



OPEN

## Construction of TUATinsecta database that integrated plant and insect database for screening phytophagous insect metabolic products with medicinal potential

Wakana Nakane<sup>1,8</sup>, Hisashi Nakamura<sup>1,8</sup>, Takeru Nakazato<sup>2,8</sup>, Natsuki Kaminaga<sup>3</sup>, Miho Nakano<sup>1</sup>, Takuma Sakamoto<sup>4,7</sup>, Maaya Nishiko<sup>4</sup>, Hidemasa Bono<sup>2</sup>, Isao Ogiwara<sup>1</sup>, Yoshikazu Kitano<sup>5</sup>, Kikuo Iwabuchi<sup>1</sup>, Kaoru Kinoshita<sup>3</sup>, Richard J. Simpson<sup>6,7</sup> & Hiroko Tabunoki<sup>1,7</sup>✉

Phytophagous insect larvae feed on plants containing secondary metabolic products with biological activity against other predatory organisms. Phytophagous insects can use their specialised metabolic systems to convert these secondary metabolic products into compounds with therapeutic properties useful to mankind. Some Asians drink tea decoctions made from phytophagous insect frass which is believed to be effective against inflammatory diseases. However, insects that can convert plant-derived secondary metabolic products into useful human therapeutic agents remain poorly studied. Here, we constructed the TUATinsecta database by integrating publicly plant/insect datasets for the purpose of selecting insect species. Using TUAT-insecta we selected the Asian swallowtail butterfly, *Papilio xuthus* larvae fed on several species of Rutaceous plants and examined whether the plant-derived secondary metabolites, especially those present in frass, were chemically altered or not. We extracted metabolic products from frass using three organic solvents with different polarities, and evaluated solvent fractions for their cytotoxic effects against several human cell lines. We found that chloroform frass extracts from *P. xuthus* larvae fed on *Poncirus trifoliata* leaves contained significant cytotoxic activity. Our findings demonstrate that screening of insect species using the 'TUATinsecta' database provides an important pipeline for discovering novel therapeutic agents that might be useful for mankind.

To date, many medicines have been extracted from natural products. Drugs from natural products comprised 42% of all approved medicines between 1981 and 2014<sup>1</sup>.

Surprisingly, 50% of approved anticancer agents are derived from natural products<sup>1</sup>. For example, paclitaxel, vincristine, and camptothecin are plant-derived anticancer agents obtained from *Taxus brevifolia*<sup>2</sup>, *Catharanthus roseus*<sup>3</sup>, and *Camptotheca acuminata*<sup>4</sup>, respectively. However, resources for plant-derived natural products are being depleted, and in pursuing sufficient resources for natural products, researchers have recognised the

<sup>1</sup>Department of Science of Biological Production, Graduate School of Agriculture, Tokyo University of Agriculture and Technology, 3-5-8 Saiwai-cho, Fuchu-shi, Tokyo 183-8509, Japan. <sup>2</sup>Database Center for Life Science (DBCLS), Joint Support-Center for Data Science Research, Research Organization of Information and Systems (ROIS), Mishima, Shizuoka 411-8540, Japan. <sup>3</sup>Department of Pharmacognosy and Phytochemistry, Meiji Pharmaceutical University, 2-522-1 Noshio, Kiyose-shi, Tokyo 204-8588, Japan. <sup>4</sup>Department of United Graduate, School of Agricultural Science, Tokyo University of Agriculture and Technology, 3-5-8 Saiwai-cho, Fuchu-shi, Tokyo 183-8509, Japan. <sup>5</sup>Department of Applied Biological Science, Tokyo University of Agriculture and Technology, 3-5-8 Saiwai-cho, Fuchu-shi, Tokyo 183-8509, Japan. <sup>6</sup>Department of Biochemistry and Genetics, La Trobe Institute for Molecular Science (LIMS), La Trobe University, Melbourne, VIC 3086, Australia. <sup>7</sup>Institute of Global Innovation Research, Tokyo University of Agriculture and Technology, 3-5-8 Saiwai-cho, Fuchu, Tokyo 183-8509, Japan. <sup>8</sup>These authors contributed equally: Wakana Nakane, Hisashi Nakamura, and Takeru Nakazato. ✉email: h\_tabuno@cc.tuat.ac.jp

need to identify other resources which include secondary metabolites (also referred to as metabolic products) for utilisation as pharmaceutical agents. The majority of resources for natural products are natural plants, many of which are difficult to breed as crops. The need to find new pharmaceutical resources is therefore pressing.

Over one million species of insects are estimated to exist worldwide<sup>5</sup>. Insects have a wide variety of feeding habitats—for example, plant leaves, plant stems, flower nectar, the bodies of trees, tree sap, timber, fruits and the blood, carcasses or excrement of other animals. Each insect species feeds on species-specific food. Insect species which eat only plants are classified phytophagous insects. Secondary metabolites, including those of plants, are known to act as factors that stimulate or inhibit feeding behaviour in insects<sup>6–8</sup>. Some plant secondary metabolites—such as alkaloids—are widely known to have toxic effects on other organisms and cannot be consumed unless they are detoxified. Therefore, the types of plants that insects are able to eat are limited, and most insects have a specific host plant.

Phytophagous insects have evolved a unique system that metabolises secondary metabolites produced by the host plant for escaping food competition<sup>9–11</sup>. Furthermore, they keep the secondary metabolites taken up from the plant in their bodies, and utilise them for escaping from predators<sup>12,13</sup>. In this way, phytophagous insects and plants have a strong coevolutionary relationship. Lepidopteran insects, in particular, may evolve a diverse metabolic process to utilise specific plants. Plant secondary metabolites absorbed into Lepidopteran insect bodies can be converted into other metabolic products with unique structures. For example, swallowtail butterflies use the variety of plants containing herbal substances such as Rutaceae, Apiaceae, Aristolochiaceae, Lauraceae, Magnoliaceae, Annonaceae, or Papaveraceae as their major host plants, converting plant-derived second metabolites into other substances in their bodies<sup>14</sup>. The tobacco hornworm (*Manduca sexta*) and tiger moth (*Arctia caja*) tolerate or even utilise plant secondary metabolites that are toxic to other organisms, to protect against predation or for synthesis of pheromones<sup>15–17</sup>. It is therefore expected that these Lepidopteran insect bodies—and the final metabolites of their frass (also called droppings)—contain plant-derived substances that have changed the chemical structure and biological activity through Lepidopteran insect-specific metabolism.

Two examples of compounds derived from Lepidopteran insects that are useful for humans are papilistatin which has anticancer activity (isolated from *Byasa polyeuctes termessus*<sup>18</sup>) and an analgesic agent, 20-hydroxyecdysone, isolated from the silkworm *Bombyx mori*<sup>18</sup>.

However, compared to the number of studies examining pharmacologically active substances in the insect body, very few studies have examined pharmacologically active substances in insect frass. The insect frass is therefore likely to contain candidates for pharmaceutical agents. It is, however, difficult to find an insect species that excretes frass containing pharmaceutical agents useful for humans and research on using insect frass as pharmaceutical resources is lacking worldwide.

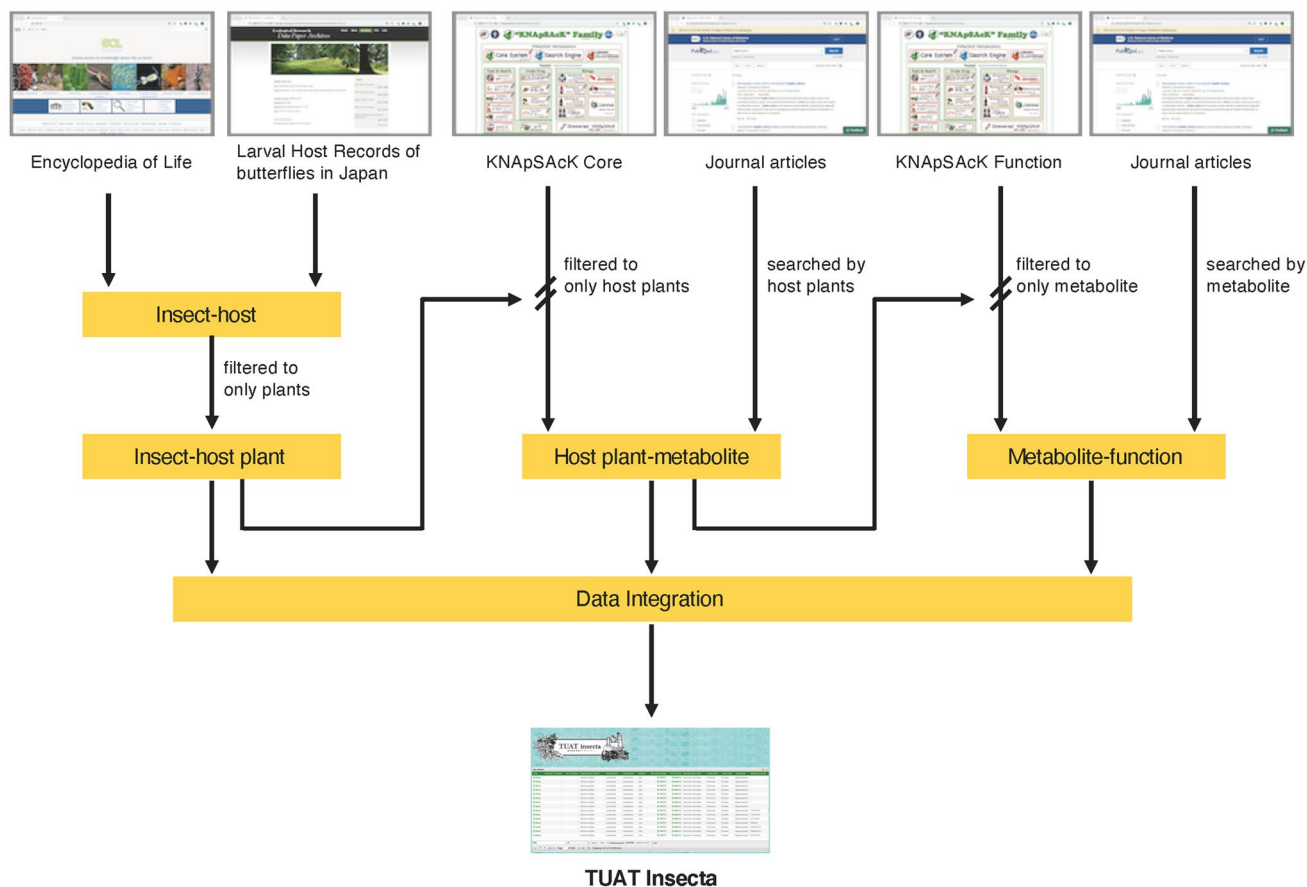
For these reasons, in this study, we constructed a database for choosing insect species that eat plants containing herbal substances or specific secondary metabolites of plants. We then focused on the butterfly *Papilio xuthus*, larvae fed on several species of Rutaceous plants. And we compared the biological activities among each substance which extracted from larval frass fed on *Poncirus trifoliata*, *Zanthoxylum piperitum*, *Citrus natsudaikai*, *Citrus junos* or *Citrus sudachi*.

## Results

**Development of the database “TUATinsecta” for choosing an insect.** We collected information about insect-host plant, host plant-secondary metabolite and metabolite-biological function interactions from public data sources and then integrated these into a database (Fig. 1). As a result, our database TUATinsecta ([https://togodb.org/db/tuat\\_insecta](https://togodb.org/db/tuat_insecta)) contained 1,092,574 relationships between 15,505 insects and 11,578 host plants, and assigned 12,807 metabolites and 2,857 biological functions to 2,831 plants. Overall, TUATinsecta provided 1,132,984 relationships among insects, host plants, metabolites in plants, and the biological activity of metabolites. Lepidopteran insects comprise 10,506 (67.8%) of the insects in TUATinsecta (Table 1). Also, the breakdown of the 11,578 registered plants species revealed that 88.5% were plants eaten by butterflies or moths: Fabales, 9.4%; Poales, 7.3%; Lamiales, 6.8%; Malpighiales, 6.6%; Rosales, 6.3%; Asterales, 5.3%; Sapindales, 4.4%; and Gentianales, 4.2% (Table 2).

Next, using TUATinsecta, we investigated the biological activity of each plant fed on Lepidopteran insects; among the available studies, we encountered reports on diabetes-related (0.18%), cardiotoxic (0.85%), antioxidant (1.5%), anti-inflammatory (2.3%), and anticancer activity (2.4%) in these plants, which had already had been used historically as folk medicine or herbal medicine. We were interested in whether or not the insect metabolic system changes second metabolites from these plants. Information in the database suggested that these plant second metabolites could change through the metabolic system in the insect. Some of the uses of Lepidopteran insects for secondary metabolites of plants included changing part of their chemical structure to become pheromones, kairomones, or defensive substances<sup>15–18</sup>. To compare the biological activity among the varieties of host plants metabolised by insect, we chose the Asian swallowtail *P. xuthus*, which eats several species of Rutaceae plants.

**Comparison of substances in different plants through swallowtail metabolic processes.** The Asian swallowtail, *P. xuthus* is Lepidopteran insect. Their life stages are shown as Fig. 2a–d. They feed on several species of Rutaceae plants during the larval developmental stage (Fig. 2b), and we examined whether secondary metabolic products including these plants may be altered further through the metabolic process in *P. xuthus*. Using three types of organic solvents of differing polarity, we extracted the metabolic products from each frass of *P. xuthus* larvae which fed on *Poncirus trifoliata* (Fig. 2e), *Zanthoxylum piperitum* (Fig. 2f), *Citrus sudachi* (Fig. 2g), *Citrus natsudaikai* (Fig. 2h), or *Citrus junos* (Fig. 2i). In Supplementary Table 1 we show the dry weight and extraction weight in each sample. The extraction dry weight of all frass methanol extracts (MeOH ext.) were higher than those of *n*-hexane extracts (Hex ext.) and chloroform extracts (CHCl<sub>3</sub> ext.) in this study.



**Figure 1.** Outline for construction of the database TUAT insecta. Information about insect-host plant, host plant-metabolite and metabolite-function interactions were collected from public data sources and then integrated into a database, called TUAT insecta.

Order	Number of entries
Lepidoptera	10506
Coleoptera	1162
Hemiptera	1029
Diptera	775
Hymenoptera	308
Trichoptera	98
Other	314

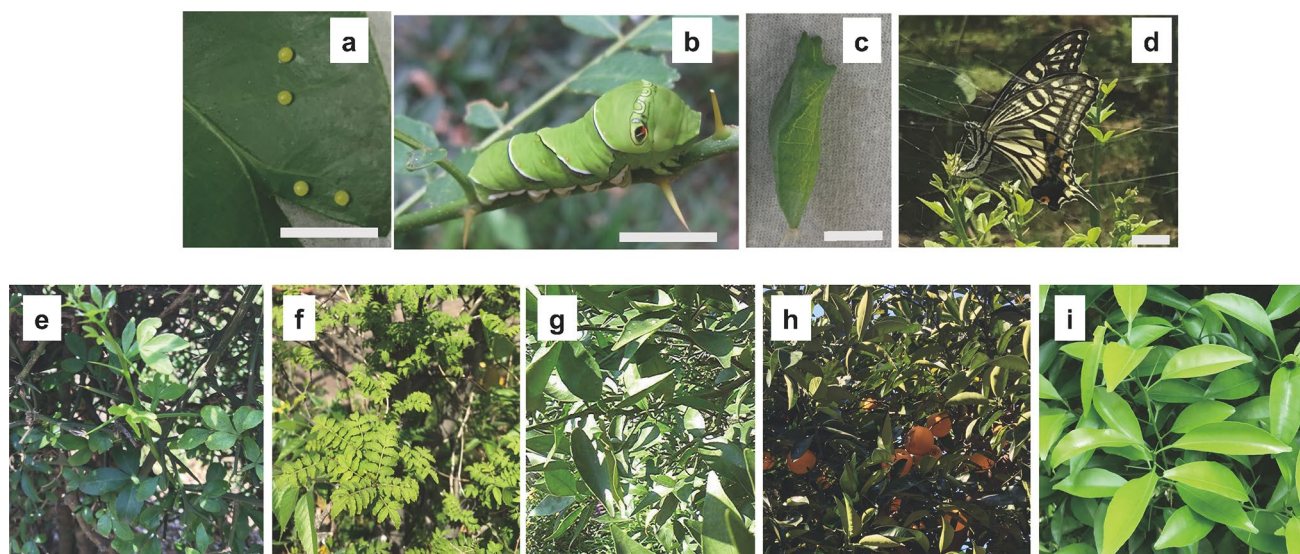
**Table 1.** Number of entries of insects to TUATinsecta.

We compared each metabolic product between leaf and frass in MeOH ext., Hex ext., or CHCl<sub>3</sub> ext., by thin layer chromatography (TLC). Consequentially, we found several spots in the frass extracts which not evident in each corresponding plant extract (Fig. 3a–c, lane 1, 4, 6, 8, 10).

**Observation of the morphological changes to human cancer cell lines.** We evaluated the cytotoxic activity of the substances extracted from each sample on three human cancer cell lines: the human liver cancer cell line (HepG2), the human uterine cancer cell line (HeLa), and the human pancreatic cancer cell line (MIA Paca-2). First, we examined morphological changes using the phenotypic change score in these human cancer cell lines (Supplementary Table 2, a–g). We found that the frass from *P. trifoliata* CHCl<sub>3</sub> ext. induced cell death in all three human cancer cell lines (Supplementary Table 2, a, c, e). We observed the strongest phenotypic change based on the phenotypic change score in the frass from *P. trifoliata* CHCl<sub>3</sub> ext. on HepG2 and MIA-Paca2 (Fig. 4a, c). The frass from *P. trifoliata* CHCl<sub>3</sub> ext. suppressed HepG2, HeLa, and MIA Paca-2 cell growth and cell viability, or induced morphological changes as assessed by the phenotypic change score. The CHCl<sub>3</sub> ext. from pupae which had been fed on *P. trifoliata* had no effect on HepG2, HeLa, or MIA-Paca2 cells (Fig. 4, Supplementary Table 2-g).

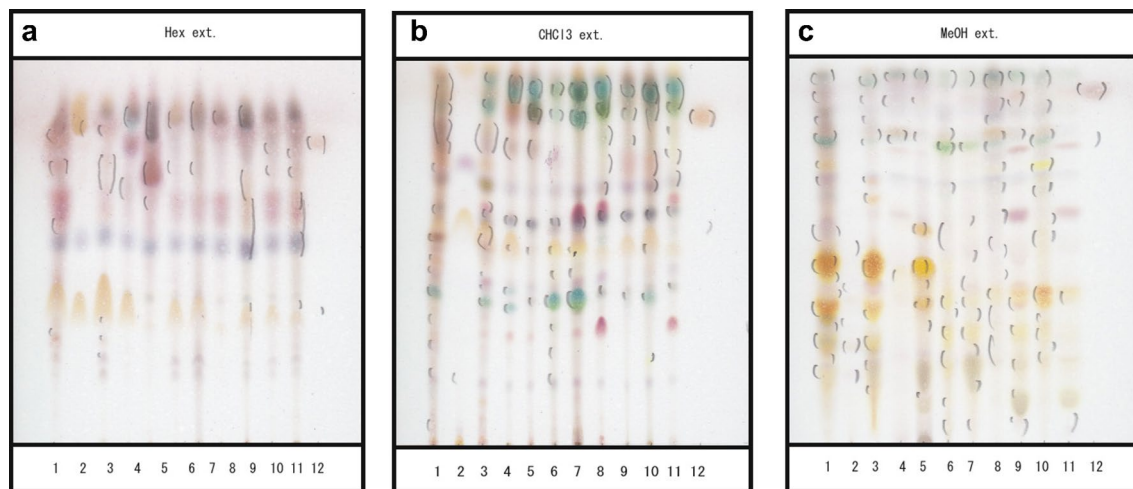
Order	Number of entries
Fabales	1088
Poales	846
Lamiales	791
Malpighiales	763
Asterales	608
Rosales	732
Sapindales	510
Gentianales	490
Caryophyllales	320
Myrtales	249
Brassicales	377
Malvales	334
Ericales	313
Apiales	306
Fagales	214
Asparagales	156
Arecales	190
Solanales	183
Saxifragales	140
Ranunculales	129
Other	1460

**Table 2.** Number of entries of plants to TUAT insecta (plants = 64 orders).

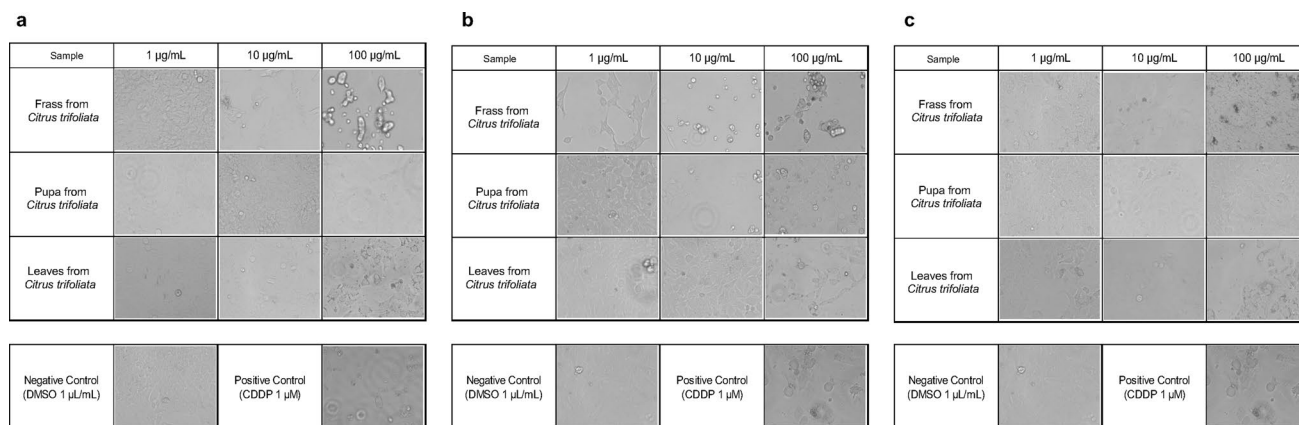


**Figure 2.** Plants and insects used in this study. (a–d) Developmental stages of *P. xuthus*. (a) *P. xuthus* egg, (b) *P. xuthus* final instar larva, (c) *P. xuthus* pupa, and (d) *P. xuthus* adult. The bars indicate 1 cm. (e–i) The five kinds of plants that are the host plants for larvae of *P. xuthus*. (e) *P. trifoliata*, (f) *Z. piperitum*, (g) *C. sudachi*, (h) *C. natsudaidai*, and (i) *C. jumos*.

**Evaluation of each extract on cell survival.** We next evaluated cell survival capability of extracts from *P. trifoliata* frass and leaf extracts and insect pupa which had been fed on *P. trifoliata* leaves. HepG2, HeLa, and MIA Paca-2 cells were treated with each extract and cell viability of each cell line was evaluated by MTT assay (Fig. 5, Supplementary Table 3, a–c). In Supplementary Table 3, a–c, we show the calculated estimated  $IC_{50}$  of each extract. The frass  $CHCl_3$  ext. derived from *P. trifoliata* induced dose-dependent cell death in HepG2, HeLa, and MIA Paca-2 (Fig. 5). We calculated estimated  $IC_{50}$  of frass  $CHCl_3$  ext. derived from *P. trifoliata* to be 11.8  $\mu\text{g}/\text{mL}$  for HepG2 cells, 51.8  $\mu\text{g}/\text{mL}$  for HeLa cells, and 3.7  $\mu\text{g}/\text{mL}$  for MIA Paca-2 cells, and the  $IC_{50}$  of leaf  $CHCl_3$  ext. derived from *P. trifoliata* was 11.7  $\mu\text{g}/\text{mL}$  in HepG2 cells, and 100  $\mu\text{g}/\text{mL}$  or more in HeLa cells and MIA Paca-2 cells. These data clearly demonstrate that frass  $CHCl_3$  ext. derived from *P. trifoliata* induced strong cell death in MIA Paca-2 and HepG2 cell lines more strongly than leaf  $CHCl_3$  ext. ( $p=0.006$  for MIA Paca-2,



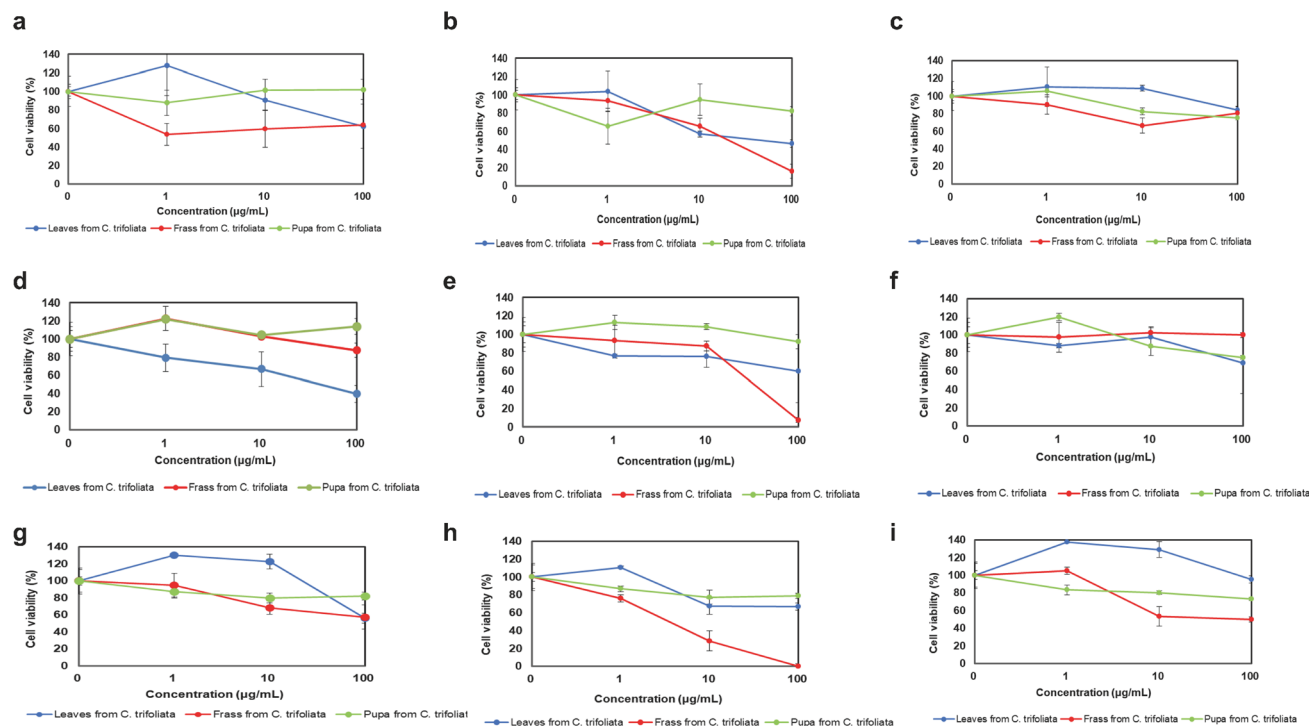
**Figure 3.** Examination of the substances from each extraction separated by TLC. We separated each substance with chloroform:methanol, 20:1 (v/v) for *n*-hexane ext., and 10:1 (v/v) for  $\text{CHCl}_3$  ext., chloroform:methanol:distilled water = 6: 4: 1 (v/v) for MeOH ext. as the developing solvents on TLC. (a) hexane extract from each sample, (b) chloroform extract, (c) methanol extract. 1, frass from *P. trifoliata*; 2, pupa from *P. trifoliata*; 3, leaves from *P. trifoliata*; 4, frass from *Z. piperitum*; 5, leaves from *Z. piperitum*; 6, frass from *C. sudachi*; 7, leaves from *C. sudachi*; 8, frass from *C. natsudaidai*; 9, leaves from *C. natsudaidai*; 10, frass from *C. junos*; 11, leaves from *C. junos*; 12, Aurraptene.



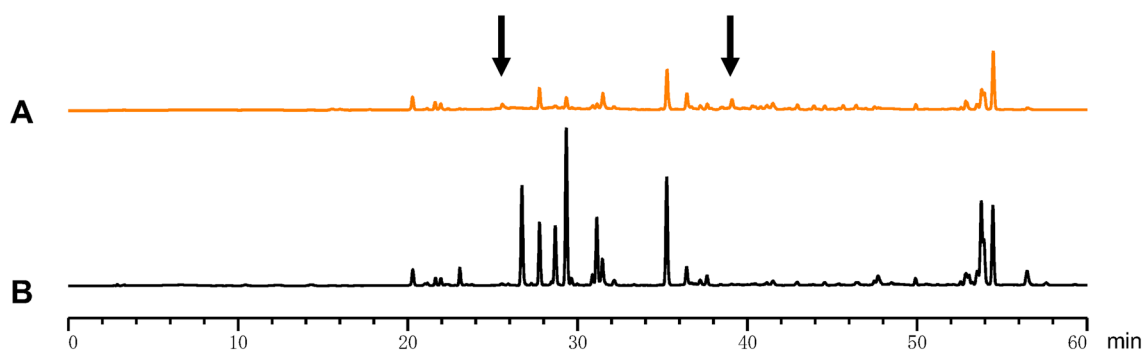
**Figure 4.** Observation of the morphological change to each human cancer cell line after treatment with each extraction. (a) HepG2 cell line; (b) HeLa cell line; (c) MIA-Paca2 cell line.

$p=0.004$  for HepG2). No statistical difference was found in the frass  $\text{CHCl}_3$  ext. derived from *P. trifoliata* in the HeLa cell line ( $p=0.054$ ) compared with that in the leaf  $\text{CHCl}_3$  ext. Furthermore, frass  $\text{CHCl}_3$  extract derived from *C. junos* also induced cell death in HepG2 ( $p=0.0004$ ), HeLa ( $p=0.0004$ ), and MIA Paca-2 ( $p=0.025$ ) cell lines to a greater extent than in the leaf  $\text{CHCl}_3$  extract (Supplementary Fig. 1 and Supplementary Tables 2b and 2c). Frass  $\text{CHCl}_3$  extract derived from *C. sudachi* induced cell death in HepG2 ( $p=0.019$ ), and MIA Paca-2 ( $p=0.0006$ ) cell lines than the leaf  $\text{CHCl}_3$  extract derived from *C. sudachi* (Supplementary Fig. 2 and Supplementary Tables 3a and 3c). Extractions from larval frass which had been fed on different types of plant leaf differed in their ability to induce cell death.

**Comparison of the metabolic products from *P. trifoliata* frass and leaves.** The frass  $\text{CHCl}_3$  ext. derived from *P. trifoliata* induced cell death to human cancer cells more strongly than that from other plants frass  $\text{CHCl}_3$  exts. Inspection the TLC plate chromatogram at long wavelength UV irradiation (365 nm) revealed seven spots with calculated relative to front ( $R_f$ ) values of 0.1 to 0.3 and 0.5 to 0.7 in the frass of *P. trifoliata*  $\text{CHCl}_3$  ext. (Supplementary Fig. 3-b lane 1, red open arrows). We did not find these spots in the leaf of *P. trifoliata*  $\text{CHCl}_3$  ext. (Supplementary Fig. 3-b lane 3). We also detected a green-coloured spot with an  $R_f$  value of 0.4 in *P. trifoliata* frass  $\text{CHCl}_3$  ext. (Supplementary Fig. 3-b lane 1, closed red arrows), and a similar green-coloured spot with an  $R_f$  value of 0.3 in *P. trifoliata*  $\text{CHCl}_3$  ext. (Supplementary Fig. 3-b lane 3, closed red arrows), and there are two distinct spots with an  $R_f$  value of 0.8 to 0.9 in the  $\text{CHCl}_3$  extract derived from leaves of *P. trifoliata* not found in



**Figure 5.** Evaluation of cell viability by MTT assay in the extracts of leaves from *P. trifoliata*. We added the extract to each cell, and then evaluated cell viability by MTT assay. (a) HepG2 treated with hexane extracts, (b) HepG2 treated with chloroform extracts, (c) HepG2 treated with methanol extracts, (d) HeLa treated with hexane extracts, (e) HeLa treated with chloroform extracts, (f) HeLa treated with methanol extracts, (g) MIA-Paca2 treated with hexane extracts, (h) MIA-Paca2 treated with chloroform extracts, (i) MIA-Paca2 treated with methanol extracts.



**Figure 6.** Comparison of substances from leaves with those from frass extract in *P. trifoliata*. We compared  $\text{CHCl}_3$  extracts from leaves of *P. trifoliata* and frass of *P. trifoliata* using HPLC.

the  $\text{CHCl}_3$  extract derived from frass of *P. trifoliata* [Supplementary Fig. 3, both green colored spots are shown with closed black arrows (right side)]. Furthermore, we detected a spot by short wavelength irradiation (254 nm) of ultraviolet irradiation with an  $R_f$  value of 0.6 in *P. trifoliata* leaf  $\text{CHCl}_3$  ext. (Supplementary Fig. 3-b lane 3 open red arrows); however, we did not detect such a spot in *P. trifoliata* frass  $\text{CHCl}_3$  ext.

Taken together, TLC analysis indicates that the composition of the metabolic products in *P. trifoliata* frass  $\text{CHCl}_3$  ext. differed from those in the *P. trifoliata* leaf  $\text{CHCl}_3$  ext.. Comparison of *P. trifoliata* frass  $\text{CHCl}_3$  ext. and *C. trifoliata* leaf  $\text{CHCl}_3$  ext. by high performance liquid chromatography (HPLC) revealed two peaks (Fig. 6a, arrows). These peaks were not observed in *P. trifoliata* leaf  $\text{CHCl}_3$  ext. (Fig. 6b). The peaks of the retention times were 26 min and 39 min (Fig. 6a, arrows).

## Discussion

Here, we constructed “TUATinsecta” by data integration to efficiently find the phytophagous insect species to screen for pharmaceutical resources and to experimentally determine whether the metabolic products from the host plants would be changed by insect metabolic system.

First, using several insect databases<sup>19,20</sup>, we entered the information regarding insect species which fed on plants, thereafter adding the information from the KNAPSAcK database<sup>21</sup> about their host-plant substances and their biological activity; we then created connections among the various information fields in the TUATinsecta database.

We found 2857 insect–host plants relationships with some biological activity on the organisms. We investigated the information on biological activity for each host plant, and discovered reports of anticancer activity for many of them. We also found reports on diabetes, cardiotoxic, antioxidant, anti-inflammatory and antibacterial effects in many host plants based on past records of folk medicines and herbal medicine.

Phytophagous insects take up metabolic products from their host plants during the larval developmental stage. These substances are mainly hydrocarbons, proteins and lipids, which are categorised as primary metabolites<sup>22</sup>. At the same time, phytophagous insects must ingest several types of secondary metabolites—such as alkaloids, terpenoids, flavonoids, polyketides or phenylpropanoids—from their host plants when they feed on them<sup>23</sup>. These secondary metabolites may have affect larval growth, development and eating behaviour<sup>24,25</sup>.

The presence of enterobacteria has been confirmed in *Papilio polytes*, but most are from plants. These bacteria do not seem to be involved in the metabolic process in *Papilio polytes*<sup>26</sup>. Furthermore, the administration of antibiotics to larvae of *Danaus chrysippus* and *Ariadne merione* does not affect their growth and survival<sup>27</sup>. Thus, we speculated that the enterobacteria are not involved in the metabolism of secondary plant metabolites in Lepidopteran insects. As a further support, the pupal extract does not affect the viability of human cancer cells. Accordingly, we considered that the larva must have a specific metabolic system when feeding on host plants for avoiding the effects of these secondary metabolites on their growth, development and eating behaviour. Presumably, host plant secondary metabolites are processed in the larval body through their special metabolic system and appear in an altered form in frass.

Cytochrome P450 (CYP) plays a role for detoxication of these secondary metabolites in phytophagous insects and has been suggested to relate to the host plant specialisation in phytophagous insects<sup>28</sup>. In the detoxication process, specific plant secondary metabolites are metabolised by CYPs, and the final metabolites are then processed to frass. Both the insect body and the larval frass may also contain metabolised substances not included in the original plant substances. For this reason, we anticipated that frass would contain biologically active agents because the secondary metabolites from the host plant could be changed through the insects' metabolic system. From this standpoint, we examined the biological activity of five kinds of frass obtained from the Asian swallowtail *P. xuthus*, which feeds on *P. trifoliata*, *Z. piperitum*, *C. natsudaoidai*, *C. junos* or *C. sudachi*.

In comparing the biological activity of each extract, we found that CHCl<sub>3</sub> ext. *P. trifoliata* frass induced strong cell death and morphological change in the HepG2, HeLa and MIA-Paca2 cell lines; CHCl<sub>3</sub> ext. from *C. junos* frass, however, induced strong cell death and morphological change only in the HeLa cell line. In contrast, CHCl<sub>3</sub> ext. from the frass of *Z. piperitum*, *C. sudachi*, and *C. natsudaoidai* did not induce strong cell death and morphological change to the HepG2, HeLa or MIA-Paca2 cell lines. We speculated plants secondary metabolite from *P. trifoliata* or *C. junos* might be altered the chemical structure by the metabolic system in *P. xuthus*.

These altered substances included in the CHCl<sub>3</sub> fraction could induce cell death. CHCl<sub>3</sub> is a less polar solvent. We thus considered these substances potentially to be lipophilic compounds of low molecular weight.

Plants in the Rutaceae include flavonoids, coumarins, carotenoids and limonoids. Some of these substances have effects on the human cancer cell line (Supplementary Table 4)<sup>29,30</sup>. The coumarin derivative, auraptene found in the *C. natsudaoidai* was reported to show anticancer activity<sup>31</sup>. We were curious as to whether the auraptene was included in the frass or the leaf extract of *P. trifoliata*; HPLC indicated an auraptene peak at 44 min at RT in the Hex ext. of frass and leaf of *P. trifoliata* (Supplementary Fig. 4a and b). However, neither of the Hex exts. induced cell death in the HepG2, HeLa or MIA-Paca2 cell lines (Supplementary Table 2, and Fig. 5a,d,g). Therefore, we hypothesised that a compound other than auraptene was altered by the *P. xuthus* metabolic system to induce cell death. According to TUATinsecta, *P. trifoliata* includes several known biologically active substances such as xanthotoxol and imperatorin. These compounds lead to proliferation in the HeLa cell line. In comparing *P. trifoliata* frass and leaf CHCl<sub>3</sub> ext. by HPLC, we identified two specific peaks at 26 min and 39 min for CHCl<sub>3</sub> ext. from *P. trifoliata* frass using HPLC (Fig. 6). However, we could not identify chemical property on these specific peaks. We predict that a very small amount of the substance or complex of several substances in the frass of *P. trifoliata* can induce cell death.

Some insect frass have a history of use as a folk medicine. In China, San-sha—the dried frass obtained from young silkworm larvae, a known as Chinese herbal medicine—has been used against inflammatory diseases such as rheumatism, joint pain, neuralgia, uterine bleeding, irregular menstruation, and conjunctivitis<sup>32</sup>. In addition, China has a tea called Chongshicha, made using the dried frass obtained from several kinds larvae of Pyralidae or Noctuidae<sup>33</sup>. Malaysia also has a tea made using dried frass obtained from stick insects that feed on guava leaves<sup>34</sup>. Some Malaysians believe this tea is effective against asthma, gastrointestinal disorders, and muscle pain<sup>35</sup>. Thus, Asian countries have a tradition of drinking tea made from frass of phytophagous insects by decoction. Silkworms eat mulberry leaves and stick insects eat guava leaves, and these plants include biologically active substances<sup>36–39</sup>. In this way, insect frass has long been used as folk medicine; however, the plant substances exhibiting pharmacological actions have, to date, not been clearly identified. Thus, if the plant contains secondary metabolic products that exhibit biological activity, it could be changed by insect metabolism into the metabolic products with even stronger biological activity. In this study, we integrated information on insect–host plants, host plant–secondary metabolites, and metabolite–biological function into TUATinsecta. TUATinsecta will assist investigators to find specific compounds, including plant-derived secondary metabolic products, in specific plants. Therefore, TUATinsecta is a helpful resource for choosing phytophagous insect species which could prove useful in identifying pharmaceutical candidates beneficial to humans.

In conclusion, we constructed a database for choosing phytophagous insect species which could convert plant-derived secondary metabolites through the insect metabolic system to metabolic products with biological

activity. With the aid of TUATinsecta we chose candidate insect species and identified *P. xuthus*, which feeds on several species of Rutaceae plants. Using three types of organic solvents, we extracted the metabolic products from the frass produced by the metabolic system of *P. xuthus* after feeding on each Rutaceae plant. We evaluated the biological activity of these extracts by cell death to the human cancer cell line, and found that *P. xuthus* produced a special substances included in the frass when *P. xuthus* fed on *P. trifoliata*. Our results demonstrate that Lepidopteran insects can alter plant metabolic products, creating new substances with different biological activity. However, identification of these biologically active substances must await development of large-scale insect frass collecting methodologies.

In future studies, we will elucidate the biologically active substance included in the frass when *P. xuthus* fed on *P. trifoliata*, and try to convert the biological activity of substances by *P. xuthus* metabolic enzymes.

Furthermore, we intend to construct a technology safe to elaborate to artificially alter the chemical structure of plant-derived secondary metabolic products by insect metabolic enzymes.

Collectively, the results of our study will contribute to the identification of phytophagous insect species that have the ability to change plant substances into potential therapeutic agents via the insect metabolic system.

## Methods

**Construction of the dataset for TUATinsecta.** To determine the relationships among insects, plants and their substances, we developed a database connecting the information among these fields of study. First, we obtained information for host plants of insects by downloading the Encyclopedia of Life database (<https://eol.org/>)<sup>19</sup> for the “eats” attribute as of August 2019. We added further information for host plants to the retrieved data from the database<sup>20</sup>. We also searched field guides and extracted host record information. We then obtained information for substances included in all insects-host plants from the KNAPSAcK database (<https://www.knapsackfamily.com/>)<sup>21</sup>, which contains species-metabolite relations for plants and the biological activities and functions of these metabolites. We also extracted descriptions of biological functions from PubMed (<https://www.ncbi.nlm.nih.gov/pubmed>). Finally, we integrated this information-including insect species, insect-host plants, plant substances, and their biological activity-using a tab-delimited formatted file. We also established hyperlinks to the corresponding web pages in the NCBI Taxonomy to facilitate the access to this information from TUATinsecta using the scientific names of insects and plants. We used TogoDB platform (Database hosting service, <https://togodb.org/>) for the visualization of the tab-delimited file. We then published the dataset as TUATinsecta on the platform. Imported data are converted to an RDF-ized format. TUATinsecta is freely available at [https://togodb.org/db/tuat\\_insecta](https://togodb.org/db/tuat_insecta) under Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0) license (<https://creativecommons.org/licenses/by-sa/4.0/>). All users can search targeted information with specific keywords derived from insect, insect-host plants, substances or biological activity. Users can also download these data in CSV, JSON and RDF format. We will update TUATinsecta when each public database used for the construction of TUATinsecta is updated.

**Insect.** We collected *Papilio xuthus* larvae from Fuchu City, Tokyo, Japan and Tokushima City, Tokushima Prefecture, Japan. A landowner permitted us to pick swallowtail butterfly *Papilio xuthus*. We maintained the Fuchu *P. xuthus* larvae on *Poncirus trifoliata*, *Zanthoxylum piperitum*, *Citrus natsudaidai* and *Citrus junos*. We harvested these plants at Fuchu campus, Tokyo University of Agriculture and Technology, Fuchu city, Tokyo, Japan. We maintained the Tokushima *P. xuthus* larvae on *Citrus sudachi*. We harvested *Citrus sudachi* at Tokushima city, Tokushima Prefecture, Japan. We maintained *P. xuthus* larvae at 25 °C with a 16-h/8-h light/dark cycle, and mated the emerged adult individuals to collect eggs, which we reared under the same conditions from larval to pupal developmental stage after hatching. We collected *P. xuthus* frass during the larval developmental stage in the course of maintenance with each plant. We used *P. xuthus* pupae maintained with *P. trifoliata* during the larval developmental stage. All samples stored at – 30 °C until extraction of substances.

**Extraction of substances from plants and several kinds of *P. xuthus* frass.** We placed each sample on a glass petri dish and freeze-dried it with lyophilizer (VD-250F, TAITECH Co., Ltd., Saitama, Japan) for 24 h. We pulverized each sample and weighed it using a Mettler electronic scale. We transferred the weighed sample to a 300 mL Erlenmeyer flask, and added 3 volumes of *n*-hexane (Hex; Wako Pure Chemical Industries, Ltd., Osaka, Japan). We left the mixture for 72 h at RT and then filtered it. We transferred the pellet to a 300 mL Erlenmeyer flask, added 3 volumes of Hex, left it for 72 h at RT, and then filtered it. We repeated this operation one more time, and transferred the filtrate to an eggplant flask for evaporation.

We transferred the pellet to a 300 mL Erlenmeyer flask, and added 3 volumes of chloroform (CHCl<sub>3</sub>; Wako Pure Chemical Industries, Ltd.) to the pellet. We then conducted the same operation with Hex.

We then transferred the pellet to a 300 mL Erlenmeyer flask and added 3 volumes of methanol (MeOH; Wako Pure Chemical Industries, Ltd.) to the pellet. We then conducted the same operation with Hex.

We concentrated these filtrates, including *n*-hexane, CHCl<sub>3</sub> or MeOH, by rotary evaporator (N-1110N, Tokyo Scientific Instruments Co., Ltd., Tokyo, Japan). Finally, we collected each extract in a screw tube (No. 7, Marem Co., Ltd., Osaka, Japan) and kept them in a vacuum desiccator for 3–5 days to completely remove each organic solvent, and then weighed the obtained extract. We shielded the screw tubes containing the extracts from light with aluminium foil and stored them at 4 °C until use. We referred to the extract with *n*-hexane as Hex ext., the extract with CHCl<sub>3</sub>, as CHCl<sub>3</sub> ext., and the extract with MeOH, as MeOH ext.

**Comparison of components of extracts by thin layer chromatography (TLC).** For each extract, we dissolved samples in appropriate organic solvents and spotted 1–2 µL on a TLC plate (silica gel 60 F 254, 8 cm long) (Merck Millipore, CA, USA) using a glass capillary. We immersed the bottom of the TLC plates in an organic



solvent until the solvent front reached approximately 7 cm from the origin. We chose chloroform:methanol, 20:1 (v/v) for *n*-hexane ext., chloroform:methanol 10:1 (v/v) for CHCl<sub>3</sub> ext. and, chloroform:methanol:distilled water = 6: 4: 1 (v/v) for MeOH ext. as the developing solvents on TLC. For visualising each spot indicating a substance, we excited the spots by ultraviolet light (short wavelength; 254 nm, long wavelength; 365 nm) using a UV irradiator (UVGL-15, Funakoshi Co., Ltd., Tokyo, Japan). Thereafter, we also sprayed the TLC plate with 50% sulfuric acid (Wako Pure Chemical Industries, Ltd., Osaka, Japan) and heat-treated it at 120 °C for 2 min after the extract had separated. We recorded the position of each spot relative to the solvent front (Rf). Rf value was calculated as the distance from baseline traveled by the solute divided by the distance from baseline traveled by the solvent (solvent front).

**Preparation of chemicals and evaluation of cell viability.** We dissolved each extract in dimethyl sulfoxide (DMSO; Sigma-Aldrich, MO, USA) to a concentration of 1 mg/mL, 10 mg/mL, and 100 mg/mL to prepare an extract solution. Each extract solution was protected from light by shading.

We used cisplatin (CDDP; Wako Pure Chemical Industries, Ltd., Osaka, Japan) as a positive control (PC). We dissolved CDDP in phosphate buffered saline (PBS; 8.1 mM NaHPO<sub>4</sub>, 137 mM NaCl, 2.7 mM KCl, 1.47 mM KH<sub>2</sub>PO<sub>4</sub>, pH 7.4), adjusted to 1 mM, and protected it from light by shading. We used an equivalent concentration (v/v) of DMSO as negative controls.

We obtained the human liver cancer cell line HepG2 (RCB1886), the human cervical cancer cell line HeLa (RCB0007), and the human pancreatic cancer cell line MIA Paca-2 (RCB2094) from RIKEN cell bank (RIKEN Tsukuba, Japan). We cultured each cell line in Dulbecco's modified Eagle's medium (Wako Pure Chemical Industries, Osaka, Japan) supplemented with 10% (v/v) fetal bovine serum (Sigma-Aldrich, MO, USA), 100 U/ml penicillin, and 100 µg/ml streptomycin solution (Wako Pure Chemical Industries, Ltd., Osaka, Japan), at 37 °C under 5% CO<sub>2</sub>. We cultured cell lines until sufficient cell stock obtained and then stored cells at -80 °C until use. The second passage number of cells was used in all experiments. To assess the effect of the extract on each cell line, we examined cell viability by MTT assay (Merck Millipore, CA, USA).

We seeded each cell line at  $2 \times 10^3$  cells/50 µL/well on a 96-well plate (Nippon Becton Dickson Co., Ltd., Tokyo, Japan), and then cultured them for 24 h at 37 °C under 5% CO<sub>2</sub>. After 24 h, we treated each well with the extract solution at final concentrations of 1 µg/mL, 10 µg/mL, and 100 µg/mL (0.1% (v/v) DMSO), respectively, and continued the culture for a further 72 h. We used cells treated with DMSO at a concentration of 0.1% (v/v) as a negative control. We added 10 µL of MTT reagent to each well of the 96 well plate after culture for 72 h, and performed the culture for a further 3 h. Thereafter, we added 100 µL of a solution of 0.04 N hydrochloric acid-isopropanol solution to each well to dissolve the formazan dye included in the cells, and then measured the absorbance at 540 nm and 620 nm using a plate reader (iMark, Bio Rad, The absorbance CA, USA). We calculated the cell viability with the following equation:

$$\text{cell viability (\%)} = \frac{([\text{sample absorbance}] - [\text{blank absorbance}])}{([\text{control absorbance}] - [\text{blank absorbance}])} \times 100$$

**Observation of cell morphology.** We observed the cell morphology of the extract on each cell line.

We seeded each cell line at  $0.5 \times 10^5$  cells/500 µL/well on a 24-well plate (IWAKI, Shizuoka, Japan) and cultured them for 24 h at 37 °C. under 5% CO<sub>2</sub>. After 24 h, we treated each well with the extract solution at final concentrations of 1 µg/mL, 10 µg/mL, and 100 µg/mL (0.1% (v/v) DMSO), respectively, and continued the culture for a further 72 h. We used CDDP at 1 mM as a PC. After 72 h, we observed the cell morphology using a stereo microscope (Fluoid Cell Imaging Station, Thermo Fisher). We categorised the cell morphology into five levels and scored it as follows: score 1: the cell proliferation and density are similar to those of the negative control group; score 2: the levels of cell proliferation and density are between those for score 1 and those for score 3; score 3: the cell density has decreased without cell morphological change; score 4: the cell density has decreased to between that for score 3 and that for score 5 by cell death; score 5: the cell density has significantly decreased, similar to that for the PC group, by cell death.

**Comparison of both extracts by high performance liquid chromatography.** We used CHCl<sub>3</sub> extracts from *P. trifoliata* leaf and frass, respectively. We adjusted the concentration of each extract to 10 µg/mL with acetonitrile (Sigma-Aldrich co. ltd., MO, USA) and then filtrated them using a 0.2 mm x φ 4 mm filter (AS ONE co. ltd. Osaka, Japan). We collected each filtrate in a 300 µL vial (Waters co. ltd. Tokyo, Japan).

We loaded these samples to a HPLC system comprising a JASCO PU-2089 (JASCO co. ltd. Tokyo, Japan), a JASCO UV-2075 (JASCO co. ltd. Tokyo, Japan), and a RHEODYNE 7125 (Rheodyne co. ltd, CA, USA) with a CAPCELL PAK C18 MG II column (4.6 φ x 250 mm) (SHISEIDO co. ltd., Tokyo Japan), pre-equilibrated with 20% acetonitrile. We set the wavelength to 220 nm, and the mobile phase was H<sub>2</sub>O: acetonitrile = 80:20 (0–10 min), 80:20 → 0:100 (10–50 min), and 0:100 (50–60 min) at a flow rate of 1.0 mL/min for 60 min. Separation on HPLC was conducted at room temperature.

**Statistical analysis.** Statistical significance was determined by a two-tailed Student's t-test using Excel (Microsoft, Redmond, WA, USA); *p* values of < 0.05 were significant.

### Data availability

TUAT insecta is freely available at [https://togodb.org/db/tuat\\_insecta](https://togodb.org/db/tuat_insecta) under Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0) license (<https://creativecommons.org/licenses/by-sa/4.0/>).

Received: 20 May 2020; Accepted: 5 October 2020

Published online: 15 October 2020

## References

- Newman, D. J. & Cragg, G. M. Natural products as sources of new drugs from 1981 to 2014. *J. Nat. Prod.* **79**, 629–661 (2016).
- Mansukhlal, W., Harold, T., Monroe, W., Philip, C. & Andrew, M. Plant antitumor agents. VI. Isolation and structure of taxol, a novel antileukemic and antitumor agent from *Taxus brevifolia*. *J. Am. Chem. Soc.* **93**, 2325–2327 (1971).
- Irving, J., James, A., Marvin, G. & Paul, B. Jr. The vinca alkaloids: a new class of oncolytic agents. *Can. Res.* **23**, 1390–1427 (1963).
- Monroe, W. *et al.* Plant antitumor agents. I. The isolation and structure of camptothecin, a novel alkaloidal leukemia and tumor inhibitor from *Camptotheca acuminata* 1,2. *J. Am. Chem. Soc.* **88**, 3888–3890 (1966).
- International Union for Conservation of Nature. Red List version 2020-1: numbers of threatened species by major groups of organisms (1996–2020) 2020\_1\_RL\_Stats\_Table\_1a.pdf at <https://www.iucnredlist.org> (2020).
- Nishida, R. & Fukami, H. Ecological adaptation of an Aristolochiaceae-feeding swallowtail butterfly, *Atrophaneura alcinous*, to aristolochic acids. *J. Chem. Ecol.* **15**, 2549–2563 (1989).
- Nishida, R. Chemical ecology of insect–plant interactions: ecological significance of plant secondary metabolites. *Biosci. Biotechnol. Biochem.* **78**, 1–13 (2014).
- Konno, K. *et al.* Papain protects papaya trees from herbivorous insects: role of cysteine proteases in latex. *Plant J.* **37**, 370–378 (2004).
- Ratzka, A., Vogel, H., Kliebenstein, D., Mitchell-Olds, T. & Kroymann, J. Disarming the mustard oil bomb. *PNAS* **99**, 11223–11228 (2002).
- Wittstock, U. *et al.* Successful herbivore attack due to metabolic diversion of a plant chemical defense. *Proc. Natl. Acad. Sci.* **101**, 4859–4864 (2004).
- Konno, K., Okada, S. & Hirayama, C. Selective secretion of free glycine, a neutralizer against a plant defense chemical, in the digestive juice of the privet moth larvae. *J. Insect Physiol.* **47**, 1451–1457 (2001).
- Kazana, E. *et al.* The cabbage aphid: a walking mustard oil bomb. *Proc. R. Soc. B* **274**, 2271–2277 (2007).
- Kumar, P., Pandit, S., Steppuhn, A. & Baldwin, I. Natural history-driven, plant-mediated RNAi-based study reveals CYP6B46's role in a nicotine-mediated antipredator herbivore defense. *Proc. Natl. Acad. Sci.* **111**, 1245–1252 (2014).
- Honnda, K. Chemical basis of differential oviposition by lepidopterous insects. *Arch. Insect Biochem. Physiol.* **30**, 1–23 (1995).
- Mark, S., Jennifer, W. & Rene, F. Metabolic fate of the allelochemical nicotine in the tobacco hornworm *Manduca sexta*. *Insect Biochem. Mol. Biol.* **24**, 837–846 (1994).
- Self, L. S., Guthrie, F. E. & Hodgson, E. Metabolism of nicotine by tobacco-feeding insects. *Nature* **204**, 300–301 (2004).
- Willer, S. J., Jacobsen, N. L. & Conner, W. E. The evolution of chemical defences and mating systems in tiger moths (Lepidoptera: Arctiidae). *Biol. J. Lin. Soc.* **68**, 557–578 (1999).
- Seabrooks, L. & Hu, L. Insects: an underrepresented resource for the discovery of biologically active natural products. *Acta Pharm. Sin. B* **7**, 409–426 (2017).
- Parr, C. S. *et al.* The Encyclopedia of Life v2: providing global access to knowledge about life on Earth. *Biodivers. Data J.* **2**, e1079 (2014).
- Saito, M. U., Jinbo, U., Yago, M., Kurashima, O. & Ito, M. Larval host records of butterflies in Japan. *Ecol. Res.* **31**, 491 (2016).
- Afendi, F. M. *et al.* KNApSACk family databases: integrated metabolite-plant species databases for multifaceted plant research. *Plant Cell Physiol.* **53**(2), e1 (2012).
- Cohen, A. C. Chapter 3, function of insect diet components. In *Insect Diet 2nd Ed. Science and Technology* (ed. Cohen, A. C.) 29–46 (CRC Press, Boca Raton, 2015).
- Nishida, R. Sequestration of defensive substances from plants by lepidoptera. *Annu. Rev. Entomol.* **47**, 57–92 (2002).
- Nishida, R. Sequestration of plant secondary compounds by butterflies and moths. *Chemoecology* **4**, 127–138 (1995).
- Scriber, J. M. Evolution of insect-plant relationships: chemical constraints, coadaptation, and concordance of insect/plant traits. *Entomol. Exp. Appl.* **104**, 217–235 (2002).
- Phalnikar, K., Kunte, K. & Agashe, D. Dietary and developmental shifts in butterfly-associated bacterial communities. *R. Soc. Open Sci.* **5**, 171559 (2018).
- Phalnikar, K., Kunte, K. & Agashe, D. Disrupting butterfly caterpillar microbiomes does not impact their survival and development. *Proc. R. Soc. B* **286**, 20192438 (2019).
- Feyereisen, R. Insect CYP genes and P450 enzymes. In *Insect Molecular Biology and Biochemistry* (ed. Gilbert, L. I.) 236–316 (Elsevier, Amsterdam, 2012).
- Kostova, I. Synthetic and natural coumarins as cytotoxic agents. *Curr. Med. Chem. Anti Cancer Agents* **5**, 29–46 (2005).
- Yano, M. Cancer prevention by citrus. *Nippon Shokuhin Kagaku Kougaku Kaishi* **49**, 139–144 (2002).
- Murakami, A. *et al.* Auraptene, a citrus coumarin, inhibits 12-O-tetradecanoylphorbol-13 acetate-induced tumor promotion in ICR mouse skin, possibly through suppression of superoxide generation in leukocytes. *Jpn. J. Cancer Res.* **88**, 443–452 (1997).
- Namba, T. *The Encyclopedia of Wakan-Yaku (Traditional Sino-Japanese Medicines) with Color Pictures*, Vol. II, 330–332 (Hoiku-sha co. ltd., Osaka, 1994).
- Shen, P.W. In *Insect Dropping Tea in Chinese Tea Business* (ed. Chen, Z.M.) 268 (Shanghai Culture Press, Shanghai, 1992) (in Chinese).
- Seow-Choen, F. Stick and leaf insect (Phasmida: Insecta) Biodiversity in the nature reserves of Singapore. *Proc. Nat. Res. Surv. Semin Gardens' Bull. Singap.* **49**, 297–312 (1997).
- Nadchattram, M. The winged stick insect, *Eurycnema versifasciata* Serville (Phasmida, Phasmatidae), with special reference to its life-history. *Malayan Nat. J.* **17**, 33–40 (1963).
- Ge, Q. *et al.* Analysis of mulberry leaf components in the treatment of diabetes using network pharmacology. *Eur. J. Pharmacol.* **833**, 50–62 (2018).
- Thaipitakwong, T., Numhom, S. & Aramwit, P. Mulberry leaves and their potential effects against cardiometabolic risks: a review of chemical compositions, biological properties and clinical efficacy. *Pharm. Biol.* **56**, 109–118 (2018).
- Luo, Y., Liu, H., Yang, S., Zeng, J. & Wu, Z. Sodium alginate-based green packaging films functionalized by guava leaf extracts and their bioactivities. *Materials (Basel)* **12**, E2923 (2019).
- Feng, X. H., Wang, Z. H., Meng, D. L. & Li, X. Cytotoxic and antioxidant constituents from the leaves of *Psidium guajava*. *Bioorg. Med. Chem. Lett.* **25**, 2193–2198 (2015).

## Acknowledgements

We thank Mr. Satoshi Uchimoto from Tokyo University of Agriculture and Technology for maintaining the host plants which eats swallowtail butterfly. We thank Dr. Ryoichi Sato, and Mr. Shun Shinzeki from Tokyo University of Agriculture and Technology for helpful discussions.

### Author contributions

Conceived and designed the experiments: Y.K., R.J.S, K.K., and H.T. Performed the experiments: W.N., N.K., Mi.N., T.S., Ma.N. Construction of TUATinsecta database: T.N., H.N.,T.S., H.B., Mi.N., and H.T. Contributed reagents/materials/analysis tools: K.K., H.B., I.O., K.I., Y.K., and H.T., Analyze data: H.N., W.N., T.N., H.B., Y.K., K.K., R.J.S and H.T. Contributed to the writing of the paper under draft version: H.N., W.N.,T.N., R.J.S., and H.T. All authors discussed data and helped with manuscript preparation. H.T. supervised the project. All authors read and approved the final manuscript.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41598-020-74590-z>.

**Correspondence** and requests for materials should be addressed to H.T.

**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) 2020