



OPEN

# Potential value and chemical characterization of gut microbiota derived nitrogen containing metabolites in feces from *Periplaneta americana* (L.) at different growth stages

Weiqli Lv<sup>1,2</sup>, Ying Cui<sup>1,2</sup>, Gen Xue<sup>1</sup>, Ziyang Wang<sup>1</sup>, Lu Niu<sup>1</sup>, Xin Chai<sup>1</sup>✉ & Yuefei Wang<sup>1</sup>✉

The American cockroach, *Periplaneta americana* (L.), is able to highly survive in various complicated environments around the globe, and often considered as a pest. In contrast, billions of *P. americana* have been massively reared in China and extensively used as a medicinal insect, due to its function for preventing and treating ulceration and heart failure. Considering the possibility that microbiota-derived metabolites could be an effective source to identify promising candidate drugs, we attempted to establish a rapid method for simultaneous determination of gut microbiota metabolites from medicinal insects. In this study, network pharmacology approach and ultra-performance liquid chromatography (UPLC) technique were employed to reveal the potential pharmacological activity and dynamics variation of nitrogen-containing metabolites (NCMs) originated from the gut microbiota of breeding *P. americana* at different growth stages. A metabolites-targets-diseases network showed that NCMs are likely to treat diseases such as ulceration and cancer. The analysis of NCMs' content with the growth pattern of *P. americana* indicated that the content of NCMs declined with *P. americana* aging. Both principal component analysis and orthogonal partial least squares discriminant analysis suggested that 8-hydroxy-2-quinolinecarboxylic acid and 8-hydroxy-3,4-dihydro-2(1H)-quinolinone are the potential differential metabolic markers for discriminating between nymphs and adults of *P. americana*. Moreover, the developed UPLC method showed an excellent linearity ( $R^2 > 0.999$ ), repeatability (RSD < 2.6%), intra- and inter-day precisions (RSD < 2.2%), and recovery (95.5%–99.0%). Collectively, the study provides a valuable strategy for analyzing gut microbiota metabolites from insects and demonstrates the prospects for discovering novel drug candidates from the feces of *P. americana*.

American cockroach, *Periplaneta americana* (L.), is an insect within the family Blattidae, which usually survives in complex and diverse habitats, ranging from the wild forest and mountains to urban environments<sup>1</sup>. It is widely known as the most common household pest that causes nuisance to peoples' homes. In China, it is used as a medicinal ingredient in traditional Chinese medicine (TCM) for treating blood stasis, chills, fever, and intra-abdominal mass, which has been recorded in two famous classical documents<sup>2</sup>, *Shen Nong Ben Cao Jing* and *Compendium of Materia Medica*. Recent studies have proved that *P. americana* is enriched with active compounds, including amino acids, polypeptides, polysaccharides, nucleic acids, lipids, alkaloids, etc<sup>3</sup>. The ethanol extract of *P. americana* exhibits beneficial pharmacological activities, including anti-inflammatory<sup>4</sup>, anti-microbial<sup>5</sup>, anti-tumor<sup>6</sup>, hepatoprotective<sup>7</sup>, cardioprotective<sup>8</sup>, wound healing<sup>9</sup>, and tissue repair promotion effects<sup>10</sup>. In addition, three Chinese patent medicines (Kangfuxin liquid, Ganlong capsule, and Xinmailong injection) have been

<sup>1</sup>State Key Laboratory of Component-Based Chinese Medicine, Tianjin Key Laboratory of TCM Chemistry and Analysis, Tianjin University of Traditional Chinese Medicine, Tianjin 301617, China. <sup>2</sup>These authors contributed equally: Weiqli Lv and Ying Cui. ✉email: chaix0622@tjutc.edu.cn; wangyf0622@tjutc.edu.cn

developed to manage or treat ulceration<sup>11</sup>, hepatic injury<sup>12</sup>, and heart failure<sup>13</sup>, respectively. These medicines prepared from *P. americana* are safe and low toxicity<sup>14</sup>.

Gut microbial metabolites from humans and other creatures are momentous sources for discovering new drugs<sup>15</sup> and TCMs<sup>16</sup>. The gut microbiota and their metabolites play a significant role in the host's overall health and well-being. Metabolites from normal gut microbiota defend against illness and boost immunity<sup>17</sup>, while the metabolites produced by dysfunctional gut microbiota often cause serious health problems<sup>18</sup>. Before Antony van Leeuwenhoek opened the door to understanding the microbial world in the seventeenth century, Chinese scholars actively utilized metabolites from gut microbiota as therapeutic agents. More than one thousand years ago, herbal doctors had achieved inspiring results by treating difficult miscellaneous diseases with medicinal animal feces. These medicinal feces include Jinzhi originated from human feces<sup>19</sup>, Cansha derived from silkworm feces<sup>20</sup>, and Ambergris produced from cetacean feces<sup>21</sup>, which have been summarized and extensively documented in *Compendium of Materia Medica*. Influenced by food source, digestive system, age, environment, and other factors, gut microbiotas in different organisms are specific about the abundance and structure, resulting in differences in the content and types of metabolites<sup>22</sup>. Modulation of beneficial gut microbiota and their biologically active metabolites have been shown to be a therapeutic way for maintaining human health<sup>23</sup>. As a result, the therapeutic approaches, such as fecal microbiota transplantation, are increasingly becoming attractive against several diseases with the proved clinical efficacy<sup>24</sup>. Moreover, in the light of the importance of gut microbiota in health, researchers have urgently developed diverse strategies and techniques to study gut microbiota structure, composition, and metabolites. Multiple techniques have been developed under genomics<sup>25</sup>, metabolomics<sup>26</sup>, proteomics<sup>27</sup>, and organic chemical analysis<sup>28</sup>. Also, health, disease, and disease-causal relationships among host, host gut microbiota, and gut microbiota metabolites are continually explored in the medical<sup>29</sup> and biomedical fraternity<sup>30</sup>.

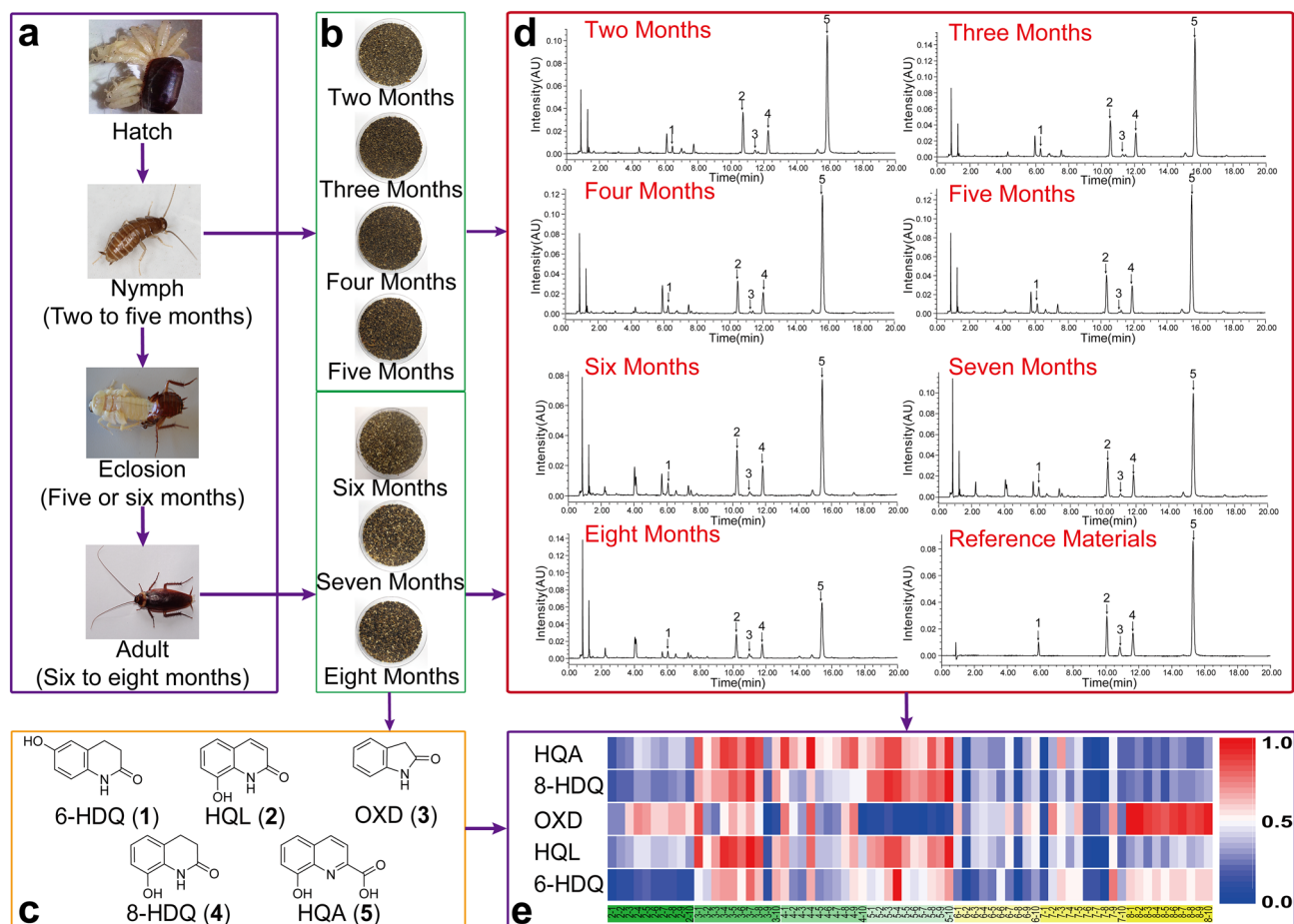
Nowadays, farmers breed *P. americana* on large scales under Good Agricultural Practices (GAP) in different farm sites in southwest China. Nevertheless, the proper disposal of feces from *P. americana* has become an intractable problem. Resource utilization of the feces from *P. americana* will strongly support environmental protection and promote economic profit. Although the gut microbial community of *P. americana* is complex and extensive<sup>31</sup>, its gut microbiota is more stable to dietary shifts than the highly responsive nature of gut microbiota from humans and other mammals<sup>32</sup>. Therefore, several active secondary metabolites produced via gut microbiota metabolism of proteins, lipids, and sugars<sup>33</sup> are speculated to be the reason for the strong vitality and environmental adaptability of *P. americana*. Bacteroidetes is one of the central microbial phylum communities of *P. americana* gut microbiota<sup>32</sup>, which is identified as the predominant proteolytic bacterial colony in feces<sup>34,35</sup>. Protein, the main ingredient of *P. americana* diet, is the primary source of essential amino acids in *P. americana*. The Blattabacterium, a member of the Bacteroidetes, is known to metabolize proteins into amino acids<sup>36</sup>. The nitrogen-containing metabolites (NCMs) that derived from amino acids, such as alkaloids and biogenic amines, are a great source of biologically active secondary metabolites<sup>37,38</sup>. The NCMs might be the key point to explain various pharmacological activities of *P. americana*. In addition, during development and growth of *P. americana*, changes are observed within its gut microbiota population, resulting in influencing production of metabolites<sup>39</sup>. Up to date, the chemical profiling, pharmacodynamics, and clinic utilization of *P. americana* have been extensively reported by numerous studies<sup>40,41</sup>. Among these studies, most of reports focused on the community composition and structural changes of *P. americana* gut microbiota<sup>42</sup>, while reports were limited on the gut microbiota metabolites from *P. americana*. Similarly, various strategies for detecting metabolites from feces by ultra-performance liquid chromatography-tandem mass spectrometry<sup>43</sup> and gas chromatography-mass spectrometry<sup>44</sup> have been reported. Nonetheless, studies on the rapid detection and precise quantitative analysis of NCMs have been rarely reported.

Herein, NCMs (Supplementary Fig. S1 and Fig. 1c) were detected with strong ultraviolet absorption, including 6-hydroxy-3,4-dihydro-1*H*-quinoline-2-one (6-HDQ), 8-hydroxy-2-(1*H*)-quinolinone (HQL), 2-oxindole (OXD), 8-hydroxy-3,4-dihydro-2(1*H*)-quinolinone (8-HDQ), and 8-hydroxy-2-quinolinecarboxylic acid (HQA). These NCMs have been revealed as active compounds responsible for amelioration of diseases through network pharmacology. Accordingly, an ultra-performance liquid chromatography (UPLC) method was established for determining NCMs from *P. americana*. Our results demonstrated that HQA and 8-HDQ are the distinguished metabolic markers for nymphs and adults of *P. americana*. In general, our study proposes a rapid detection method for NCMs and reveals variations of NCMs' content with the different growth stages of *P. americana*.

## Results and discussion

**Exploration of potential pharmacological value of NCMs.** The NCMs separated from *P. americana* feces were identified as nitrogen-containing compounds with a benzene ring structure, which were classified to be alkaloids as previously described. This suggests that NCMs might be active compounds with potential therapeutic benefits. Therefore, a network pharmacology approach was used to explore the possible pharmacological activities and action mechanisms of NCMs.

After eliminating duplicates, 124 targets associated with NCMs were obtained from the Swiss Target Prediction database. Then, the targets were uploaded to the STRING database to investigate the functional partnerships based on protein-protein interactions (PPI) network<sup>45</sup>. After hiding the disconnected nodes, PPI network (Fig. 2a) with a minimum interaction score of 0.4 was imported into Cytoscape software to visualize, which comprised 112 nodes and 570 edges. Moreover, the topological parameters of the node degree, which revealed the number of interactions, were calculated by the network analyzer function of Cytoscape software. Finally, a total of 19 proteins (Table 1) were regarded as key target proteins of the network with the degrees of 20 or higher. Among them, AKT1, SRC, CASP3, COMT, and DRD2 had significantly higher values in degrees, betweenness centrality, and closeness centrality than other proteins in the network, indicating that these proteins are essential in biochemical processes. AKT1, a serine/threonine-specific protein kinase, is the central node of signaling



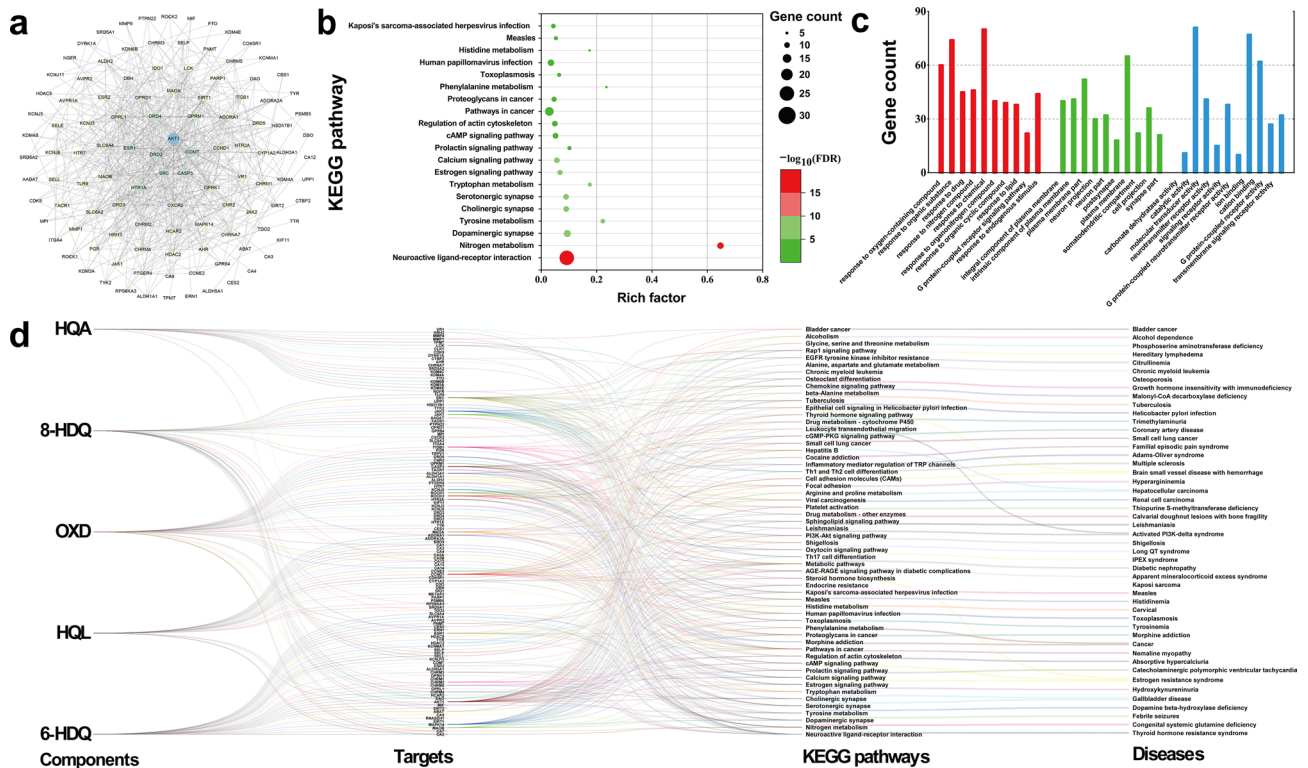
**Figure 1.** Collection process and the results of fecal samples analysis of *P. americana* at the different growth stages. (a) *P. americana* at the different growth stages. (b) Collected feces from *P. americana* at 2–8 months. (c) Chemical structures of five nitrogen-containing metabolites (NCMs). (d) The results of UPLC analysis. (e) Heat map for NCMs content in fecal samples. Each column represents one sample and each row represents one of NCMs.

pathways regulating cell survival<sup>46</sup>. In the PPI network, ATK1 was identified as the foremost target. SRC<sup>47</sup> and CASP3<sup>48</sup> have been reported as key targets for regulating cell proliferation and apoptosis, while COMT<sup>49</sup> and DRD2<sup>50</sup> have crucial roles in dopamine synthesis and secretion.

The focused potential targets were submitted to STRING for further enrichment analysis. The Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway annotation showed that the potential targets were significantly enriched in 58 pathways (false discovery rate (FDR)  $\leq 0.01$ ) (Supplementary Table S1). The top 20 terms with lower FDR are displayed in bubble charts as shown in Fig. 2b, including neuroactive ligand-receptor interaction, nitrogen metabolism, dopaminergic synapse, tyrosine metabolism, cholinergic synapse, and other biologic mechanisms. The FDR and gene counts for each term correspond to color and size scales of the dots. Each FDR of Gene Ontology (GO) enrichment term was calculated (FDR  $\leq 0.01$ ) in terms of the order from small to large. Overall, the 124 targets were enriched into 916 terms under the three main categories of Biological Process (BP), Cellular Component (CC), and Molecular Function (MF). The top ten terms of the three main categories are shown in Fig. 2c. They are represented by red, green, and blue bars, respectively. The height of the bars represents the number of genes in each term. Within BP category, response to chemical, response to organic substance, and regulation of biological quality were the most enriched. The enrichment terms of CC category involved cell part, plasma membrane, and plasma membrane part, etc. Binding, catalytic activity, and ion binding were the highly represented categories in MF. The above results help to predict possible pharmacological activity and action mechanisms of NCMs.

The KEGG pathway and GO enrichment analysis showed primary metabolic pathways and biological functions of the potential targets. The constructive compounds-targets-pathways-diseases network shown in Fig. 2d displays the predicted relationship among NCMs, potential targets, KEGG pathways, and related diseases. Collectively, these results indicate that NCMs are worthy of being studied for their potential pharmacological activities.

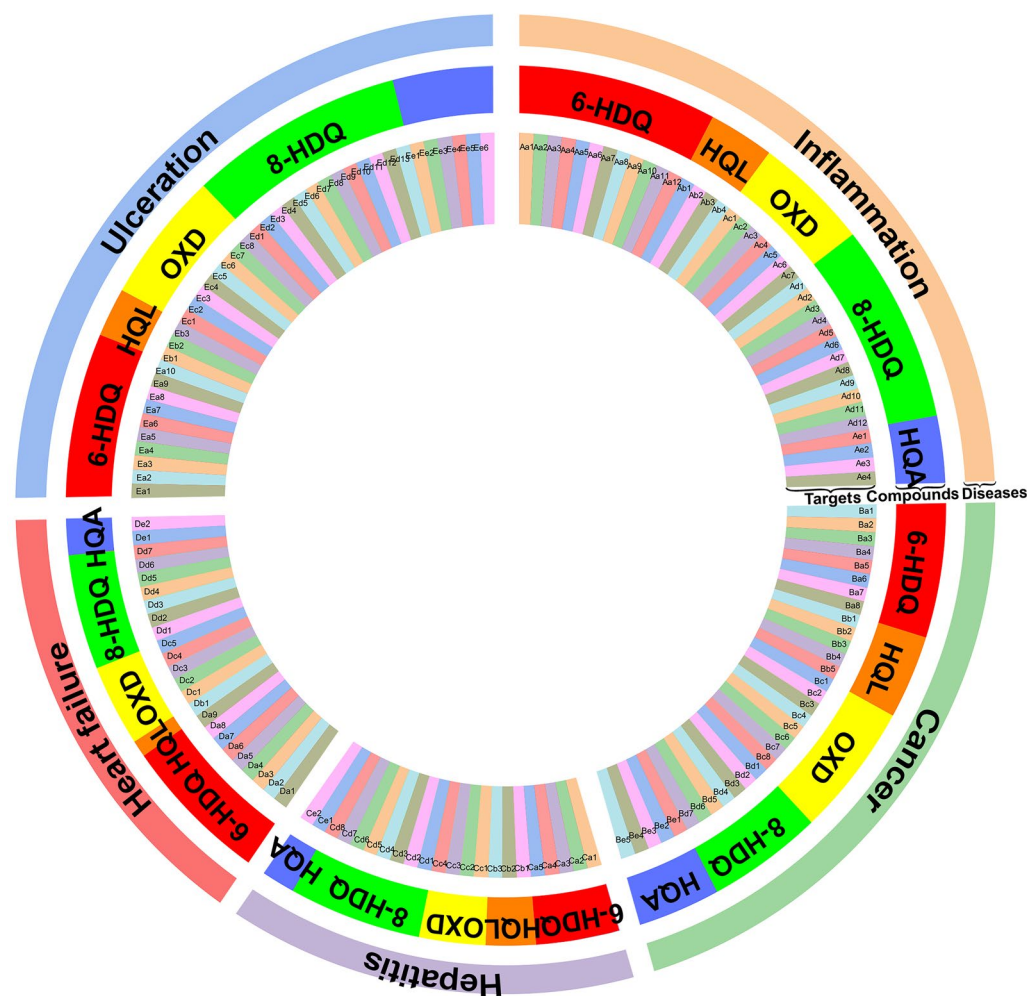
The doughnut chart (Fig. 3) visualized the predictable relation between the *P. americana*-treated diseases and NCMs. The doughnut chart is composed of three doughnuts that respectively represent *P. americana*-treated diseases (the outer ring of doughnut chart), NCMs (the center ring of doughnut chart), and their intersecting targets (the inner ring of doughnut chart). More intersecting targets (Supplementary Table S2) under the



**Figure 2.** The results of network pharmacology analysis of NCMs. **(a)** Protein–protein interaction (PPI) network for the predicted targets of NCMs. **(b)** Top 20 of Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment terms for NCMs. **(c)** Gene Ontology (GO) enrichment terms of NCMs involving Biological Process (BP), Cellular Component (CC), and Molecular Function (MF). **(d)** Compounds–targets–pathways–diseases network represents the potential relationships among NCMs, targets, action mechanisms, and diseases.

No	Name	Degree	Betweenness	Closeness
1	AKT1	45	0.238	0.587
2	SRC	29	0.076	0.516
3	CASP3	29	0.084	0.496
4	COMT	29	0.079	0.489
5	DRD2	29	0.042	0.489
6	ESR1	27	0.109	0.509
7	HTR1A	27	0.038	0.464
8	DRD4	26	0.024	0.474
9	OPRM1	24	0.023	0.481
10	CCND1	23	0.043	0.472
11	OPRK1	22	0.012	0.472
12	CXCR2	22	0.028	0.451
13	MAOB	21	0.041	0.498
14	DRD3	21	0.007	0.459
15	CHRM2	21	0.009	0.424
16	MAOA	20	0.026	0.470
17	OPRD1	20	0.006	0.430
18	OPRL1	20	0.008	0.464
19	SLC6A4	20	0.012	0.437

**Table 1.** The 19 key protein targets of NCMs selected based on three topological parameters, including degree, betweenness, and closeness.



**Figure 3.** The relationships among *P. americana*-treated diseases (ulceration, heart failure, hepatitis, cancer, and inflammation, shown by the outer ring of doughnut chart), NCMs (6-HDQ, HQL, OXD, 8-HDQ, and HQA, shown by the center ring of doughnut chart) and their intersecting targets (Aa1 to Ee6, shown by the inner ring of doughnut chart).

doughnuts mean closer connections between *P. americana*-treated diseases and compounds. This indicates that NCMs are potential active compounds responsible for alleviating diseases such as ulceration and inflammation.

**Optimization of sample extraction and UPLC conditions.** Because NCMs have strong ultraviolet absorption, UPLC equipped with UV detector was selected as a suitable analysis strategy. The extraction conditions were optimized to obtain maximum extraction efficiency with minimum expense. The optimization of UPLC conditions was aimed at acquiring better resolution and sensitivity in a shorter time for quantification of NCMs.

Three processing variables, including methanol concentration, solid–liquid ratio, and time of ultrasonic extraction, were tested in the preliminary experiment to determine the effects on extraction yield of NCMs under the different extraction conditions. The levels of three independent variables were investigated as follow: solvents (30%, 50%, and 70% methanol aqueous solution), solid–liquid ratio (1:20 g/mL, 1:50 g/mL, and 1:100 g/mL), and time of ultrasonic extraction (15 min, 30 min, 45 min, and 60 min). The results proved that the optimal extraction conditions of *P. americana* fecal samples were 70% methanol aqueous solution, 1:50 g/mL solid–liquid ratio, and 30 min ultrasonic extraction.

Several UPLC parameters, including mobile phase, column temperature, and detection wavelength, were modified to obtain optimal separation and sensitivity in a shorter time. Addition of formic acid into mobile phase led to better separation and symmetric peak shape. Meanwhile, it has been proved that temperature of chromatographic column is conducive for producing sharper peaks at 50 °C. Chromatographic peaks could be sensitively obtained at 254 nm.

Compounds	LOD ( $\mu\text{g}/\text{mL}$ )	LOQ ( $\mu\text{g}/\text{mL}$ )	Linear range ( $\mu\text{g}/\text{mL}$ )	Determination coefficient ( $R^2$ )	Precision (RSD%)		Repeatability (RSD%, $n=6$ )	Stability (RSD%, $n=6$ )	Recovery ( $n=6$ )	
					Intra-day ( $n=6$ )	Inter-day ( $n=3$ )			Recovery (%)	RSD (%)
6-HDQ	0.054	0.179	0.179–10.3	0.9999	0.9	1.0	0.7	0.8	99.0	1.0
HQL	0.077	0.257	0.771–49.3	0.9999	0.2	1.3	0.8	0.3	96.5	0.6
OXD	0.080	0.267	0.267–15.4	0.9999	1.1	0.9	0.6	1.1	97.4	1.1
8-HDQ	0.058	0.193	0.579–37.0	0.9999	0.6	1.0	0.5	0.4	95.5	0.8
HQA	0.134	0.447	1.34–85.8	0.9998	2.2	1.5	2.6	0.8	98.1	1.7

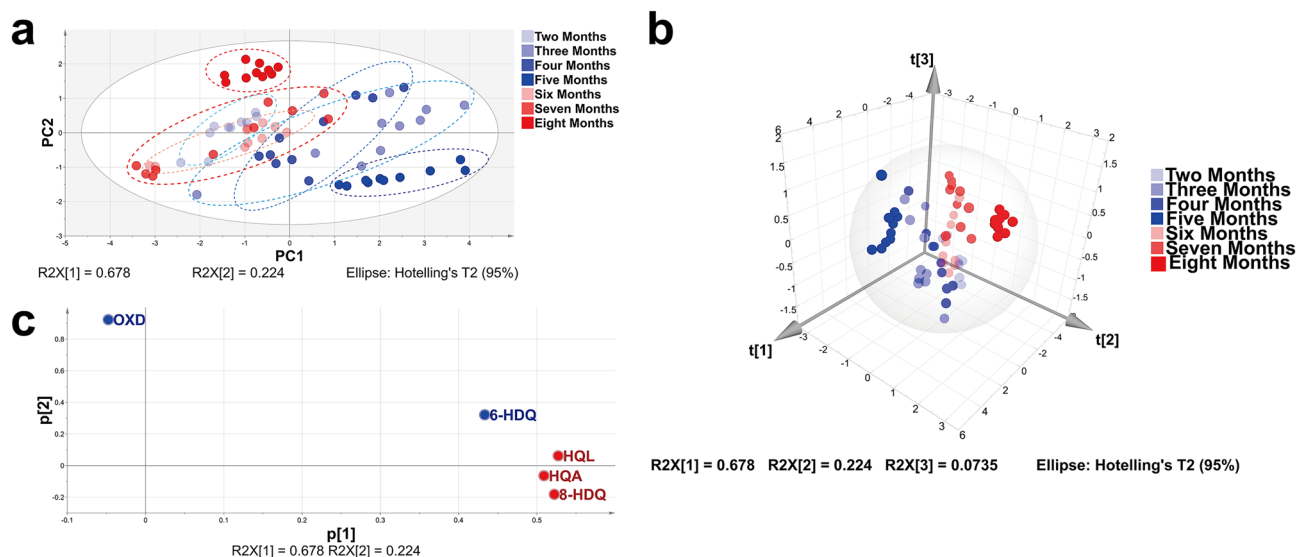
**Table 2.** The validation results of the UPLC method including LOD, LOQ, linear range, precision, repeatability, stability, and recovery.

**Methodological validation of UPLC analysis.** The linear range, limit of detection (LOD), limit of quantification (LOQ), and determination coefficient of the regression equation of NCMs are shown in Table 2. The results indicate that the calibration curves have satisfactory linearity with determination coefficients from 0.9998 to 0.9999. The LODs and LOQs are in the range of 0.054–0.134  $\mu\text{g}/\text{mL}$  and 0.179–0.447  $\mu\text{g}/\text{mL}$ , respectively, suggesting that the developed method has high sensitivity. The results of precision, repeatability, stability, and recovery rate are also listed in Table 2. The RSDs of intra- and inter-day precisions are below 2.2%, indicating that the precision of the quantitative method is acceptable. The RSDs of repeatability are all less than 2.6%, suggesting that the method is repeatable enough for the quantitative evaluation of NCMs. The RSDs of stability range from 0.3% to 1.5% within 12 h at room temperature, indicating that NCMs in the sample solutions show high stability. As shown in Table 2, the average recoveries of NCMs range from 95.5% to 99.0%, and RSDs are in the range of 0.4% and 1.7%, revealing that the method has been validated with excellent recovery and accuracy.

**Content determination and heat map analysis of NCMs from *P. americana* fecal samples.** In order to understand the possible variations of the five metabolites of *P. americana* at different growth stages, fecal samples collected at 2–8 months were employed for the study. The validated UPLC method was subsequently applied to obtain the chromatograms and detect NCMs in 70 batches fecal samples of *P. americana*. The representative chromatograms of reference materials and samples from the different stages are successively presented in Fig. 1d.

In order to exhibit differences in the metabolic profiles of *P. americana* feces at the seven growth stages, the content database of NCMs was used to construct a heat map, in which the brighter red color indicated the higher content of metabolites, while the brighter blue color showed the lower content. As shown in Fig. 1e, the relative content of 6-HDQ, HQL, 8-HDQ, and HQA is high in samples collected at 3–5 months but decreased with an increased age. However, sudden increase of OXD and 6-HDQ in their content levels is observed in samples collected at 8 months. Generally, these results imply that high content of NCMs is detected during the nymph stage of *P. americana* than the adult stage, which could be likely attributed to changes occurring in *P. americana* compositional state of gut microbiota as the insect aging. It indicates that the quantification of NCMs is essential for the further investigation of quality control and pharmacodynamics of metabolites from *P. americana* gut microbiota. Therefore, the differential metabolites in samples collected at the different months are worth exploring further.

**Principal components analysis (PCA) and orthogonal partial least squares discriminant analysis (OPLS-DA) of gut microbiota metabolites from *P. americana* at the different growth stages.** PCA, an unsupervised pattern recognition technique, is used to reduce the dimensionality of the data<sup>51</sup> and summarize the main characteristics of variables for building predictive models that can be visualized to ascertain relatedness among samples<sup>52</sup>. This tool provides the key information from chemical content data, with a depth overview of clustering patterns. In order to further locate the possible outliers and visualize the potential separation tendency among samples from *P. americana* with the different ages, PCA analysis involving unit variance scaling was performed on the content data of NCMs. The results indicated that the total variance explained by PC1 and PC2 accounted for 90.2% accumulation contribution rate. The PCA score plot (Fig. 4a) showed that the 70 batches samples were divided into seven groups at 95% confidence level without an outlier. All samples except for the samples from 5 and 8 months had intersection areas in the plot, which were not fully separated from the other samples. However, as shown in Fig. 4a, the samples from adult (6–8 months) represented by various shades of red circular dots tend to cluster to the left part, while the samples from nymph (2–5 months) represented by various shades of blue circular dots scatter to the right, depicting that the gut microbial metabolites of nymphs are approximately separated from adults. The three-dimensional (3D) score plot (Fig. 4b) shows the rough spatial separation for nymphs and adults. This separation likely suggests that the gut microbiota compositional state and their dependent metabolic activities change with the aging of *P. americana*. The loading score plot (Fig. 4c) displays that HQL, HQA, and 8-HDQ demonstrate greater contributions in PC1, which are primary factors that influence the separation. Moreover, PCA suggests that NCMs change with growth time, showcasing that NCMs could be used as reference biomarkers for evaluating *P. americana* gut microbiota-dependent metabolites at the different growth stages.



**Figure 4.** The results of principal components analysis (PCA) for NCMs from *P. americana* at the different growth stages. (a) Score plot of PCA model for differentiating samples at the different growth stages. (b) 3D score plot of PCA model based on the three principal components. (c) Loading score plot from PCA model.

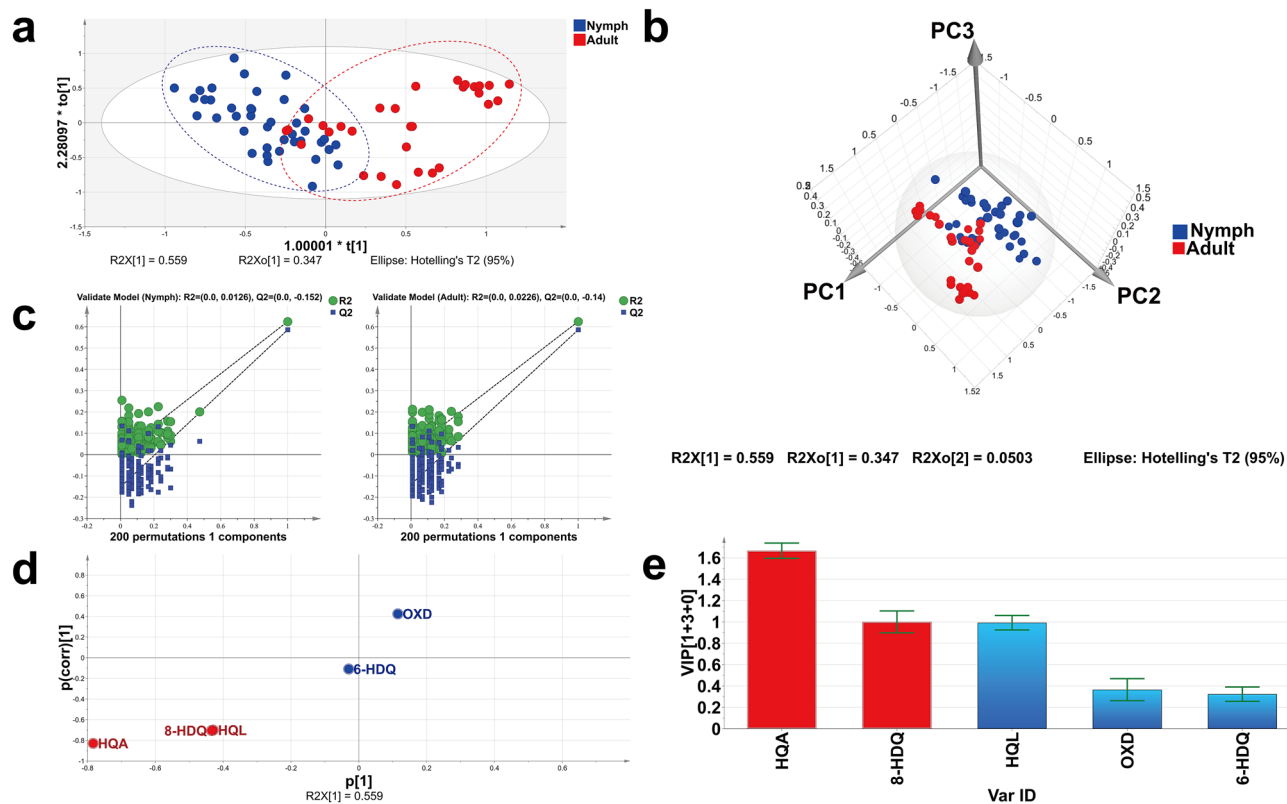
The OPLS-DA model, which incorporates an inbuilt orthogonal signal correction filter, can effectively extract principal variables from the known chemical variables to distinguish samples from different sources<sup>53</sup>. Comparing with PCA model, the OPLS-DA model provides a better effective separation performance. In order to find out the different significant metabolites in the feces of nymphs and adults, the quantitative data of NCMs for nymphs aged from 2 to 5 months and adults aged from 6 to 8 months were further analyzed by OPLS-DA. The data matrix was dealt with Pareto (Par) scaling. Compared with PCA model, interpretability ( $R_{\text{cumulated}}^2 = 0.625$ ) and prediction performance ( $Q^2_{\text{cumulated}} = 0.586$ ) of the OPLS-DA model with one predictive and three orthogonal (1 + 3) components were enhanced. As shown in Fig. 5a, the adult and nymph samples are clustered in the right and left parts of the plot, respectively. An obvious separation between nymphs and adults is also found in 3D score plot (Fig. 5b). The  $p$ -value ( $2.599 \times 10^{-9}$ ) acquired by CV-ANOVA tests is less than 0.05, which testify the predictive capability of this model. After that, the permutation test ( $n = 200$ ) was performed to validate the model performance. As shown in Fig. 5c, the values of  $R^2$  ( $= 0.0, 0.0126$ ) and  $Q^2$  ( $= 0.0, -0.152$ ) for nymphs and  $R^2$  ( $= 0.0, 0.0226$ ) and  $Q^2$  ( $= 0.0, -0.140$ ) for adults indicate that the OPLS-DA model has a low risk of overfitting. It is manifested that HQA, 8-HDQ, and HQL are the important classification elements as shown in S-plot (Fig. 5d). Also, the two variables (HQA and 8-HDQ) are selected as the differential metabolites with important discrimination capability by virtue of variable importance in projection (VIP)  $\geq 1$  (Fig. 5e). The result reveals that HQA and 8-HDQ are the distinguished metabolic markers for nymphs and adults. Accordingly, a similar variation tendency for HQA and 8-HDQ in different stages is observed in heat map plots, indicating that the aging of *P. americana* influences the relative amount of HQA and 8-HDQ. Additionally, the two differential metabolites are closely related to ulceration than other diseases, as shown in Fig. 3 and Supplementary Table S2.

## Conclusion

In this study, a network pharmacology approach was devoted to elucidating possible pharmacological activity and action mechanisms of NCMs. Meanwhile, a fast and stable UPLC method was established to simultaneously quantify NCMs from *P. americana* at the different growth stages. Multivariate statistical methods and a network pharmacology approach were used to analyze the differences between nymphs and adults of *P. americana*. Our findings show that HQA and 8-HDQ can be used as the distinguished metabolic markers between nymphs and adults. The research contributes to analyzing metabolites from insects' gut microbiota and provides a theoretical basis for developing drugs related to *P. americana*.

## Materials and methods

**Network pharmacology analysis.** In this study, NCMs' connections with potential targets, pathways, and related diseases were explored via a network pharmacology analysis. The potential targets of NCMs were obtained from the Swiss Target Prediction database platform (<http://www.swisstargetprediction.ch/>). After deletion of repetitive terms, target symbols of each NCM were merged. Subsequently, the focused targets were imported into the STRING database platform (<https://string-db.org/>) for PPI analysis, GO terms enrichment analysis, and KEGG pathway enrichment analysis. Furthermore, the PPI network was visualized by Cytoscape 3.7.1 software after deletion of disconnected targets. The degrees, betweenness centrality, and closeness centrality, which indicated the topological importance of targets in the PPI network, were analyzed using Network Analyzer function of Cytoscape software. The terms of GO and KEGG pathways were selected by taking FDR into account. Then the related diseases associated with the filtered KEGG pathways were collected by KEGG database platform (<https://www.kegg.jp/kegg/>)<sup>54</sup>. The GO terms enrichment analysis was carried out to reveal the GO



**Figure 5.** The results of orthogonal partial least squares discriminant analysis (OPLS-DA) for NCMs from nymphs and adults. (a) Score plot of OPLS-DA model for differentiating adults from nymphs. (b) 3D score plot of OPLS-DA model based on the three principal components. (c) Permutation plots of OPLS-DA model with 200 permutation tests. (d) S-plot of OPLS-DA model. (e) Variable importance for the projection (VIP) of each compound.

terms, which were most enriched in BP, CC, and MF. The visualization of KEGG pathways and GO terms were performed by Origin 2019b software (Originlab Corp., Northampton, MA, USA). Finally, with the obtained data, Origin software was used to construct the compounds-targets-pathways-diseases network, displaying the predictable clinical values and action mechanisms of NCMs.

With the notion that NCMs might have similar pharmacological activities to the extracts of *P. americana*, we tried to explore the relationship between NCMs and *P. americana*-treated diseases. Five *P. americana*-treated diseases (including ulceration, heart failure, hepatitis, cancer, and inflammation) were chosen to collect their related targets, which were obtained by capturing the top 1000 closely relevant disease targets exported from GeneCards database platform (<https://www.genecards.org/>). The targets of NCMs were obtained from the Swiss Target Prediction database platform as described above. The intersection between targets of NCMs and *P. americana*-treated diseases was performed by Venn platform (<https://bioinfo.gp.cnb.csic.es/tools/venny/>) and visualized by a form of chart composed of three doughnuts. It was tried to expound the relational degrees between NCMs and the treated diseases by counting intersecting targets.

**Materials and reagents.** The fecal samples of *P. americana*, a gift from Yunnan Tengchong Pharmaceutical Co., Ltd, were collected from Stone Mountain GAP breeding base of *P. americana*. *P. americana* was raised strictly according to the GAP standards, and their main diet was soybean powder. Fecal samples of 70 batches were collected at seven growth stages (including 2, 3, 4, 5, 6, 7, and 8 months), and ten batches of each growth stage were collected from the different workshops, which were numbered from 2-1~2-10 to 8-1~8-10. The samples collected are shown in Fig. 1a and b. All samples were stored at  $-80^\circ\text{C}$ .

As shown in Fig. 1c, five NCMs including 6-HDQ, HQL, OXD, 8-HDQ, and HQA were obtained with high purity ( $\geq 95\%$ ) from the feces of *P. americana* by our laboratory. Methanol, acetonitrile, and formic acid were purchased from Sigma-Aldrich Company (St. Louis, MO, USA). Ultrapure water was supplied by a Milli-Q system (Millipore, Billerica, MA, USA). All chemicals were of chromatographic grade.

**Preparation of standard solution.** The NCMs were used as the reference materials and accurately weighed. 6-HDQ and HQL were respectively dissolved in 70% methanol aqueous solution ( $v/v$ ) to obtain their stock solutions (0.515 and 0.548 mg/mL). Similarly, OXD, 8-HDQ, and HQA were dissolved in 30% methanol aqueous solution to prepare their stock solutions at the concentrations of 0.513, 0.529, and 1.072 mg/mL, respectively. Afterward, the five stock solutions were mixed and diluted with 30% methanol aqueous solution to prepare the mixed stock solution, whose concentrations were 10.3  $\mu\text{g/mL}$  for 6-HDQ, 49.3  $\mu\text{g/mL}$  for HQL,



15.4 µg/mL for OXD, 37.0 µg/mL for 8-HDQ, and 85.8 µg/mL for HQA, respectively. All reference solutions were stored at 4 °C before determination.

**Preparation of sample solution.** The powdered sample (0.5 g) was transferred into 25 mL volumetric flask and ultrasonically extracted by 70% methanol aqueous solution (540 W, 35 kHz) for 30 min at room temperature. The extracted solution was adjusted to the scale by adding 70% methanol aqueous solution. An aliquot (2.0 mL) of each sample was centrifuged at 14,000 rpm for 10 min, and then the supernatant was filtered through 0.45 mm PTFE membranes. An aliquot (1.0 mL) was diluted with 1.0 mL ultrapure water and mixed. All the samples were carried out in triplicates.

**UPLC analysis.** All samples were analyzed on a Waters ACQUITY UPLC system equipped with a PDA detector (Waters, Milford, MA, USA). The mobile phase system was composed of 0.1% formic acid in water (A) and acetonitrile (B). Samples were separated on a Waters ACQUITY UPLC® HSS T3 chromatographic column (2.1 × 100 mm, 1.8 µm, Waters). The gradient program was performed as follows: 0–5 min, 98%–93% A; 5–9 min, 93%–90% A; 9–19 min, 90%–78% A; and 19–20 min, 78%–98% A. The injection volume was 2 µL and the flow rate was maintained at 0.3 mL/min. The column temperature was set at 50 °C and the detection wavelength was 254 nm.

**Methodological validation of the quantitative analysis.** Series of seven calibration solutions were prepared to construct regression equations, including 0.179–10.3 µg/mL for 6-HDQ, 0.771–49.3 µg/mL for HQL, 0.267–15.4 µg/mL for OXD, 0.579–37.0 µg/mL for 8-HDQ, and 1.34–85.8 µg/mL for HQA. Calibration curves were constructed by plotting the peak areas versus the corresponding concentration of the tested compounds. The slopes, intercepts, and determination coefficients of the calibration curves were calculated by linear regression analysis. The sensitivity was expressed by LOD and LOQ for the tested compounds. In general, the LODs and LOQs were determined at a signal–noise ratio (S/N) of about 3 and 10 by serially diluting the calibration solution. The intra- and inter-day precisions were calculated by determining the selected sample in six replicates during a single day and by duplicating the experiments for three consecutive days. The repeatability was described by analyzing six samples from the same batch using the same method. The stability was also investigated at 0, 2, 4, 6, 8, 10, and 12 h after initial storage at room temperature. The precision, repeatability, and stability were expressed by RSD. The recovery was validated by adding the standard solution to six parallel sample solutions with known concentrations previously. The mixture was analyzed by the normal method and performed in six replicates. The recovery percentages and RSD were calculated by the formula: Recovery (%) = (Detected amount – Original amount)/Spiked amount × 100%, RSD (%) = (SD/mean) × 100%.

**Multivariate statistical analysis.** Data matrix composed of 70 observations and five variables were constructed and analyzed by multivariate statistical analysis. The observations and variables severally represented samples and NCMs. The preprocessing of data, the calculation of parameters, and the quantitative analysis of NCMs were carried out with Microsoft Excel (Microsoft Corp., Redmond, WA, USA). The heat map was used to describe the variations in content, and the visual image was exported by Origin software. SIMCA-P 14.1 software (Umetrics, Umea, Sweden) was used for multivariate analysis, including PCA and OPLS-DA.

## Data availability

The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Received: 26 July 2021; Accepted: 7 October 2021

Published online: 27 October 2021

## References

1. Tee, H. S., Saad, A. R. & Lee, C. Y. Population ecology and movement of the American cockroach (Dictyoptera: Blattidae) in sewers. *J. Med. Entomol.* **48**, 797–805 (2011).
2. Zeng, C. *et al.* The role of *Periplaneta americana* (Blattodea: Blattidae) in modern versus traditional Chinese medicine. *J. Med. Entomol.* **56**, 1522–1526 (2019).
3. Li, Q. J., Wang, Z. G., Liu, Q., Xie, Y. & Hu, H. L. Research status of *Periplaneta americana* with analyses and prospects of key issues. *Chin. J. Chin. Mater. Med.* **43**, 1507–1516 (2018).
4. Nguyen, T. *et al.* Antipyretic, anti-inflammatory and analgesic activities of *Periplaneta americana* extract and underlying mechanisms. *Biomed. Pharmacother.* **123**, 109753 (2020).
5. Basseri, H. R., Dadi-Khoeni, A., Bakhtiari, R., Abolhassani, M. & Hajhosseini-Baghdadabadi, R. Isolation and purification of an antibacterial protein from immune induced Haemolymph of American Cockroach, *Periplaneta americana*. *J. Arthropod. Borne. Dis.* **10**, 519–527 (2016).
6. Zhao, Y., Yang, A., Tu, P. & Hu, Z. Anti-tumor effects of the American cockroach, *Periplaneta americana*. *Chin. Med.* **12**, 26 (2017).
7. Li, D. *et al.* Anti-fibrotic role and mechanism of *Periplaneta americana* extracts in CCl<sub>4</sub>-induced hepatic fibrosis in rats. *Acta Biochim Biophys Sin (Shanghai)*. **50**, 491–498 (2018).
8. Li, J., Shi, W., Zhang, J. & Ren, L. To explore the protective mechanism of PTEN-induced Kinase 1 (PINK1)/Parkin mitophagy-mediated extract of *Periplaneta americana* on lipopolysaccharide-induced cardiomyocyte injury. *Med. Sci. Monit.* **25**, 1383–1391 (2019).
9. Li, L. J. *et al.* The crude ethanol extract of *Periplaneta americana* L. stimulates wound healing in vitro & In vivo. *Chin Med.* **14**, 33 (2019).
10. Zou, Y. *et al.* *Periplaneta americana* extracts accelerate liver regeneration via a complex network of pathways. *Front. Pharmacol.* **11**, 1174 (2020).

11. Li, H. B. *et al.* Efficacy and safety of kangfuxin liquid combined with aminosalicic acid for the treatment of ulcerative colitis: a systematic review and meta-analysis. *Medicine (Baltimore)*. **97**, e10807 (2018).
12. Zhang, X. Q. *et al.* Preventive effect of ganlong capsule on chronic alcoholic hepatic injury in rats. *Chin. J. Chin. Mater. Med.* **38**, 2197–2201 (2013).
13. Qi, J. *et al.* Mechanisms of Chinese medicine xinmailong's protection against heart failure in pressure-overloaded mice and cultured cardiomyocytes. *Sci. Rep.* **7**, 42843 (2017).
14. Lu, X. *et al.* Clinical efficacy and safety of xinmailong injection for the treatment of chronic heart failure: a meta-analysis. *Front. Pharmacol.* **9**, 810 (2018).
15. Saha, S., Rajpal, D. K. & Brown, J. R. Human microbial metabolites as a source of new drugs. *Drug Discov. Today* **21**, 692–698 (2016).
16. Wang, Y. *et al.* Intestinal microbiota-associated metabolites: crucial factors in the effectiveness of herbal medicines and diet therapies. *Front. Physiol.* **10**, 1343 (2019).
17. Friedland, R. P. & Chapman, M. R. The role of microbial amyloid in neurodegeneration. *Plos Pathog.* **13**, e1006654 (2017).
18. Zuo, T. *et al.* Alterations in fecal fungal microbiome of patients with COVID-19 during time of hospitalization until discharge. *Gastroenterology* **159**, 1302–1310 (2020).
19. Xu, J. *et al.* Jinzhi protects lipopolysaccharide-treated mice against mortality by repairing intestinal mucosal barrier damage and intestinal microecology. *Biomed. Pharmacother.* **123**, 109749 (2020).
20. Zheng, T. *et al.* Metabolomic analysis of biochemical changes in the serum and urine of Freund's adjuvant-induced arthritis in rats after treatment with silkworm excrement. *Molecules* **23**, 1490 (2018).
21. Brito, C., Jordao, V. L. & Pierce, G. J. Ambergris as an overlooked historical marine resource: its biology and role as a global economic commodity. *J. Mar. Biol. Assoc. UK*. **96**, 585–596 (2016).
22. Van Treuren, W. & Dodd, D. Microbial contribution to the human metabolome: implications for health and disease. *Annu. Rev. Pathol.* **15**, 345–369 (2019).
23. Wu, Y., Wang, C. Z., Wan, J. Y., Yao, H. & Yuan, C. S. Dissecting the interplay mechanism between epigenetics and gut microbiota: health maintenance and disease prevention. *Int. J. Mol. Sci.* **22**, 6933 (2021).
24. Allegretti, J. R., Mullish, B. H., Kelly, C. & Fischer, M. The evolution of the use of faecal microbiota transplantation and emerging therapeutic indications. *Lancet* **394**, 420–431 (2019).
25. Wang, L., Ravichandran, V., Yin, Y., Yin, J. & Zhang, Y. Natural products from Mammalian Gut Microbiota. *Trends Biotechnol.* **37**, 492–504 (2019).
26. Ohira, H., Tsutsui, W. & Fujioka, Y. Are short chain fatty acids in gut microbiota defensive players for inflammation and atherosclerosis?. *J. Atheroscler. Thromb.* **24**, 660–672 (2017).
27. Tomomura, S., Ihara, M. & Friedland, R. P. Microbiota in cerebrovascular disease: a key player and future therapeutic target. *J. Cereb. Blood Flow Metab.* **40**, 1368–1380 (2020).
28. Yu, H. *et al.* Facile fluorescence monitoring of gut microbial metabolite trimethylamine *n*-oxide via molecular recognition of Guanidinium-Modified calixarene. *Theranostics*. **9**, 4624–4632 (2019).
29. Lin, P. B. *et al.* Prodigiosin isolated from *Serratia marcescens* in the *Periplaneta americana* gut and its apoptosis-inducing activity in HeLa cells. *Oncol. Rep.* **41**, 3377–3385 (2019).
30. Lynch, S. V. & Pedersen, O. The human intestinal microbiome in health and disease. *N. Engl. J. Med.* **375**, 2369–2379 (2016).
31. Vicente, C. S., Ozawa, S. & Hasegawa, K. Composition of the cockroach gut microbiome in the presence of parasitic nematodes. *Microbes Environ.* **31**, 314–320 (2016).
32. Tinker, K. A. & Ottesen, E. A. The core gut microbiome of the American cockroach, *Periplaneta americana*, is stable and resilient to dietary shifts. *Appl. Environ. Microbiol.* **82**, 6603–6610 (2016).
33. Porter, N. T. & Martens, E. C. The critical roles of polysaccharides in gut microbial ecology and physiology. *Annu. Rev. Microbiol.* **71**, 349–369 (2017).
34. Zhao, Y. *et al.* Dietary protein levels and amino acid supplementation patterns alter the composition and functions of colonic microbiota in pigs. *Anim. Nutr.* **6**, 143–151 (2020).
35. Macfarlane, G. T., Cummings, J. H. & Allison, C. Protein degradation by human intestinal bacteria. *J. Gen. Microbiol.* **132**, 1647–1656 (1986).
36. Sabree, Z. L., Kambhampati, S. & Moran, N. A. Nitrogen recycling and nutritional provisioning by *Blattabacterium*, the cockroach endosymbiont. *Proc. Natl. Acad. Sci. USA* **106**, 19521–19526 (2019).
37. Del, R. B. *et al.* *Enterococcus faecalis* bacteriophage 156 is an effective biotechnological tool for reducing the presence of tyramine and putrescine in an experimental cheese model. *Front. Microbiol.* **10**, 566 (2019).
38. Souza, C., Bezerra, W. P. & Souto, J. T. Marine alkaloids with anti-inflammatory activity: current knowledge and future perspectives. *Mar. Drugs*. **18**, 147 (2020).
39. Yang, B. *et al.* Analysis of gene expression in the midgut of *Bombyx mori* during the larval molting stage. *BMC Genom.* **17**, 866 (2016).
40. Zhu, J. J. *et al.* Bioactivity-guided screening of wound-healing active constituents from American Cockroach (*Periplaneta americana*). *Molecules* **23**, 101 (2018).
41. Li, L. J. *et al.* *Periplaneta Americana* L. as a novel therapeutics accelerates wound repair and regeneration. *Biomed. Pharmacother.* **114**, 108858 (2019).
42. Zhang, J., Zhang, Y., Li, J., Liu, M. & Liu, Z. Midgut transcriptome of the cockroach *Periplaneta americana* and its microbiota: digestion, detoxification and oxidative stress response. *Plos One*. **11**, e0155254 (2016).
43. Wang, Z. *et al.* Pharmacokinetic and metabolism studies of 12-riboside-pseudoginsengin DQ by UPLC-MS/MS and UPLC-QTOF-MSE. *Molecules* **23**, 2499 (2018).
44. Babakobi, M. D. *et al.* Effect of maternal diet and milk lipid composition on the infant gut and maternal milk microbiomes. *Nutrients* **12**, 2539 (2020).
45. Szklarczyk, D. *et al.* STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* **43**, 447–452 (2015).
46. Balasuriya, N. *et al.* Phosphorylation-dependent substrate selectivity of protein kinase B (AKT1). *J. Biol. Chem.* **295**, 8120–8134 (2020).
47. Matozaki, T., Kotani, T., Murata, Y. & Saito, Y. Roles of Src family kinase, Ras, and mTOR signaling in intestinal epithelial homeostasis and tumorigenesis. *Cancer Sci.* **112**, 16–21 (2020).
48. Ma, X. Y. *et al.* Ursolic acid reduces hepatocellular apoptosis and alleviates alcohol-induced liver injury via irreversible inhibition of CASP3 in vivo. *Acta Pharmacol. Sin.* **42**, 1101–1110 (2011).
49. Phoswa, W. N. Dopamine in the pathophysiology of preeclampsia and gestational hypertension: monoamine oxidase (MAO) and catechol-o-methyl transferase (COMT) as possible mechanisms. *Oxid. Med. Cell. Longev.* **2019**, 3546294 (2019).
50. He, Y. *et al.* Epidermal growth factor receptor as a molecular determinant of glioblastoma response to dopamine receptor D2 inhibitors. *Neuro Oncol.* **22**, 6 (2020).
51. Wang, C., Bi, H. & Xie, J. Visualization of the distance among fishes by MALDI MS for rapid determination of the taxonomic status of fish fillets. *J. Agric. Food Chem.* **68**, 8438–8446 (2020).

52. Clark, N. R. & Ma'Ayan, A. Introduction to statistical methods to analyze large data sets: principal components analysis. *Sci. Signal.* **4**, r3 (2011).
53. Boccard, J. & Rutledge, D. N. A consensus orthogonal partial least squares discriminant analysis (OPLS-DA) strategy for multiblock Omics data fusion. *Anal. Chim. Acta.* **769**, 30–39 (2013).
54. Kanehisa, M. & Goto, S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).

## Acknowledgements

We would like to thank Dan Wu and Xiangyue Liu for their assistance in separation of NCMs. This work was supported by grants from the Science and Technology Projects of Tianjin, China (No. 20ZYJJC00070), Graduate Research Innovation Project of Tianjin, China (No. 2019YJSB143), Graduate Research Innovation Project of TUTCM, China (No. YJSKC-20191004), and the Innovation Group of Component-based Chinese Medicine and Intelligent Manufacturing with Multi-crossed Disciplines.

## Author contributions

Conceptualization: W.L., X.C., Y.C., and Y.W. Sample treatment: W.L., G.X., Z.W., and L.N. Data acquisition and analysis: W.L., G.X., and Y.W. Original draft preparation: W.L., X.C., Y.C., and Y.W. Review and editing: W.L., X.C., Y.C., and Y.W. Project administration: W.L., X.C., Y.C., and Y.W. Funding acquisition: W.L., X.C., and Y.W. All authors have read and agreed to the published version of the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-00182-0>.

**Correspondence** and requests for materials should be addressed to X.C. or Y.W.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021