regional identification biomarkers of *Thesium chinense* was established. In Anhui, RF algorithm identified 15 compounds, while Lasso algorithm identified 5 specific biomarkers, resulting in 4 overlapping compounds: Kaempferol 3-O-beta-sophoroside, 13-alpha-(21)-Epoxyeurycomanone, Myristicin, and Glechomafuran.

For Henan, RF algorithm identified 15 marker compounds, Lasso algorithm identified 10 characteristic components, resulting in 6 overlapping compounds: 7-Hydroxy-3-(2-hydroxy-propyl)-5-methyl-isochromen-1-one, Myristoylcarnitine, Kaempferol 3-O-beta-sophoroside, Kaempferol 7-neohesperidoside, Glechomafuran, and 4-(Tridecanoylamino)benzoic acid.

For Shanxi, RF algorithm identified 15 biomarkers, Lasso algorithm identified 6 specific compounds, resulting in 3 overlapping compounds: Rhamnetin 3-sophoroside, 6-Pentyl-2H-pyran-2-one, and Myristoylcarnitine. Kaempferol 3-O-beta-sophoroside, Glechomafuran, and Myristoylcarnitine were commonly identified in Anhui and Henan models, while Myristoylcarnitine was commonly identified in Shanxi and Henan models.

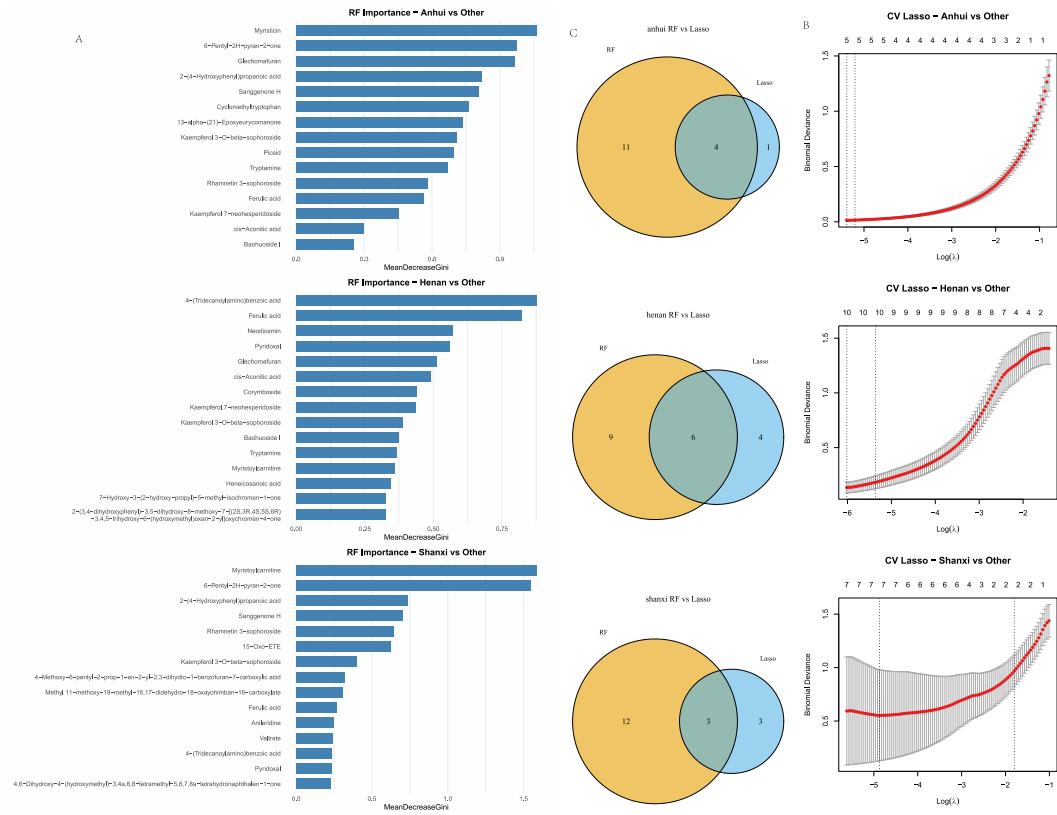


Figure 4. Machine learning-based screening of region-specific biomarkers in Bairyicao. (A) RF importance ranking; (B) LASSO coefficient profiles for three regions; (C) Venn diagrams of RF-LASSO overlapping biomarkers.

3.4. Pharmacological Mechanism Network of Geographically Specific Components

To elucidate the potential pharmacological mechanisms of region-specific biomarkers in *Thesium chinense*, key protein targets were predicted, and their biological functions were analyzed based on GO functional annotation and KEGG pathway enrichment analysis. GO analysis (Figure 5A) indicated that molecular function (MF) was mainly associated with histone H3Y41 kinase activity and G protein-coupled 5-hydroxytryptamine receptor activity, cellular component (CC) was concentrated in the plasma membrane and synapse, and biological process (BP) was dominated by G protein-coupled receptor signaling pathway and inflammatory response.

The 489 targets (Figure 5B) identified were significantly enriched in neuroactive ligand-receptor interaction, calcium signaling pathway, and PI3K-Akt signaling pathway, with key receptor genes such as CHRM3, HTR2A, and DRD2 involved in the neuroactive pathway.

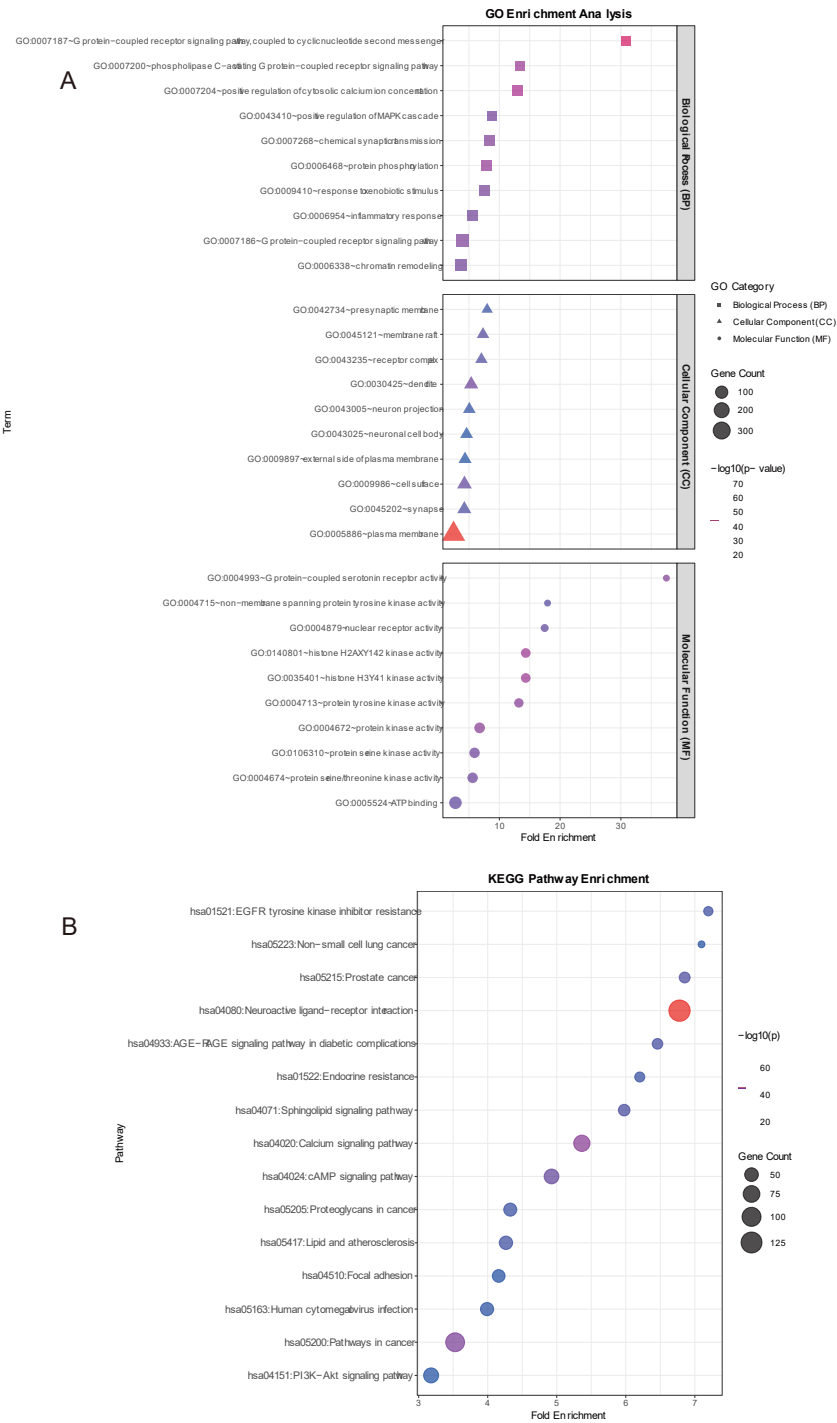


Figure 5. Functional enrichment analysis of Baihuicao biomarkers. (A) GO terms categorized into biological process (BP), cellular component (CC), and molecular function (MF); (B) KEGG pathway enrichment.

The construction of region-specific “component-target-pathway” networks (Figure 6) revealed that in Anhui samples, Myristicin and Glechomafuran significantly regulated the cholinergic receptor genes CHRM2/CHRM3/CHRM4, leading to enrichment in neuroactive ligand-receptor interaction pathways and calcium signaling pathways (Figure 6A). In Henan samples, Kaempferol 7-neohesperidoside and Myristoylcarnitine targeted ADORA1/PTGFR/PTGS2, showing a close association with lipid metabolism-related disease pathways (Figure 6B). Meanwhile, in Shanxi

samples, Rhamnetin 3-sophoroside and 6-Pentyl-2H-pyran-2-one modulated kinase genes such as AKT1/PRKACA/GSK3B, specifically influencing EGFR tyrosine kinase inhibitor resistance pathways and the PI3K-Akt signaling pathway (Figure 6C). Notably, the gene NOS3 played a central role in pathway enrichment across all three regions (accounting for 62.3% of cross-regional targets), but its regulatory weight in the diabetes complication-related AGE-RAGE pathway exhibited significant regional differences (Anhui: 34 targets, Henan: 21 targets, Shanxi: 16 targets). Geographically specific compounds exert unique pharmacological effects by differentially regulating neurotransmitter receptors and metabolism-related pathways, with NOS3 potentially serving as a key regulatory hub for cross-regional synergistic effects.

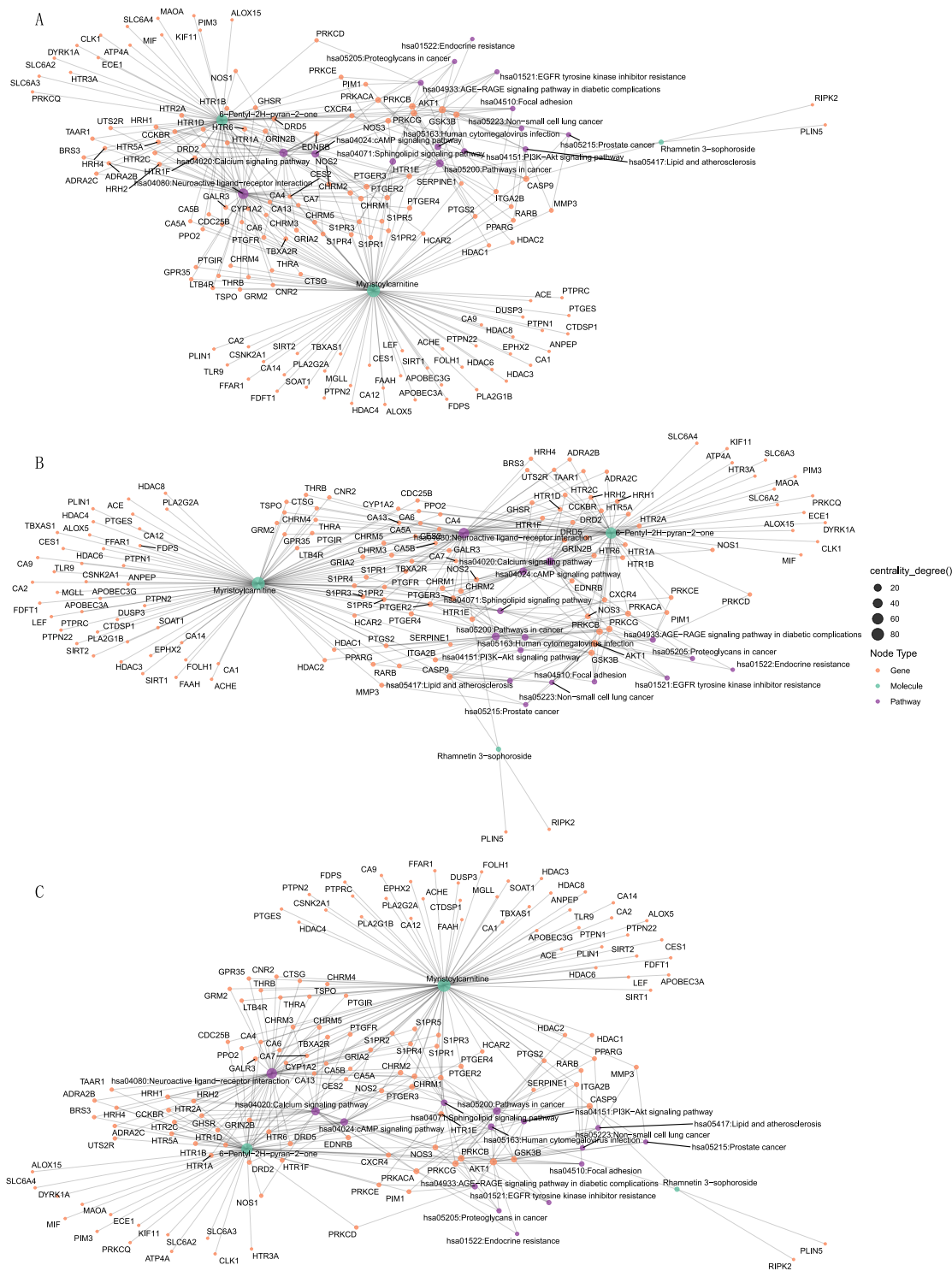


Figure 6. Region-specific compound-target-pathway networks: Anhui (A), Henan (B), and Shanxi (C).

3.5. Ecological-Activity Correlation Model Construction and Validation

To reveal the geographical differentiation patterns of antioxidant activity in *Thesium chinense* samples and the ecological regulation mechanisms of climate factors, antioxidant capacity was quantitatively assessed through DPPH and hydroxyl radical scavenging experiments. Anhui samples showed significantly higher DPPH scavenging rates (Figure 7A) and hydroxyl radical inhibition rates (Figure 7B) than Shanxi and Henan ($p < 0.01$), with Rhamnetin 3-sophoroside and 6-Pentyl-2H-pyran-2-one strongly positively correlated with antioxidant indices.

RDA (Figure 7C) indicated that temperature-related variables (bio1, bio8, bio9) and DPPH showed significant negative correlation on RDA1, suggesting a close relationship between low-temperature environments and high DPPH activity. Precipitation indicators (bio4 temperature seasonality) dominated negatively on RDA2, suggesting that precipitation fluctuations may suppress antioxidant activity. Sample distribution showed Anhui samples concentrated on the negative axis of RDA1, consistent with their higher DPPH values, while Shanxi samples were dispersed along the negative axis of RDA2, corresponding to low hydroxyl activity.

The Pearson network (Figure 7D) revealed that Myristoylcarnitine in Shanxi was positively correlated with altitude ($r = 0.63$) and slope ($r = 0.59$), while Kaempferol 7-neohesperidoside in Henan was significantly negatively correlated with diurnal temperature range. The key component Kaempferol 3-O-beta-sophoroside in Anhui was extremely significantly positively correlated with DPPH activity, regulated positively by annual precipitation (bio12) and precipitation of the wettest month (bio13). Rhamnetin 3-sophoroside in Shanxi was significantly positively correlated with hydroxyl activity, indicating geographical adaptive metabolic characteristics. The correlation between 7-Hydroxy-3-(2-hydroxy-propyl)-5-methyl-isochromen-1-one in Henan and annual temperature range (bio7) was weak, suggesting limited impact of annual temperature fluctuations on its synthesis.

Geographical differentiation of antioxidant activity is mainly regulated by temperature fluctuations and precipitation patterns, with mean annual temperature and seasonal precipitation being key factors driving the geographical differentiation of antioxidant activity in *Thesium chinense*. Low-temperature stable environments are conducive to enhancing DPPH activity, while drastic precipitation fluctuations in Shanxi may lead to reduced hydroxyl activity.

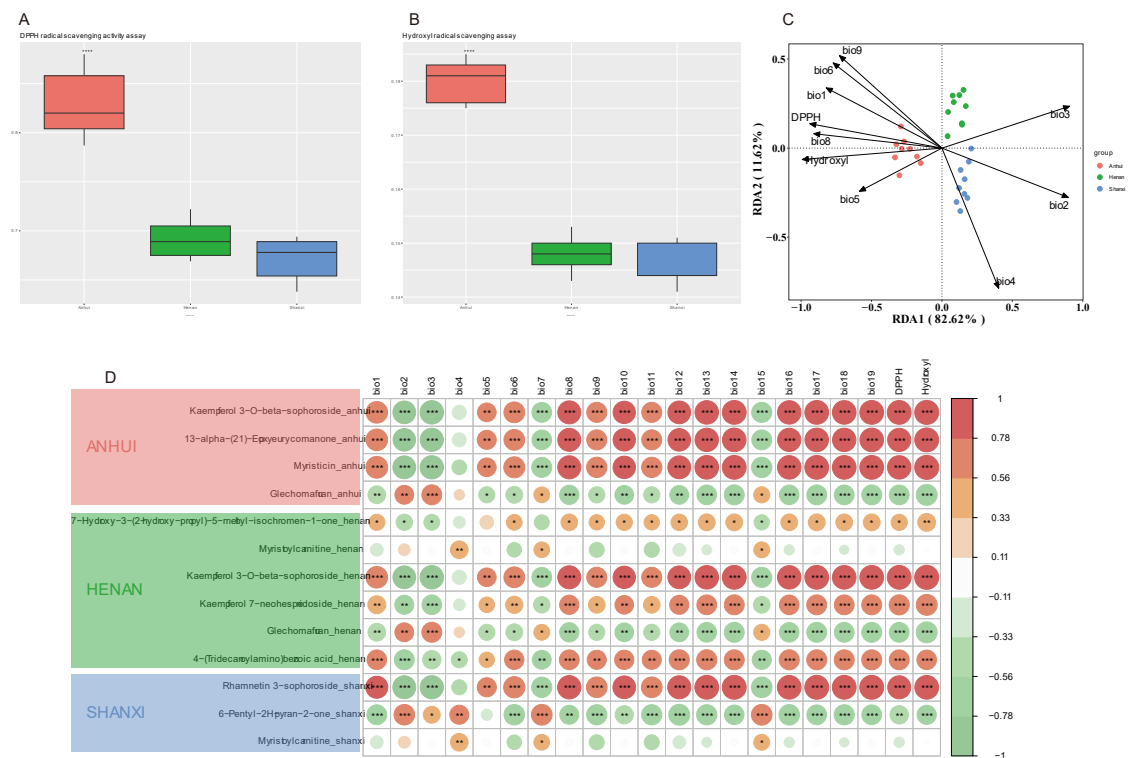


Figure 7. Antioxidant activity and eco-geographical correlation analysis of Bairuicao.

- (A) DPPH radical scavenging rates;
(B) Hydroxyl radical inhibition rates;
(C) RDA triplot of climatic factors, antioxidant activities, and biomarkers, The climatic variables used in the redundancy analysis (RDA) are based on the WorldClim Bioclimatic Variables and include:
bio1: Annual mean temperature (°C),
bio2: Mean diurnal range (mean of monthly (max temp – min temp)) (°C),
bio3: Isothermality (bio2/bio7) (*100),
bio4: Temperature seasonality (standard deviation *100) (%),
bio5: Max temperature of warmest month (°C),
bio6: Min temperature of coldest month (°C),
bio8: Mean temperature of wettest quarter (°C),
bio9: Mean temperature of driest quarter (°C);
(D) Pearson correlation heatmap of region-specific metabolites with geo-climatic variables and antioxidant indices.

4. Discussion

This study systematically elucidates the geographical differentiation patterns of secondary metabolites in *Thesium chinense* Turcz. from Anhui, Henan, and Shanxi provinces in China, as well as the underlying interaction mechanisms with ecological environmental factors, through the integration of multidimensional analytical methods. The metabolomic analysis based on LC-MS combined with machine learning algorithms not only confirms the decisive role of geographical environments in shaping the chemical characteristics of medicinal plants, as proposed by the traditional “geo-authentic medicinal materials” theory, but also provides a quantifiable biomarker screening strategy for modern quality control systems of traditional Chinese medicine (TCM). The findings reveal the molecular mechanisms by which complex ecological factors regulate key metabolic pathways to influence plant chemical phenotypes, offering new theoretical insights into the environmental adaptability and evolutionary processes of medicinal plants.

The metabolomic analysis conducted in both positive and negative ion modes detected a total of 506 secondary metabolites, with 289 identified in positive ion mode and 217 in negative ion mode. PCA demonstrated a gradient distribution pattern of samples from the three production regions in the metabolic space. Samples from Henan and Shanxi partially overlapped along the PC1 and PC2 axes, indicating that shared ecological pressures or genetic homogeneity in these adjacent regions, consistent with reports on geographically proximate plant populations exhibiting convergent metabolic traits. From a climatic perspective, the sample collection sites in Anhui Province are situated south of the North-South climatic boundary, corresponding to a subtropical monsoon climate regime. In contrast, Henan and Shanxi Provinces are characterized by temperate monsoon climate conditions, characterized by comparable temperature ranges and precipitation patterns, which may homogenize flavonoid and alkaloid biosynthesis [21–23]. Notably, samples from Anhui exhibited distinct spatial clustering in positive ion mode, clearly distinguishing them from other regions. PLS-DA further validated these findings. This observation aligns with previous studies demonstrating that environmental factors such as altitude, temperature, and soil properties significantly influence secondary metabolite biosynthesis in medicinal plants.

The heatmap and enrichment analysis further identified 43 differential metabolites, dominated by flavonoids (19), alkaloids (12), and phenylpropanoids (8), and ten enriched pathways, with flavone/flavonol biosynthesis ranking highest. Flavonoids are well-documented for their antioxidant, anti-inflammatory, and neuroprotective properties [24–26], which correlate with the traditional use of *Thesium chinense* in treating inflammatory disorders. The prominence of indole alkaloid pathways is equally significant, as these compounds exhibit antitumor and antimicrobial activities. Interestingly, the bidirectional regulation of kaempferol glycosides (e.g., kaempferol 7-neohesperidoside vs. kaempferol 3-O-beta-sophoroside) in Anhui samples implies tissue-specific enzymatic modifications, possibly influenced by regional soil micronutrients such as zinc or selenium, which modulate glycosyltransferase activity. This phenomenon underscores the plasticity of plant secondary metabolism in adapting to localized environmental cues, a phenomenon documented in other medicinal plants [27] such as *Ginkgo biloba*.

The comparative PLS-DA models (Anhui vs. others, Shanxi vs. others, Henan vs. others) and metabolite set enrichment analysis identified distinct regional metabolic signatures, with Anhui samples exhibiting the most unique chemical characteristics. Myristicin—a phenylpropene with known acetylcholinesterase inhibitory effects [28]—was enriched in Anhui but downregulated in Shanxi. This contrast may reflect differences in soil nitrogen availability, as phenylpropanoid biosynthesis is nitrogen-dependent [29]. Similarly, the reversal of ferulic acid trends between Anhui (downregulated) and Shanxi (upregulated) could be attributed to ultraviolet (UV) radiation intensity, which stimulates ferulic acid synthesis as a photoprotective mechanism [30]. Shanxi's higher altitude and stronger UV exposure align with this hypothesis. Glechomafuran and Cyclomethyltryptophan were downregulated in Anhui samples but upregulated in Shanxi populations. This cross-regional expression divergence likely reflects adaptive responses to local climatic conditions, particularly temperature fluctuations and precipitation patterns. Similar geographical regulation of secondary metabolites has been observed in *Saussurea involucreata* leaves, where low-temperature environments enhance phenolic compound accumulation [31]. The reverse regulation of 6-Pentyl-2H-pyran-2-one in Anhui and Shanxi samples highlights the complexity of ecological adaptation. This compound, known for its antimicrobial properties [32], may serve as a stress-response metabolite under contrasting environmental pressures. The isomer-specific regulation of kaempferol glycosides further highlights the role of environmental factors in steering metabolic branching. The 82% decrease in kaempferol 7-neohesperidoside and 3.2-fold increase in its isomer in Anhui samples suggest that temperature fluctuations may influence glycosyltransferase specificity. Additionally, the consistent significance of 13- α -(21)-Epoxyeurycomanone across all regions positions it as a universal biomarker for *Thesium chinense*.

The integration of Random Forest and Lasso regression algorithms identified overlapping biomarkers with high diagnostic value for geographical discrimination. This dual-algorithm

approach outperforms traditional univariate analyses by capturing nonlinear relationships and reducing overfitting risks. Among the Anhui markers, Kaempferol 3-O-beta-sophoroside, 13-alpha-(21)-Epoxyeurycomanone, Myristicin, and Glechomafuran exhibited significant regional specificity. Myristicin and Glechomafuran were significantly enriched in neuroactive ligand-receptor interaction and calcium signaling pathways, suggesting their roles in modulating neurotransmitter regulation and calcium homeostasis through genes like CHRM2/CHRM3/CHRM4. Among the six intersection compounds in Henan, the specific accumulation of 7-Hydroxy-3-(2-hydroxy-propyl)-5-methyl-isochromen-1-one and 4-(Tridecanoylamino)benzoic acid may reflect the diversification of secondary metabolic pathways due to complex terrain. Kaempferol 7-neohesperidoside and Myristoylcarnitine were primarily enriched in lipid metabolism-related pathways, implicating their roles in regulating lipid metabolism via targets such as ADORA1/PTGFR/PTGS2. This aligns with the lower annual precipitation in Henan, as drought stress often elevates lipid-derived metabolites to maintain membrane integrity. Among the three intersection compounds in Shanxi, the significant upregulation of Rhamnetin 3-sophoroside and 6-Pentyl-2H-pyran-2-one affected EGFR tyrosine kinase inhibitor resistance and PI3K-Akt signaling pathways by modulating kinase genes like AKT1/PRKACA/GSK3B, likely driven by high-altitude oxidative stress. Cross-regional comparisons revealed that Kaempferol 3-O-beta-sophoroside, Glechomafuran, and Myristoylcarnitine were consistently identified across multiple models, demonstrating robust regional differentiation capabilities. In the multi-regional “component-target-pathway” integrated network, the NOS3 gene played a central role in pathway enrichment across all three regions, accounting for 62.3% of cross-regional targets. However, its regulatory weight in the AGE-RAGE pathway associated with diabetic complications varied significantly across regions: 34 targets in Anhui, 21 in Henan, and 16 in Shanxi. This variation suggests differences in the functional mechanisms of NOS3 among regions, but its status as a potential key regulatory hub for cross-regional synergy remains noteworthy.

RDA and Pearson correlation models elucidated the climatic and topographical drivers of antioxidant activity. Anhui’s superior DPPH and hydroxyl radical scavenging rates were tightly linked to Rhamnetin 3-sophoroside and 6-Pentyl-2H-pyran-2-one, which correlated positively with annual precipitation (bio12) and negatively with temperature seasonality (bio4). This mirrors findings in *Camellia sinensis*, where stable rainfall enhances phenolic accumulation [33]. Conversely, Shanxi’s reduced hydroxyl activity under precipitation fluctuations (bio4) may reflect trade-offs between antioxidant synthesis and drought resilience mechanisms. The inverse correlation between temperature variables (bio1, bio8) and DPPH activity underscores the role of cool climates in promoting flavonoid biosynthesis, as low temperatures upregulate phenylalanine ammonia-lyase (PAL) activity [34]. Similarly, the altitude-dependent enrichment of Myristoylcarnitine in Shanxi ($r = 0.63$) suggests adaptive lipid metabolism to counter hypoxia-induced oxidative stress, a phenomenon observed in high-altitude medicinal plants. The negative correlation between Kaempferol 7-neohesperidoside in Henan samples and diurnal temperature variation suggests the regulatory effect of temperature stability on flavonoid glycosyltransferase activity.

The quantitative integration of climate variables with specific metabolites provides predictive models for cultivation optimization, enhanced by the RF-LASSO pipeline’s superior specificity in reducing false-positive biomarker identification. However, limitations include restricted sampling across three provinces and single-year data collection, necessitating broader geographic and longitudinal studies to assess biomarker universality and climate change impacts on metabolite stability, alongside incorporation of rhizosphere microbiome analyses through metagenomics to elucidate plant-microbe co-metabolism effects. For quality control and cultivation optimization, actionable strategies emerge: Anhui Province should prioritize low-temperature/high-precipitation microclimates to maximize neuroactive flavonoids (e.g., Kaempferol glycosides), Shanxi Province can enhance alkamide production (Myristoylcarnitine) by targeting high-soil organic carbon zones, and Henan Province requires diurnal temperature management to balance lipid-regulating and antioxidant metabolites, with stable biomarkers like 13-alpha-Epoxyeurycomanone serving as key quality indicators.

4. Conclusions

This study systematically elucidates the geographical chemodiversity of *Thesium chinense* and its ecological determinants through an integrated approach combining metabolomics, machine learning, and ecological modeling. Using liquid chromatography-mass spectrometry coupled with multivariate statistical analysis, researchers revealed significant metabolic profile differentiation among samples from three production regions and identified region-specific biomarkers that establish quantitative authentication indices for geographical origin tracing. Machine learning algorithms screened core metabolic markers enabling precise differentiation of cultivation sources, while RDA models and Pearson correlation analyses demonstrated strong associations between metabolomic profiles and regional climatic/edaphic factors, particularly uncovering climate-driven regulation of antioxidant activity. These findings not only validate the scientific basis of traditional geo-authentic medicinal material theory but also provide a molecular framework for precision evaluation and targeted utilization of TCM resources. The integration of spatial metabolomics with environmental variables establishes a systems biology precedent for geo-herbalism research, offering theoretical support for molecular breeding programs and quality control standards. Future studies should employ multi-omics approaches to decipher environmental stress response mechanisms in metabolic pathways, thereby enhancing resource development strategies for ecologically sustainable utilization of medicinal plants.

Author Contributions: Conceptualization, Junsong Wang; methodology, Cong Wang and Ke Che; validation, Cong Wang and Guanglei Zhang; formal analysis, Cong Wang; writing—original draft preparation, Cong Wang and Hao Yu; writing—review and editing, Cong Wang and Junsong Wang; visualization, Cong Wang and Junsong Wang; supervision, Junsong Wang; funding acquisition, Cong Wang and Junsong Wang. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by the National Natural Science Foundation of China (Grant No. 81773857) and the Postgraduate Research & Practice Innovation Program of Jiangsu Province (KYCX24_0689).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in the study are included in the article; further inquiries can be directed to the corresponding author.

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

Abbreviations

The following abbreviations are used in this manuscript:

TCM	Traditional Chinese Medicine
LC-MS	Liquid Chromatography-Mass Spectrometry
UPLC-QTOF-MS	Ultra-Performance Liquid Chromatography-Quadrupole Time-of-Flight Mass Spectrometry
PCA	Principal Component Analysis
PLS-DA	Partial Least Squares Discriminant Analysis
VIP	Variable Importance in Projection
RF	Random Forest
LASSO	Least Absolute Shrinkage and Selection Operator
RDA	Redundancy Analysis
DPPH	2
HMDB	Human Metabolome Database
KEGG	Kyoto Encyclopedia of Genes and Genomes
GO	Gene Ontology
ANOVA	Analysis of Variance

ROS	Reactive Oxygen Species
NOS3	Nitric Oxide Synthase 3
AGE-RAGE	Advanced Glycation End Products-Receptor for Advanced Glycation End Products
EGFR	Epidermal Growth Factor Receptor
PI3K-Akt	Phosphatidylinositol 3-Kinase-Protein Kinase B
ADORA1	Adenosine A1 Receptor
PTGFR	Prostaglandin F2α Receptor
PTGS2	Prostaglandin-Endoperoxide Synthase 2
AKT1	Protein Kinase B Alpha
PRKACA	Protein Kinase A Catalytic Subunit Alpha
GSK3B	Glycogen Synthase Kinase 3 Beta
CHRM2/CHRM3 /CHRM4	Cholinergic Receptor Muscarinic 2/3/4
HTR2A	5-Hydroxytryptamine Receptor 2A
DRD2	Dopamine Receptor D2
CHRM3	Cholinergic Receptor Muscarinic 3

References

1. Li G, Guan Y, Xu L, et al. Total alkaloids from *Thesium chinense* inhibit lipopolysaccharide-induced respiratory inflammation by modulating Nrf2/NF-κB/NLRP3 signaling pathway. *Chin J Nat Med.* 2025;23(4):421-430.

2. Xu H, Zhang F, Che Y, et al. *Thesium chinense* Turcz. and its compound astragaloside alleviate lipopolysaccharide-induced acute lung injury via the PI3K/AKT/p53 signaling pathway. *J Ethnopharmacol.* 2025;346:119691.

3. Xu H, Zhang F, Huang Y, et al. *Thesium chinense* Turcz. alleviates antibiotic-associated diarrhea in mice by modulating gut microbiota structure and regulating the EGFR/PI3K/Akt signaling pathway. *Nan Fang Yi Ke Da Xue Xue Bao.* 2025;45(2):285-295.

4. Shao L, Ling B, Jiang H, et al. The emulsifying capacity and mechanism of *Thesium chinense* Turcz. polysaccharides based on synergistic effect of insoluble and soluble fractions. *Int J Biol Macromol.* 2025;302:140545.

5. Zhang F, Zhang G, Wang C, et al. Geographical variation in metabolite profiles and bioactivity of *Thesium chinense* Turcz. revealed by UPLC-Q-TOF-MS-based metabolomics. *Front Plant Sci.* 2025;15:1471729.

6. Zhang S, Chen H, Hua J, et al. Cytotoxic activity of phenolic compounds in *Bairui* Granules obtained from the Chinese medicinal plant *Thesium chinense*. *Front Chem.* 2024;12:1506792.

7. Ling B, Shao L, Jiang H, et al. Wide pH, Adaptable High Internal Phase Pickering Emulsion Stabilized by a Crude Polysaccharide from *Thesium chinense* Turcz. *Molecules.* 2024;29(18):4312.

8. Liu MJ, Xu ZP, Guan YQ, et al. Ethyl acetate fraction of *Thesium chinense* Turcz. alleviates chronic obstructive pulmonary disease through inhibition of ferroptosis mediated by activating Nrf2/SLC7A11/GPX4 axis. *J Ethnopharmacol.* 2025;337(Pt 1):118776.

9. Wang R, Ding A, Wang J, et al. Astragaloside from *Thesium chinense*: A Novel Anti-Aging and Antioxidant Agent Targeting IGF1R/CD38/Sirtuins. *Antioxidants (Basel).* 2024;13(7):859.

10. Peng GC, Hao JH, Guan YQ, et al. Systematic investigation of the material basis, effectiveness and safety of *Thesium chinense* Turcz. and its preparation *Bairui* Granules against lung inflammation. *Chin Med.* 2024;19(1):67.

11. Lê Cao KA, Boitard S, Besse P. Sparse PLS discriminant analysis: biologically relevant feature selection and graphical displays for multiclass problems. *BMC Bioinformatics.* 2011;12:253.

12. Shen H, Huang J Z. Sparse principal component analysis via regularized low rank matrix approximation. *Journal of multivariate analysis,* 2008, 99(6): 1015-1034.

13. Tiozon RJN, Sreenivasulu N, Alseekh S, et al. Metabolomics and machine learning technique revealed that germination enhances the multi-nutritional properties of pigmented rice. *Commun Biol.* 2023;6(1):1000.

14. Vergara-Diaz O, Vatter T, Vicente R, et al. Metabolome Profiling Supports the Key Role of the Spike in Wheat Yield Performance. *Cells*. 2020;9(4):1025.
15. Pang Z, Chong J, Zhou G, et al. MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. *Nucleic Acids Res*. 2021;49(W1):W388-W396.
16. Kanehisa M, Furumichi M, Sato Y, et al. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Res*. 2023;51(D1):D587-D592.
17. Nan YD, Mu BD, Ge CX, et al. Exploring the novel antioxidant peptides in low-salt dry-cured ham: Preparation, purification, identification and molecular docking. *Food Chem*. 2024;446:138697.
18. Yan M, Zhao Y, Feng S, et al. Hydroxyl group-induced enhancement of antioxidant activity of resveratrol over pterostilbene by binding to lactoferrin. *Food Chem*. 2024;441:138356.
19. Fick S E, Hijmans R J. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International journal of climatology*, 2017, 37(12): 4302-4315.
20. Booth T H, Nix H A, Busby J R, et al. BIOCLIM: the first species distribution modelling package, its early applications and relevance to most current MAXENT studies. *Diversity and distributions*, 2014, 20(1): 1-9.
21. Wang X, Cao J, Tian L, et al. Elucidating Flavonoid and Antioxidant Activity in Edible and Medicinal Herbs *Woodwardia japonica* (L.f.) Sm. Based on HPLC-ESI-TOF-MS and Artificial Neural Network Model: Response to Climatic Factors. *Molecules*. 2023;28(4):1985.
22. Tong L, Jiang Y, Zhang X, et al. Metabolic and molecular basis of flavonoid biosynthesis in *Lycii fructus*: An integration of metabolomic and transcriptomic analysis. *J Pharm Biomed Anal*. 2025;255:116653.
23. Guo J, Zhou X, Wang T, Wang G, Cao F. Regulation of flavonoid metabolism in ginkgo leaves in response to different day-night temperature combinations. *Plant Physiol Biochem*. 2020;147:133-140.
24. Maleki SJ, Crespo JF, Cabanillas B. Anti-inflammatory effects of flavonoids. *Food Chem*. 2019;299:125124.
25. Shen N, Wang T, Gan Q, et al. Plant flavonoids: Classification, distribution, biosynthesis, and antioxidant activity. *Food Chem*. 2022;383:132531.
26. Calderaro A, Patanè GT, Tellone E, et al. The Neuroprotective Potentiality of Flavonoids on Alzheimer's Disease. *Int J Mol Sci*. 2022;23(23):14835.
27. Su X, Shen G, Di S, et al. Characterization of UGT716A1 as a Multi-substrate UDP:Flavonoid Glucosyltransferase Gene in *Ginkgo biloba*. *Front Plant Sci*. 2017;8:2085.
28. Seo SM, Jung CS, Kang J, et al. Larvicidal and acetylcholinesterase inhibitory activities of apiaceae plant essential oils and their constituents against *aedes albopictus* and formulation development. *J Agric Food Chem*. 2015;63(45):9977-9986.
29. Qiao M, Sun R, Wang Z, et al. Legume rhizodeposition promotes nitrogen fixation by soil microbiota under crop diversification. *Nat Commun*. 2024;15(1):2924.
30. Bhatia N. Polypodium leucotomos: a potential new photoprotective agent. *Am J Clin Dermatol*. 2015;16(2):73-79.
31. Sun Q, Ma L, Zhu X. Metabolomics-based exploration the response mechanisms of *Saussurea involucrata* leaves under different levels of low temperature stress. *BMC Genomics*. 2023;24(1):297.
32. Liu M, Niu Q, Wang Z, et al. Comparative physiological and transcriptome analysis provide insights into the inhibitory effect of 6-pentyl-2H-pyran-2-one on *Clarireedia jacksonii*. *Pestic Biochem Physiol*. 2023;193:105456.
33. Yao L, Caffin N, D'Arcy B, et al. Seasonal variations of phenolic compounds in Australia-grown tea (*Camellia sinensis*). *J Agric Food Chem*. 2005;53(16):6477-6483.
34. Wang N, Zhu Y, Zhou Y, Gao F, Cui S. Transcriptome Analysis Reveals the Crucial Role of Phenylalanine Ammonia-Lyase in Low Temperature Response in *Ammopiptanthus mongolicus*. *Genes (Basel)*. 2024;15(11):1465.

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.