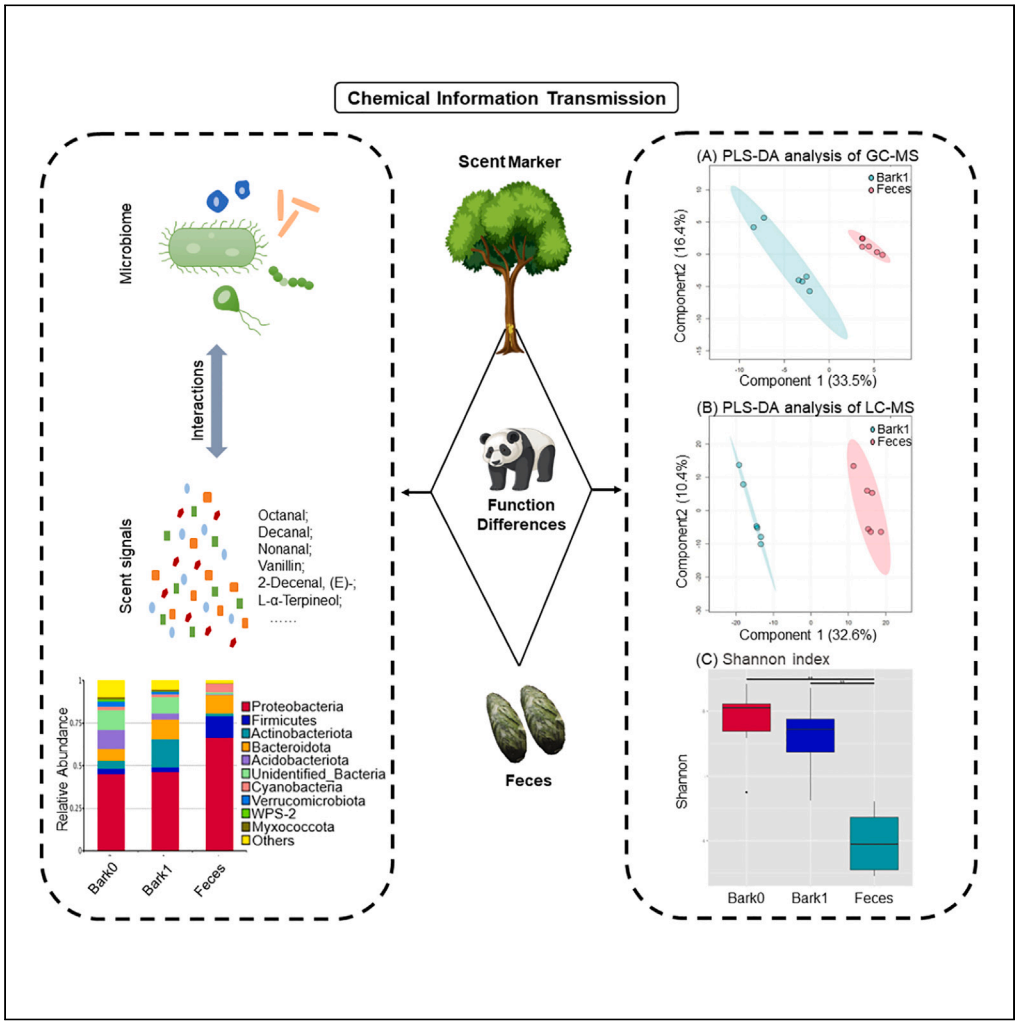
iScience

CellPress
OPEN ACCESS

Article

Characterizing the metabolome and microbiome at giant panda scent marking sites during the mating season



Jin Hou, Vanessa Hull, Masanori Fujimoto, ..., Thomas Connor, Dunwu Qi, Jindong Zhang

zhangjd224@cwnu.edu.cn

Highlights

Characteristics of metabolome and microbiome of scent markers were revealed

The feces of pandas do not share the same functions as scent markers

Crucial scent compounds and microbes were unrevealed

Hou et al., iScience 27, 110051
June 21, 2024
https://doi.org/10.1016/j.isci.2024.110051

iScience

CellPress
OPEN ACCESS

Article

Characterizing the metabolome and microbiome at giant panda scent marking sites during the mating season

Jin Hou,[1,2] Vanessa Hull,[3] Masanori Fujimoto,[4] Zejun Zhang,[1] Xiaoyuan Chen,[1] Shiyu Chen,[1] Rui Chen,[1] Thomas Connor,[5] Dunwu Qi,[6] and Jindong Zhang[1,7,*]

SUMMARY

Scent marking sites served as a primary means of chemical communication for giant pandas, enabling intraspecific communication. We integrated metabolomics and high-throughput sequencing techniques to examine the non-targeted metabolome and microbial community structure of scent marking sites and feces in the field. Integrative analysis revealed a more comprehensive array of chemical compounds compared to previous investigations, including ketones, acids, heterocycles, alcohols, and aldehydes. Notably, specific compounds such as 2-decenal, (E)-, octanal, decanal, L-α-terpineol, vanillin, and nonanal emerged as potential key players in scent signaling. Intriguingly, our study of the microbial domain identified dominant bacterial species from the Actinobacteria, Bacteroidetes, and Proteobacteria phyla, likely orchestrating metabolic processes at scent marking sites. Comparative analyses showed, for the first time, that feces do not share the same functions as scent markers, indicating distinct functional roles. This research deepens scientific understanding of chemical communication in wild pandas.

INTRODUCTION

Olfactory signaling plays a vital role in communication within the natural world. Animals employ these chemical cues to convey information about a broad range of behaviors and physiological states, including mating, aggression, territoriality, and alarm.[1,2] This mode of information transfer is particularly significant for solitary species as it enables communication across long distances and in environments where visual or auditory signals may be challenging to detect.[3] Generally, the chemical odorants responsible for chemical communication can be produced as by-products of biochemical pathways by the sender. However, multiple animal studies have demonstrated that host associated microbial symbionts indeed mediate signals synthesis.[4–7] A variety of chemical components are associated with potential pheromones, including volatile fatty acids, hydrocarbons, peptides, phenols, ketones, and others. These studies strongly support the fermentative hypothesis and underscore the irreplaceable role of microbes.[2,8]

In the case of the threatened solitary species, the giant panda (*Ailuropoda melanoleuca*), olfactory communication further enhances social connections among individuals, enabling them to successfully navigate essential life processes.[9] Research on scent communication by giant pandas has garnered significant attention. Existing studies have revealed the chemical composition profile of scent marks.[10–12] Pandas can discern age, sex, individual identity, and even kinship through volatile compounds.[13–17] Furthermore, diverse behavioral strategies have been reported in their scent marking, showcasing their adaptive capacity to the environment and signal detection.[18,19] In summary, a substantial body of relevant knowledge indicates that chemical communication is a key component in social interactions among giant pandas.

Reproductive activity undoubtedly ranks among the primary concerns for panda conservation. Studies conducted on giant pandas have indicated that reproductive success is greatly influenced by chemical signals and odors emitted by the opposite sex during the mating season.[20,21] Similar to many other species, symbiotic bacteria present in samples collected during the mating period have been found to participate in the production of volatile components in giant panda scent marks, although the causal relationship between microorganisms and the compounds involved is still not fully understood.[12,21]

Based on previous prospective studies, however, three aspects warrant in-depth consideration in the context of wild pandas. Firstly, in general, giant panda deposit chemical signals on bark surfaces using two types of scent marks: anogenital gland secretions (AGS) and

[1]Key Laboratory of Southwest China Wildlife Resources Conservation (Ministry of Education), China West Normal University, Nanchong, Sichuan Province 637002, China
[2]College of Life Sciences, Beijing Normal University, Beijing 100875, China
[3]Department of Wildlife Ecology and Conservation, University of Florida, Gainesville, FL 32611, USA
[4]Soil and Water Sciences Department, University of Florida, Gainesville, FL 32611, USA
[5]Department of Environmental Science, Policy, and Management, University of California, Berkeley, Berkeley, CA 94720, USA
[6]Chengdu Research Base of Giant Panda Breeding, Chengdu, Sichuan Province 610081, China
[7]Lead contact
*Correspondence: zhangjd224@cwnu.edu.cn
https://doi.org/10.1016/j.isci.2024.110051

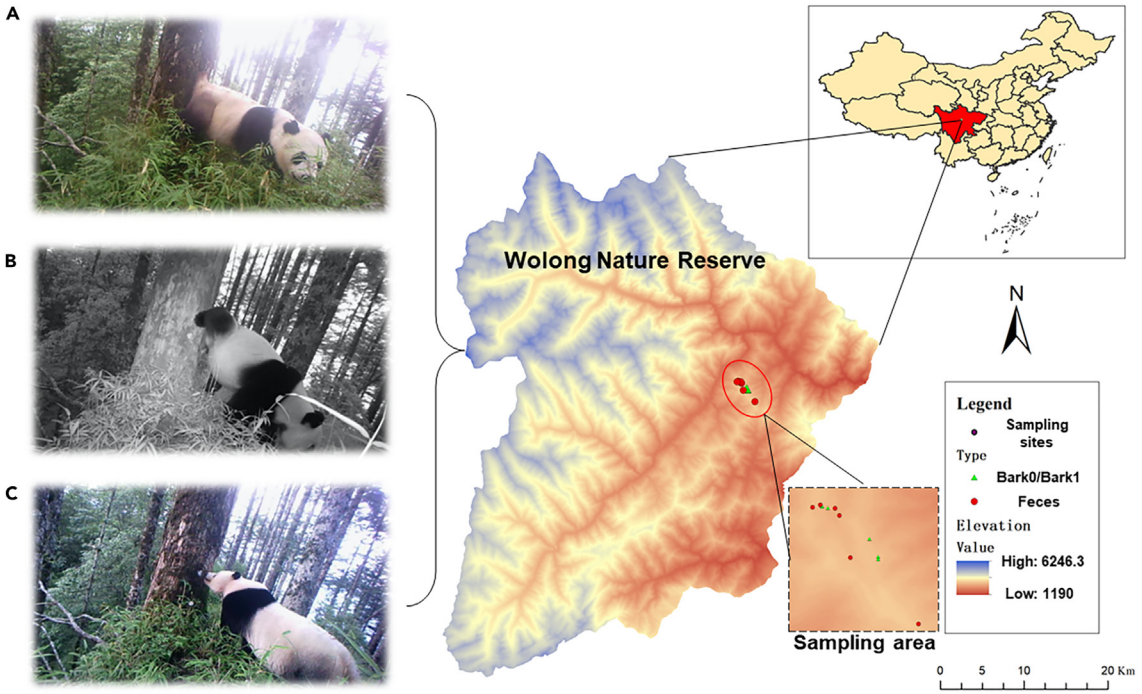


Figure 1. The study area in Wolong Nature Reserve, China

(A–C) represent three different scent communication behaviors by giant pandas at a same scent marking site, including anogenital gland secretion marking, urine marking, and olfactory investigation, respectively.

urine.[18,19] Therefore, the chemical components present at scent marking sites are derived from two sources in wild pandas, which is a common phenomenon (Figure 1). This implies that the receiver of these signals must extract information from the mixture of AGS and urine rather than from one of them individually in the field (Figure 1). This phenomenon can be confirmed by infrared camera traps (unpublished data). As shown in Figure 1 in images derived from camera traps, the same individual panda will employ different postures and signaling substances (AGS and urine) to scent mark at the same scent marking site. However, current studies have focused more on identifying compounds in single AGS or urine samples, and the microbial composition and their roles in scent communications at wild scent marking sites remains poorly understood. Secondly, while feces are often considered as a scent mark in other carnivores,[22] past observations suggest that they do not serve as a chemical signal source for giant pandas.[9,18] Nevertheless, our own unpublished data indicates that wild pandas do indeed investigate the feces of completely unfamiliar individuals for several seconds. It is intriguing to further investigate whether the function of feces is consistent with the function of scent marks for giant pandas by exploring the differences in metabolites and microbial community composition between feces and the two known chemical signal carriers (i.e., AGS and urine). Thirdly, since chemical signals are metabolites that have originated from the sender and are also mediated by bacteria, with the gland and the environment acting as sources,[8] an overall exploration of metabolites on scent marking sites is necessary to better understand the characteristics of scent marks as they are perceived by other individuals. Furthermore, previous studies have primarily focused on volatile compounds while neglecting semi-volatile and non-volatile compounds due to the influence of different research objectives and different extraction and injection methods for single gas chromatography-mass spectrometry (GC-MS) or liquid chromatography-mass spectrometry (LC-MS). In reality, both the production and emission of volatile compounds are closely related to the two types of constituents. For instance, non-volatile compounds can be converted into volatile compounds through a process known as derivatization, which can occur independently or rely on microbial activity.[23] Derivatization involves chemically modifying a non-volatile compound to form a new compound that is more volatile and can be analyzed using gas chromatography (GC) or other techniques that rely on the analyte's volatility. Limited knowledge exists regarding compounds that are difficult to identify using a single method. Given these two gaps in understanding, a more comprehensive study is warranted.

In this study, we have employed an integrated approach utilizing GC-MS/LC-MS and 16S rRNA amplicon sequencing to shed light on both metabolites and microbial community composition of scent marking sites and feces of wild giant pandas during the mating season. The findings from this study will contribute new knowledge to our understanding of scent communication in this species of concern.

RESULTS

Composition of chemical constituents in scent marks and feces identified by GC-MS and LC-MS

We tentatively identified a total of 23 compounds in the Bark1 group using GC-MS, while LC-MS analysis revealed 82 different compounds. Among the compounds identified with GC-MS in the scent marking samples, there were 7 aldehydes, 5 alcohols, 3 ketones, 2 alkenes, 2

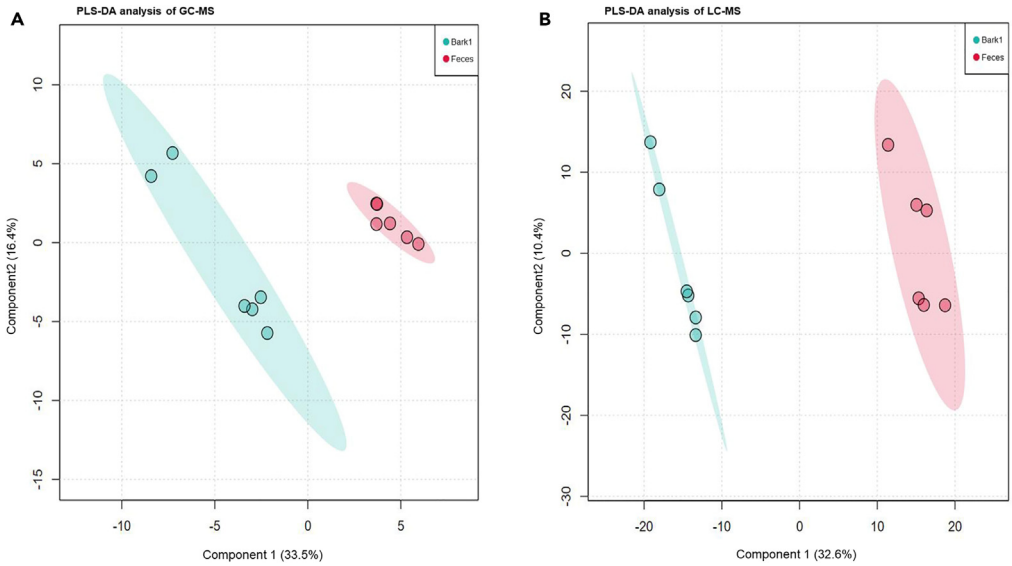


Figure 2. Partial least squares discriminant analysis of metabolites

(A) Partial least squares discriminant analysis of metabolites detected by GC-MS.

(B) Partial least squares discriminant analysis of metabolites detected by LC-MS. Bark1: samples with scents; feces: panda feces collected in the same area as the scent markings.

alkanes, and 1 heterocycle (Tables S1 and S2). In the Bark1 group, completely different compounds were found using LC-MS, with a total of 8 types of constituents. We identified 30 ketones, 18 acids, 16 heterocycles, 8 alcohols, 4 esters, 3 aldehydes, 2 phenols, and 1 alkene in the Bark1 group. It is worth noting that all of the compounds identified in the GC-MS analysis were volatile compounds, whereas the LC-MS analysis detected 42 volatile compounds and 40 semi-volatile compounds (Tables S1 and S2).

Regarding the feces group, a total of 29 compounds were detected using GC-MS, while LC-MS analysis revealed 207 compounds. The GC-MS analysis identified 9 types of chemical compounds, including 8 ketones, 7 esters, 6 heterocycles, 2 alcohols, 2 terpenoids, 1 alkene, 1 alkane, 1 phenol, and 1 acid. On the other hand, the LC-MS analysis identified 8 types of chemical compounds, which consisted of 68 acids, 61 heterocycles, 32 ketones, 16 esters, 15 alcohols, 8 phenols, 6 aldehydes, and 1 amine. In terms of volatility, the GC-MS analysis detected 19 volatile compounds and 10 semi-volatile compounds, while the LC-MS analysis detected 138 volatile compounds and 69 semi-volatile compounds (Tables S3 and S4).

The results of the partial least squares discriminant analysis (PLS-DA) analysis demonstrated significant differences between the aggregation positions of the metabolites collected from the scent marks and feces, as depicted in Figure 2.

Distribution of scent mark and feces microbial community structure

The operational taxonomic units (OTU) cluster analysis and Venn diagram revealed the number of OTUs in common among the three groups was 1,614. The number of unique OTUs in Bark1 was 1,134 compared to Bark0 and was 1,878 in Bark1 compared to feces. The number of OTUs unique in feces was 671 compared to Bark0 and 595 in feces compared to Bark1. The number of common OTUs shared between feces and Bark1 was higher than that shared between feces and Bark0 (Figure 3A). As shown in Figure 3B, the rarefaction curves of different samples tended to be smooth, indicating that the amount of sequencing data was reasonable and reflected the microbial diversity in the samples. The principal co-ordinates analysis (PCoA) results showed distinct changes in the aggregation positions of the Bark0 group, Bark1 group, and feces group, while the aggregation positions of Bark0 group and Bark1 group changed only slightly compared with the feces group (Figure 3C).

Alpha diversity analysis reflected that the microbial structure of bark changed after giant pandas deposited their chemosignals. Shannon index of the feces group was significantly lower than that of Bark0 and Bark1 groups, while Shannon index of the Bark1 group was lower than that of Bark0 (Figure 3D). The Simpson index also showed the similar trend (Figure 3E).

The Anosim analysis showed significant difference between feces vs. Bark1($R = 0.84$, $p < 0.01$), Bark0 vs. Bark1($R = 0.17$, $p < 0.05$), and Bark0 vs. Feces ($R = 0.72$, $p < 0.01$).

The stacked bar plot results of the top 10 taxa at the phylum classification level in the different groups are shown in Figure 3F. The *Proteobacteria*, *Acidobacteria*, and *unidentified_Bacteria* were most abundant in Bark0 group. The *Proteobacteria*, *Actinobacteria*, and *Bacteroidetes* were dominant species in the Bark1 group. The microbial species in the feces group mainly consisted of *Proteobacteria*, *Firmicutes*, and *Bacteroidetes*. The unweighted pair-group method of analysis (UPGMA) cluster analysis revealed that a higher structure similarity was found in the Bark0 and Bark1 groups than the feces group (Figure 3F).

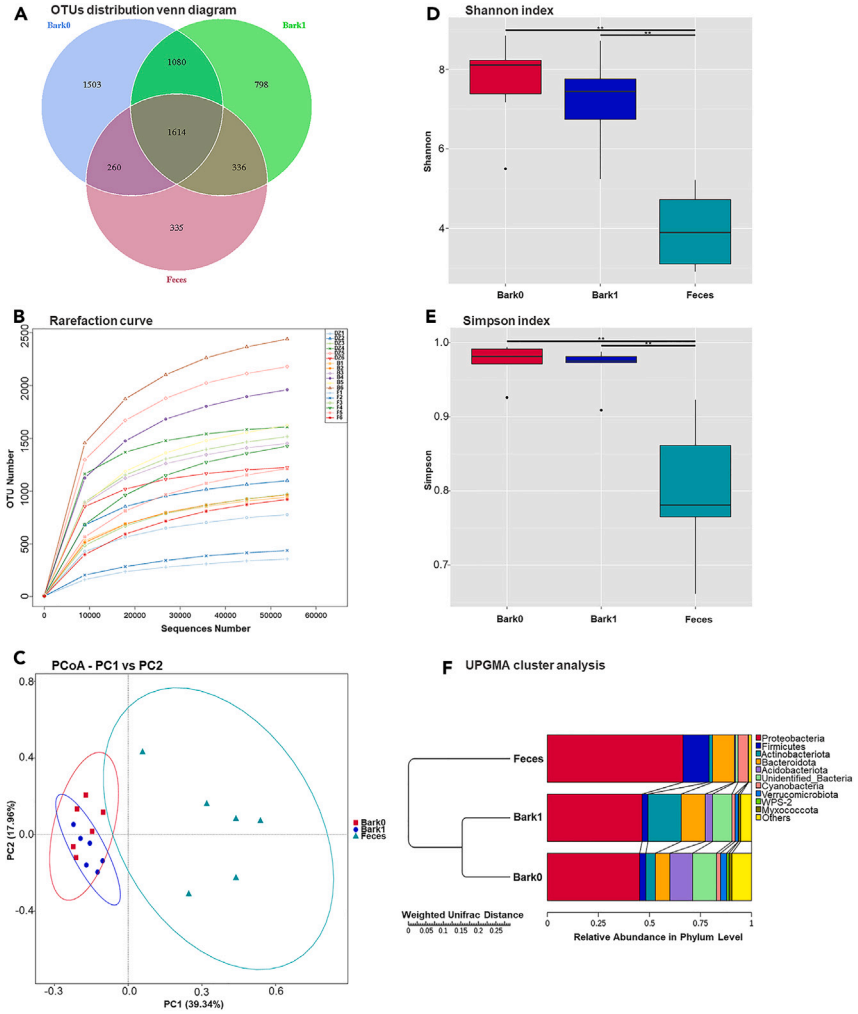


Figure 3. Microbial community diversity difference analysis

(A) Venn diagram of the quantity of operational taxonomic units (OTUs) in bacteria from Bark0, Bark1, and feces groups.

(B) Rarefaction curve of bacterial community.

(C) Principal co-ordinates analysis of bacterial community diversity. Microbial diversity alpha analysis.

(D and E) (D) Shannon index and (E) Simpson index.

(F) Analysis of unweighted pair-group method with arithmetic mean. Note: Bark0 represents controlled samples without scent marks. Bark1 represents samples from scent marking sites. Feces represents panda feces collected from the surrounding area of scent marking sites.

Dominant microorganisms in scent marks and feces

The result of the analysis of the contribution of different species to the difference is shown in Figure 4 (linear discriminant analysis (LDA) score >4). The Bark0 group contained 19 rich bacterial branches: *Chitinophagaceae* (class), *Rhizobiales* (order), *Acidobacteriae* (class), *Deltaproteobacteria* (class), *Beijerinckiaceae* (family), *Alphaproteobacteria* (class), *Xanthobacteraceae* (family), *Bradyrhizobium* (genus), *Methylobacterium*_Methylorubrum (genus), *Verrucomicrobiota* (phylum), *Verrucomicrobiae* (class), *Caulobacterales* (order), *Caulobacteraceae* (family), *Acetobacterales* (order), *Acetobacteraceae* (family), *Chthoniobacterales* (order), *Chthoniobacteraceae* (family), *Chitinophagales* (order), and *Chitinophagaceae* (family).

The Bark1 group contained 18 abundant bacterial branches: *Actinobacteriota* (phylum), *Actinobacteria* (class), *Micrococcales* (order), *Sphingomonadaceae* (family), *Sphingomonadales* (order), *Micrococcaceae* (family), *Sphingomonas* (genus), *Arthrobacter* (genus), *Xanthomonadales* (order), *Rhodanobacteraceae* (family), *Rhodanobacter* (genus), *Mucilaginibacter* (genus), *Bacteroidia* (class), *Rhodanobacter*_sp (species), *Chitinophagales* (order), *Chitinophagaceae* (family), *Polaromonas* (genus), and *Cytophagales* (order).

The feces group contained 10 abundant bacterial branches: *Gammaproteobacteria* (class), *Pseudomonadales* (order), *Pseudomonas* (genus), *Pseudomonadaceae* (family), *Enterobacterales* (order), *Yersiniaceae* (family), *Yersinia* (genus), *Yersinia*_enterocolitica (species), *Pedobacter* (genus), and *Pedobacter*_sp_PACM_27299 (species).

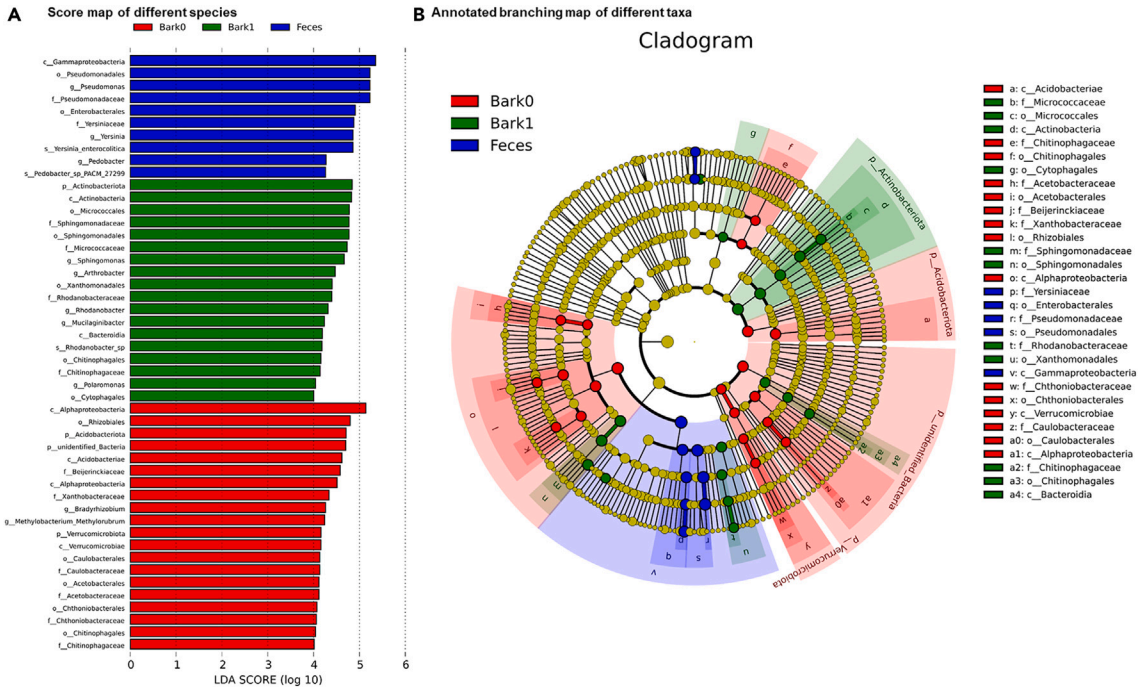


Figure 4. LEfSe analysis results of bacteria

(A) Score map of different taxa.

(B) Annotated branching map of different taxa. Note: In Figure 4A, red, green, and blue color bars represent taxa with a relatively high abundance in Bark0, Bark1, and feces groups, respectively. In Figure 4B, the yellow node represents the taxa with no significant difference in different samples. The node diameter is proportional to the relative abundance, and each layer node represents phylum, class, order, family, and genus from the inside to outside. Bark0: controlled samples without scent marks, Bark1: samples from scent marking sites, Feces: panda feces collected in the same area as scent marks.

Microbes and metabolites correlation analysis in scent marks

In Figure 5, the GC-MS/bacteria pair revealed significant correlations between metabolites and bacterial taxa. *Ahniella* exhibited a positive correlation with vanillin and a negative correlation with nonanal. *Collimonas* showed positive correlations with 3-(3-Hydroxy-4-methyl-phenyl)-3,4,4-trimethyl-cyclopentanone, iso-isovelleral, and decanal. *Emticicia* exhibited a positive correlation with vanillin and a negative correlation with 2-decenal, (E)-. *Lysobacter* displayed a negative correlation with 3-ethyl-2,6,10-trimethylundecane. Additionally, *unidentified_RBG-13-46-9* showed negative correlations with 3-ethyl-2,6,10-trimethylundecane and nonanal, while *unidentified_TSAC18* exhibited negative correlations with nonanal and 3-ethyl-2,6,10-trimethylundecane. *Dokdonella* displayed positive correlations with iso-isovelleral, 3-(3-hydroxy-4-methyl-phenyl)-3,4,4-trimethyl-cyclopentanone, and decanal. Furthermore, *Allobranchiibius* exhibited a positive correlation with 3-ethyl-2,6,10-trimethylundecane, while *Rhodanobacter* displayed positive correlations with 3-ethyl-2,6,10-trimethylundecane and 4-(3,3-dimethyl-but-1-ynyl)-4-hydroxy-2,6,6-trimethylcyclohex-2-enone. Lastly, *Aeromicrobium* exhibited a positive correlation with 4-(4-hydroxy-phenyl)-2-butanone.

In the LC-MS/bacteria pair, significant correlations between metabolites and bacterial taxa were also observed (Figure 6). *Brevundimonas* displayed positive correlations with 11β-hydroxyandrosterone, idebenone, 13,14-dihydro-15-keto prostaglandin A2, 1,4-dihydroxy-1,4-dimethyl-7-(propan-2-ylidene)-decahydroazulen-6-one, 6 β-hydroxycortisol, coenzyme Q2, and epitestosterone. *Collimonas* exhibited positive correlations with tetranor-12(S)-HETE and 8,8-dimethyl-2-phenyl-4H,8H-pyrano[2,3-h] chromen-4-one. *Aeromicrobium* showed a positive correlation with 2,3-dihydroxypropyl 12-methyltridecanoate. Additionally, *unidentified_RBG-13-46-9* exhibited a positive correlation with 4-methoxybenzaldehyde, while *Nitrosospira* displayed a positive correlation with 13,14-dihydro-15-keto-tetranor prostaglandin E2. *P3OB-42* exhibited a positive correlation with (5ξ,9ξ,16ξ)-17-hydroxykauran-19-oic acid. Furthermore, *Dokdonella* displayed a positive correlation with tetranor-12(S)-HETE, and *Actinoplanes* exhibited a positive correlation with glycerol-3-phosphate. Finally, *Halomonas* showed a negative correlation with 2,3-dihydro xypropyl 12-methyltridecanoate.

DISCUSSION

For solitary giant pandas, chemical communication plays an essential role in conveying conspecific information and fostering inter-individual relationships to mitigate the challenges posed by spatial isolation and to attract partners during the mating season.[18] This study represents the first attempt to integrate comprehensive metabolome analysis and high-throughput sequencing to investigate the characteristics of scent marking sites during the mating season. We emphasize that scent marking sites contain a mixture of two different types of scent markers, and

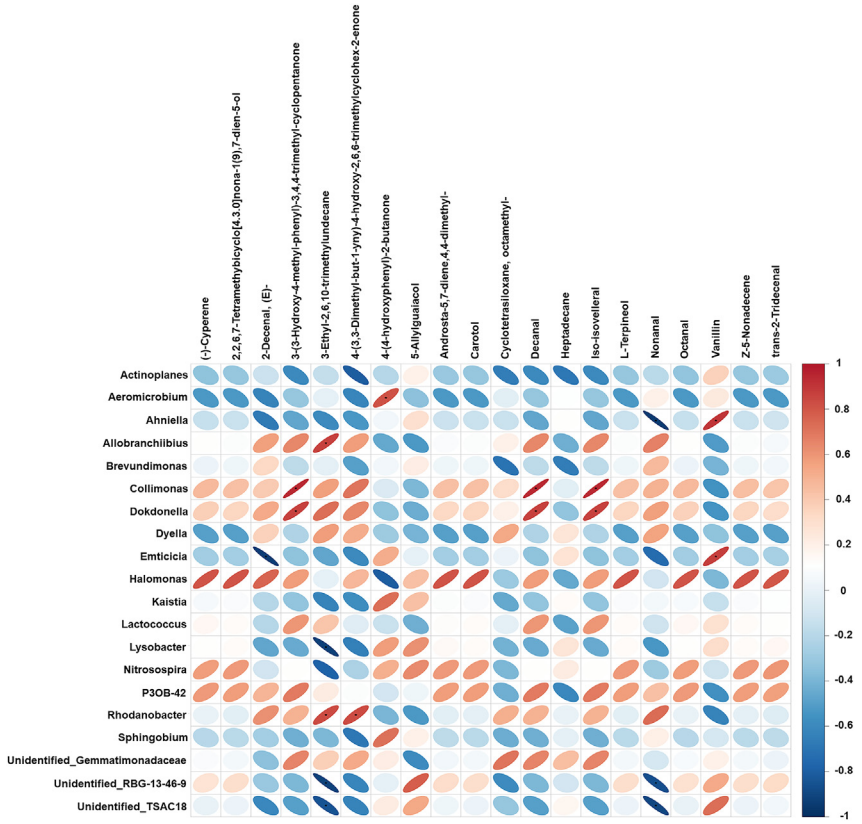


Figure 5. The Pearson correlation heatmap results depicting the top 20 differentially dominant bacteria at the genus level and compounds in GC-MS analysis

In the figure, the vertical axis represents different bacterial genera, while the horizontal axis represents different metabolites. The legend on the right side indicates the correlation coefficient, with stronger positive correlation indicated by a redder color and stronger negative correlation indicated by a bluer color. The flatter the ellipse, the higher the absolute value of correlation it represents. The sections marked with an "*" in the resulting figure indicate statistical significance at a level of $p < 0.05$.

understanding the characteristics of this mixture is more likely to reflect the actual chemical information extracted by giant pandas and in turn promote conservation efforts. The study yielded insights into the compositional features of scent marking sites, including the identification of over 100 metabolites and key microorganisms. Additionally, by comparing the micro characteristics of scent marking sites and feces, we identified significant differences between the two groups. These findings provide valuable insights into chemical communication and the conservation of giant pandas.

Metabolome analysis using both GC-MS and LC-MS revealed extensive metabolite profiles of scent marks. Following the integration of LC-MS analytical techniques, we have identified metabolites in quantities approximately 4-fold greater than those detected by GC-MS. This significantly enriches the chemical composition profile of scent marking trees, making a substantial contribution to the establishment of a more comprehensive metabolomic library of scent markers for the giant panda. The compounds identified include not only those reflecting the physiological status of the giant panda but also substances known as chemical pheromones in other animal studies (e.g., 4-methoxybenzaldehyde and cyclotetrasiloxane octamethyl-),[24,25] which have not been reported in previous research and may represent potential sources of important chemical pheromones specific to the giant panda. Additionally, we discovered a suite of about 40 semi-volatile compounds, which constitute roughly half of the profile, that may be directly detected and utilized by the giant panda or might serve as species-specific signals through further metabolism by microbes.[12] The predominant compound types in scent marks were ketones, acids, heterocycles, alcohols, and aldehydes, with semi-volatile compounds mainly comprising ketones, acids, heterocycles, alcohols, and esters. Previous studies have indicated that airborne pheromones typically consist of 5–20 carbon atoms and can volatilize to reach the receiver.[26] In line with this, our results identified 91 compounds with 5–20 carbon atoms from a total of 105 compounds, which are suitable for signal transmission. Notably, the main compound types found in our study are important sources of chemosignals reported in other species (e.g., aldehydes and heterocycles).[2] Furthermore, we found several compounds that were also identified in previous studies on panda scent marks. Examples include 2-decenal, (E)-, octanal, decanal, L-α-terpineol, vanillin, nonanal, N-methyl (−) ephedrine, and styrene, which have been associated with seasonal variation in scent marks and cues for both individual and sex identification.[13–17] Therefore, these compounds, which have been reported in multiple studies, may be important candidates for chemosignaling. In addition to the compounds found in

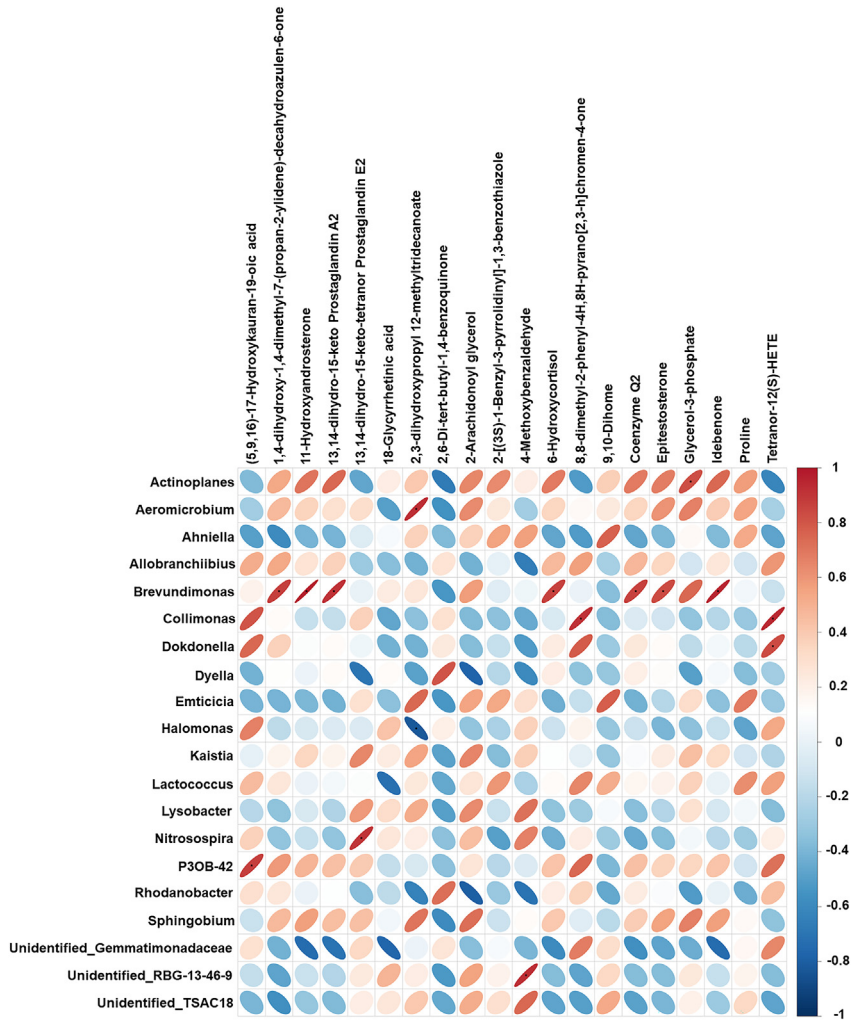


Figure 6. The Pearson correlation heatmap results for the top 20 differentially dominant bacteria at the genus level and compounds in LC-MS analysis

In the figure, the vertical axis represents different bacterial genera, while the horizontal axis represents different metabolites. The legend on the right side indicates the correlation coefficient, with stronger positive correlation indicated by a redder color and stronger negative correlation indicated by a bluer color. The flatter the ellipse, the higher the absolute value of correlation it represents. The sections marked with an "*" in the resulting figure indicate statistical significance at a level of $p < 0.05$.

previous studies, some metabolites may play potential roles in breeding and chemical communication. For example, carotol can be metabolized by microbes to produce a compound with a musky scent.[27] Cyclotetrasiloxane, octamethyl-, has been identified as a bioindicator for estrus detection in cattle.[24] *trans*-Verbenol has been characterized as an aggregation pheromone in the mountain pine beetle (*Dendroctonus ponderosae*) and has also been reported in water buffalo (*Bubalus bubalis*) as a compound that facilitates estrus.[28,29] In addition, recent research by Poirier et al.[25] has established the role of 4-methoxybenzaldehyde as a prominent source of pheromones in primates. An extensive analysis of the Kyoto Encyclopedia of Genes and Genomes (KEGG) metabolic pathway has revealed the intricate involvement of numerous metabolites in the biosynthesis of steroid hormones, including 16α-hydroxy estrone, 16α-hydroxy dehydroepiandrosterone, cortisol, testosterone, estrone, and androsterone. These significant findings indicate that the chemical composition of scent marks can serve as a metabolic indicator of an animal's breeding status, providing valuable insights into their reproductive physiology and behavior.

Some intriguing findings were also observed in the analysis of microorganisms. At the phylum level, Actinobacteriota, Bacteroidetes, and Proteobacteria were found to be the most abundant groups, consistent with previous studies on scent marks in captive pandas.[12,21,30] Microbes from the Actinobacteriota, Bacteroidetes, and Proteobacteria phyla have also been implicated in generating scent signals in the chemical communication of other species. For example, in a study of wild meerkats (*Suricata suricatta*), Proteobacteria contributed to the formation of social odors.[31] Microbes have the ability to metabolize various organic compounds into more volatile and odorous metabolites.[8,23] Proteobacteria can produce odorous stimulants, such as C9 to C18 carboxylic acids and some corresponding methyl esters,[32] which aligns with the presence of ketones and carboxylic acids found in our study, suggesting they could serve as substrates. The final metabolism

products of fatty acids can include aldehydes and alcohols. Bacteria also participate in decarboxylation processes, leading to the formation of alkanes and methyl ketones.[33] These possible processes are reasonably inferred from our findings. Odorless steroids (e.g., epitestosterone, testosterone, and estrone in our study) could potentially be transformed into odorous derivatives by Proteobacteria and Actinobacteriota. The correlation analysis further revealed that a substantial proportion (77%) of the significantly correlated pairs of metabolites and bacteria displayed a positive correlation, indicating a consistent trend between the production of metabolites and microbial activity. Therefore, many odorless organic compounds may be metabolized by symbiotic and environmental microorganisms to form scent signals. Considering the constraints of field studies, we excluded some microorganism candidates from Xanthomonadales, which are typical plant pathogens.

In general, this study reveals significant differences in the chemical and microbial composition of scent marks and feces in giant pandas, suggesting that these two signaling modes may have different functional roles in intraspecies communication. Analysis of volatile compounds in scent marks revealed a completely different odorant composition compared to feces. Furthermore, analysis of the microbial communities through 16S rRNA gene sequencing showed significant differences between the microbial populations in feces and those in scent marks. This may indicate that the gut microbiota influencing fecal composition is more adapted to digestive and metabolic functions rather than communication.[9,12] In contrast, bacteria present in scent marks may specifically produce signal molecules that influence olfactory cues. These pieces of evidence collectively point to an intriguing phenomenon of functional differentiation between scent marks and feces in giant pandas. The former appears to be involved in conveying specific information, potentially related to reproductive status or territorial boundaries, while the latter may serve different purposes.[18] For example, one purpose could be to inform individual giant pandas of the presence of conspecifics, promoting resource partitioning in space and avoiding competition. Additionally, observations from infrared cameras also revealed that wild giant pandas spend significantly less time investigating feces from unknown individuals (samples collected from another area and experimentally placed under scent marking trees) compared to investigating naturally occurring scent marks deposited by local (and presumably known) individuals (unpublished data). Therefore, we are inclined to hypothesize that giant panda feces may not transmit the same type of information as AGS and urine but that feces may serve different functions that are worth exploring in future research (such as visual cues indicating the presence of other giant panda individuals in the wild). Alternatively, feces could simply be the metabolic excretion of giant pandas and lack a specific role in information transmission. These findings reassess the communicative significance of chemical signals in giant pandas and emphasize the need for further research on how these animals integrate multiple modes of communication to facilitate social interactions and survival in their natural habitat.

Through a comprehensive analysis of our findings and other studies, we may conclude that the production of a wider range of odors relies on symbiotic and environmental microbes capable of metabolizing the scent marks. In future research, it would be valuable to explore the differences in microbial composition between panda individuals and populations. This would provide further insights into the role of symbiotic and environmental microorganisms in shaping unique scent characteristics. In the context of giant panda conservation, understanding the differences in microorganisms between isolated populations and leveraging microorganisms from suitable habitat environments could be beneficial for the long-term survival of the species. By employing comprehensive omics techniques to study panda chemical communication, we have revealed, for the first time, the metabolome and microbial profiles from multiple perspectives. These findings illuminate previously elusive aspects concerning the characteristics of metabolites and microorganism structures among scent marking sites within the field.

Limitations of the study

In this study, we did not further verify the function of each microbial taxa. The limitation is mainly because of the use of 16S rRNA sequencing, which is mostly used to characterize microbial composition. Therefore, the mechanism of production of scent signals cannot be explored from the gene function. Scent signals are extremely complex, mainly due to their intricate combinations and unclear biological functions. The specific role of each metabolite was unrevealed although we know some candidates functioning as scent signals. Another noteworthy limitation is that as the exposure time of scent marks by giant pandas in the environment increases, the chemical composition within them may undergo changes. This could potentially result in alterations of crucial chemical substances that become difficult to detect.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- METHOD DETAILS
  - Sample collection and extraction
  - Untargeted GC-MS metabonomics analysis
  - Untargeted HPLC-MS metabonomics analysis
  - DNA extraction and database construction
  - Sequencing data processing
  - Metabonomics data processing

- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Analysis of metabonomics
  - Analysis of 16S rRNA sequencing data
  - Correlation analysis of microorganisms and metabolites

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2024.110051.

ACKNOWLEDGMENTS

The authors wish to thank Y. Hong, T. Cai, and Z. Mao from China west normal university for data collection. We thank Mr. X. Zhong for his guidance on the data analysis. We appreciated W. Yang and H. Yang help to conduct field work for this subject. We are deeply grateful for the support and assistance provided by the "Southwest Mountain Region Ecological and Sustainable Development Exchange Center" in this research. This study was financially supported by the sources of funding: the National Natural Science Foundation of China (U21A20193; 42071279; and 32270551), the Second Tibetan Plateau Scientific Expedition and Research Program (STEP) (grant no. 2019QZKK0402), the Key Laboratory of Southwest China Wildlife Resources Conservation (China West Normal University), Ministry of Education (Grant no. XNYB9-01), 2022 College students' innovative entrepreneurial training plan program of China West Normal University (X. Chen).

AUTHOR CONTRIBUTIONS

Conceptualization, J.Z. and V.H ; investigation, J.H. and R.C.; writing – original draft, J.H.; methodology, M.F.; writing – reviewing and editing, J.Z., V.H., M.F., and T.C.; resources, J.Z., Z.Z., and D.Q.; formal analysis, J.H., X.C., and S.C.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: December 5, 2023
Revised: March 2, 2024
Accepted: May 17, 2024
Published: May 23, 2024

REFERENCES

1. Wyatt, T.D. (2014). Pheromones and Animal Behavior: Chemical Signals and Signatures (Cambridge University Press).
2. Apps, P.J., Weldon, P.J., and Kramer, M. (2015). Chemical signals in terrestrial vertebrates: search for design features. Nat. Prod. Rep. *32*, 1131–1153. https://doi.org/10.1039/C5NP00029G.
3. Campbell-Palmer, R., and Rosell, F. (2011). The importance of chemical communication studies to mammalian conservation biology: A review. Biol. Conserv. *144*, 1919–1930. https://doi.org/10.1016/j.biocon.2011.04.028.
4. Albone, E.S., and Perry, G.C. (1976). Anal sac secretion of the red fox, *Vulpes vulpes*; volatile fatty acids and diamines: Implications for a fermentation hypothesis of chemical recognition. J. Chem. Ecol. *2*, 101–111. https://doi.org/10.1007/BF00988029.
5. Theis, K.R., Venkataraman, A., Dycus, J.A., Koonter, K.D., Schmitt-Matzen, E.N., Wagner, A.P., Holekamp, K.E., and Schmidt, T.M. (2013). Symbiotic bacteria appear to mediate hyena social odors. Proc. Natl. Acad. Sci. USA *110*, 19832–19837. https://doi.org/10.1073/pnas.1306477110.
6. Brunetti, A.E., Lyra, M.L., Melo, W.G.P., Andrade, L.E., Palacios-Rodríguez, P., Prado, B.M., Haddad, C.F.B., Pupo, M.T., and Lopes, N.P. (2019). Symbiotic skin bacteria as a source for sex-specific scents in frogs. Proc. Natl. Acad. Sci. USA *116*, 2124–2129. https://doi.org/10.1073/pnas.1806834116.
7. Brunetti, A.E., Lyra, M.L., Bauermeister, A., Bunk, B., Boedeker, C., Müsken, M., Neto, F.C., Mendonça, J.N., Caraballo-Rodríguez, A.M., Melo, W.G.P., et al. (2023). Host macrocyclic acylcarnitines mediate symbiotic interactions between frogs and their skin microbiome. iScience *26*, 108109. https://doi.org/10.1016/j.isci.2023.108109.
8. Gosling, L.M., and Roberts, S.C. (2001). Testing ideas about the function of scent marks in territories from spatial patterns. Anim. Behav. *62*, F7–F10. https://doi.org/10.1006/anbe.2001.1802.
9. Hu, J., Xia, L., Pan, W., and Zhu, J. (1985). Wolong's Giant Panda (Sichuan Science and Technology Press). (In Chinese).
10. Zhou, W., Nie, Y., Hu, Y., Swaisgood, R.R., Zhang, Y., Liu, D., and Wei, F. (2019). Seasonal and reproductive variation in chemical constituents of scent signals in wild giant pandas. Sci. China Life Sci. *62*, 648–660. https://doi.org/10.1007/s11427-018-9388-9.
11. Wilson, A.E., Sparks, D.L., Knott, K.K., Willard, S., Brown, A., Connor, T., and Zhang, Z. (2019). Field air analysis of volatile compounds from free-ranging giant pandas. Ursus *29*, 75–81. https://doi.org/10.2192/URSUS-D-18-00009.1.
12. Zhou, W., Qi, D., Swaisgood, R.R., Wang, L., Jin, Y., Wu, Q., Wei, F., and Nie, Y. (2021). Symbiotic bacteria mediate volatile chemical signal synthesis in a large solitary mammal species. ISME J. *15*, 2070–2080. https://doi.org/10.1038/s41396-021-00905-1.
13. Swaisgood, R.R., Lindburg, D.G., and Zhou, X. (1999). Giant pandas discriminate individual differences in conspecific scent. Anim. Behav. *57*, 1045–1053. https://doi.org/10.1006/anbe.1998.1070.
14. White, A.M., Swaisgood, R.R., and Zhang, H. (2002). The highs and lows of chemical communication in giant pandas (*Ailuropoda melanoleuca*): effect of scent deposition height on signal discrimination. Behav. Ecol. Sociobiol. *51*, 519–529. https://doi.org/10.1007/s00265-002-0473-3.
15. Hagey, L., and MacDonald, E. (2003). Chemical Cues Identify Gender and Individuality in Giant Pandas (*Ailuropoda melanoleuca*). J. Chem. Ecol. *29*, 1479–1488.
16. Yuan, H., Liu, D., Sun, L., Wei, R., Zhang, G., and Sun, R. (2004). Anogenital gland secretions code for sex and age in the giant panda, *Ailuropoda melanoleuca*. Can. J. Zool. *82*, 1596–1604. https://doi.org/10.1139/z04-144.
17. Gilad, O., Swaisgood, R.R., Owen, M.A., and Zhou, X. (2016). Giant pandas use odor cues to discriminate kin from nonkin. Curr. Zool. *62*, 333–336. https://doi.org/10.1093/cz/zow025.
18. Nie, Y., Swaisgood, R.R., Zhang, Z., Hu, Y., Ma, Y., and Wei, F. (2012). Giant panda scent-marking strategies in the wild: role of season, sex and marking surface. Anim. Behav. *84*, 39–44. https://doi.org/10.1016/j.anbehav.2012.03.026.

19. Hou, J., Hull, V., Connor, T., Yang, H., Gao, J., Zhao, F., Liao, Y., Chen, S., Huang, J., Zeng, Y., et al. (2021). Scent communication behavior by giant pandas. Glob. Ecol. Conserv. *25*, e01431. https://doi.org/10.1016/j.gecco.2020.e01431.
20. Bian, X., Liu, D., Zeng, H., Zhang, G., Wei, R., and Hou, R. (2013). Exposure to odors of rivals enhances sexual motivation in male giant pandas. PLoS One *8*, e69889. https://doi.org/10.1371/journal.pone.0069889.
21. Ma, R., Zheng, W., Guo, J., Hou, R., Huang, H., Xue, F., Zhou, Y., Wu, W., Huang, C., Gu, J., et al. (2022). Symbiotic microbiota and odor ensure mating in time for giant pandas. Front. Microbiol. *13*, 1015513. https://doi.org/10.3389/fmicb.2022.1015513.
22. Macdonald, D.W. (1980). Pattern of scent marking with urine and feces amongst social communities. Symp. Zool. Soc. Lond. *45*, 107–139.
23. Gao, R., Liu, H., Li, Y., Liu, H., Zhou, Y., and Yuan, L. (2023). Correlation between dominant bacterial community and non-volatile organic compounds during the fermentation of shrimp sauces. Food Sci. Hum. Well. *12*, 233–241. https://doi.org/10.1016/j.fshw.2022.07.002.
24. Anitaş, Ö., and Göncü, S. (2020). Investigation of body secretions as bioindicators in cattle estrus detection. Turk. J. Vet. Anim. Sci. *44*, 1070–1086. https://doi.org/10.3906/vet-2002-74.
25. Poirier, A.C., Waterhouse, J.S., Watsa, M., Erkenswick, G.A., Moreira, L.A.A., Tang, J., Dunn, J.C., Melin, A.D., and Smith, A.C. (2021). On the trail of primate scent signals: A field analysis of callitrichid scent-gland secretions by portable gas chromatography-mass spectrometry. Am. J. Primatol. *83*, e23236. https://doi.org/10.1002/ajp.23236.
26. Bradbury, J.W., and Vehrencamp, S.L. (2011). Principles of Animal Communication (Sinauer Associates Inc).
27. Kula, J., Bonikowski, R., Staniszewska, M., Krakowiak, A., Wieczorek, M., Majzner, W.R., and Bujacz, G.D. (2002). Transformation of Carotol into the Hydroindane-Derived Musk Odorant. Eur. J. Org. Chem. *2002*, 1826–1829. https://doi.org/10.1002/1099-0690(200206)2002:11<1826::AID-EJOC1826>3.0.CO;2-A.
28. Chiu, C.C., Keeling, C.I., and Bohlmann, J. (2018). Monoterpenyl esters in juvenile mountain pine beetle and sex-specific release of the aggregation pheromone trans-verbenol. Proc. Natl. Acad. Sci. USA *115*, 3652–3657. https://doi.org/10.1073/pnas.1722380115.
29. Karthikeyan, K., Muniasamy, S., Sankarganesh, D., Achiraman, S., Ramesh Saravanakumar, V., and Archunan, G. (2013). Faecal chemical cues in water buffalo that facilitate estrus detection. Anim. Reprod. Sci. *138*, 163–167. https://doi.org/10.1016/j.anireprosci.2013.02.017.
30. Guo, J.L. (2020). Correlation Analysis of Chemical Components and Flora Structure of Perianal Gland Markers in Pandas during Estrus (China West Normal University). (in Chinese). https://doi.org/10.27859/d.cnki.gxhsf.2020.000158.
31. Leclaire, S., Jacob, S., Greene, L.K., Dubay, G.R., and Drea, C.M. (2017). Social odours covary with bacterial community in the anal secretions of wild meerkats. Sci. Rep. *7*, 3240. https://doi.org/10.1038/s41598-017-03356-x.
32. Ponnusamy, L., Xu, N., Nojima, S., Wesson, D.M., Schal, C., and Apperson, C.S. (2008). Identification of bacteria and bacteria-associated chemical cues that mediate oviposition site preferences by Aedes aegypti. Proc. Natl. Acad. Sci. USA *105*, 9262–9267. https://doi.org/10.1073/pnas.0802505105.
33. Yamaguchi, M.S., Ganz, H.H., Cho, A.W., Zaw, T.H., Jospin, G., McCartney, M.M., Davis, C.E., Eisen, J.A., and Coil, D.A. (2019). Bacteria isolated from Bengal cat (*Felis catus* × *Prionailurus bengalensis*) anal sac secretions produce volatile compounds potentially associated with animal signaling. PLoS One *14*, e0216846. https://doi.org/10.1371/journal.pone.0216846.
34. Magoč, T., and Salzberg, S.L. (2011). Flash: fast length adjustment of short reads to improve genome assemblies. Bioinformatics *27*, 2957–2963. https://doi.org/10.1093/bioinformatics/btr507.
35. Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Peña, A.G., Goodrich, J.K., Gordon, J.I., et al. (2010). Qiime allows analysis of high-throughput community sequencing data. Nat. Met. *7*, 335–336. https://doi.org/10.1038/nmeth.f.303.
36. Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C., and Knight, R. (2011). Uchime improves sensitivity and speed of chimera detection. Bioinformatics *27*, 2194–2200. https://doi.org/10.1093/bioinformatics/btr381.
37. Edgar, R.C. (2013). Uparse: highly accurate otu sequences from microbial amplicon reads. Nat. Methods *10*, 996–998. https://doi.org/10.1038/NMETH.2604.
38. Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., and Glöckner, F.O. (2013). The silva ribosomal rna gene database project: improved data processing and web-based tools. Nucleic Acids Res. *41*, D590–D596. https://doi.org/10.1093/nar/gks1219.
39. Qiao, M., Connor, T., Shi, X., Huang, J., Huang, Y., Zhang, H., and Ran, J. (2019). Population genetics reveals high connectivity of giant panda populations across human disturbance features in key nature reserve. Ecol. Evol. *9*, 1809–1819. https://doi.org/10.1002/ece3.4869.
40. Zhang, J., Li, Y., Wang, Y., Ouyang, Z., and Qiao, M. (2017). Comparison of two investigation methods for population of wild giant pandas. Chin. J. Appl. Environ. Biol. *23*, 1142–1147. (In Chinese). https://doi.org/10.3724/SP.J.1145.2017.01010.
41. Want, E.J., Masson, P., Michopoulos, F., Wilson, I.D., Theodoridis, G., Plumb, R.S., Shockcor, J., Loftus, N., Holmes, E., and Nicholson, J.K. (2013). Global metabolic profiling of animal and human tissues via UPLC-MS. Nat. Protoc. *8*, 17–32. https://doi.org/10.1038/nprot.2012.135.
42. Wilson, E.O. (1963). Pheromones. Sci. Am. *208*, 100–114. https://www.jstor.org/stable/24936151.
43. Rognes, T., Flouri, T., Nichols, B., Quince, C., and Mahé, F. (2016). VSEARCH: a versatile open source tool for metagenomics. PeerJ *4*, e2584. https://doi.org/10.7717/peerj.2584.
44. Huang, B., Long, J., Liao, H., Liu, L., Li, J., Zhang, J., Li, Y., Wang, X., and Yang, R. (2019). Characteristics of bacterial community and function in paddy soil profile around antimony mine and its response to antimony and arsenic contamination. Int. J. Environ. Res. Public Health *16*, 4883. https://doi.org/10.3390/ijerph16244883.
45. Paruch, L., Paruch, A.M., Eiken, H.G., and Sørheim, R. (2019). Aquatic microbial diversity associated with faecal pollution of Norwegian waterbodies characterized by 16S rRNA gene amplicon deep sequencing. Microb. Biotechnol. *12*, 1487–1491. https://doi.org/10.1111/1751-7915.13461.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Original metabolomics and 16s rRNA data of scent markers and feces	This manuscript	http://cloud.cwnu.edu.cn/share?id=57r3v8133988
Software and algorithms		
FLASH (V1.2.7)	Mago et al.[34]	http://ccb.jhu.edu/software/FLASH/
Compound Discoverer 3.1	ThermoFisher Scientific	N/A
R: A language and environment for statistical computing v4.0.2	R Core Team	http://www.r-project.org
QIIME v1.9.1	Caporaso et al.[35]	http://qiime.org/scripts/split_libraries_fastq.html
UCHIME	Edgar et al.[36]	http://www.drive5.com/usearch/manual/uchime_algo.html
Uparse v7.0.1001	Edgar et al.[37]	http://drive5.com/uparse/
Other		
Vanquish UPLC system	ThermoFisher Scientific	N/A
Orbitrap Q Exactive™ HF-X mass spectrometer	ThermoFisher Scientific	N/A
mzCloud	ThermoFisher Scientific	https://www.mzcloud.org/
mzVault database	ThermoFisher Scientific	N/A
MassList database	ThermoFisher Scientific	N/A
Nist 2017 Library	Agilent Technology Co., Ltd	N/A
Silva Database	Quast et al.[38]	http://www.arb-silva.de/

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Jindong Zhang (zhangjd224@cwnu.edu.cn).

Materials availability

This study did not generate new unique reagents or materials.

Data and code availability

- The metabolomics data and 16S rRNA sequencing data are available at http://cloud.cwnu.edu.cn/share?id=57r3v8133988. Datasets can be publicly available at the date of publication.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

METHOD DETAILS

Sample collection and extraction

We collected a total of six giant panda scent marking bark samples and six fecal samples from different sites between February 2021 and May 2021. Our sampling was conducted in the Wolong Nature Reserve, Sichuan Province, China, covering an area of 2000 km^2. The scent marking trees were initially identified based on changes in color and odor following panda marking. Infrared cameras were then installed in these potential marking sites to further confirm the trees that were marked by giant pandas.[19] All scent marks analyzed in this study were collected within 10 days of being deposited, as determined by infrared camera monitoring. At the same scent marking tree, different individual giant pandas utilize both AGS and urine for scent marking. Therefore, it is common for a single scent marking tree to contain both scent marks. Additionally, these scent marks can persist in the natural environment for up to six months.[19] Hence, the tree bark samples we

collected actually represent a mixture of two scent marks, with the most recent mark being deposited within a maximum of 10 days prior to collection. During the collection process, collectors wearing latex gloves used forceps and a disposable surgical knife to collect the tree bark.

Based on their characteristics, we divided our samples into three groups. The collected bark samples containing giant panda odor marks (Bark1 group) were placed in 20 ml glass bottles with polytetrafluoroethylene caps. As a background control, an unmarked tree bark sample (Bark0 group), which was not marked by giant pandas, was simultaneously collected from the same tree at a distance of approximately 1.5 meters away. The fresh fecal samples were collected within approximately 500 meters around the scent marked trees, approximately one week after excretion (Feces group). Giant panda feces can be dated using standard, established protocols in the field that reliably identify fresh feces based on color and consistency. Upon arrival at the laboratory, all samples were sealed and stored in a -20°C freezer until analysis conducted within a maximum of one week later.[10]

While our sample size is modest, it remains representative. Our research is centered in Hetaoping, a region in the Wolong Nature Reserve known for its high giant panda densities.[39] Using molecular biology techniques, we identified 13 individuals from 17 fresh giant panda fecal samples collected exclusively from this area.[40] Extensive field monitoring confirms a stable giant panda population in the study area, ensuring an adequate sample size.[9,19,39] In our sampling approach, we targeted scent-marking trees frequented by giant pandas, leveraging their information exchange habits, especially during the breeding season, to increase our chances of collecting scent marks from a larger number of individuals. The distance between each sampling point ranges from 200-600m. Despite the localized distribution of these trees, the overlapping home ranges of wild pandas in Wolong facilitate visits from individuals in the region and neighboring areas.[40] Most importantly, based on the monitoring data obtained from the infrared cameras shortly before our sampling (Unpublished data), we observed an estimated 3-6 distinct giant panda individuals engaging in scent marking at the sampling points, identified by their physical characteristics. Sampling conducted within ten days of the most recent scent marking activity yielded samples containing a mix of fresh and partially degraded scent marks. This approach ensures the collection of scent marks from various past individuals, enriching the chemical diversity of the collected scent marks to align with the diverse olfactory cues vital for wild giant pandas.

Prior to GC-MS analysis, the bark and feces were removed from the glass bottles using clean disposable medical-grade plastic tweezers sterilized with ethylene oxide and placed into separate clean 20 ml glass bottles. Dichloromethane was added to each bottle at a ratio of 1g/600 μl solvent. After 12 hours, the samples were thoroughly mixed and refrigerated at 4°C for GC-MS analysis. All experimenters involved in these procedures were required to wear gloves to prevent odor contamination.[10]

For the analysis, 100 mg of each sample was individually ground with liquid nitrogen, and the homogenate was resuspended in prechilled 80% methanol and 0.1% formic acid using a vortex. The samples were then incubated on ice for 5 minutes and centrifuged at 15,000 g, 4°C for 20 minutes. Some of the supernatant was diluted with LC-MS grade water to a final concentration containing 53% methanol. The samples were transferred to fresh Eppendorf tubes and centrifuged again at 15,000 g, 4°C for 20 minutes. Finally, the supernatant was injected into the LC-MS/MS system for analysis.[41]

During the entire process, we minimized the cross-contamination between samples as much as possible. Firstly, a solvent blank sample was run every 2 samples. Secondly, before each new sample was introduced into the instrument, the injection needle cleaning procedure was carried out using dichloromethane as the cleaning agent. Thirdly, the analytical temperature of the injection port was very high (230°C), which ensured that there were no residual chemicals from the previous sample.

Untargeted GC-MS metabonomics analysis

GC-MS analysis was performed with an Agilent Technologies Network 6890 N gas chromatograph system equipped with a 30 m DB - 5 MS glass capillary column (250 μm × 0.25 μm film thickness) coupled with 5973 Mass Selective Detector. Helium gas was set to constant flow (1.0 ml/min) using the splitless mode. The injector port temperature was set at 250°C. We ran the following temperature protocol after a 3 min solvent delay: initial oven temperature was set to 35°C with 1 min held; 40°C to 280°C ramped at 10°C/min (hold at 280°C for 10 min); the entire run lasted 35 min. Transfer line temperature was 280°C. Scanning mass ranged from 50 to 550 amu, and the 1 μl sample was injected using the splitless mode. The compounds were tentatively identified by matching the mass spectra with structures available in the NIST 2017 library (Agilent Technologies 2017, USA). Relative mass fraction of each chemical compounds was calculated by peak area normalization method.[10]

Untargeted HPLC-MS metabonomics analysis

The ultra-high performance liquid chromatography (UHPLC) coupled with high-resolution mass spectrometry, also known as the liquid chromatography-mass spectrometry (UHPLC-MS) technique, has been increasingly applied in untargeted metabolomics studies due to its wide detection range for different chemicals, high sensitivity, low cross-contamination between samples, and simple pretreatment procedures. UHPLC-MS analyses were performed using a Vanquish UHPLC system (ThermoFisher, Germany) coupled with an Orbitrap Q Exactive™ HF-X mass spectrometer (Thermo Fisher, Germany). Samples were injected onto a Hypesil Gold column (100 × 2.1 mm, 1.9μm) using a 17-min linear gradient at a flow rate of 0.2 mL/min. The eluents for the positive polarity mode were eluent A (0.1% FA in Water) and eluent B (Methanol). The solvent gradient was set as follows: 2% B, 1.5 min; 2-100% B, 12.0 min; 100% B, 14.0 min; 100-2% B, 14.1 min; 2% B, 17 min. Q Exactive™ HF-X mass spectrometer was operated in positive polarity mode with spray voltage of 3.2 kV, capillary temperature of 320C, sheath gas flow rate of 40 arb and aux gasflow rate of 10 arb.

DNA extraction and database construction

Total genome DNA from samples was extracted using Cetyltrimethylammonium Bromide (CTAB) method. DNA concentration and purity were monitored on 1% agarose gels. According to the concentration, DNA was diluted to 1ng/μL using sterile water. The V3-V4 region of 16S rRNA were amplified using specific primers 341F and 806R. All PCR reactions were carried out with 15 μL of Phusion® High-Fidelity PCR Master Mix (New England Biolabs), 2 μM of forward and reverse primers, and about 10 ng template DNA. Thermal cycling consisted of initial denaturation at 98°C for 1 min, followed by 30 cycles of denaturation at 98°C for 10 s, annealing at 50°C for 30 s, and elongation at 72°C for 30 s and 72°C for 5 min. We mixed the same volume of 1X loading buffer (contained SYB green) with PCR products and operated electrophoresis on a 2% agarose gel for detection. PCR products were mixed in equidensity ratios. Then, the mixture of PCR products was purified with a Qiagen Gel Extraction Kit (Qiagen, Germany). Sequencing libraries were generated using a TruSeq® DNA PCR-Free Sample Preparation Kit (Illumina, USA) following the manufacturer's recommendations and index codes were added. The library quality was assessed on the Qubit@ 2.0 Fluorometer (Thermo Scientific) and Agilent Bioanalyzer 2100 system. In the last step, the library was sequenced on an Illumina NovaSeq platform and 250 bp paired-end reads were generated.

Sequencing data processing

Paired-end reads were assigned to samples based on their unique barcode and truncated by cutting off the barcode and primer sequence. Paired-end reads were merged using FLASH (V1.2.7, http://ccb.jhu.edu/software/FLASH/), a very fast and accurate analysis tool, which was designed to merge paired-end reads when at least some of the reads overlap with the read generated from the opposite end of the same DNA fragment, and the splicing sequences were called raw tags. Quality filtering on the raw tags was performed under specific filtering conditions to obtain the high-quality clean tags according to the QIIME (V1.9.1, http://qiime.org/scripts/split_libraries_fastq.html) quality controlled process. The tags were compared with the reference database (Silva database, using UCHIME algorithm (http://www.drive5.com/usearch/manual/uchime_algo.html)) to detect chimera sequences, and then the chimera sequences were removed. Effective tags were thus obtained.

Metabonomics data processing

In processing and analysis of GC-MS data, after the original data of each sample were obtained from the NIST 2017 library, the components with relative concentration percentage less than 0.1% were first eliminated, and the compounds with very low response were filtered out. The environmental background impurities (Bark0) were removed from these data, and then the components that appeared two or more times were retained to avoid component contamination caused by accidental errors and experimental operation errors.

The raw data files generated by UHPLC-MS/MS were processed using the Compound Discoverer 3.1 (CD3.1, ThermoFisher) to perform peak alignment, peak picking, and quantitation for each metabolite. The main parameters were set as follows: retention time tolerance, 0.2 minutes; actual mass tolerance, 5ppm; signal intensity tolerance, 30%; signal/noise ratio, 3; and minimum intensity. Afterward, peak intensities were normalized to the total spectral intensity. Normalized data were used to predict the molecular formula based on additive ions, molecular ion peaks and fragment ions. Then, peaks were matched with the mzCloud (https://www.mzcloud.org/), mzVault and MassList databases to obtain accurate qualitative and relative quantitative results. Bark control was used by blank for background compounds removal. If the same compound was identified in both the sample and blank, and the value of peak area of the compound in the sample with a scent mark (the largest peak area in all samples) divided by the peak area in blank was less than 5, the compound would be considered a background compound.

QUANTIFICATION AND STATISTICAL ANALYSIS

Analysis of metabonomics

Partial Least Squares Discriminant Analysis (PLS-DA) was used to show the profile differences between the Bark1 group and Feces group. Statistical analyses were performed using the statistical software R (R version R-3.4.3). When metabolite data were not normally distributed, Log2 transformation was employed for normalizing the data.

In this study, compounds with a relative molecular mass of less than 300 were classified as "volatile" and those with a relative molecular mass of more than 300 were classified as "semi-volatile".[16,42]

Analysis of 16S rRNA sequencing data

Sequencing analysis was performed by Uparse software (Uparse v7.0.1001, http://drive5.com/uparse/). Sequences with ≥97% similarity were assigned to the same OTUs and a Venn diagram was drawn to analyze the common and unique numbers of OTUs of different groups.[43] We screened representative sequences of each OTU for further annotation. For each representative sequence, the Silva Database (http://www.arb-silva.de/) was used to annotate taxonomic information based on the Mothur algorithm. OTU abundance information was normalized using a standard of the sequence number corresponding to the sample with the least sequences. Subsequent analysis of alpha diversity (Simpson index and Shannon index) and beta diversity (Anosim analysis based on Bray-Curtis distance) were all performed basing on this normalized data output. In the results of the Anosim analysis, the R-value ranges between -1 and 1. A positive R-value indicates significant differences between groups. On the other hand, a negative R-value suggests that the within-group differences are greater than the between-group differences. The statistical significance is typically represented by the P-value, where $P < 0.05$ indicates statistical significance.

The sequencing quantity was evaluated as reasonable by comparing the rarefaction curves of different samples. We used the principal coordinates analysis (PCoA based on weighted unifrac distance) to analyze the composition differences among multiple groups of data. A stacked bar plot was used to show the order of the top 10 abundances of taxa at phylum levels. The composition of different species in microbial communities among different groups was analyzed by linear discriminant analysis (LDA) coupled with effect size measurements (LEfSe) analysis,[44] including the analysis of the contribution of different species to the differences, the annotation analysis on different species, and the relative abundance of different species in each sample.[45]

Correlation analysis of microorganisms and metabolites

The total abundance of the top 20 metabolites and the top 20 species in the Bark1 and Feces groups, respectively were selected at the genus level. To exclude the species of Bark0, we selected the top 20 significantly differential species as the profile of the Bark1 group using MetaStat analysis (P<0.05). The Pearson correlation coefficient between microflora and metabolites was analyzed using the "corrplot" package in R (version 4.0.2). A significant correlation was shown with $P < 0.05$.